The Hadronic Final State in Deep Inelastic Scattering at HERA

DISSERTATION

zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

Vorgelegt von

Peter Lanius

aus Berlin

Hamburg
1994

Gutachter der Dissertation:	Prof. Dr. B. Naroska Prof. Dr. G. Buschhorn
Gutachter der Disputation:	Prof. Dr. Prof. Dr.
Datum der Disputation:	x. September 1994
Sprecher des Fachbereichs Physik und Vorsitzender des Promotionsausschusses:	Prof. Dr. E. Lohrmann

Don't try technological fixes
for sociological problems

– Bjarne Stroustrup
(designer of C++*)*

Abstract

Global properties of the hadronic final state of deep inelastic scattering events recorded in 1992 with the H1 detector at HERA, are investigated. The data are corrected for detector effects and can be compared directly with QCD phenomenology and calculations based on BFKL dynamics. The measurement of the energy flows in the laboratory frame and in the hadronic centre of mass system reveal large discrepancies between the data and the different model predictions, indicating the failure of models based on Altarelli-Parisi evolution at low x. However, these energy flow results are found to agree fairly well with theoretical predictions derived from Lipatov (BFKL) evolution. In the hadronic centre of mass frame the longitudinal and transverse momentum components of charged particles are measured. The longitudinal component exhibits scaling with W and allows comparison with lower energy lepton-nucleon scattering data as well as with e^+e^- data from LEP. For the 1993 data, studies of the charged particle energy spectra in the Breit frame are undertaken. This measurement allows a first tentative look at predictions from the Modified Leading Logarithmic Approximation for the target region, a region that to-date unexplored has been unexplored.

Kurzfassung

In der vorliegenden Arbeit werden globale Eigenschaften des hadronischen Endzustands in der tief-unelastischen Elektron-Proton Streuung bei HERA untersucht. Dabei werden Daten der Runperiode von 1992 des H1 Experiments zugrunde gelegt. Nach der Korrektur von Detektoreffekten können die Daten direkt mit phänomenologischen QCD Modellen bzw. mit theoretischen Rechnungen, basierend auf der BFKL Dynamik bei kleinen x, verglichen werden. Die gemessenen Energieflüsse im Laborsystem und im hadronischen Schwerpunktssystem zeigen große Differenzen zwischen den Daten und Modellen, die auf Altarelli-Parisi Evolution basieren, auf. Andererseits stimmen die theoretischen Vorhersagen der Lipatov (BFKL) Evolution relativ gut mit den Daten überein. Die longitudinalen und transversalen Impulskomponenten geladener Teilchen wurden im hadronischen Schwerpunktsystem untersucht. Dabei zeigt die longitudinale Komponente Skalenverhalten mit W und die Messungen können mit tief-unelastischen Experimenten bei niedrigerem W sowie mit e^+e^- Daten von LEP verglichen werden. Für die Daten von 1993 wurden Spektren geladener Hadronen im Breit-System untersucht. Diese Messungen erlauben einen ersten, vorläufigen Blick in die bislang unerforschte Targetregion und einen Vergleich mit theoretischen Vorhersagen, basierend auf der MLLA (Modified Leading Logarithmic Approximation).

to Lynn

Acknowledgements

Firstly, I would like to thank my wonderful wife, Lynn, for all her love and support, for her invaluable help in overcoming the troughs. It has been the greatest miracle to meet her in a place where you probably least expect to meet such a wonderful person. I am grateful to her for her outstanding help in discussing, writing and proofreading this thesis.

I would like to offer special thanks to Michael Kuhlen, for his supervison, guidance and advise. It would have been much more difficult to complete my thesis and impossible to write a publication on the subject without his close collaboration and his patience.

I would like to thank Guenter Grindhammer for giving me the opportunity to carry out this PhD and for introducing me to the wide field of Monte Carlo generators.

I have enjoyed the friendly companionship of my fellow students, Konrad, Christian, Robert and Thomas, particularly for our lunchtime laughter and its disruptive effect on the refectory.

Special thanks also to the people I have shared an office with, to Stefan and Matthias for opening my eyes and to Arno, Lionel, Fabian and Jobst.

Finally, I am grateful to Gunnar Ingelman, Mike Seymour, Jiri Chyla and Leif Loennblad for many useful conversations and their help as I struggled through their Monte Carlos.

Contents

Chapter 1

Introduction

This means either that the universe is more full
of wonders than we can hope to understand or, more probably,
that scientists make things up as they go along.

– Terry Pratchett

Deep inelastic scattering (DIS) experiments have always been a fruitful testing ground for the investigation of the theory of strong interactions, QCD. In fact, Quantum Chromo Dynamics owes a lot to DIS experiments, as they provided the first dynamical evidence for the existence of quarks. Also, the quark parton model (QPM) was initiated and developed by studying deep inelastic scattering data.

A major step in the history of DIS was the measurement of the proton structure functions and then the observation of scaling violations in the structure function evolution. Such investigations enable the strong coupling constant, α_s, the fundamental parameter of QCD, to be measured.

With the advent of high energy fixed target experiments, such as EMC and E665, investigations of the hadronic final state in DIS became more popular. Studies of QCD effects, such as jets or energy flows, were still difficult since the hadronic mass W of these experiments was rather low; and so the measured quantities were dominated by non-perturbative effects, such as the intrinsic transverse momentum of the partons in the proton and hadronisation.

HERA is the first electron proton collider, with a centre of mass energy of currently 296 GeV. This opens up a completely new kinematic domain for DIS studies, namely the low x, high W region. Compared to the EMC measurements, at $<$W$>$ = 14 GeV, the available W is up to 20 times larger, resulting in much more QCD radiation of gluons and quarks. In addition to higher QCD activity, the low x ($x < 10^{-3}$) domain is interesting; at low x a new dynamic behaviour of the proton's partons can be anticipated because $\ln \frac{1}{x}$ terms in the QCD evolution become large and need to be resummed [3].

The QCD evolution of structure functions is usually based on the Altarelli-Parisi (or DGLAP) equations [1]. In this scheme a set of starting distributions at a given Q_0^2

is evolved up in Q^2. The Altarelli-Parisi equations resum the leading order $\alpha_s \ln Q^2$ contributions. At small x this picture is no longer valid, $\ln \frac{1}{x}$ contributions become large and need to be taken into account in the resummation. An evolution equation that is valid in the small x, moderate Q^2 ($Q^2 \gtrsim Q_0^2$) region, is the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation [2].

Many of the investigations at HERA are devoted to observing the first signs of this anticipated new behaviour at low x. However, the inclusive nature of the structure functions and the freedom in choosing the (non-perturbative) input distributions, allow Altarelli-Parisi evolution to mimic the steep rise of the gluon distribution at low x prediticted by the BFKL dynamics. Due to this, it is very difficult to discriminate Altarelli-Parisi and BFKL evolution from the Q^2 dependence of the structure functions at fixed x [3].

More discriminative power can only be achieved by studying the hadronic final state in deep inelastic scattering events. One possibility is to look for jets close to the proton remnant direction, a process for which the cross section predicted by BFKL is much larger than for Altarelli-Parisi evolution, if appropriate cuts are made [4]. The reason for this is that BFKL evolution produces much higher transverse energies in the region between the proton remnant and the current jet. This discriminative feature can also be exploited by measuring transverse energy flows, the main topic of this thesis.

In e^+e^- interactions the initial state does not contain strongly interacting particles, only leptons. This is different in ep collisions, where there is a proton in the initial state. The partons in the proton can give rise to so-called initial state QCD radiation, and studies at HERA will therefore concentrate on possible new effects in the low x range and on initial state QCD radiation.

Another reason to investigate QCD effects, in terms of energy flows and particle spectra, is hadronisation and its relation to perturbative QCD. Although the theory has only one free parameter, α_s, perturbation theory cannot explain the transition of partons into hadrons. For this low momentum transfer process, α_s is large and perturbative QCD is not applicable, so phenomenological models have to take over. Hence, a knowledge of α_s alone is not sufficient to interpret the measured results in terms of QCD. Comprehensive estimates of the influence of the size of non-perturbative effects, by the EMC experiment, revealed that their measurements of transverse energies were dominated by non-perturbative contributions. The perturbative part of QCD has been extensively tested, thanks to the high precision measurements of α_s and related quantities at LEP and previous e^+e^- experiments. However, for a full exploitation of the strong predictive power of the perturbative approach, it is essential to understand which features of the perturbative partonic evolution are preserved through the non-perturbative hadronisation process and which are not.

In this thesis energy flows and charged particle spectra in deep inelastic scattering at HERA are investigated, with the aim of looking for new QCD effects in the low

x region. Monte Carlo generators, which model QCD in different approximations, form the basis of the comparison to data. An analytical calculation of the transverse energy flow is now available and will also be compared to the data. The analysis is carried out using the 22.5 nb^{-1} of DIS events collected by H1 during the 1992 running period. The 1993 data, with their much higher statistics, are still not sufficiently understood to attempt a correction for detector effects. Also, there are many systematics studies required for the 1993 data which are ongoing.

However, the 1993 data have been employed for a first look at charged particle spectra in the Breit frame, with the aim of investigating QCD coherence effects and the validity of Local Parton-Hadron Duality. Such studies could not be carried out with the 1992 data as the low statistics does not allow for the fine binning in x and Q^2, which is crucial for this purpose. Also, the forward tracker was not available for the analysis of the 1992 data, so, during this period, the angular coverage for charged tracks studies does not extend sufficiently into the forward region.

The organisation of this thesis is as follows: In the second chapter a brief description of the H1 detector is given. After presenting QCD models for the hadronic final state in chapter 3, the Monte Carlo generators used for comparisons with data are introduced. In chapter 5 the event selection procedure is outlined, along with the reconstruction of the kinematic variables. Following a discussion of the correction for detector effects, the results on energy flows are presented in chapter 6. Charged particle spectra in different Lorentz frames are investigated, before the final chapter summarises the results.

One final remark concerning the prospects of high energy physics may (or may not) be appropriate here, though it would be pointless to attempt to discuss this important subject in detail within a PhD thesis. The current developments in accelerators, the exclusive hunt for higher and higher energies without accompanying breakthroughs in accelerator technology, appear to be single-minded to me. The turn down of the SSC, though highly controversial, it may, I think, offer a good opportunity to take a break, re-think issues and may be we will not have the flash of inspiration coming too late, as nicely illustrated by Bill Watterson in fig. 1.1...

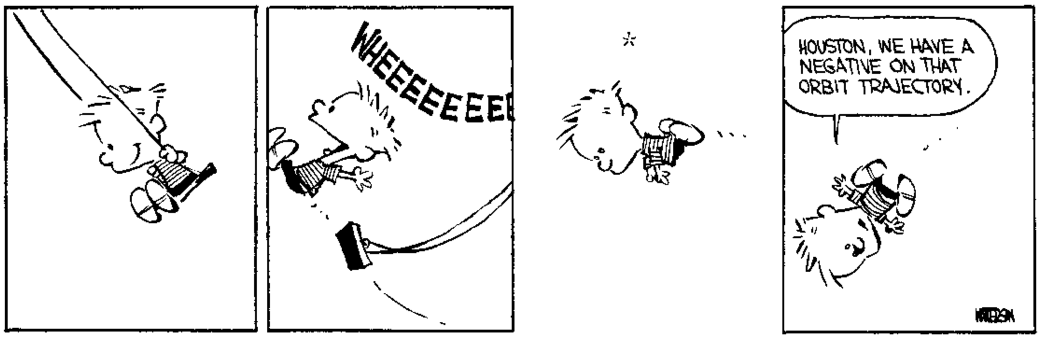


Figure 1.1: *Are we flying too high?*

Chapter 2

HERA and the H1 Detector

Die so enstandenen Säuger gelangten durch
Lahmen und Hinken zum funktionierenden Gang
und dadurch, daß sie sich gegenseitig das Leben
schwer machten, zur Vernunft.

– Stanisław Lem

2.1 HERA

HERA ('Hadron Elektron Ring Anlage') is an electron proton collider facility situated at DESY in Hamburg, Germany. The design specifications for this 6.3 km long accelerator enable 30 GeV electrons to be collided on 820 GeV protons. However, during the 1992 and 1993 running periods the actual electron energy was 26.7 GeV. Colliding 84 electron bunches on 84 proton bunches (1993), HERA was able to deliver an integrated luminosity of ≈ 1 pb^{-1}.

One of the many important experimental conditions is the average bunch length of the proton bunches, in 1993 this was ≈ 25 cm. This factor defines the length of the interaction zone in the detector, illustrating that the nominal collision point is at the center of a smeared vertex distribution.

2.2 Detector Description

A detailed discussion of the H1 apparatus can be found elsewhere [5]. Here, the components of the detector relevant to the present analysis are briefly described; the calorimeters and the central and forward tracking systems. The main components of H1 are illustrated in fig. 2.1.

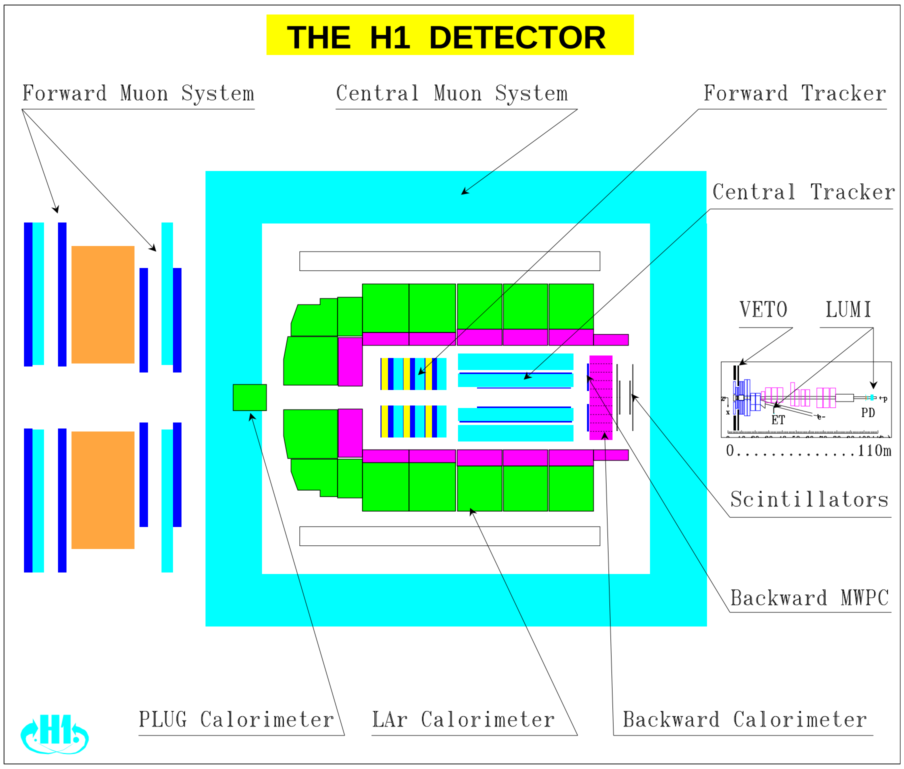


Figure 2.1: *Vertical cross section (through the beam pipe) of the H1 detector.*

2.2.1 The calorimeters

The hadronic energy flow and the scattered electron are measured in the liquid argon (LAr) calorimeter and the backward electromagnetic calorimeter (BEMC). Leaking hadronic showers are measured in the surrounding tail catcher (TC), which is incorporated into the return yoke of the magnet.

The liquid argon (LAr) calorimeter [6] extends over a polar angle range of $4° < \theta < 153°$ with full azimuthal coverage. The polar angle, θ, is defined with respect to the proton beam direction ($+z$ axis). The calorimeter consists of an electromagnetic section with lead absorbers and a hadronic section with steel absorbers. Both calorimeters are highly segmented in both the transverse and longitudinal direction. The total depth of the electromagnetic part varies between 20 and 30 radiation lengths, while the total depth of both calorimeters varies between 4.5 and 8 hadronic interaction lengths. The electronic noise per channel is typically between 10 and 30 MeV, corresponding to 1σ equivalent energy.

Test beam measurements of the LAr calorimeter modules give an energy resolution of $\sigma_E/E \approx 0.12/\sqrt{E\ [\ \mathrm{GeV}]} \oplus 0.01$ for electrons and $\sigma_E/E \approx 0.50/\sqrt{E\ [\ \mathrm{GeV}]} \oplus 0.02$ for charged pions [5, 6, 7]. The hadronic energy measurement is performed by applying a weighting technique [8] in order to account for the non-compensating behaviour of the calorimeter. Both the energy scales and resolutions have been verified in the H1 experiment. The electromagnetic energy scale is known to a precision of 3%, a result obtained by comparing the measured track momentum of electrons and positrons with their corresponding energy deposition in the LAr calorimeter [9]. The overall calibration of the hadronic energy is presently known to 6%, as determined from studies of the transverse momentum ($p_\perp$) balance in DIS events.

Inbetween the trackers and the LAr calorimeter, electronics, cables and support structures (as the LAr cryostat) give rise to so-called dead material, i.e. material in which hadrons might shower, but which is not equipped with active detector elements. Other sources of dead material are the beam pipe and the beam collimator inside the beam pipe in the region of the forward tracker.

The lead-scintillator calorimeter BEMC, with a depth of 22.5 radiation lengths or 1 interaction length, covers the backward region of the detector, in the polar angular region $151° < \theta < 177°$. A major task of the BEMC is to trigger and to measure scattered electrons in DIS events having a Q^2 value ranging from 5 to 100 GeV^2. The absolute energy scale is determined by comparing the measured electron energy in the BEMC with the one calculated from the angles of the hadronic system and the electron. The BEMC energy scale for electrons is known to an accuracy of 2%. and it has a resolution $\sigma_E/E = 0.10/\sqrt{E\text{ [GeV]}} \oplus 0.42/E\text{[GeV]} \oplus 0.03$ [10].

The calorimeters are surrounded by a superconducting solenoid providing a uniform magnetic field of 1.15 T parallel to the beam axis over the entire tracking region. The return yoke surrounding the coil is fully instrumented with a tail catcher (TC) consisting of streamer tubes, with analog read-out to measure the leakage of hadronic showers. The TC resolution is $\sigma_E/E \approx 1.0/\sqrt{E\text{ [GeV]}}$.

2.2.2 Tracking chambers

H1 has two tracking systems, a forward and a central tracker, shown in a vertical cross section in fig. 2.2. Both devices measure charged particle tracks, and provide particle identification via transition radiation detection in the forward tracker and dE/dx-measurements in the central tracker.

Tracking in the forward region, $5° < \theta < 30°$, is provided by the forward tracking detector (FTD). The FTD consists of 3 nearly identical supermodules, each containing a planar drift chamber, a MWPC, a transition radiation detector and a radial drift chamber. The planar modules provide homogeneous spatial resolution in the plane perpendicular to the beam direction, while the radials serve for precise momentum measurement.

Charged tracks in the central region are measured, in the main, in two concentric jet drift chamber modules (CJC), covering the polar angle range $15° < \theta < 165°$. The inner ring consists of 30 segments with 24 sense wires each, while the outer ring

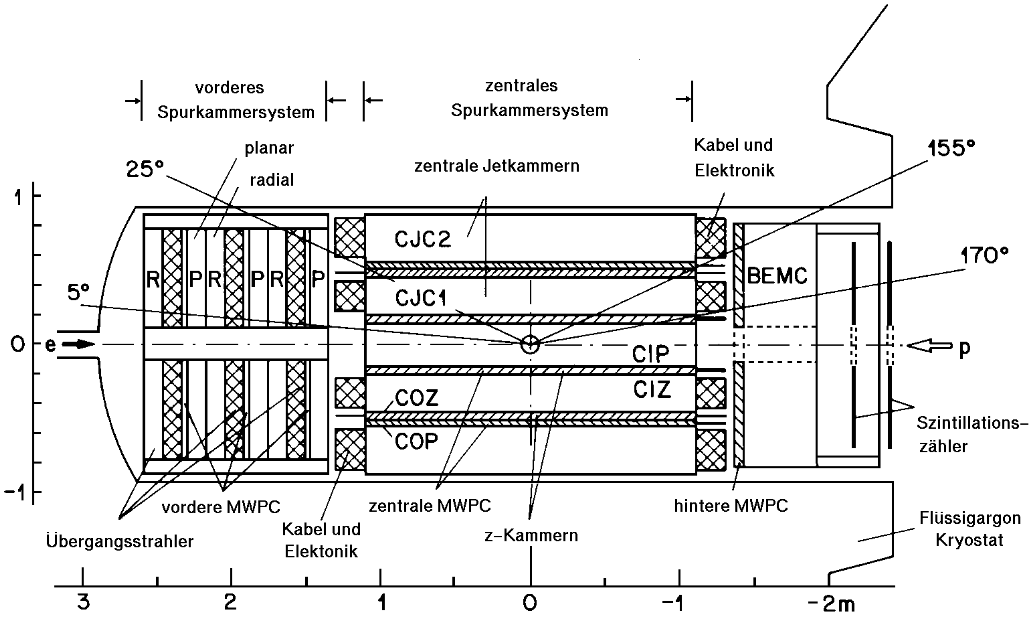


Figure 2.2: *Side view of the H1 tracking system. The separation into a forward tracker and the central tracker is clearly seen.*

is divided into 60 segments equipped with 32 sense wires. Through the magnetic field, only tracks with $\gtrsim$ 0.15 GeV transverse momentum will pass through both chambers. The resolutions achieved are $\sigma_{p_\perp}/p_\perp \approx 0.0086 \cdot p_\perp$ [GeV] + 0.015 and $\sigma_\theta = 21$ mrad.

A backward proportional chamber (BPC), located in front of the BEMC, has an angular acceptance of $155.5° < \theta < 174.5°$ and serves to identify electrons and to precisely measure their direction. Using information from the BPC, the BEMC and the reconstructed event vertex, the polar angle of the scattered electron is known to within 5 mrad.

A scintillator hodoscope (ToF) is situated behind the BEMC. It is used to veto proton-induced background events, based on their early time of arrival compared with the desired electron-proton collisions.

Chapter 3

QCD Models of the Hadronic Final State

The telephone operator ℗, when acting on the vacuum, produces a dialatone state.

– V. Gates et al.[11]

In this chapter, different models for the description of hadronic final states are introduced. They are based on perturbative QCD, leaving the domain of soft processes, such as hadronisation, to phenomenological models which are described in the last section. To ensure that perturbative methods are applicable for the analysis, a sufficiently high momentum transfer from the electron to the proton is required. Despite of this, the running coupling constant, α_s, is still not a small number; this means that higher orders in the perturbation series expansion need to be taken into account or resummation techniques are required to obtain approximations to all orders.

3.1 The Quark Parton Model

3.1.1 Leptoproduction Kinematics

Deep inelastic scattering involves a high momentum transfer (> 1 GeV^2) from the electron to the proton. In neutral current (NC) interactions this momentum is carried by the exchanged γ or Z^0 boson. The basic NC leptoproduction scattering

$$e^- + p \rightarrow e^- + X$$

is schematically shown in fig. 3.1. The notation introduced will be used throughout:

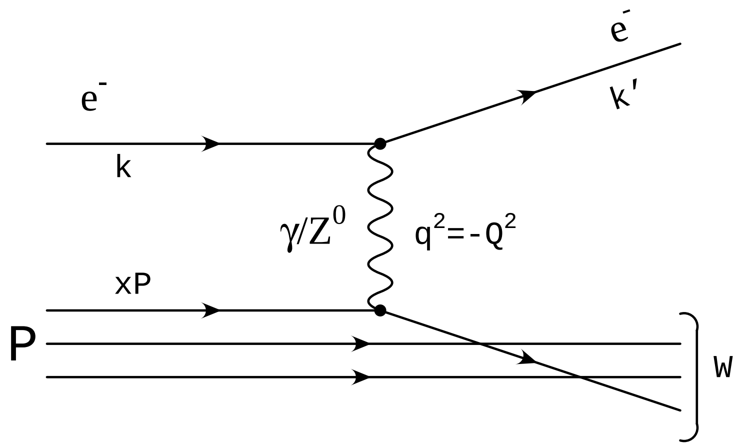


Figure 3.1: *The basic scattering diagram for DIS in the approximation of the QPM.*

k, k'	4-vector of the incoming and scattered electron
P	4-vector of the incoming proton
$q = k - k'$	4-vector of the virtual boson
E, E'	energy of the incoming and scattered electrons
E_p	energy of the incoming proton
θ_e	electron scattering angle
W	mass of the hadronic system X

These quantities are defined with respect to the H1 coordinate system, here the proton goes along the positive z-axis in the laboratory system.

Neglecting the lepton mass, the usual kinematical variables for leptoproduction are defined as:

$$s = (k+P)^2 = \frac{Q^2}{xy} + M_P^2 = 4EE_p \tag{3.1}$$

$$Q^2 = -q^2 = -(k-k')^2 = 4EE' \sin^2 \frac{\theta_e}{2} \tag{3.2}$$

$$x = \frac{Q^2}{q \cdot P} \tag{3.3}$$

$$W^2 = (P+q)^2 \approx Q^2 \frac{1-x}{x} \tag{3.4}$$

$$y = \frac{q \cdot P}{k \cdot P} \approx \frac{Q^2}{s \cdot x} \tag{3.5}$$

The $\approx$ sign in the last two equations indicates that the proton mass has been neglected too, a very good approximation at HERA, since the hadronic masses W^2 involved are much bigger than the proton mass. Since the beam energies of the proton and the electron are known, only two of the kinematic variables are independent.

3.1.2 The Quark Parton Model Approximations

The quark parton model (QPM) provides the simplest description of the partonic final state in DIS. Although various approximations limit the models scope of applicability considerably, it is very useful for studying basic features of the scattering process. Deep inelastic scattering in the QPM is viewed as the scattering of a single valence or a sea quark in the proton, which carries a fraction x of the proton momentum P, neglecting QCD radiation. It corresponds to the $0th$ order in the expansion in α_s in QCD. In this approximation the partonic final state contains two coloured objects, namely the scattered quark and the proton remnant. The proton remnant is a di-quark in the simplest case and, as a spectator, it does not take part in the interaction.

The model rests upon two physical concepts, the Lorentz contraction and time dilation of the parton states in the proton and the long time nature of hadronisation. Due to the time dilation of the fast moving proton the initial state interactions between partons in the proton happen too early to influence the scattering and hadronisation happens too late. The time of the scattering itself is shortened due to the Lorentz contraction of the proton, which transforms the proton into a disc in the plane transverse to the direction of motion. Hence the time scales involved are very different and the scattering can be viewed as scattering off an individual and free parton, with the cross section being proportional to the densities of partons in the proton.

The consequences of this model are most obvious in the rest system of the proton and the boson, the so-called hadronic centre of mass system (CMS). In this frame the struck quark and the proton remnant fly apart, in opposite directions along the z-axis, after the scattering; both particles have the energy $\frac{W}{2}$. The hemisphere of the scattered quark ($z > 0$) is called the current region, the $z < 0$ direction is referred to as the target region. This subdivision of phase space is strictly only correct in the approximation of the QPM. When including QCD radiation, hadronisation and other non-perturbative effects, such as intrinsic $p_\perp$ of partons in the proton, this simple picture is destroyed. Fragmentation will produce a continuous chain of hadrons, thus there is no clean separation between what is still classified as current and target region.

3.1.3 Parton Density Functions and Factorisation

The differential cross section for NC leptoproduction in terms of the kinematical variables Q^2 and x is given by:

$$\frac{d^2\sigma}{dx dQ^2} = \frac{2\pi\alpha^2}{Q^4 x}\left(2\ (1-y)\ +\ \frac{y^2}{1+R}\right) F_2(x,Q^2), \tag{3.6}$$

This is the Born term of the cross section, i.e. excluding QED radiative corrections and is only valid for $Q^2 << M_Z^2$. Here, scattering off a proton is expressed in terms of the structure function F_2 and the photoabsorbtion cross section ratio of longitudinally and transversely polarised photons, $R = \sigma_L/\sigma_T$. Perturbative QCD predicts the magnitude of R, as it is a result of gluon radiation. Again, in the language of the QPM, where $R = 0$, the structure function F_2 can be expressed by the underlying parton density functions, $\phi(x, Q^2)$:

$$F_2(x, Q^2) = x \sum_f Q_f^2 \phi_f(x, Q^2) \tag{3.7}$$

where the sum goes over all flavours f and eQ_f is the electric charge of the quark of the flavour f. The parton density functions give the probabilities to find a valence or sea quark of the flavour f, with the momentum fraction x, in the proton.

Factorisation is a generalisation of the parton model ideas. Basically, in perturbative QCD, one assumes that the cross section can be divided into a hard, perturbatively calculable, part which is associated with short distance phenomena and a soft, long distance part, that is non-perturbative. Thus the differential cross section in higher orders of QCD is a convolution of the hard subprocess cross section with the parton density functions. Utilizing factorisation means that one can absorb part of the singularities that occur in QCD calculations into the parton density functions. Similarly, the non-perturbative hadronisation step can be factorised from the hard subprocess.

All Monte Carlo generators used to generate hadronic final states subdivide the process of event generation into the two stages mentioned above. The parton density functions are assumed to be independent of the hard subprocess and are given in the form of parametrisations obtained from fits to DIS scattering data. Hadronisation models are assumed to be, at least asymptotically, independent of the hard subprocess.

3.2 Matrix Elements

Within the framework of perturbation theory, matrix elements calculated in fixed order in α_s offer an exact description of parton emissions. Although matrix elements for DIS up to second order have been calculated recently [12], to date only the first order matrix elements have been incorporated into the Monte Carlo generators. In this analysis the data are corrected for detector effects, so the measurements correspond to measuring hadrons. Hence the experimental distributions can only be compared with generations of the full hadronic final state. Since this is only possible using Monte Carlo generators, only the first order matrix elements will be discussed here. In first order QCD two partons are generated in the final state. The basic processes are QCD Compton scattering (QCDC), fig. 3.2 and boson-gluon-fusion (BGF), fig. 3.3.

One problem arises as the matrix elements are divergent in the region of soft and collinear emissions. These singularities are partly canceled by higher order virtual

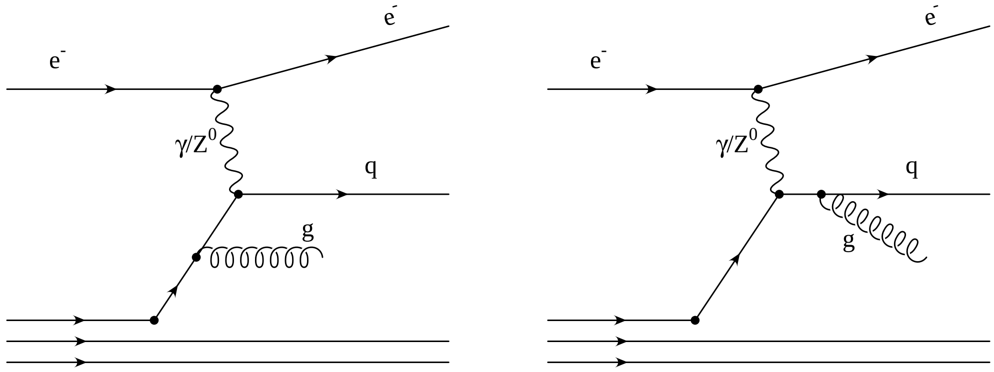


Figure 3.2: *The Feynman diagrams for the QCD Compton process*

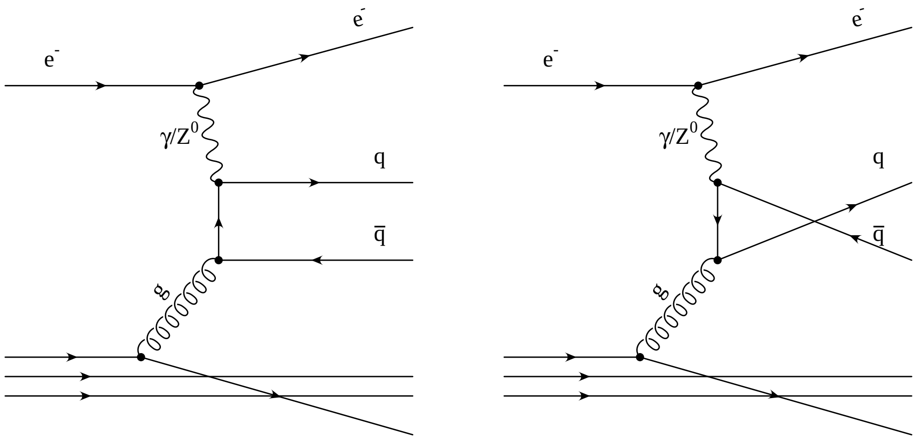


Figure 3.3: *Feynman diagrams for the BGF process*

corrections and partly absorbed into the parton density functions. However, for Monte Carlo generators it is necessary to impose a cut-off, y_{cut}, to avoid the divergence region. This can be done by requiring that the parton pair produced has a minimum invariant mass, $m_{ij} > m_{ij_{min}}$. An alternative is to scale the invariant mass m_{ij} with W^2, i.e. using $\frac{m_{ij}{}^2}{W^2} > y_{cut}$. The latter procedure follows more the shape of the divergence region over the full W^2 range. For practical Monte Carlo simulations it is preferable to keep the cut-off as low as feasible to retain as large a region of phase space as possible.

3.3 Leading Logarithmic Parton Showers

In DIS the struck quark can emit gluons both before and after the boson vertex. In the parton shower approach this gives rise to initial and final state parton showers, as illustrated in fig. 3.4. For the initial cascade a parton typically close to mass-shell in the incoming proton is the starting point. If a gluon is radiated off this parton, it becomes (further) off-shell. So after each branching the parton becomes increasingly off-shell with a spacelike virtuality ($m^2 < 0$). In the basic quark scattering diagram the resulting spacelike quark is then struck by the boson and turned into a quark that is either on-shell or has a timelike virtuality ($m^2 > 0$). This quark can radiate further according to its off-shellness and gives rise to the final state parton shower. In contrast to the initial cascade, in the final state shower the virtuality of the quark decreases after each branching.

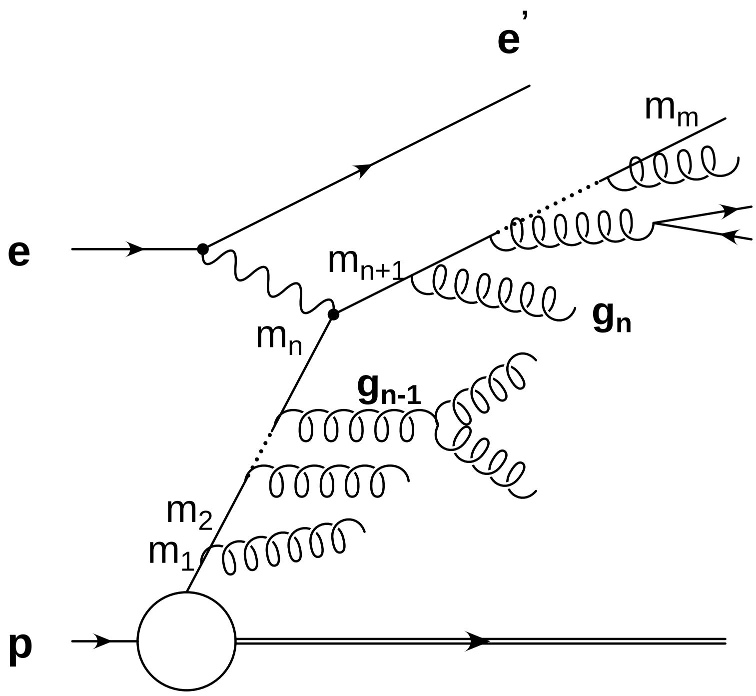


Figure 3.4: *The parton shower picture for DIS. Gluon emission of the scattered quark before the boson vertex is referred to as initial state radiation. After the boson vertex the struck quark gives rise to the final state cascade.*

In both cases a cut-off on the virtuality of the order of a hadron's mass terminates the branching process. This infrared cut-off is not the same as the cut-off used to avoid divergences of the matrix elements, which is used mainly to avoid collinear singularities. For the final state cascade the parton shower cut-off can be rather low ($\lesssim 1$ GeV^2). Since the parton density functions are not very well known for

$Q^2 < 4$ GeV^2 and the initial cascade has to be matched with the parton density functions, the cut-off in this case tends to be higher (1-4 GeV^2).

The parton shower picture is an approximation for higher order QCD processes, in fact to all orders, but there are various possible approaches. Usually the parton showers are implemented in the leading logarithmic approximation, with the Altarelli-Parisi equations [1] and the Sudakov form factor [13] governing the evolution. In terms of QCD this represents a resummation of the leading collinear divergences. The Altarelli-Parisi equations comprise the contribution of the real diagrams, whereas the Sudakov form factor sums the virtual graphs and unresolvable real contributions, so reducing the amount of emission. The distribution of the energy-momentum fractions in each branching is given by the Altarelli-Parisi splitting kernels and the distribution of parton virtualities is controlled by the Sudakov form factor.

The QCD approximation, as provided by parton showers, can be improved by implementing next-to-leading order evolution equations. In an attempt to take at least part of the higher orders (namely the large logarithms due to infrared divergences) into account, one can impose QCD coherence and allow for a running α_s. QCD coherence corresponds to infrared singularities and leads to the reduction of soft gluon emission, in certain angular regions, due to destructive interference. The Monte Carlo generators used for this analysis all utilise leading orderAltarelli-Parisi splitting. However, the methods to modify it in implementing QCD coherence and a running α_s, are quite different.

The use of the Altarelli-Parisi splitting kernels, instead of the true matrix elements for the quark or gluon branchings, determines the sort of approximation used in the parton shower picture. The basic $q \rightarrow q + g$ splitting, in terms of Altarelli-Parisi evolution, corresponds to the collinear limit of the first order QCD matrix elements for this process. Extending the validity of the branching formulas, beyond collinear emission, to the full phase space is an assumption of the parton shower model. This shows that a parton shower will nicely describe soft and collinear branchings, but is not well suited for hard, large angle gluon emission. For these reasons it is advantageous to build Monte Carlo generators with matching between exact QCD matrix elements (up to some fixed order in α_s) and parton showers. In such matched generators the first emission(s) in the parton shower is adjusted to reproduce the matrix elements result.

To fully implement QCD coherence into a parton shower Monte Carlo three different factors must be accounted for:

1. Coherence in the final state cascade; reducing the phase space for gluon emissions by enforcing strong angular ordering of their respective emission angles will properly take into account leading infrared divergences [14]. Gluons emitted at a later stage (further away from the boson vertex) in the final state parton shower are forced to have smaller emission angles, as indicated in fig. 3.4.

2. Coherence in the initial parton shower can be obeyed by ordering the emissions in the variable $E\theta$ (in the limit of small emission angles) [15]. Here E is the energy of the emitting quark before the branching has occured and θ is the angle between the quark and the emitted gluon.

3. Interference between the initial and final state parton shower can be taken into account by imposing angular cuts on the emission angle of the first gluon in the final parton shower (g_n in fig. 3.4) and the last emission of the initial cascade (g_{n-1} in fig. 3.4). Making an angular cut corresponds to a limit on the maximum of the evolution variable. Such a limit can be derived from the structure of the colour flow between the partons of the hard subprocess in the initial and final state [15].

If a parton shower model obeys all these three colour coherence requirements, then the parton showering process is a subdivision of the simplified gauge-invariant matrix elements into two stages, the initial and final state parton shower. With the initial state singularities collected in the first stage and the final state ones in the last stage, the sum of the two will still include interference. However, interference between the initial and final state relies on the correct treatment of soft emissions, so QCD coherence must be fully obeyed.

For the initial state parton shower the evolution must not only be driven by the Altarelli-Parisi equations and the Sudakov form factor, the parton density functions of the proton have also to be taken into account. To achieve the necessary matching between the event kinematics (which is generated first, according to the differential cross section) and the parton density functions, the initial state parton shower is evolved backwards, from the boson vertex towards the proton [16]. Every space-like branching changes the x and Q arguments in the parton density functions, $\phi(x, Q^2)$, leading to a sequence of decreasing Q^2: $Q, Q_1, \ldots, Q_n, Q_s$ and increasing x: $x, x_1 \ldots, x_n$ from the boson vertex to the proton. A modified Sudakov form factor, including a term $\frac{\phi(x_k, Q^2_{k+1})}{\phi(x_k, Q^2_k)}$, will ensure that after a backward evolution of the cascade x_n is distributed according to the structure function $\phi(x_n, Q_s^2)$ [15].

If the maximum of the evolution variable is not derived from the colour flow between initial and final state (as described earlier in QCD coherence condition 3), then the problem of choosing the correct maximum for DIS can only be arrived at experimentally. The degree of hardness of the initial and final state radiation depends on the virtuality of the quark at the point just before (m_n in fig. 3.4) and just after (m_{n+1}) the boson vertex. The phase space limit for this virtuality is given by W^2. However, in principle any function of Q^2 and W^2 is possible in DIS, there is no unique scale as there is in e^+e^-. This problem becomes most striking in the very small x region, where Q^2 can be rather small (≈ 10 GeV^2) and W^2 very large ($10^4 - 10^5$ GeV^2). Several different scales have been suggested based on the knowledge one can gain from studying the first order QCD matrix element or the colour dipole model [17, 18]. W^2 will overestimate the parton radiation because, for small x, the proton remnant takes an energy fraction $(1-x)$. This energy is part of W^2, but it will not be available for parton radiation. From the asymptotic parton transverse momentum distribution for $x \to 0$ and $x \to 1$, as given by the matrix elements, one can deduce a scale $Q^2(1-x)\max(1, \ln\frac{1}{x})$ by interpolating between the low and high x region [17].

3.4 The Colour Dipole Model

3.4.1 The Basics

In the colour dipole model (CDM) QCD radiation of quarks and gluons is described by a chain of radiating colour dipoles. In the simplest case, the emission of a gluon from a $q\bar{q}$ system, as in e^+e^-, can be described by the dipole formula [19]:

$$d\sigma = \frac{\alpha_s}{4\pi^2} N_c \frac{dp_\perp^2}{p_\perp^2} dy d\phi \qquad (3.8)$$

where N_c is the number of colours, $p_\perp$ denotes the transverse momentum of the emitted gluon, y is its rapidity and ϕ the azimuthal angle. All these quantities are taken in the CMS of the dipole, which corresponds to the CMS in e^+e^- before the first gluon emission. As a consequence of energy-momentum conservation the available emission region is restricted to:

$$|y| \leq \ln \frac{\sqrt{s}}{p_\perp}. \qquad (3.9)$$

This phase space corresponds to the triangular region in fig. 3.5.

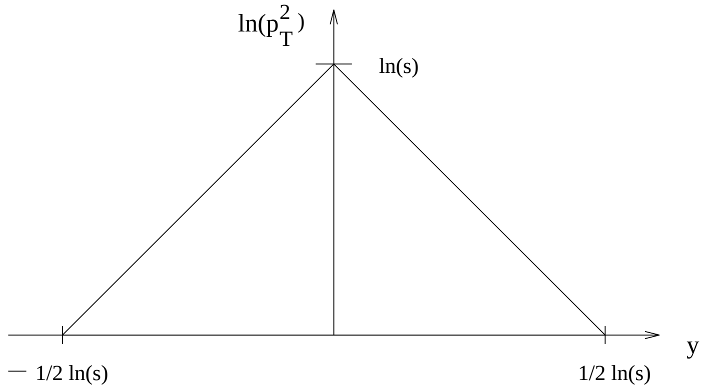


Figure 3.5: *Phase space boundaries for dipole emission in e^+e^- in the dipole's CMS. s denotes the CMS energy of the dipole. The rapidity axis corresponds to the axis of the $q\bar{q}$ pair.*

The produced state after the first gluon emission, according to the above formula, contains three coloured objects, namely a quark, an antiquark and a gluon. Gluon emission of this state can be treated to a very good approximation as emission of two independent dipoles, provided the transverse momenta of the first and second gluon are ordered $p_{\perp 1} > p_{\perp 2}$ [20]. This new state then continues the emission process, ordered in decreasing $p_\perp$, and resulting in a chain of independently emitting dipoles, connected via the gluons. Two stages of this process are illustrated in fig. 3.6.

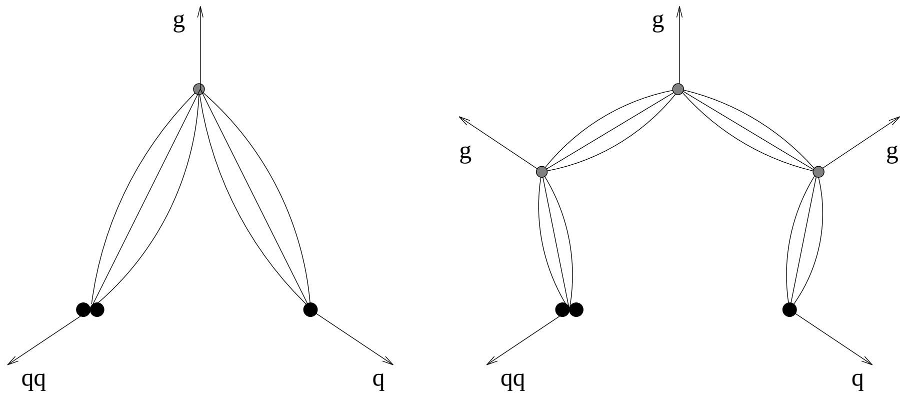


Figure 3.6: *Two stages of the dipole chain development. In the Lund model a colour field can be specified either in terms of gluons or in terms of the dipoles stretched between the gluons or the gluons and the quarks. The field behaves like a string stretched from the di-quark via the gluons to the quark.*

The emission distributions, in $p_\perp$ and y, of each individual dipole are given by the dipole formula, in the rest frame of this dipole. This leads to a kinematical constraint for the allowed rapidity of each emitted gluon, which corresponds to the strong angular ordering prescription in parton showers. Soft gluon interference is therefore automatically taken into account in the CDM.

3.4.2 Extensions for DIS

DIS requires an extension to the above picture, since the target remnant is an extended coloured object. In contrast to the parton shower model the CDM does not divide the process into an initial and a final cascade. One can instead argue [21], that all radiation can be described as radiation from the colour dipole formed between the struck quark after the scattering and the proton remnant. The underlying basic assumption is that an undisturbed colour singlet system, i.e. the proton before the scattering, does not emit any radiation. However, one can still divide the phase space within the triangle in fig. 3.7 into regions that would correspond to the initial parton shower (shaded) and to the final state cascade (blank) respectively, in the parton shower language [22]. In this figure the phase space boundaries for DIS are given in terms of lightcone momenta[1], see [23]. These can be related to the normal

[1]Lightcone coordinates of a 4-vector p are defined by $p = (p^+, p^-, p^1, p^2)$ with $p^\pm = (p^0 \pm p^3)/\sqrt{2}$

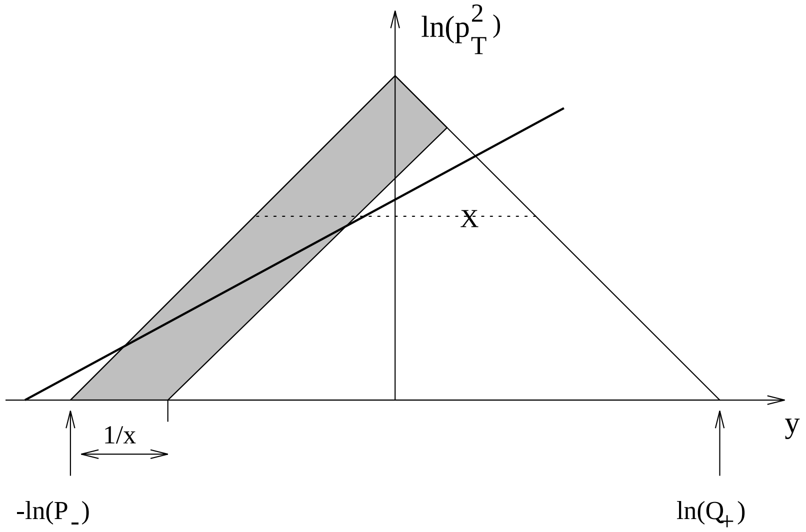


Figure 3.7: *Phase space boundaries for dipole emission in DIS in the dipole's CMS. The thick line indicates the reduction due to the extended proton remnant (see text). In the parton shower language the shaded area corresponds to the initial state parton shower. The positive y-axis gives the direction of the struck quark. The* X *marks a possible phase space point for the first gluon, further gluons then can only be radiated below the dotted line.*

DIS quantities via: $q = (Q_+, -Q_-)$ and $x = Q_-/P_-$, with P_- being the protons negative lightcone component.

It is a well known fact that the emission of wavelengths shorter than the size of the emitting dipole is suppressed. This can be implemented in the CDM by allowing only a fraction of the proton remnant to take part in the emissions. This fraction, $\mu/p_\perp$, governs the suppression of emissions in the target region, which is indicated by the thick line in fig. 3.7. $1/\mu$ denotes the transverse size of the target remnant and should be of the order of a hadrons size, and $p_\perp$ is the transverse momentum of the emitted gluon. As a direct consequence, [21], the maximum value of the transverse momentum of the first gluon is given by $p_\perp^2 = \left(\frac{W}{2}\right)^{\frac{4}{3}} \mu^{\frac{2}{3}}$. In the CDM the transverse momentum acts as the evolution variable. The colour coherence effects leading to this suppression of emission in the target region have no correspondence in the parton shower approach.

3.4.3 Inclusion of Boson-Gluon Fusion

One drawback of the scheme, so far, is the lack of a description of the BGF process. As the CDM only looks at radiation from the dipole between the struck quark and the proton remnant, it will not allow for a situation where an initial state splitting of the incoming gluon takes place. BGF events also require the proton remnant to be left in a colour octet state. So one has to split the remnant into two colour triplet parts and to form two dipoles with the oppositely coloured quark and antiquark, respectively.

The problem can be overcome by generating BGF processes using the first order matrix elements and add subsequent emissions using dipole radiation. However, for theoretical considerations [24], emissions of the $q\bar{q}$ pair with a $p_\perp$ larger than the $p_\perp$ of the pair should be highly suppressed. This leads to the necessity to impose a cut-off on the maximum allowed $p_\perp$ for gluon emission that is not inherent to the model.

Two further technical complications arise from the inclusion of the BGF process, when adding dipole radiation to the matrix elements. Firstly, a dependence on the cut-off, y_{cut}, for the matrix elements is introduced into the model. Secondly, one has to distinguish sea-quark interactions from BGF events [18]. Both can be viewed, in terms of the Lund string picture, as resolving a gluon (a kink on the string) into a $q\bar{q}$ pair, but the time scales involved are quite different. In case of BGF the pair has not lived long enough to be considered independent, both quarks receive a momentum transfer. A suitable method for this distinction can be implemented in a Monte Carlo generator by requiring the virtuality of the pair, which is taken to be their $p_\perp^2$, to be larger than Q^2.

An alternative approach to implement the BGF process is to correct the colour dipole cascade to match the first order matrix elements in the first emission [25]. In this way the matrix element divergences are cured by introducing a competition between gluon emission and ‘emitting’ an (anti)quark, leading to a BGF event.

3.4.4 Recoils

In contrast to the ordinary parton shower approach, in the CDM one must also consider the distribution of transverse recoils, which are necessary for transverse momentum compensation. When a single parton radiates a gluon the transverse momentum of the emitted gluon will be balanced by a transverse momentum (the recoil) given to the emitting parton. Having two emitters means that the transverse recoil must be split between the two. To preserve the assumption that the dipoles are independent emitters, the CDM postulates that the recoils are distributed such that neighbouring dipoles are minimally disturbed. However, changing this recoil strategy may have a considerable effect on the partonic final state, especially in DIS where there is the extra complication due to only part of the extended proton remnant taking part in the emissions [18].

3.5 Hadronisation

Hadronisation describes the process of transforming a parton into a jet of hadrons. This process can be technically divided into the perturbative phase of gluon emission, as in the parton shower model, and the subsequent, non-perturbative, hadronisation phase[2]. Although this implies a clear distinction, both phases are connected and should not be viewed separately. Depending on which description of the hadronisation phase is used, different features of the perturbative phase become important. For example, soft gluon coherence is of major importance for a cluster fragmentation model, but of little importance in case of string fragmentation, since the larger number of soft gluons will not produce a great deal of extra hadrons. In the same context the cut-off for the parton shower evolution is of differing importance in these two approaches to hadronisation.

Eventually, non-perturbative techniques may provide sufficient understanding of confinement to permit calculations of hadron distributions in jets from first principles. At present, one is far from such an understanding and it is necessary to resort to phenomenological models, containing large numbers of adjustable parameters. Most of these free parameters concern heavy quark production and decays and the mechanism of baryon production. Since all distributions studied in this analysis are inclusive, very little dependence on most of the parameters is assumed. In addition, hadronisation should, by construction, describe LEP and HERA data simultaneously; the parameters used here have been tuned to describe the high precision LEP data.

The Monte Carlo generators used in this thesis to compare with data use either the string or cluster fragmentation models. Whereas the local parton-hadron duality approach can be used to compare corrected data directly to theoretical calculations on the parton level, the independent fragmentation model is described, since it has been the first model used for hadronisation.

3.5.1 Local Parton-Hadron Duality

The simplest assumption for hadronisation is local parton-hadron duality (LPHD) [26]. In this model it is assumed that the general features of the partonic final state, such as energy flows, multiplicities and energy correlations, can be related to the spectra of hadrons simply by normalisation constants. These constants have to be determined by experiment. This model was a consequence of the unexpectedly close correspondence between some analytically described features of partonic systems and measured final state hadron characteristics at high energy interactions. The striking advantage is that one can directly compare results of analytical calculations obtained in a suitable approximation, such as the modified leading logarithmic approximation (MLLA, see section 7.2) to measured hadron distributions, without invoking any hadronisation model. This approach tries to take perturbative QCD as near as possible to the limits posed by confinement. The existence of LPHD is not yet verified experimentally, although studies at LEP show very good agreement with predictions based on this assumption [27, 28, 29].

[2]There is no general consens on the exact meaning of the terms hadronisation and fragmentation. Sometimes the perturbative phase is explicitly included in what is termed hadronisation, however, more often both words mean exclusively the non-perturbative stage.

3.5.2 Independent Fragmentation

The original model which attempted to describe the way partons fragment into hadrons was proposed by Feynman and Field in 1978 [30]. The so-called independent fragmentation model is based on the scaling hypothesis and on meson production. The latter assumes collinear $q\bar{q}$ pairs popping out of the vacuum till the energy of the original parton is used up. One basic deficiency of this algorithm is that it cannot conserve both energy and longitudinal momentum simultaneously. The original parton mass is zero, but the resulting jet has a mass of several GeV. This has to be overcome by making small adjustments to the energy of the jets when the process is finished. Although there are still some independent fragmentation models on the market, they have been shown to fail to describe some interference and coherence phenomena [31] and are therefore not widely used anymore.

3.5.3 String Fragmentation

Another way of looking at the process of fragmentation is the Lund String Model [32], which is the de facto standard for Monte Carlo generators. In this picture the colour force spans a narrow flux tube, the string, between two colour charges. If the particles are rapidly separating from each other the string is stretched. This requires energy, since the colour force rises with increasing separation (confinement). At high enough energies a $q\bar{q}$ pair can tunnel out of the vacuum and the string will break into shorter substrings, still rapidly separating. The process is sketched in fig. 3.8. In this way the original string will subsequently break up into smaller string pieces, corresponding to conventional on-shell hadrons. The string splits according to the so-called Lund left-right symmetric fragmentation function, which contains two free parameters a and b [32]. During the process of fragmentation the hadrons acquire some transverse momentum according to a Gaussian distribution, with the width given by the second parameter, b.

In the string model, gluons can be viewed as string excitations, this is visualized as kinks on the otherwise straight string. Since the gluons carry energy and momentum the different segments of the string will move with respect to the overall centre of mass frame. It is interesting to note that the string segments move the same way as the colour dipoles spanned between the colour charges in the CDM. This can be viewed as a realisation of LPHD [33]. Also the string fragmentation scheme is infrared safe with respect to soft and collinear gluon emission, i.e. the hadronic final state produced will not change when extra soft gluons are added. This makes generators incorporating the string model less sensitive to the correct inclusion of QCD coherence effects.

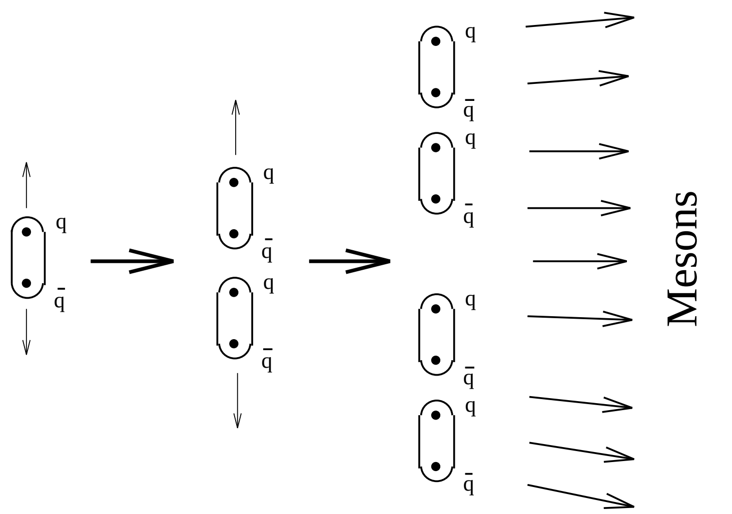


Figure 3.8: *String fragmentation model. A string piece (here between a $q\bar{q}$ pair) is stretched and breaks subsequently into smaller pieces. Finally the low energy string pieces give rise to mesons and baryons.*

3.5.4 Cluster Fragmentation

Parton shower evolution stops at a certain cut-off in the evolution variable that is chosen such that α_s does not become too large, i.e. perturbation theory still holds. The resulting quarks and gluons can now be combined into colour singlet clusters, by joining colour connected partons. The so formed clusters do not resemble hadrons themselves, but subsequently decay into hadrons, according to the available phase space. The colour connections are only unique in the so-called planar approximation, which corresponds to an expansion in $\frac{1}{N_C}$, N_C being the number of colours. The process in shown schematically in fig. 3.9.

In this cluster fragmentation approach [34], gluons are forced to split non-perturbatively into quark antiquark pairs, after the parton shower development has stopped. Joining nearby quarks and antiquarks into hadrons employs the 'preconfinement' property of QCD, i.e. the tendency of the partons generated in the branching process to be arranged in colour singlet clusters with limited extension in both coordinate and momentum space. The cluster model assumes that these clusters are the basic units out of which hadrons arise non-perturbatively. Therefore, the produced clusters preferably have low masses, of the order of the cut-off at which the perturbative branching is terminated. This process is asymptotically independent of the hard subprocess.

Cluster fragmentation models have the attractive feature of introducing very few free parameters for the process of hadronisation. For instance, the transverse momentum generated in hadronisation is determined by the average energy release in cluster decay and not by an additional parameter as in the string model. Apart from Λ_{QCD}

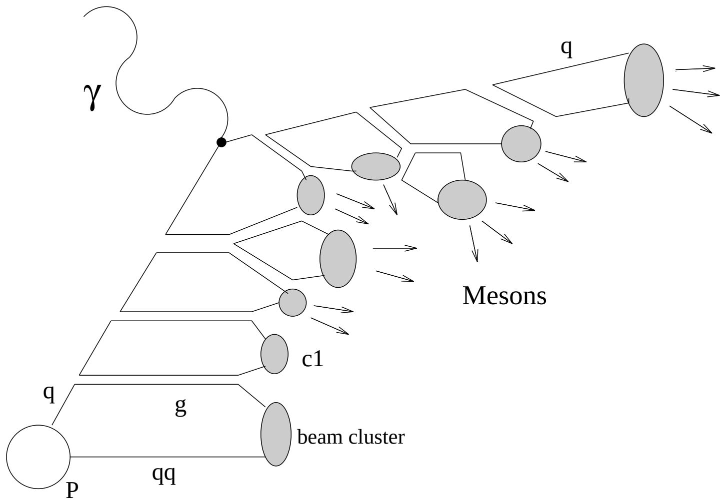


Figure 3.9: *A cluster fragmentation model drawn in the so-called planar approximation. In this approximation gluons are represented by double lines, corresponding to the colour and anticolour they carry. The gluons are forced to split non-perturbatively into $q\bar{q}$-pairs that then form colour singlet clusters.*

and the cut-off for the parton shower, the fission threshold is the most important parameter of the model. When a cluster has a mass higher than this threshold it is not assumed to decay isotropically, but will rather undergo an anisotropic fission, similar to a 'symmetrical string breaking'.

3.5.5 Remnant Fragmentation

The DIS process requires modeling of the fragmentation of the proton remnant, a field that to-date has many unknowns and needs more experimental input. Two approaches are mentioned here, simple quark counting rules and the soft underlying event.

In DIS there are three possible remnant scenarios: scattering of a valence quark, of a sea quark and boson-gluon fusion.

- When the struck quark is a valence quark the proton remnant is simply a di-quark composed of the two remaining valence quarks. This system forms a colour triplet and colour anti-triplet state.

- In case of sea quark (or antiquark) scattering, the remnant contains the three valence quarks and the partner of the $q\bar{q}$ pair. Depending on if the left-over sea quark is a quark or an antiquark, a meson or a baryon can be formed, respectively, at random from the possible flavour compositions. The splitting process into the meson or baryon can be assumed similar to hadronisation and is done according to a longitudinal splitting function, also generating a small Gaussian $p_\perp$.

- If a boson-gluon fusion process takes place, the target remnant is left in a colour octet state. The remnant is split into a quark and a di-quark chosen with random flavours and using the same longitudinal momentum sharing and Gaussian $p_\perp$ as in the sea quark case.

3.5.6 Soft Underlying Event

Another approach to model the remnant fragmentation and the interaction of the remnant with the rest of the event is the so-called soft underlying event, as described in [35]. Here, a soft interaction of the remnant with a parton of the initial state parton shower is modeled, inspired by minimum bias $p\bar{p}$ interactions. In fig. 3.9 the cluster labeled *beam cluster* would undergo a soft collision with the cluster labeled *c1*. Although both move parallel to the proton, they will still collide in their centre of mass frame.

The disadvantage of this model is that it contains many parameters and has therefore little predictive power. The original model comes from $p\bar{p}$ experiments, in which two beam remnants exist and may collide in a soft interaction. To make the model applicable to DIS required considerable modification and only experiments can verify or dismiss this approach.

Chapter 4

Generators

There was the faintest of sounds,
like the rattle of old teeth.
...It sounded like dice.

– *Terry Pratchett*

Monte Carlo generators are computer programs that, starting from a given interaction type like e^+e^- or ep, generate the full development of the process through to the final state of hadrons and leptons. The 'first name' Monte Carlo emerged since these generators use (pseudo)random numbers wherever certain variables have to be distributed according to a known distribution. Examples of such variables are the x and Q^2 distributions in ep, which have to be in concordance with the differential cross section, or the longitudinal momentum fraction of an emitted gluon in a parton shower, which is distributed according to the Altarelli-Parisi splitting probabilities.

The whole process of Monte Carlo event generation and detector simulation has four steps, illustrated in fig. 4.1. First the basic ep scattering is generated, then QCD radiation is added, e.g. in the form of parton showers; this gives the partonic final state. The third generator step turns partons into jets of hadrons, using a fragmentation model as discussed in section 3.5. The final step is the H1 detector simulation, provided by the H1SIM program. A model representation of the H1 detector is used and the generated hadrons are tracked through this 'detector', giving rise to particle decays, secondary interactions, drift chamber hits and finally calorimetric energy depositions. At this stage one has Monte Carlo events which are fed into the same reconstruction program as real events.

In the following sections the three Monte Carlo generators, LEPTO, ARIADNE and HERWIG, are presented with respect to their application in ep scattering at HERA. While both, LEPTO and HERWIG, are able to generate all three stages of the scattering process (differential cross section, QCD radiation and hadronisation), ARIADNE can only generate the QCD radiation part and so must be interfaced to, e.g., the LEPTO Monte Carlo generator to account for the basic scattering and the hadronisation phase.

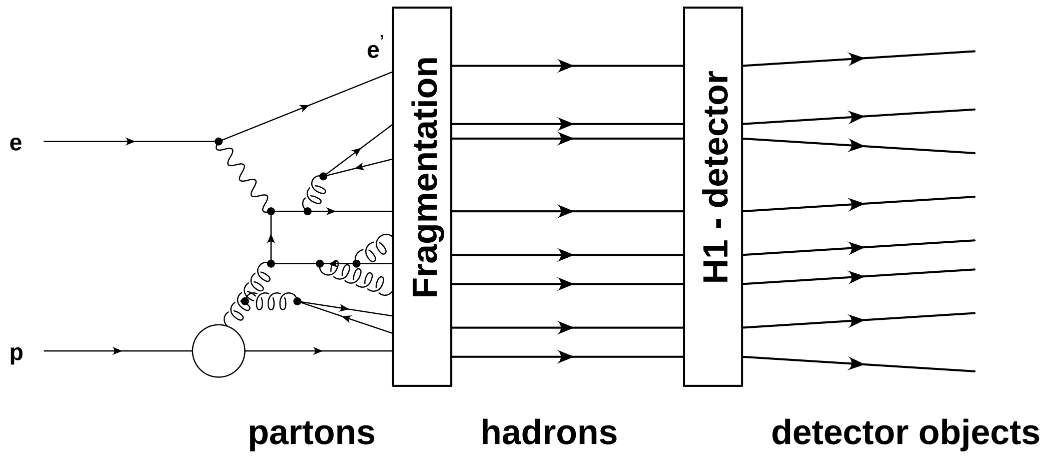


Figure 4.1: *The simulation chain for Monte Carlo generators and H1 detector simulation. The generator covers the first three steps, generating the basic scattering, partons and finally hadrons, while the detector simulation (H1SIM) provides particle decays, secondary interactions and energy depositions in a model representation of the H1 detector.*

4.1 LEPTO

The LEPTO 6.10 Monte Carlo generator [17] simulates deep inelastic lepton-nucleon scattering, without QED radiative corrections. The parton level interaction is based on the standard model electroweak cross sections and parametrisations of parton density functions. First order QCD matrix elements are implemented and higher order QCD radiation can be modeled using parton showers. Hadronisation is performed using the JETSET 7.3 Monte Carlo [32], implementing the Lund string fragmentation model.

In modeling QCD radiation, 4 distinct options are available in LEPTO, steered by the LST(8) parameter:

- no QCD radiation
- first order matrix elements
- leading logarithmic parton showers, with a free choice of the maximum value of the evolution variable
- first order matrix elements merged with leading logarithmic parton showers

These options are discussed in more detail in the following subsections with particular emphasis on the dependence on the LEPTO steering parameters.

4.1.1 No QCD Radiation

LEPTO allows QCD radiation to be switched off completely, generating only a scattered quark and the proton remnant in the partonic final state. This can be viewed as an implementation of the QPM, however, since hadronisation is required to compare to the data, there are two non-perturbative sources of $p_\perp$ present in LEPTO that are not part of the naive QPM. Firstly, LEPTO generates an intrinsic $p_\perp$ for the partons in the proton according to a Gaussian distribution, with a width given by PARL(3). Secondly, the already mentioned $p_\perp$ introduced in the hadronisation phase also contributes.

Although a danger of a mix-up regarding the name exists, the model is referred to as the QPM in the comparison with data. One should always bear in mind that the hadronic final state generated with this option comes from the hadronisation of a string connecting the struck quark and the remnant. One cannot expect the QPM model to describe the data, however, the model is still used, as comparing the data to the most simple model will help towards a deeper insight into the nature of QCD effects.

4.1.2 Matrix Elements

Choosing the matrix elements (ME) option to generate QCD radiation has the advantage that the first (and only) emission is exactly modeled. Since high energy reactions are known to require multi parton emission to get an acceptable description of hadronic final states, the range of applicability of the ME option is clearly limited. For HERA physics the matrix elements option can help in deriving α_s from the the 2+1 jet rate, if a cut on the hard subprocess centre of mass energy is applied [36].

The matrix elements are divergent in the limit of soft and collinear gluon radiation. A cut-off, to avoid the singular regions, can be imposed in three different ways in LEPTO:

- A cut-off $y_{cut} = m_{ij}^2/W^2$, which is the PARL(8) steering parameter of LEPTO. LEPTO internally checks for each point in x and Q^2 if the value requested is below what the program considers to be the cut-off value for y where the ME diverges. In such a case the user supplied value is adjusted to avoid the divergent regions.

- A direct cut-off on the invariant mass m_{ij}, being PARL(9) in the LEPTO steering, if PARL(9)> 0.

- Or, if PARL(9)< 0, a certain distance in invariant mass from the divergence region[1]. As in the first case, the divergence limit y_{cut} for a given x, Q^2 is calculated by comparing the leading order and next-to-leading order cross

[1] This option is not part of LEPTO, but has been implemented in the H1 version of this generator [37].

sections and then a value of $-\mathrm{PARL}(9)/W^2$ is added to this y_{cut} to keep away from the divergent region. This option follows the shape of the divergence region as function of the kinematic variables even closer as the first option.

It should be noted that LEPTO changes JETSET parameters internally when either the ME or the 'no QCD radiation' options are chosen. These changes, that are hidden from the user, affect the b parameter of the symmetric Lund fragmentation function (see section 3.5) and thus the amount of $p_\perp$ generated in the hadronisation step.

4.1.3 Parton Showers

With the parton shower option, LEPTO generates leading logarithmic parton showers as discussed in section 3.3 with the following features:

- the evolution variable is the virtuality (the invariant mass squared, m^2) of the struck quark
- no interference between initial and final state parton shower
- free choice of the maximum value for the evolution variable
- strong angular ordering in the final state cascade

The maximum virtuality m^2_{max}, for the initial and final state parton shower is set by default to $Q^2(1-x)\max(1, \ln\frac{1}{x})$. The actual value of m^2, used in a specific generation of an event, is chosen randomly between m^2_{min} and m^2_{max}, according to a $1/m^2$ distribution. Other possible scale choices include $W\sqrt{Q^2}$ and $W^{\frac{4}{3}}$. In LEPTO the parameter LST(9) is used to select the parton shower scale.

It should be stressed again that one cannot expect the parton shower model to give a reasonable description of all the properties of the hadronic final state. For certain observables and/or kinematical regions hard (large angle) QCD emissions become important, which are not well described by leading logarithmic parton showers. This problem is overcome by matching the parton shower to the first order matrix elements.

LEPTO parton showers have been studied in this analysis with the following parameter settings.

parameter	meaning	default value	value used
LST(8)	use PS instead of ME+PS	12	2
LST(9)	PS scale	5	3

Throughout, the number of flavours has been set to 5, MRS $\mathrm{D}^{-'}$(DIS) [39] parton distributions are used, and Λ_{QCD} has been set to 190 MeV. Since the setting for the parton shower scale corresponds to $W\sqrt{Q^2}$ as maximum for the evolution variable, this model will be referred to as PSWQ.

4.1.4 Matrix Elements + Parton Showers

Although LEPTO *does not* offer matching, as described in section 3.3, between matrix elements and parton showers, an attempt is made to add parton showers on top of the first emission generated by the matrix elements. Double counting of Feynman graphs when generating further emissions is avoided by limiting the maximum of the evolution variable according to the output of the matrix elements. The matrix elements give three possible cases:

a) q-event (no emission by the matrix elements),
b) qg-event (QCD Compton),
c) $q\bar{q}$-event (boson-gluon fusion).

In case (a) the maximum virtuality, m^2_{max}, for the final state parton shower is the same as the cut-off value for the matrix elements, i.e. $y_{cut}W^2$. The maximum virtuality for the initial state parton shower will be slightly different, due to the kinematical constraints from the exchanged boson. Evaluating the matrix elements for this particular event does not lead to first order QCD emission, hence the maximum possible value for parton shower emissions must be restricted to be less than the lower limit for ME emissions.

For case (b) and (c) the maximum virtuality, m^2_{max}, for the *final* state parton shower is calculated from the invariant mass of the two outgoing partons, given by the matrix elements. For the *initial* parton shower, when looking at the BGF diagrams, it seems natural that the virtuality of the quark propagator between the gluon and the boson gives the maximum virtuality. There are two Feynman diagrams for $q\bar{q}$-events, see fig. 3.3, and due to interference it is not possible to decide which particular diagram led to a specific event. Therefore the maximum virtuality for the initial state parton shower is not unambiguously defined. The current approach in LEPTO is to calculate the squared 4-vector difference between the boson and both quarks individually and take the *maximum* of the two as the virtuality scale m^2_{max}. Similar arguments apply for the initial state parton shower in case (b).

It can be derived from the above implementation details that the results of the ME+PS option will strongly depend on the cut-off for the matrix elements. If this cut-off is too high, the adding approach is invalidated, as all events generated will be q-events. This would result in a pure parton shower Monte Carlo generator, with W^2 as the maximum of the evolution variable. However, if the cut-off is chosen too low, too little hadronic energy will be generated. This is a consequence of the adding procedure and is explained in the following. Firstly, the cross section for first order processes is overestimated due to the lack of NLO corrections and this causes too many events of the type (b) and (c), for example in the kinematic range $Q^2 \approx 25$ GeV^2 and $x \approx 10^{-3}$:

1. PARL(8) = 0.015:
   $\approx 6\%$ of all events are qg or $q\bar{q}$-events,

2. PARL(9) = -2.0:
   $\approx$ 35% of all events are qg or $q\bar{q}$-events.

Since at a y_{cut} of 0.015 the leading order and next-to-leading order cross ections for 2+1 jet events agree [38], the 35% obtained with the low cut-off option indicate a serious overestimation of the number of 2+1 jet events. Secondly, the parton showers that are added on top of the first emission from the matrix elements use the invariant mass of the two partons from the ME as the maximum virtuality. This invariant mass will be very low for low cut-off values, allowing few gluons to be emitted. Both effects together thus result in too low a hadronic activity.

The matrix elements+parton shower model has also been compared to the data in this analysis. The model, with the parameter settings given below, will be referred to as MEPS.

parameter	meaning	default value	value used
PARL(8)	y_{cut}	0.015	0.015

As for the PSWQ model, 5 flavours, the MRS D$^{-'}$(DIS) parton distributions and Λ_{QCD} = 190 MeV are used.

4.1.5 The Event Record

Within the H1 generator framework the GTR bank[2], which contains the generated particles, is a straight translation of the LUJETS common block utilised internally by LEPTO. The GTR bank obeys the standard HEPEVT definitions for particle names, codes and status flags, for more details see [40, 41]. The following extract from a MEPS event may serve to explain the basic features of the event record.

```
------------------------------------------------------------------------
    PARTICLE  PARTICLE     PARENT          MOMENTUM (GEV)        ENERGY
    TYPE      FLAG        PARTICLE    PX         PY         PZ      (GEV)
------------------------------------------------------------------------

 1  e-       beam part.  0   0      .000       .000    -26.700     26.700
 2  p+       beam part.  0   0      .000       .000    820.000    820.001
 3  Z0        document.  1   0    -2.612      -.611    -16.084     15.751
 4  e-        undecayed  1   0     2.612       .611    -10.616     10.949
 5  u#        document.  2   0      .307       .173      8.884      8.891
 6  u#        document.  5   0      .162       .072      8.855      8.857
 7  u#        document.  6   0    -2.450      -.539     -7.228     24.607
 8  Delta++  decay/frag  2   0     -.027      -.008    498.773    498.775
 9  u#       decay/frag  0   0     -.447       .173     -4.028      4.056
10  g        decay/frag  0   0    -1.308      -.933     -5.530      5.759
```

[2]Banks are briefly explained in chapter 5.

11	g	decay/frag	0	0	-1.706	-1.021	-3.579	4.094
12	g	decay/frag	0	0	-.541	.937	-2.100	2.363
13	g	decay/frag	0	0	.148	.161	.041	.222
14	g	decay/frag	0	0	.158	.205	1.018	1.051
15	g	decay/frag	0	0	1.249	-.061	6.950	7.062
16	d	decay/frag	2	0	-.135	-.064	312.370	312.370
17	p+	undecayed	8	0	-.032	.080	473.488	473.488
18	pi+	undecayed	8	0	.004	-.089	25.286	25.286
19	string	decay/frag	9	16	-2.585	-.602	305.143	336.977

The first two lines contain the beam particles. On line three is the exchanged boson. For NC interactions, i.e. γ/Z° interference, the particle type is always set to be a Z°. Next comes the scattered electron. The following three documentation lines[3] contain the struck quark (a $\bar{u}$ is denoted by `u#`) at different stages of the event generation. On line 5 is the struck quark (a sea quark in this example) before the initial state parton shower. The next entry is the same quark after the initial parton shower and directly before the boson vertex; line 7 is the quark directly after the boson vertex. On line 9 is the struck quark after the final state parton shower, i.e. this is the quark belonging to the partonic final state. The proton remnant is made of a Δ^{++} and a d on line 8 and 16 respectively; note that both have the incoming proton (line 2) as their parent particle. Lines 10-15 contain gluons generated from parton showering; with no distinction being made between initial and final state gluons.

The example given was a q-event in the MEPS language. If a qg-event is generated by the matrix elements line 8 will contain the gluon produced in the hard subprocess as a documentation line. Similarly for $q\bar{q}$-events, the lines 7 and 8 contain the quark antiquark pair before parton showering as documentation entries. Unfortunately this structure of the event record is only true for the MEPS option. If ME or no QCD radiation are selected a different layout for the first (important) lines is chosen.

4.2 ARIADNE

The ARIADNE 4.04 program [42] provides an implementation of the CDM with the necessary additions for DIS, as discussed in section 3.4. ARIADNE is interfaced to the LEPTO Monte Carlo generator, which generates the electroweak part of the scattering leaving ARIADNE generate the QCD cascade process afterwards. Hadronisation is subsequently performed by JETSET.

The plain CDM, as used in e^+e^-, has very few free parameters. Basically the treatment of recoils is a steerable process, but the default, the 'minimum disturbance of the colour flow', is the best physically motivated one.

[3] Documentation lines indicate intermediate states that are not part of the partonic or hadronic final state.

Having extended sources, such as the proton remnant in DIS, complicates the picture and introduces new free parameters to the model. The remnants extension (μ = PARA(11) = 0.6 GeV) can be varied as long as it is of the order of a proton's mass. In addition, the dimensionality of the extended source, α = PARA(10) = 1.0, is not constrained to a certain value by physical arguments. Both parameters together vary the position of the thick line in fig. 3.7, so reducing or increasing the suppressed emission from the remnant. Also, the distribution of transverse recoils is different in DIS compared to e^+e^-. Since only a part of the remnant takes part in the emission, only this part should be allowed to recoil. This strategy can be changed using the MSTA(17) and MSTA(18) parameters.

The most important aspect of the current interface of ARIADNE to LEPTO is the inclusion of the BGF events. As the CDM looks only at radiation from the dipole between the struck quark and the remnant it will completely neglect the BGF events. For this reason the first order matrix elements are used to generate $q\bar{q}$-events, with further emissions added by ARIADNE[4]. To ensure a proper evaluation of the matrix elements, LEPTO generates all events with the matrix elements option. This again gives the three possible cases discussed in the previous section:

- q-event

  ARIADNE will form a dipole between the scattered quark and the remnant and will generate dipole radiation in the normal way.

- qg-event

  In case of a qg-event the gluon will simply be stripped off the event record by ARIADNE. CDM radiation will be generated from dipole formed by the remnant and the scattered quark before the gluon emission. The claim is that ARIADNE will reproduce hard gluon emission in concordance with the matrix elements [43].

- $q\bar{q}$-event

  For $q\bar{q}$-events both quarks form a dipole with the remnant and radiate independently. To avoid gluon emissions, with a $p_\perp$ larger than the $p_\perp$ of the $q\bar{q}$-pair (and thus avoiding double counting), the gluon $p_\perp$ can be limited using the PARA(6) parameter internally. This approach has been implemented in the H1 version of ARIADNE (IBGF = 1), however, since the effect on energy flows in the investigated kinematical regions is only of the order of 5%, it is not essential.

Using the matrix elements introduces the previously discussed dependence on the cut-off also into the context of the CDM. It seems preferable to choose a low cut-off

[4]The alternative approach briefly discussed in section 3.4.3 is used in the latest version of ARIADNE, 4.05. The results for the energy flows differ only slightly from the distributions obtained with ARIADNE 4.04, see [25]

and to let ARIADNE decide whether or not a given event is a sea quark interaction or a BGF event. ARIADNE bases this decision on the virtuality of the $q\bar{q}$-pair, taken to be their $p_\perp^2$. If this $p_\perp^2$ is larger than Q^2 multiplied by PARA(20) the event is treated as a BGF event, spanning a dipole between each (anti)quark and the remnant.

In case of ARIADNE the event record is slightly modified with respect to that for LEPTO. The first 4 lines are as in the LEPTO case. The struck quark and the remnant are usually reported on line 5 and 6, in their state before dipole radiation, and then again further down after the QCD cascading has taken place. In contrast to LEPTO, line 5 and 6 are not documentation lines, and one has to check the mother/daughter relationship in order to find the final quarks and gluons. Partons having the beam proton as their parent again stem from the proton remnant splitting process.

The term CDM is used throughout this thesis for both, the colour dipole model and for the actual ARIADNE implementation used to compare with data. The parameter settings are given in the following table:

parameter	meaning	default value	value used
PARL(8)	y_{cut}	0.015	0.015
PARA(11)	μ (GeV)	0.6	0.6
PARA(20)	BGF cut-off	0.5	0.5

Also all other parameters are set to their defaults according to the 'EMC/DELPHI' tuning set provided with ARIADNE. Again 5 flavours, MRS $\mathrm{D}^{-'}$(DIS) parton distributions and $\Lambda_{QCD} = 190$ MeV are used.

4.3 HERWIG

HERWIG 5.7 [35, 44] is the only multi-purpose event generator used in this analysis. Amongst several possiblities for e^+e^- and $p\bar{p}$ it provides two options for DIS, neither of which take QED radiative corrections into account:

- generating only 2+1 jet events using the first order matrix elements with merged parton showers, or
- generating events with leading logarithmic parton showers.

The first option only gives events with hard QCD radiation, according to a cut-off applied to the matrix elements. This is different from the LEPTO MEPS model, where also events without an emission from the ME are taken into account, according to their fraction of the cross section. No such version, where the matrix elements are fully merged with parton showers, is available yet for HERWIG. Hence, only the parton shower option is reviewed here.

The major differences between HERWIG and LEPTO, concerning the parton shower option, can be summarized as follows:

- HERWIG parton showers obey all three colour coherence conditions listed in section 3.3,
- a cluster fragmentation model is used in HERWIG, the Lund string model is used by LEPTO,
- HERWIG can generate soft underlying events, as discussed in section 3.5.6.

Although a first look in the manual might reveal a confusing number of parameters, in fact very few free parameters are available for DIS. The important ones will be discussed subsequently, a more comprehensive discussion can be found in [44, 45].

The soft underlying event (s.u.e.) option turns out to be one of the most sensitive HERWIG parameters when studying energy flows in DIS. As already briefly described in section 3.5.6, this option provides soft underlying events similar to $p\bar{p}$ minimum bias interactions. The details of the specific HERWIG implementation can be found in [35, 15]. It must be pointed out that this model is a very empirical approach, it contains 11 parameters, which have been mainly fitted to UA5 data. Hence it is not surprising that the two free parameters available for DIS can be freely varied by the user. The basic parameter for the s.u.e. is PRSOF, the probability for generating a s.u.e. in a given event. HERWIG will decide on a event by event basis whether to generate a s.u.e., or not. The default for PRSOF is 1.0, i.e. every event will have a soft underlying event. Furthermore the multiplicity in the s.u.e., which is normally chosen according to $\sqrt{s}$ of the soft collision, can be enhanced by setting ENSOF>1, or decreased by using ENSOF<1, the default again being 1.0.

Another parameter available for user steering, since version 5.7 of HERWIG, is BTCLM, the fission threshold parameter for the beam remnant cluster. The default for this parameter, after some tuning to H1 data, is 1.0, in contrast to the default of 3.0 in versions prior to 5.7. In HERWIG the default intrinsic $p_{\perp}$ of partons in the proton is set to 0. A value of 0.7 for the width of the Gaussian, obtained from fits to EMC data, is used in this study. More important is QSPAC, the cut-off for initial state parton shower development. Energy flows generated with HERWIG are slightly dependent on its value and it should be chosen as low as possible. CLMAX is the fission threshold for high mass clusters. Clusters heavier than the quark masses plus 3.0 GeV will undergo a fission process, instead of an isotropical decay to hadrons.

The model will be referred to as HERWIG with the following parameter settings[5]:

[5]These settings are the result of a tuning effort, to the H1 data discussed in this thesis, by J. Chyla, one of the authors of HERWIG.

parameter	meaning	default value	value used
IPROC	process number	none	9000 (DIS with s.u.e.)
PRSOF	prob. for s.u.e.	1.0	1.0
CLMAX	max. cluster mass (GeV)	3.5	3.0
PSPLT	cluster split spectrum	1.0	0.5
QSPAC	cut-off for initial PS	2.5	2.0
PTRMS	intrinsic k_T (GeV)	0.0	0.7
BTCLM	max. beam cluster mass (GeV)	1.0	1.0

As for all other models, 5 flavours, the MRS $D^{-'}$(DIS) parton distributions and $\Lambda_{QCD} = 190$ MeV are used.

Chapter 5

Event Selection and Kinematics

Aladdin said, a trifle comic,
"If you're the djinn then were's the tonic?"

– Roald Dahl

In this chapter the details of the event selection procedure are presented, along with further cuts to home in on the events desired for this analysis. The approach used here is to enable the potential reader to reproduce the results, to this end names of the banks containing the relevant information within the H1 software, are given. All information concerning events is stored in so-called BOS banks [46], which are named using a four letter code. The system represents an entity-relationship model, i.e. the information is arranged in tables and relationships between different tables, for more information see [47, 48].

5.1 Event Selection

The events used for the analysis were selected from a preselection of DIS events provided in the form of a data summary tape (DST). This preselection makes use of an electron signature in the BEMC to tag low, i.e. $Q^2 < 100$ GeV^2, NC events. Events are required to have a cluster in the BEMC with an energy of more than 4 GeV, no hit inside the background time window of the ToF behind the cluster, and to have at least one hit in the BPC, with a distance less than 15 cm to the cluster centre of gravity. In addition the event classification requires a vertex reconstructed from a forward or central track.

A second selection step refines the above electron cuts and makes some additional cuts to ensure a clean sample of low Q^2 NC events. A description of these cuts is given in the following list. The cuts are applied to both, the 1992 and 1993 data.

1. $|z_{vertex}| < 50$ cm, determined from a track

   At least one good (see next section) central track is required in each event to ensure a well measured event vertex. The z-position of the event vertex is restricted to ± 50 cm from the nominal interaction point, according to the proton bunch length of 50 cm in the 1992 running period. The bunch length was 25 cm for the 1993 data.

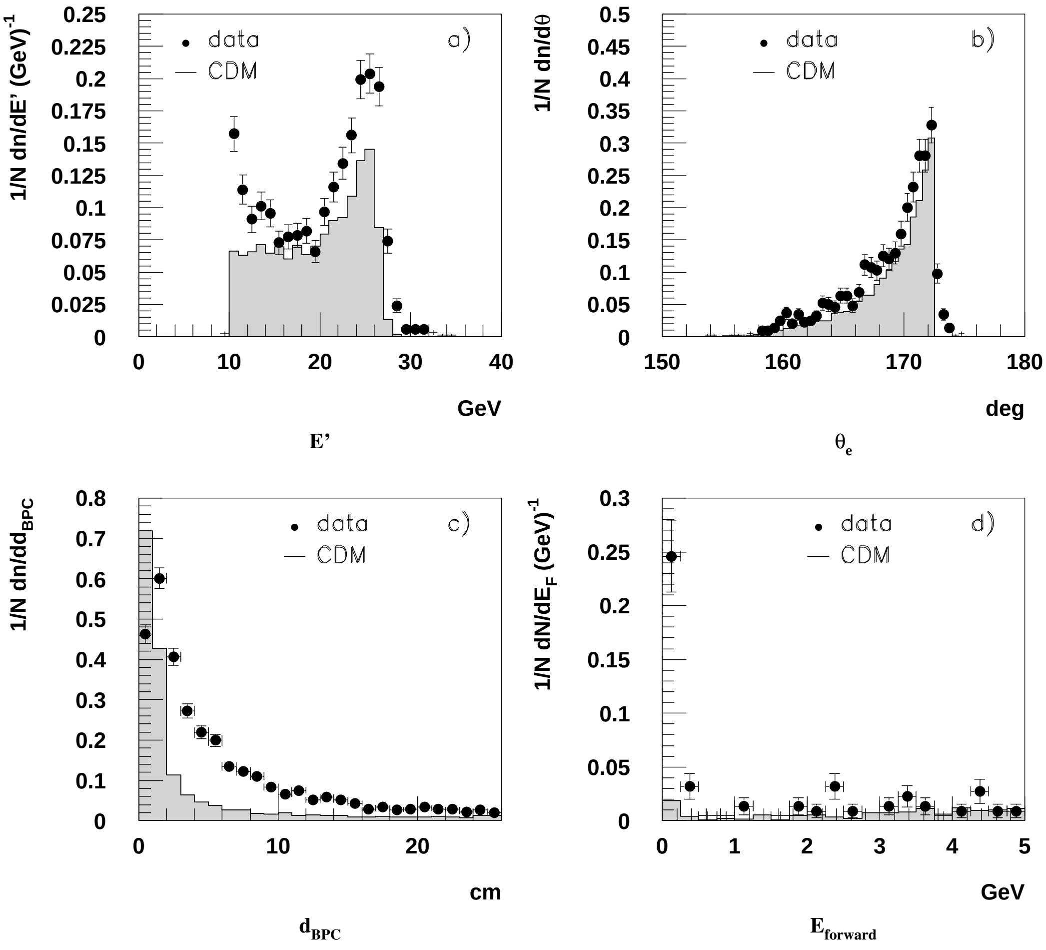


Figure 5.1: *Distributions of event properties that result in important event selection cuts. Compared are data with reconstructed CDM Monte Carlo. The plots are made before the actual cut is applied, but after the preceding cuts. a) Electron energy, b) electron angle, c) distance of e^--cluster centre of gravity from BPC hit, and d) energy in $3.5 < \eta < 2.0$. All plots are normalised to the number of events, N.*

2. $E' > 14$ GeV, $\theta_e < 172.5°$

   The electron energy cut, using the energy from the BEFR bank, is raised to 14 GeV, to reduce the photoproduction background. Studies have shown [10] that this background is negligible above 16 GeV and is at the 5% level between 14 and 16 GeV. The energy distribution for the data, before any cuts are applied, is compared to the CDM Monte Carlo in fig. 5.1a. Photoproduction background is responsible for the rise towards small electron energies, whereas the kinematic peak is clearly seen at around 26 GeV. The θ_e of the electron, determined from the BPC hit discussed below, fig. 5.1b, has to be smaller than 172.5^o and the cluster centre of gravity should be outside the region of the triangular stacks in the BEMC. Both cuts make sure that the cluster is well contained in the BEMC, so the energy measurement is reliable and no leakage is expected.

3. BPC hit within 5 cm, $r_{BPC} < 60$ cm

   To ensure a better measurement of the electron angle and to further reduce photoproduction background, the BPC hit is required to be within 5 cm of the cluster centre of gravity. If there is more than one hit fulfilling this criterion the closest one to the cluster is taken. The distribution before the cut is shown in fig. 5.1c, the difference between data and the Monte Carlo is due to the presence of photoproduction background at this stage of the event selection. After all cuts are applied the remaining photoproduction background in the 1992 sample is estimated to be less than 0.5%, as determined in a Monte Carlo study of the hard photoproduction background. An upper cut on the radius of the BPC hit, corresponding to an angular cut $\theta_e > 157°$ for events at the nominal vertex position, ensures the electron to be away from the outer edge of the BEMC.

4. $W^2 > 3000$ GeV^2

   To have sufficient hadronic activity in the detector and to avoid the problematic low y region, where QED radiative corrections are large, a cut at $W^2 > 3000$ GeV^2 is made.

5. Veto wall

   Events containing muons travelling parallel to the beam are rejected, based on their early time of arrival, compared to the interaction time. This time difference is deduced from the trigger elements of the veto wall.

6. $E(4.4° < \theta < 15°) > 0.5$ GeV

   In fig. 5.1d, the energy in the forward region of the calorimeter is compared to CDM Monte Carlo predictions. The big spike in the data at $\lesssim 0.5$ GeV is not present in the Monte Carlo. This difference stems from so-called rapidity gap events [49], which due to their diffractive nature deposit very little energy in the forward region. Since in this analysis the data are compared to standard DIS events from Monte Carlo generators that do not contain diffractive processes, these events have been removed by requiring at least 500 MeV energy in the forward region.

Finally 877 events survive all selection criteria. For the 1993 data the selection yields 15239 events. The selected events have an average Q^2 of 25 GeV^2 and lie in the x-range from $10^{-4} - 10^{-2}$.

5.2 Track Selection

5.2.1 Central Track Selection

Charged particles are mainly measured in the central drift chamber (CJC). The tracks are used to define the event vertex, and also determine the event kinematics, together with the calorimeters; they are also used for measuring energy flows and charged particle spectra. The track information is extracted from the CTKR bank. z-chamber information has not been used, due to the low efficiencies in the linking for the 1992 data. The selection criteria described below have been applied for both, the '92 and the '93 data.

Well measured charged tracks in a fiducial volume, with large and homogeneous efficiency, are selected by the following cuts.

1. 0.15 GeV $< p_\perp <$ 20 GeV.

   The lower cut in transverse momentum, $p_\perp$, selects tracks which can traverse both CJC rings and do not curl back over. Below that cut the track finding efficiency is found to deteriorate when studying single track Monte Carlo events. The higher cut protects against rare cases of grossly mismeasured tracks with apparent large momenta that distort the whole event if they are not excluded. The $p_\perp$-distribtuion is shown in fig. 5.2a.

2. $22° < \theta < 158°$.

   The track polar angle θ is restricted to an efficiency plateau, determined in a Monte Carlo study, comparing simulated and reconstructed tracks. Selected tracks are at least fully contained in the inner part of the CJC (CJC1, see fig 2.2). The drop of the efficiency outside this range can be seen in fig. 5.2b. For low Q^2 events, with the electron measured in the BEMC, this cut also eliminates the track of the scattered electron.

3. $|z_0 - z_{\mathrm{vertex}}| <$ 20cm

   This cut selects tracks that come from the primary vertex, the distribution is presented in fig. 5.2c. It will deselect parts of curling tracks, backscattered tracks and tracks from decays that coincidently point back to the main vertex in $r - \phi$.

4. $DCA <$ 3cm.

   Tracks from the primary vertex are selected by restricting the distance of closest approach (DCA), shown in fig. 5.2d, to the primary vertex in the $xy-$plane.

5. $L >$ 10cm.

   The track length has to be large enough to allow a measurement of track parameters to some precision. This cut is almost equivalent to a cut in the

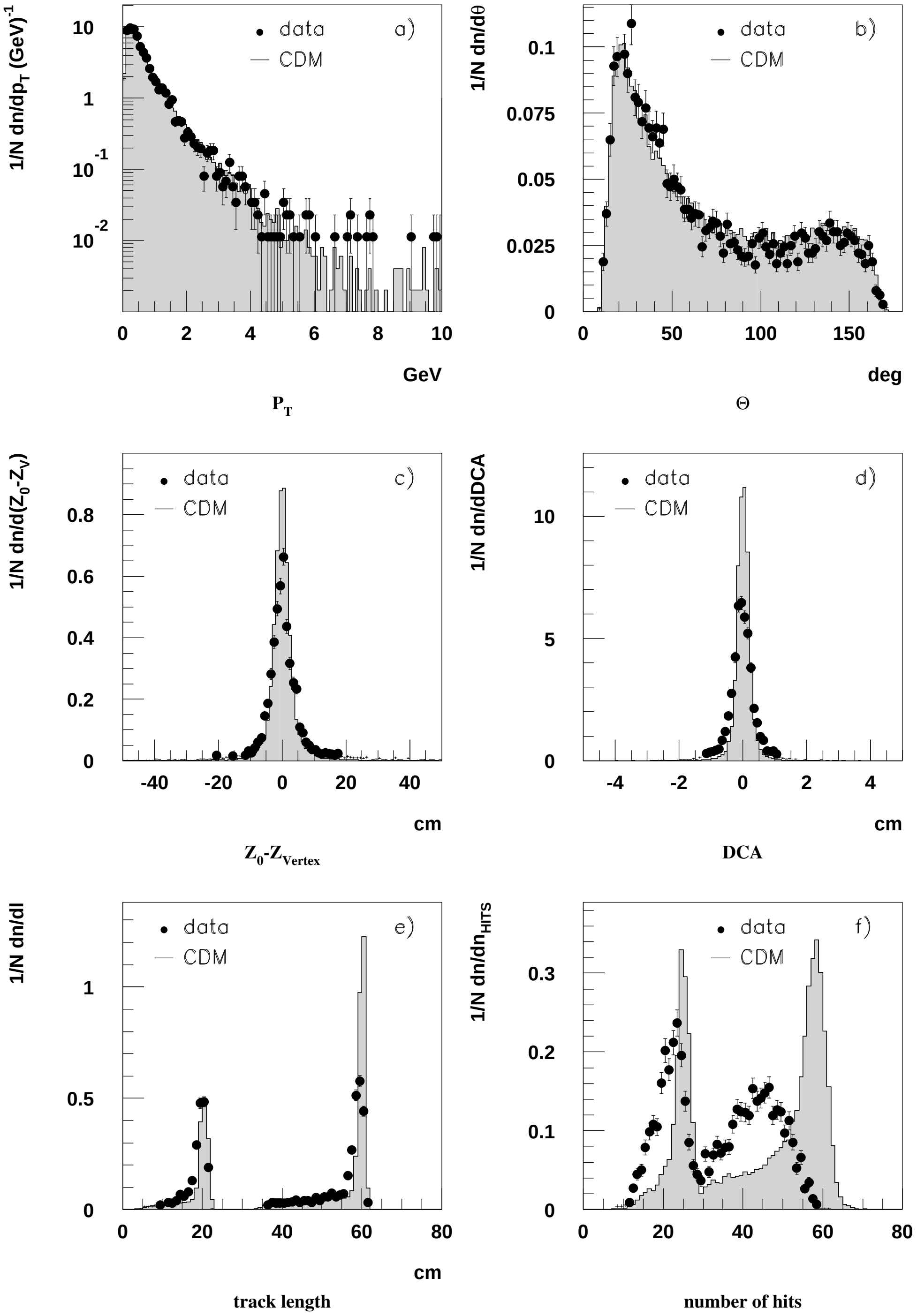


Figure 5.2: *Distributions of central track variables in which cuts are applied. Compared are 1992 data with reconstructed CDM Monte Carlo. The plots are made before the actual cut is applied, but after preceding track cuts. The order of figures corresponds to the order of the cuts.*

number of hits, $n_{\rm hit}$, of the track. The two peaks in fig. 5.2e represent tracks that traverse only CJC1 (left) or both chambers (right).

6. $n_{\rm hit} > 10$.

   A cut in the number of hits is applied to remove some rare tracks that pass the length cut, but have less than 10 hits. The difference in the hit distribution between data and MC, fig. 5.2f, is due to electronics malfunctioning and other hardware effects which are not accounted for in the detector simulation.

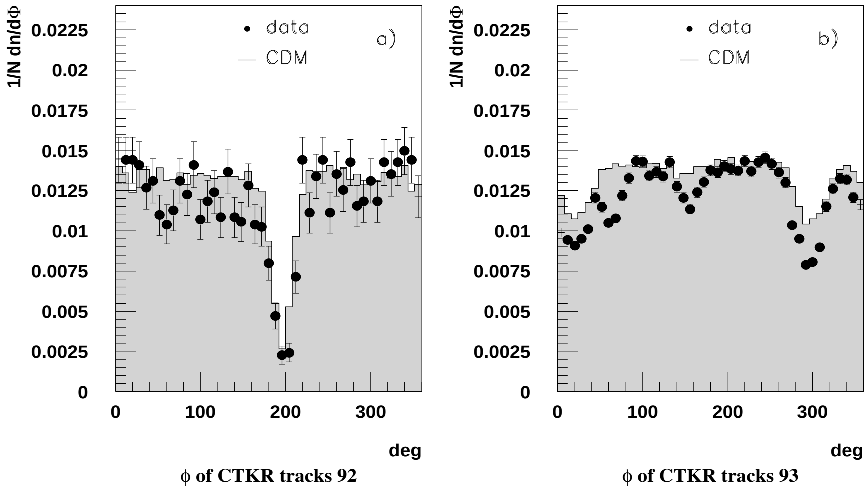


Figure 5.3: *The ϕ distribution for central tracks after all selection cuts for a) 1992 and b) 1993. Compared are data with reconstructed CDM Monte Carlo. In a) a factor 0.92 has been applied to rescale the MC, this reflects the lower track reconstruction efficiency in the data. The holes are due to dead sectors in the CJC 1 that are not perfectly reproduced by the event simulation.*

Wires in CJC1 which had a reduced efficiency are set to 'dead' in the Monte Carlo and also in the data. This is necessary since inefficient cells in CJC1, at ϕ around $190°$ in 1992, cause holes in the ϕ distribution of charged tracks, see fig. 5.3.

The track cuts and hence the plots are made during the event selection procedure. The basic electron identification and electron cuts have been made at this stage, but the W^2 cut and subsequent cuts are still to come (see section 5.1). The quantities are shown before the cut is applied, thus all but the cuts 5,6,7 have been applied when the DCA distribution is plotted.

An important issue is the track finding efficiency. It is defined as the probability of a charged particle, which satisfies the acceptance cuts $(p_\perp, \theta)$, to give rise to a reconstructed CJC track satisfying the selection criteria. From a scan of the data, counting visible particles (aligned hits), which do not lead to a reconstructed track, this track finding efficiency was determined to be 0.92 ± 0.03. Reasons for lost tracks are hardware problems, like occasional malfunctioning of electronics. These problems are not present in the Monte Carlo. There is some indication that negative tracks are more affected than positive ones, and that mostly tracks between 0.1 GeV and 0.5 GeV are affected.

5.2.2 Forward Track Selection

Due to the incomplete state of the readout electronics the forward tracker had a very low efficiency in the first data taking period of the H1 experiment in 1992. Thus the forward track selection discussed here is applied exclusively to the 1993 data. When selecting forward tracks special attention has to be paid to the problem of secondary tracks. Due to dead material in front of the forward tracker secondary interactions give rise to high numbers of secondary tracks. Depending on the track cuts used, of the order of 50% of the tracks can stem from secondary interactions or decays.

The following cuts were chosen to select well measured, primary tracks in the forward tracker. A more detailed discussion can be found in [50].

1. $8° < \theta < 22°$

   There is a sharp drop in the reconstruction efficiency of forward tracks below 8-9°, as tracks get shorter, due to the limited angular acceptance. Since the bulk of the tracks are at low angles, the lower of the two numbers has been chosen as the cut-off. The θ-distribution, shown in fig. 5.4a, extends beyond the upper cut-off of 22°. To avoid double counting of non-linked tracks between the central and the forward tracker, forward tracks are not used above 22°, this is the polar angle where the efficiency of the central tracker is known to be sufficiently high for central tracks to be used.

2. $p > 0.6$ GeV

   To obtain a clean sample of tracks coming from the primary event vertex a cut in the track momentum is used. Below 0.6 GeV, the contamination with secondary tracks is very high and the track reconstruction efficiency is quite low. A higher cut in p (≈ 1 GeV), would reduce the amount of secondary tracks even further, however, since the momentum distribution shown in fig. 5.4b peaks at low momenta, the loss of tracks would be too high, and so this stricter cut is not applied.

3. $\chi^2_{\text{vertex fit}} < 20$

   Tracks that do not originate from the primary vertex will only coincidentally obtain a good, i.e. low, χ^2 of the track vertex fit. Thus, a cut in this quantity will reduce the amount of selected secondary tracks. The cut chosen is very loose, since a stronger cut would also result in a severe loss of good tracks.

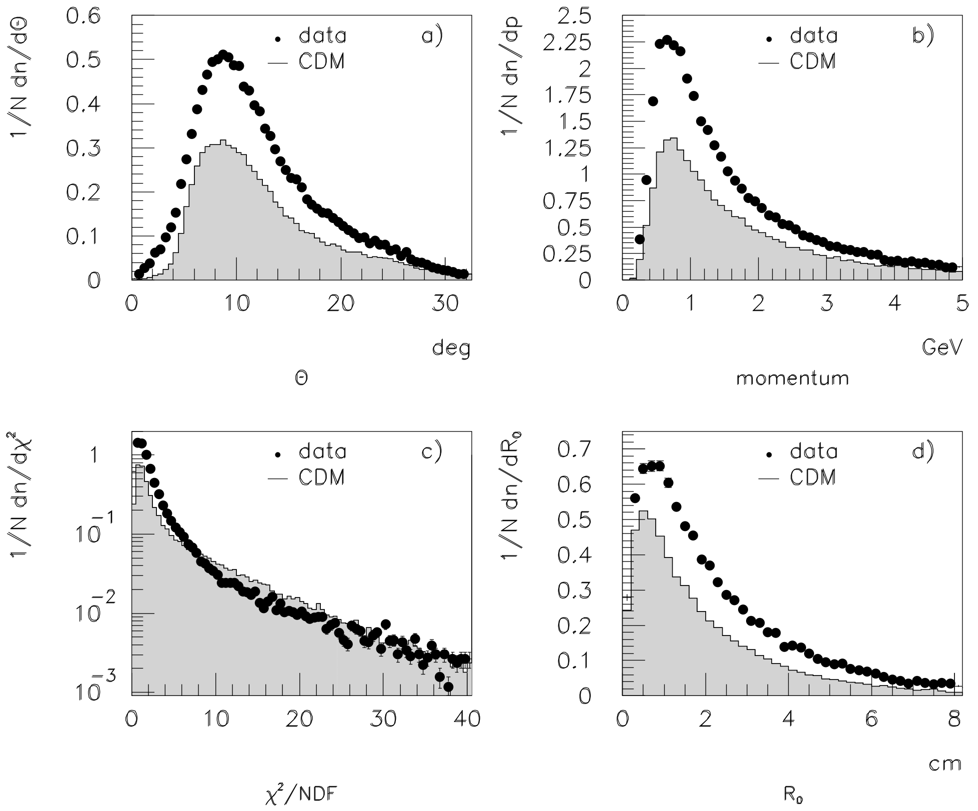


Figure 5.4: *Distributions of forward track variables in which cuts are applied. Compared are 93 data with reconstructed CDM Monte Carlo. The plots are made before the actual cut is applied, but after the preceding cuts. The order of figures a)-d) corresponds to the order of the cuts.*

4. $\chi^2_{\text{track fit}}/\text{NDF} < 10$

   This cut is to deselect badly measured or badly reconstructed tracks which have unreliable track parameters. The cut is again loose, due to the current lack of in depth knowledge of the forward tracker performance and reconstruction quality. The number of degrees of freedom (NDF) corresponds to the number of hits of a track. The χ^2 distribution is shown in fig. 5.4c.

5. $R_0 < 5.0$ cm

   A final cut, to maximise the probability that the tracks are coming from the primary vertex, is to make a cut on the distance of the track, in the x-y-plane, from the vertex, R_0. A compromise between a low contamination with secondary tracks and acceptable losses of good tracks is the chosen value of 5 cm. The distribution of R_0 is shown in fig. 5.4d.

After these cuts there is still a significant difference between the number of forward tracks in the data and in the Monte Carlo. The CDM Monte Carlo predicts $\approx 20\%$ fewer tracks than observed in the data. From the distributions shown in fig. 5.4, there is no indication that low momentum or small angle tracks, i.e. tracks in the regions of poor efficiency of the tracker, cause this difference. One possible reason is that tracks consisting only of one planar segment are also accepted, if they passed the selection cuts. Such tracks have poor momentum resolution, an effect that could be insufficiently described by the detector simulation. Further investigations are needed to understand the reasons for the observed discrepancies.

5.3 Calorimetric Measurements

In this low Q^2 analysis the scattered electron is detected exclusively in the BEMC. Calorimetric measurements of the hadronic final state are based mainly on the liquid argon calorimeter (LAr), but include also the tail catcher and the BEMC. Clusters of energy depositions are used for energy flow measurements, as obtained from the generalised clustering step in the reconstruction. For the BEMC, the cells are used directly, because clustering in the BEMC was optimised for electromagnetic showers in the analysis of the 1992 data. The clusters are supposed to represent the best energy estimate, including weighting and dead material corrections. For electromagnetic particles, one particle, in general corresponds to one cluster, while for hadrons usually one particle gives rise to many clusters.

Electronic noise in the LAr calorimeter is suppressed using a three stage procedure. Firstly, on-line electronic noise suppression is provided on a channel-by-channel basis of the electronic read-out channels. Here, for each channel, the distribution of the pedestal is measured then fitted by a Gaussian and a cut at the level of ± 2 standard deviations (σ_{noise}) is applied. During the reconstruction step a topological noise cut is applied to the calorimeter cells. Cells with an energy above $4\sigma_{noise}$ are kept as are cells having energies above $2\sigma_{noise}$, provided that they are neighbours to a cell with an energy $> 4\sigma_{noise}$. The level of noise differs over the calorimeter range, with $\sigma_{noise} = 10 \ldots 30$ MeV equivalent energy.

Due to its non-compensating nature, the signal of the LAr is $\approx 30\%$ smaller for hadrons than for electrons of the same energy. A software weighting technique [8], based on the identification of electromagnetic clusters due to their higher energy density, attemps to correct for this. Weighting constants, derived from single particle test beam measurements, are extended with the help of Monte Carlo simulation to be applicable for jets of particles, too. The weighting is only applied to so-called prominent clusters, i.e. to clusters with $(\sum_{cells}(E_i/\sigma_{noise})^2)^{1/2} > 8$.

The only cut applied to LAr clusters during the event selection concerns the so called ETNS (event topological noise suppression), the last stage of the noise suppression procedure. Clusters containing cells marked by this algorithm as noise are rejected. ETNS determines regions with hadronic activity by looking for prominent clusters and suppresses cells outside these regions (more than 50cm away). For the BEMC cells a noise cut of 0.5 GeV is applied (equivalent to $3\sigma_{noise}$), and they are scaled with an ad hoc factor of 1.6, accounting for the fact that one tries to measure hadrons in an electromagnetic calorimeter [51].

In fig. 5.5 some basic properties of clusters are shown. There is very good agreement between data and MC, indicating a high level of understanding of the calorimeter simulation. It is worth stressing that the Monte Carlo even gives a good description of quantities like the number of cells in a cluster and the energy fraction in the LAr.

Most important for the calorimetric measurements is the absolute calibration. Here the understanding of the '92 data taking period is summarised. The BEMC has been calibrated to $\pm 2\%$ using the kinematic peak of DIS events. The electromagnetic scale of the LAr has been determined to $\pm 3\%$ with electrons and positrons, from cosmic induced electromagnetic showers, comparing the calorimetric energy measurements with the track momenta. The $p_\perp$ balance of the hadronic final state measured in the LAr, along with the electron measured in the BEMC, enables the hadronic scale to be determined to a precision of $\pm 6\%$.

From the low Q^2 DIS sample cross checks of the calorimeter calibration and performance are possible. A basic quantity, that can also be used to select DIS events, is the $E - p_z$ of the event. From energy-momentum conservation this should be two times the electron beam energy, 53.4 GeV, since the electron adds $1 - y$ to the formula 5.1, given below. Fig. 5.6a, which shows $E - p_z$ determined from calorimetric measurements, signals good agreement between the data and the Monte Carlo. The $p_\perp$ balance of the electron vs. the hadronic system, shown in fig. 5.6b, provides a check of the overall hadronic scale of the liquid argon calorimeter. The ${p_\perp}^h/{p_\perp}^e$ has a mean of less than 1, for both data and MC. This can be attributed to low energy particles which do not reach the sensitive part of the calorimeter. From a study of this distribution a scaling factor of 1.06 for the energies measured in the LAr has been derived for the 1992 data, which has already been applied in the plots shown.

5.4 Kinematics Reconstruction

At HERA the event kinematics can be measured from the electron alone, from the hadronic system, or a combination of the two. Since the beam energies are known, only two kinematic variables are independent, thus the kinematics is overconstrained. Any combination of hadronic and electron measurements can be used to determine the kinematic quantities of an event. When comparing different methods, the following points are used to judge their performance:

- good resolution
- insensitivity to QED radiative corrections,

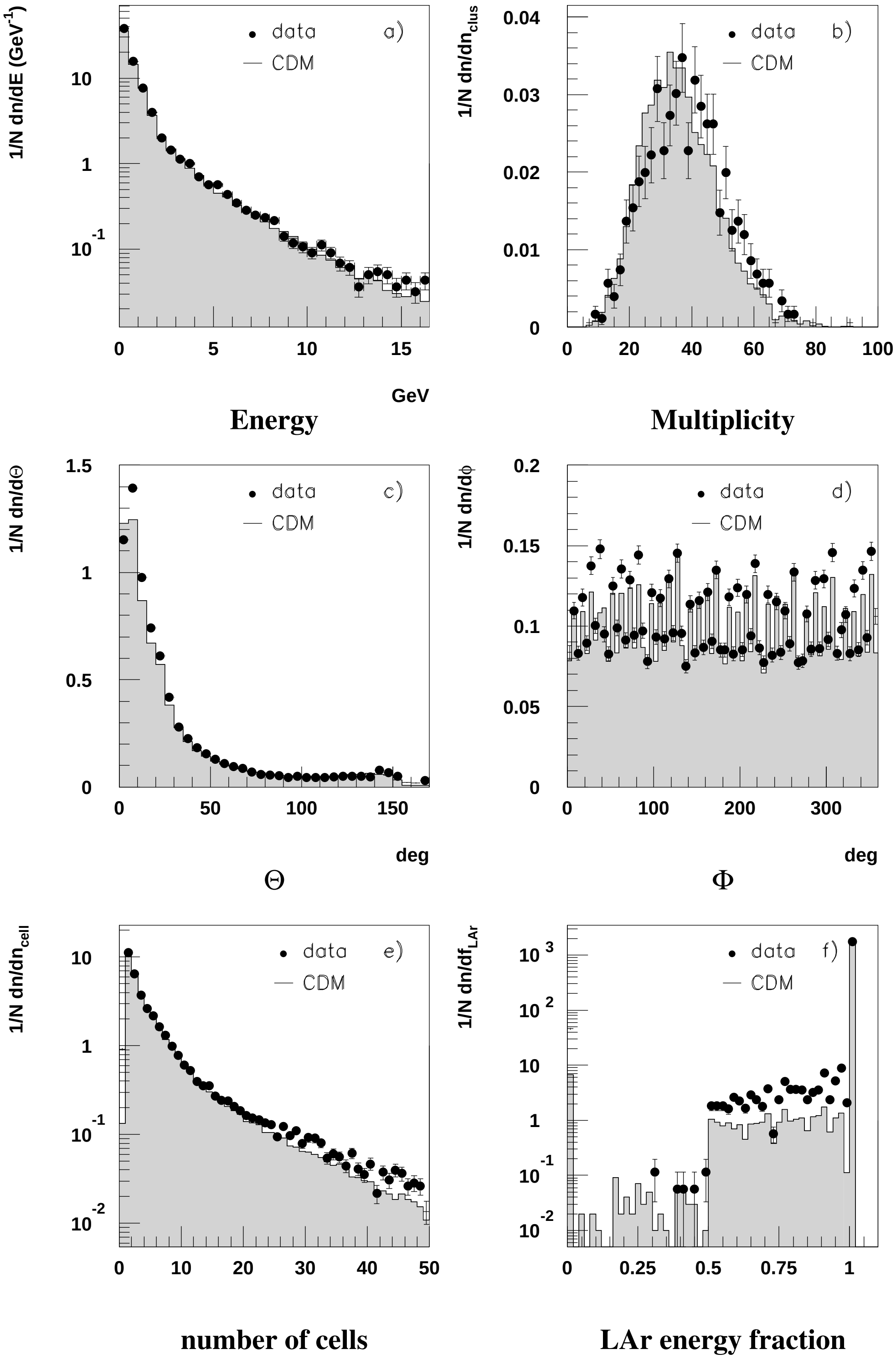


Figure 5.5: *Distributions of cluster properties from 1992 data compared to reconstructed CDM Monte Carlo. a) Cluster energy, b) Cluster multiplicity, c) Polar angle, d) Azimuthal angle, e) Number of cells in a cluster, f) Energy in LAr divided by the total cluster energy (in all calorimeters).*

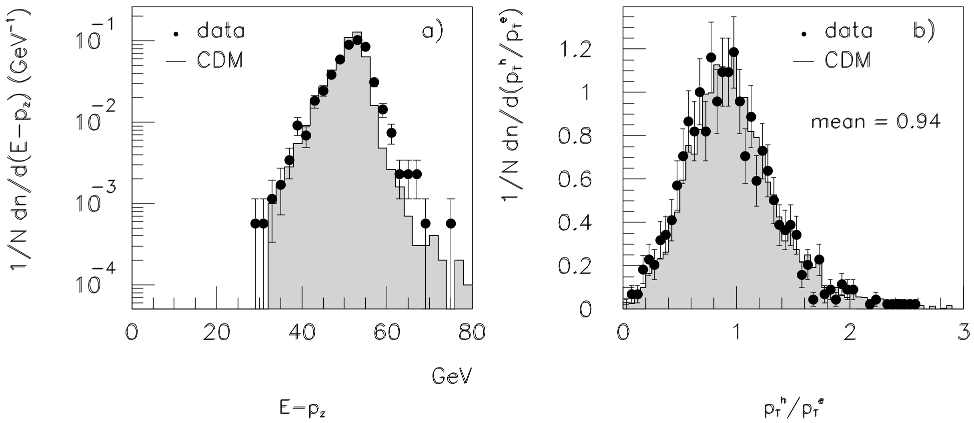


Figure 5.6: *Calorimetric checks for two important event quantities; 1992 data compared to CDM Monte Carlo. a) $E - p_z$ from calorimeter alone, b) ${p_\perp}^h/{p_\perp}^e$ from calorimeter.*

- insensitivity to the calorimeter calibration,
- insensitivity to the limited hadronic measurement, due to the beam pipe holes in the angular coverage.

The last point has been tackled by Jacquet and Blondel [52], who proposed to measure y according to the following formula:

$$y = \sum_{hadrons} \frac{E_h - p_{z,h}}{2E_e}. \tag{5.1}$$

Since most of the proton remnant escapes undetected in the beam pipe, this method provides a nice way out of the 'impossibility' of a kinematic measurement using the hadronic final state. Proton remnant particles have very low values of $E - P_z$, thus the influence of the undetected remnant on y is considerably diminished. However, particles escaping through the backward beam hole deteriorate the y measurement.

The kinematic variables determined from the electron alone have a very high precision, but suffer significantly from QED radiative effects. The cross section for QED radiation from the incoming and scattered electron at HERA is very high, especially in the region of low and high y. Fortunately, radiation of the scattered electron is predominately collinear with the electron, thus the measured cluster in the BEMC in the vast majority of the cases will contain both, the electron and the γ, so the kinematics stays unaffected. No such easy way out exists for the initial state QED radiation, and this is the main reason to combine electron and hadron kinematics, which is insensitive to QED radiation.

Combining the Jacquet-Blondel method for the hadron kinematics with, for example, the Q^2 as determined from the electron will address the second and last criterion mentioned above, but the third one remains unmatched. This was the main reason for the development of the double angle method (see e.g. [53]). Since in this approach only the electron angle and the angle of the hadronic system are used, the method will be insensitive to the absolute calorimeter calibration. Using an extension of the original method, it can also be made insensitive to QED radiative corrections.

Which method is best suited will always depend on the analysis goals. If the highest possible precision is required in a limited region, then the electron measurement is superior at $0.1 \lesssim y \lesssim 0.7$. For the present analysis the main emphasis was on having a method that provides good precision in a large range of y and is insensitive to QED radiative effects. Combining y from Jacquet-Blondel with the Q^2 from the electron, the so-called mixed method, fulfills both requirements and was therefore chosen. The calorimeter calibration is well understood for the 1992 data, thus it seems preferable to use the mixed method as it provides a slightly better resolution than the double angle method.

To determine the $E - p_z$ of the hadronic system, both calorimeter information and tracks measured in the central tracker are used. Double counting of energy is avoided by masking calorimeter cells behind a track in a cylinder of 15(25) cm in the electromagnetic (hadronic) calorimeter, The contribution of the tracks to the y is about 40%, reducing the energy scale uncertainty and improving the y resolution.

Fig. 5.7 shows the reconstruction quality of the kinematic variables as derived from the CDM Monte Carlo after full detector simulation and reconstruction. The achieved resolutions are 5% for Q^2 and 21% for x and y.

To obtain a better understanding of the forthcoming analysis details, it proves useful to look at dependencies of quantities characterising the final state, on the kinematical variables. Choosing the x-Q^2-plane as the two independent kinematic variables, the electron energy and angle, W and the direction of the current jet in the QPM are presented in fig. 5.8; isolines of y are also given. With these pictures in mind it is easier to interpret the features of the event selection and the hadronic final state. For instance it can be seen that a cut at $W^2 > 3000$ GeV^2 corresponds to a cut $y > 0.03$ and that the direction of the current jet strongly depends on x.

5.5 The Transformation to the Hadronic CMS

The hadronic centre of mass frame (CMS) is the rest system of the beam proton and the exchanged boson. Both particles collide in the hadronic CMS head on along

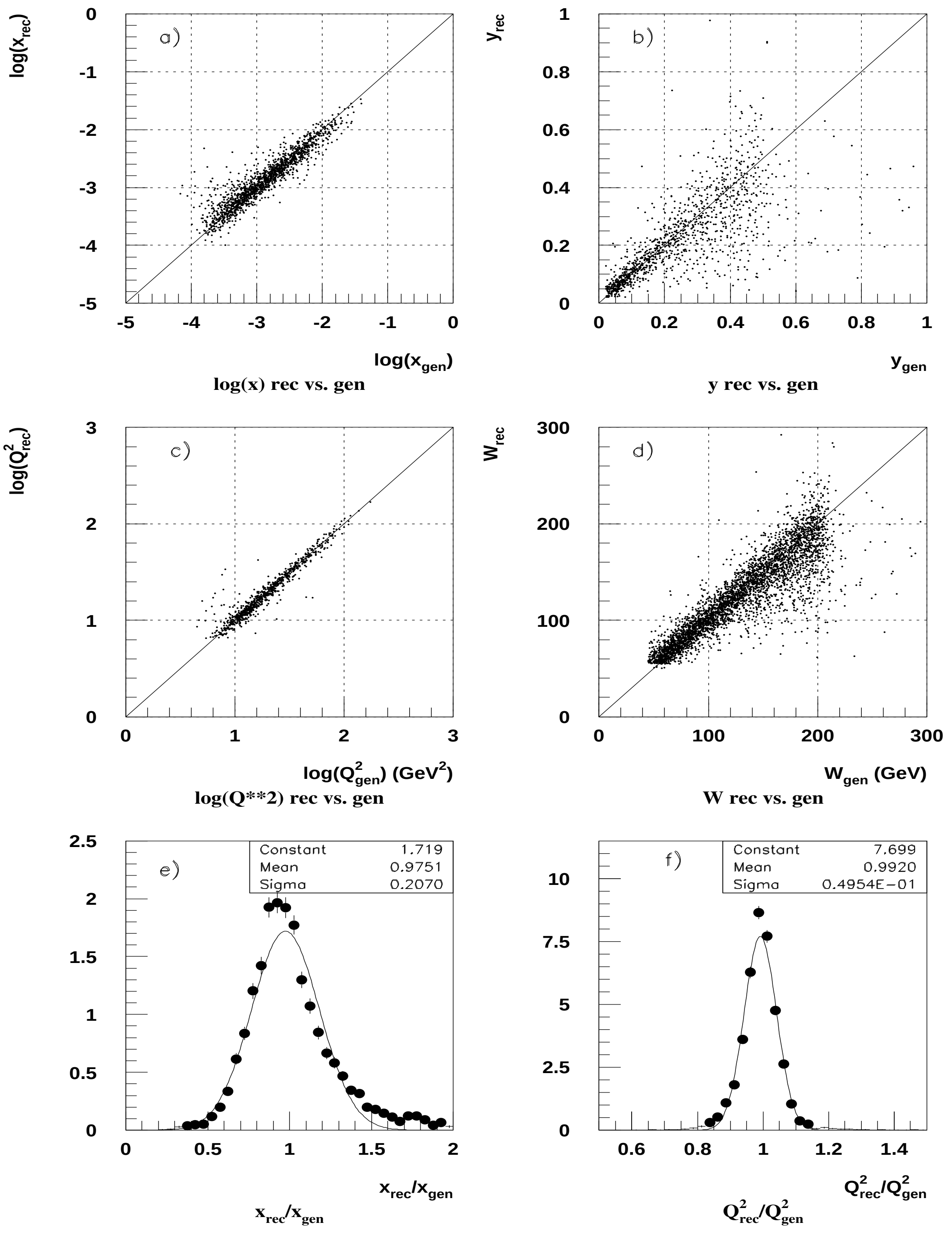


Figure 5.7: *Quality of the kinematics reconstruction using the mixed method and CDM Monte Carlo. Compared are the generated kinematic variables to the reconstructed ones. Note that CDM does not include QED radiative corrections, but studies with radiative Monte Carlos yield very similar results.*

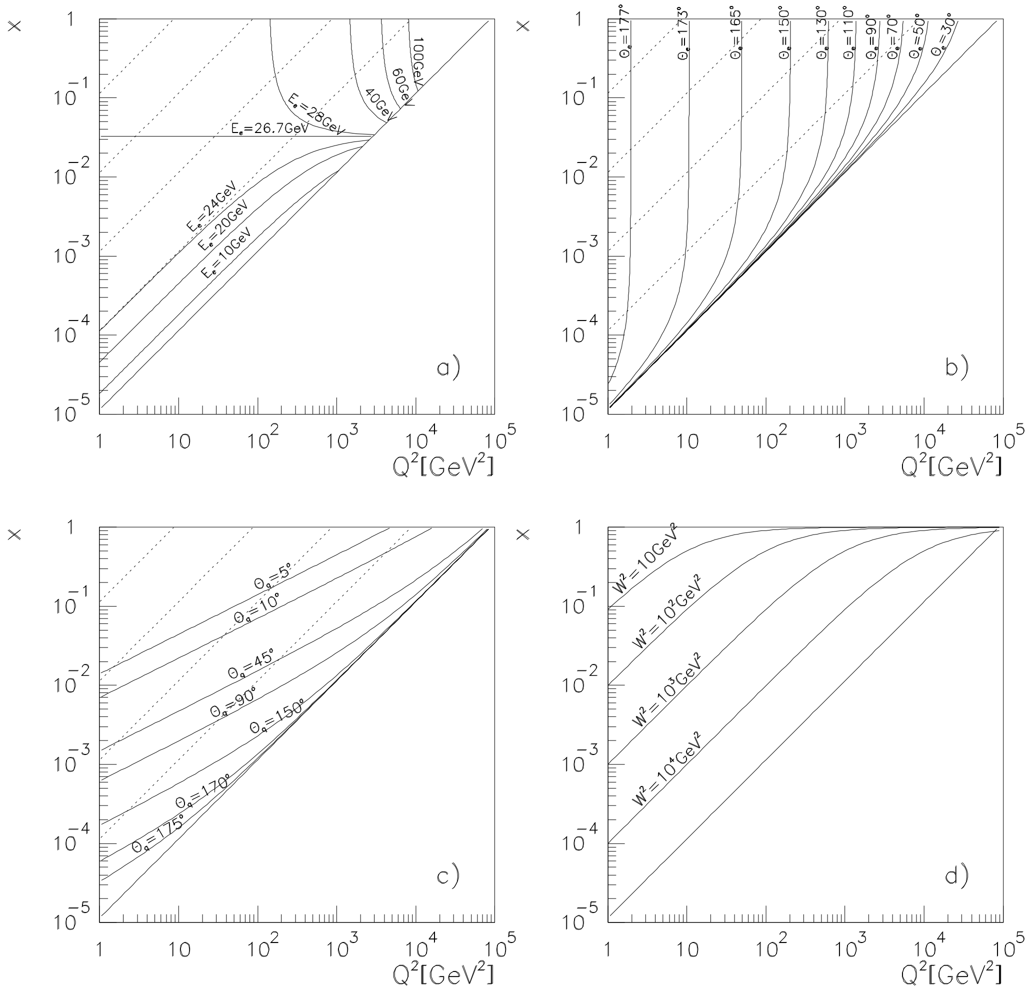


Figure 5.8: *Kinematical quantities for HERA in the basic x-Q^2-plane. The dotted lines are isolines in y, in logarithmic steps, where the kinematic border corresponds to $y = 1$. Shown are a) the energy of the scattered electron, b) the polar angle of the scattered electron, c) the polar angle of the scattered quark in the QPM, d) W^2.*

the z-axis, giving rise to a scattered quark and a proton remnant flying apart in opposite directions. Obviously the transformation to this frame requires knowledge of the vector of the exchanged boson. In the outlined ideal world the boson is known from the scattered electron, there is no reconstruction uncertainty, which gives rise to a boson tilted to the 'true' z-axis, in the hadronic CMS. One aim of the transformation must therefore be to minimize effects arising due to the non-ideal environment, when one deals with reconstructed quantities. After outlining the strategy chosen for this analysis, the influence of the effects will be discussed briefly.

The transformation to the hadronic CMS makes use of the mixed kinematic variables, explained above. The boson is defined as the 4-vector difference between the 'outgoing' and incoming electron. This 'outgoing' electron is calculated from the mixed kinematic variables instead of using the measured quantities directly. Using this method, QED radiative corrections, that otherwise would result in wrongly reconstructed bosons, can be partially corrected for. The achieved resolution in reconstructing the boson can be found in fig. 5.9. The resolution of the transverse component is better than that of the longitudinal one. For these plots it makes no difference if the boost is built as described or the measured electron alone is used.

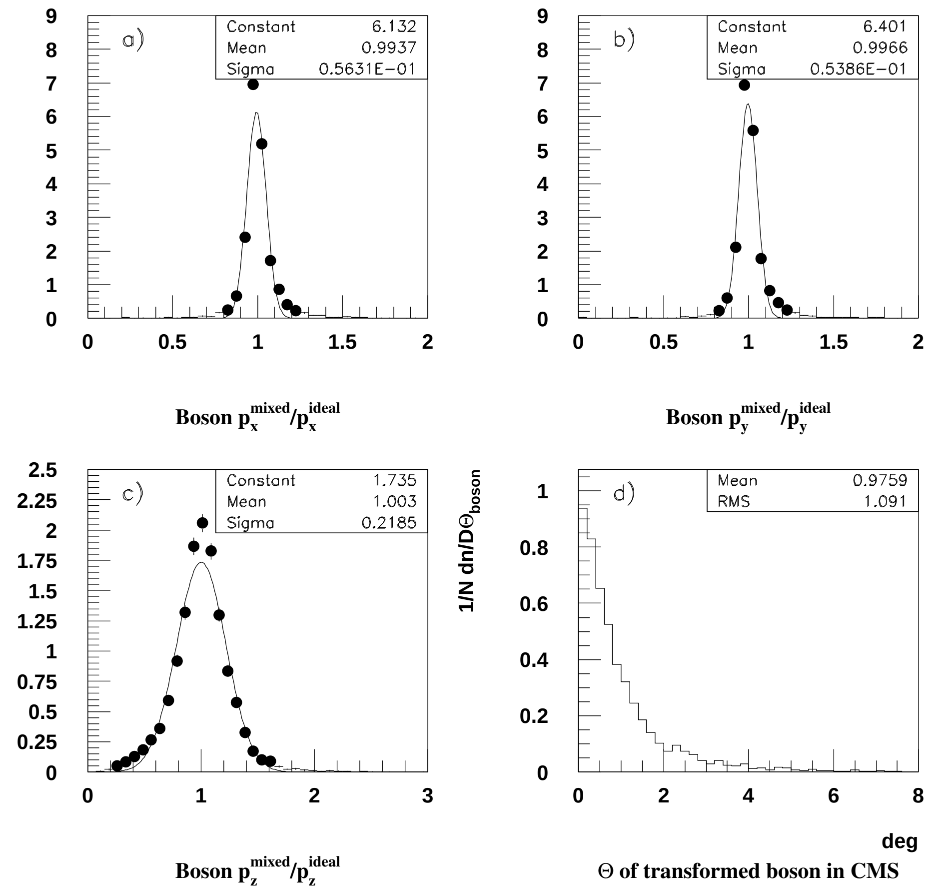


Figure 5.9: *Achieved resolutions for the reconstructed boson from the mixed method, for the CDM Monte Carlo. For a)-c) the reconstructed boson is compared to the ideal one, a) P_x, b) P_y c) P_z. d) θ of the ideal boson in the hadronic CMS. Transforming the ideal boson to the CMS, using the mixed method, gives the angle by which the hadronic CMS is tilted to the ideal z-axis.*

A simple, but very effective, test of the quality of the transformation, is to transform the ideal boson itself to the CMS, using the described mixed method for the boost. Ideally the boson should be along the z-axis, thus its polar angle should be zero. That this really is the case has been proven using the ideal boson for the boost. Utilising the reconstructed boost results in a non-zero polar angle of the ideal boson in the CMS, which can be interpreted as the angle by which the CMS is tilted to its nominal direction. This effect gives rise to additional $p_\perp$ of particles in the CMS, solely as an effect of the boost. The main reason for the angle being so small (see fig. 5.9d), is the low Q^2. The boost is almost entirely longitudinal, and so the error on the transverse component is small.

Fig. 5.10 shows that the mixed method is very well suited for the transformation to the hadronic CMS. The difference of the two superimposed histograms illustrates that on average very little $p_\perp$ is gained by using the reconstructed boost instead of the ideal (generated) boson. Since the gain in $p_\perp$ is considerably less when using the mixed method than when using the electron alone for the transformation, this method has been chosen. The resolution in p_z^* in the hadronic CMS[1], shown in fig. 5.11a, is of the order of 13%, although the shape is not Gaussian. When combined with the resolution in W of 20% this gives the resolution for x_F, being defined as $2p_z^*/W$. The resolution in ${p_\perp}^*$ is given in fig. 5.11b, here only a small error is introduced by the boost to the CMS. Both resolutions cover only the error introduced by reconstructing the boson from the mixed kinematic variables. For a given Monte Carlo event the generated particles have been transformed to the hadronic CMS, using the ideal boson, and again, using the boson derived from the mixed kinematics (after detector simulation and reconstruction). So additional effects, due to the $p_\perp$ resolution of the tracker, are not accounted for in the numbers quoted, but they are small compared to the effect of the boost.

[1]Quantities in the hadronic CMS are identified with a '*' superscript if the quantity also exists in the laboratory frame.

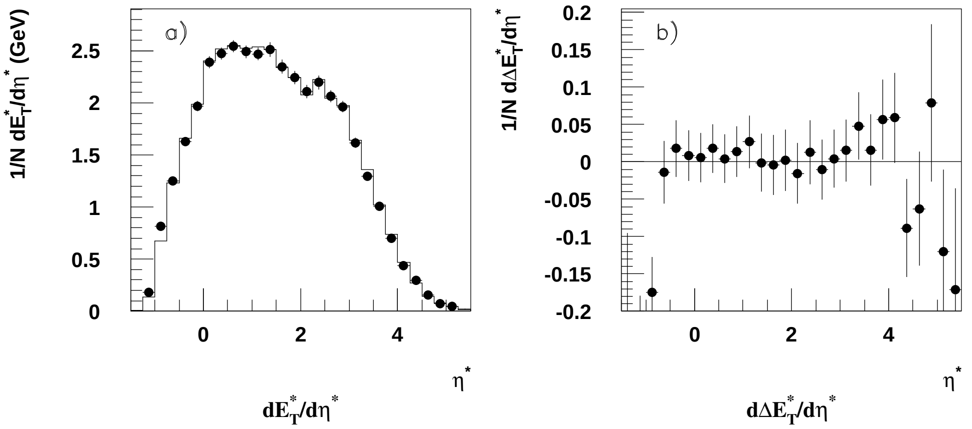


Figure 5.10: *Energy flows in the hadronic CMS as function of the pseudorapidity, η^*, comparing the mixed transformation method and the ideal boost, CDM Monte Carlo. The '*' denotes quantities in the hadronic CMS. a) E_T-flow using the ideal boost (full line) compared to flow obtained using the mixed method (dots), b) difference of the two, normalised to the former.*

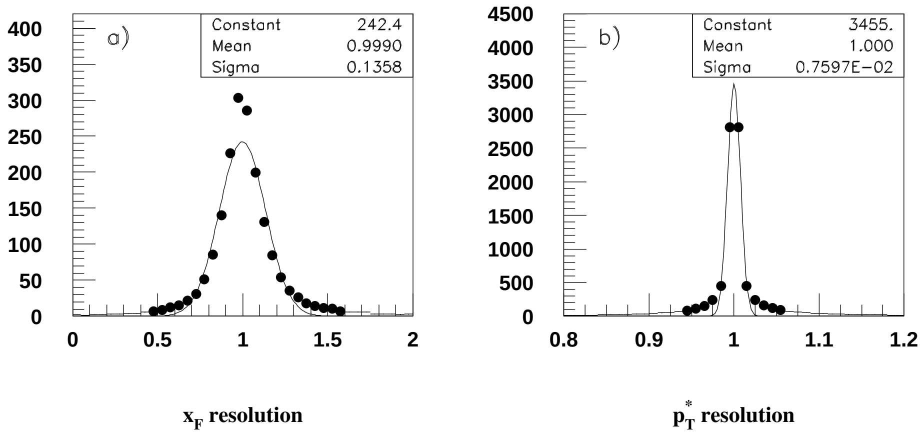


Figure 5.11: *Resolutions for x_F and $p_\perp{}^*$ by comparing particles transformed with the mixed boost to particles transformed with the ideal boost. Note that this resolution does only cover detector effects on the boson, as generated particles have been used. a) x_F resolution, b) $p_\perp{}^*$ resolution.*

Chapter 6

Energy Flows

Die Mehrheit der Irdioten aber setze Ihr Vertrauen in die Wissenschaft, die ungerührt und unverdrossen jedes Phänomen untersuchte. Sie war formalisiert, ordnete alles ein und wunderte sich über nichts.

– Stanisław Lem

In the quark parton model, the hadronic final state consists of a current quark, scattered at large angles into the detector, and a remnant di-quark, parallel to the incident proton. The transverse momentum of the scattered electron is essentially balanced by the current quark. QCD radiated soft and hard gluons populate the region between the struck quark and the di-quark. Hard parton emission in QCD Compton and boson-gluon fusion processes leads to two, well separated, partons, in addition to the target remnant, a so-called 2+1 jet event. Non-perturbative hadronisation and particle decays tend to wash out the underlying parton topology. However, perturbative QCD effects should still be observable, particularly in the region between the current quark and the di-quark. The effects of the initial state parton shower should be observable towards the remnant. Very little knowledge exists so far about remnant fragmentation; just one of the many good reasons to take a closer look into the region between the current and the remnant jets.

In the following, QCD effects in DIS events are studied in terms of the flow of transverse energy of the hadronic final state in both the laboratory frame and the hadronic CMS. The comparisons are made exclusively with the data from the 1992 run period. Since the event topology strongly depends on the kinematic variables x and Q^2, energy flows are investigated separately for events with $x > 10^{-3}$ (2/3 of the events in the sample) and $x < 10^{-3}$. The latter region is of particular interest as it has not yet been explored much experimentally.

The data are corrected for detector effects using a bin-by-bin correction method. This allows a direct comparison with flows obtained using hadrons generated by one of the generators listed in table 6.1, after the appropriate kinematic cuts are made. Thus one can easily study the influence of variations of generator parameters

Table 6.1: *The Monte Carlo generators used to compare with the corrected data.*

Abbreviation	Model Features	Generator	See Sect.
QPM	no QCD radiation, string fragmentation	LEPTO 6.10	4.1.1
PSWQ	parton shower model with $W\sqrt{Q^2}$ scale, string fragmentation	LEPTO 6.10	4.1.3
MEPS	$O(\alpha_s)$ matrix elements with added parton showers, string fragmentation	LEPTO 6.10	4.1.4
CDM	colour dipole model with interface to $O(\alpha_s)$ matrix elements, string fragmentation	ARIADNE 4.04	4.2
HERWIG	coherent parton showers, cluster fragmentation, soft underlying event	HERWIG 5.7	4.3

(tuning) without the need for time consuming detector simulation. Some of the features of the hadronic final state have been calculated recently at the parton level [54], a comparison to these calculations is also given.

After presenting, in detail, the correction procedure applied to the data and verifying that it gives the desired results, in the second section the resulting energy flows are presented. The emphasis of the study lies on the QCD effects in the region between the current and the remnant jets. Dependencies on x and W^2 are investigated in an attempt to prove the existence of a new QCD dynamics at low x. The final sections deal with the systematic errors of the distributions and the effects of varying model parameters.

6.1 The Correction for Detector Effects

There are several methods available to correct experimental distributions for detector effects. The matrix unfolding provides the correct mathematical approach to tackle the problem, but suffers from the need of very high Monte Carlo statistics, to determine the usually sparsely populated, off-diagonal regions of the unfolding matrix. Under normal circumstances the iterative bin-by-bin correction method provides sufficient precision and fast enough convergence. This analysis follows closely the approach of the OPAL collaboration, described in detail in [55]. A bin-by-bin correction method, taking into account experimental resolutions, is used and it is shown that one iteration is sufficient at the current level of statistical and systematical errors.

To unfold the measured distributions, bin-by-bin correction constants are derived for each distribution individually. They take into account the effects of finite detector resolution and acceptance, remaining electronic noise, losses in dead material, as well as the scattering of particles from secondary interactions into the fiducial volume. To obtain these constants two Monte Carlo samples are required, with and without detector simulation. In the former case event selection cuts as described in section 5.1 are applied. For the pure generator sample only the physics cuts are applied:

1. $E' > 14$ GeV
2. $157° < \theta_e < 172.5°$
3. $W^2 > 3000$ GeV^2
4. $E_F(4.4° < \theta < 15°) > 0.5$ GeV

On the generator level all particles with a lifetime greater than $8ns$ are considered as stable (H1 convention). The term 'generator particles' includes all stable charged and neutral particles, including neutrinos. Furthermore, the generator particles are restricted to the same angular region as the calorimeter ($4.4° < \theta < 174°$). This provides more direct and better control over the correction constants; otherwise one could not disentangle the effect of the additional energy observed in the forward region from the losses of energy due to the limited acceptance, these effects will be discussed in section 6.1.2. Applying no acceptance cuts on the generator level, would correspond to an extrapolation of the measurements based on the Monte Carlo predictions, which requires better understanding of the corrections and a better description of the data by the Monte Carlo.

6.1.1 The Unfolding Procedure

Let REC_i^{MC} be the contents of bin i of a particular distribution, as obtained using the Monte Carlo sample with detector simulation and N_{rec}^{MC} the number of events in this sample. For the generator sample the corresponding quantities are GEN_i^{MC} and N_{gen}^{MC}. Then the correction constant for bin i of this particular distribution is given by:

$$C_i = \frac{(REC_i^{MC}/N_{rec}^{MC})}{(GEN_i^{MC}/N_{gen}^{MC})} \tag{6.1}$$

Dividing the experimentally derived distribution bin-by-bin, with the factors obtained, results in the unfolded (or corrected) experimental values[1].

To ensure this correction procedure to work properly one important requirement is to obtain a good description of the experimental distributions by the Monte Carlo sample that has undergone the detector simulation. The H1 detector simulation is based on the GEANT [56] package and contains a detailed description of the detector geometry and of the materials used. Electromagnetic and hadronic showers in the

[1]Note that this definition differs from the usual one, where numerator and denominator are interchanged, resulting in multiplicative factors.

calorimeter are simulated using fast parametrisations for the shower development [57]. The response of the calorimeters has been extensively tuned using test beam and ep data [58]. The simulated events are subject to the same reconstruction as experimental data. As will be shown later, the Monte Carlo used for the correction describes the energy flows very well.

For the method to work it is also crucial that the bin-to-bin migrations are small. In this case the method resembles the standard matrix unfolding procedure with a diagonal matrix. Speaking in experimental terms, one has to adjust the bin widths to the experimental resolutions in the quantity plotted, otherwise a significant bias can be introduced into the unfolded distributions, as demonstrated in [55]. The resolutions will be discussed below.

Last, but not least, the method should be applied iteratively, i.e. after obtaining correction factors using the Monte Carlo that describes the data best one has to apply these factors to the data and repeat the detector simulation using these corrected distributions as the input. From this second step new corrections factors will be obtained and after correcting the data with these new factors the old and the new corrected distributions should be compared. If the two are equal the method has converged, otherwise further iterations are to be carried out. The need for iteration is a consequence of the detector effects being a function of the input distributions of the hadronic final state. Different particle density and energy distributions will result in a different detector response.

Since the iteration method described requires a lot of computer time, an estimate for the convergence has been used instead. The correction factors obtained from the CDM have been applied to MEPS Monte Carlo after detector simulation. Comparing MEPS, corrected with the CDM, to the generated MEPS distributions gives an estimate of the error when only one iteration is used. Studies have shown that this error is indeed very small and so one iteration is sufficient for the correction of the energy flows presented.

6.1.2 The Corrections for the Energy Flow

The above mentioned requirements will now be investigated, using as an example the energy flow in the laboratory system. The correction constants are shown as a function of the pseudorapidity, $\eta = -\ln\tan\frac{\theta}{2}$, in a histogram form, fig. 6.1c,d, for the two different x ranges. The Monte Carlo used for deriving the corrections is CDM, since it provides the best description of the data, as seen in fig. 6.1a,b and explored further in the next section.

The vertical lines in the figures indicate the range of the detector without hadron calorimetry (BBE and BEMC). This explains the large correction constants in the backward region. In the barrel range ($-1 < \eta < 1$) the corrections are almost 1, indicating very small detector effects. The rise in the forward region requires special consideration. Systematic studies of this effect revealed [59] that the increase in

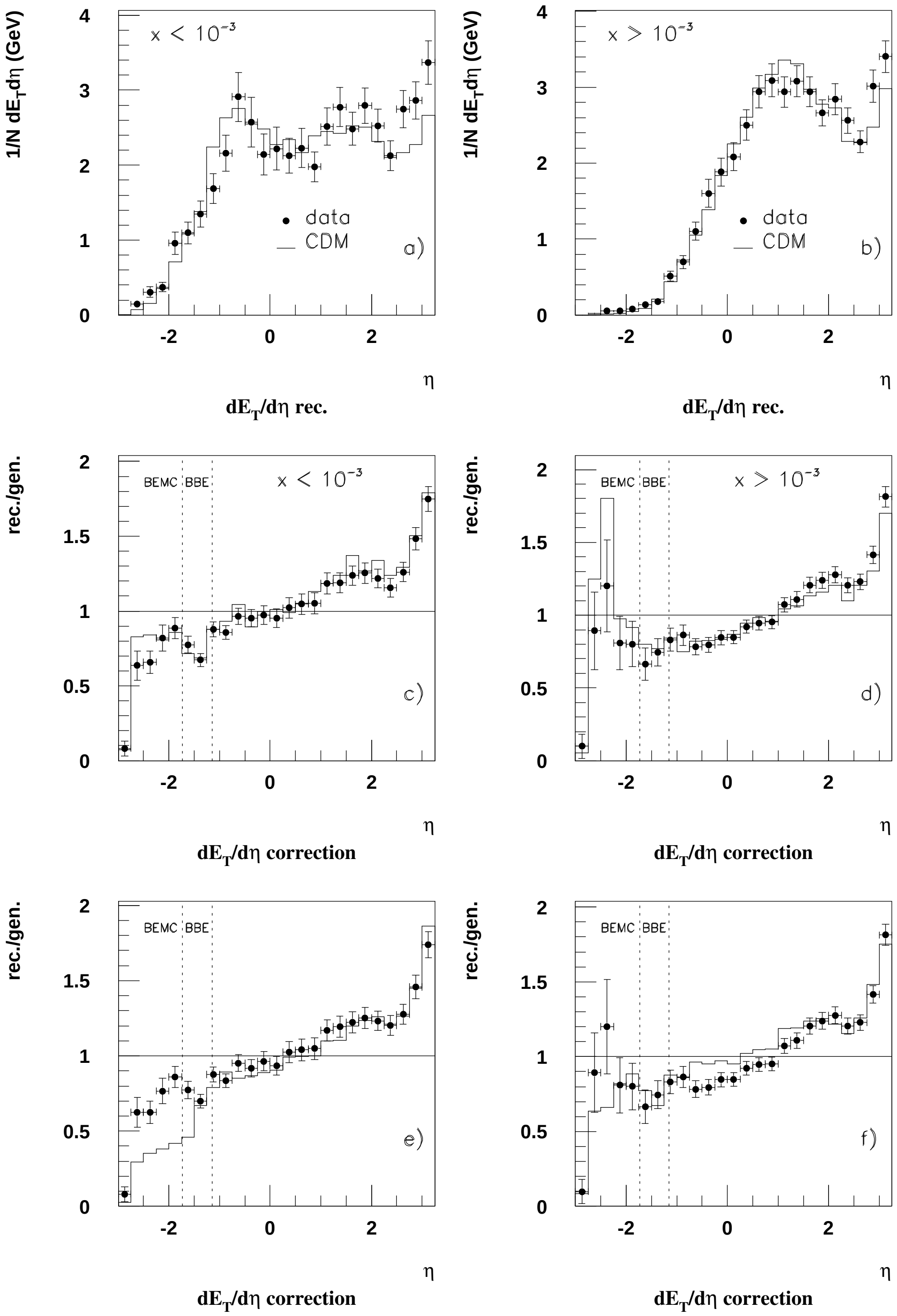


Figure 6.1: *a),b) Uncorrected transverse energy flow in the lab, with statistical errors only. c),d) Correction constants for the $E_\perp$-flow comparing CDM (circles) with MEPS (full line). e),f) Corrections with and without migrations for the CDM Monte Carlo (see text).*

the forward region is due to secondary interactions in dead material in front of the calorimeter. In the angular range from 1 to 4° the dead material amounts up to 5 hadronic interaction length. Also backscattering, off the cryostat walls, into the calorimeter and spread of secondaries out of the beam pipe, contribute to the effect. It is very satisfying that the size of the effect is similar when using the CDM and the MEPS Monte Carlo, demonstrating that the effect is fairly independent of the model used. These models otherwise have a very different prediction for the energy flow in the important forward region.

An important question is whether the large corrections in the forward region are independent of the kinematical range investigated. When subdividing the sample into two x bins, the correction factors account for two effects:

- possible differences in the detector resolution and response for different input distributions (energy flows), and
- possible migrations between the x bins, due to the finite resolution in the reconstructed x.

To investigate solely the first effect, the following modification to the correction procedure has been employed: Instead of using the generator sample, with its generated kinematics, the generated energy flow of events, selected according to the kinematics *after* detector simulation, has been used in the denominator of equation 6.1. This separates the effect of the detector resolution from the migrations due to the reconstruction of the kinematics. If the observed effect is indeed independent of the kinematical range, overlaying the modified corrections for $x < 10^{-3}$ and $x > 10^{-3}$ should give the same result. Fig. 6.1e proves that this is the case for the CDM Monte Carlo. The difference in the backward region can be attributed to the different direction of the current jet, for low and high x, resulting in very little energies in this region for $x > 10^{-3}$. Thus the errors on the full line are very large in the backward region.

In fig. 6.1f, the effect of the migrations in x has been investigated for the low x bin. Compared are the normal correction functions from fig. 6.1c and d. The additional differences between the two curves, compared to fig. 6.1e, are due to migrations between the low and the high x sample.

6.1.3 The Angular Resolution

The angular resolution of the calorimeter has to be studied for the bin-by-bin correction method and bin sizes have to be adjusted to the experimental resolutions. In order to explore the detector resolution for hadrons, single pions with energies between 0.5 GeV and 5.0 GeV have been generated by Monte Carlo and traced through the detector simulation. During the reconstruction step these pions were overlaid with electronic noise as recorded in random trigger events.

A pion on average gives rise to 3 calorimeter clusters. In addition such an event will have $\approx$ 3 noise clusters. Noise suppression is obtained by using the ETNS method described in section 5.3. Pseudorapidity resolutions were then studied in the laboratory frame; 25 bins in η are used, in the range $-3. < \eta < 3.25$. The quantity $\frac{\mathrm{d}E_\perp}{\mathrm{d}(\eta_{gen}-\eta_{rec})}$, i.e. the resolution in the pseudorapidity weighted with the transverse energy is shown in fig. 6.2a-c, for three angular regions. η_{gen} is the simulated pions

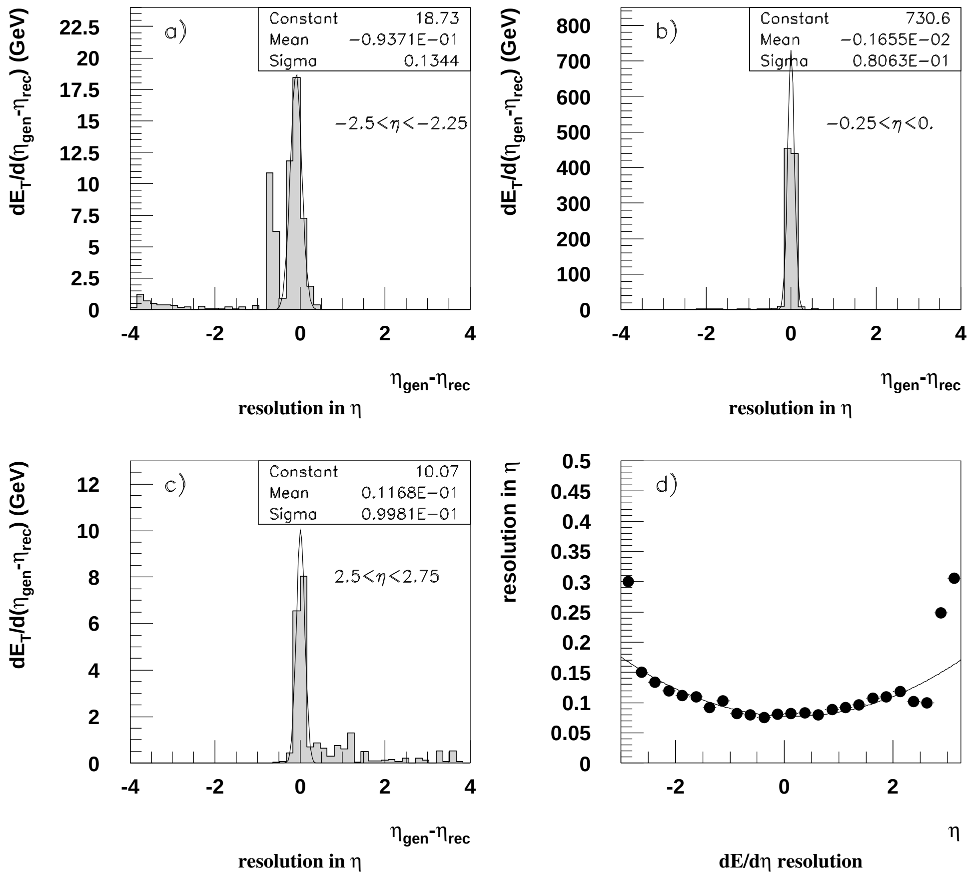


Figure 6.2: *Angular resolution of the calorimeter obtained from single pions. The energy weighted resolutions are shown for 3 angular regions a)-c) and the values of σ obtained from Gaussian fits are given in d) as a function of η.*

pseudorapidity and η_{rec} is the measured pseudorapidity of a calorimeter cluster. Each pion gives rise to several clusters and will thus make several entries in the plots. The distributions are very narrow and have almost no tails in the central region of the calorimeter, as shown in fig. 6.2b. In the backward region, fig. 6.2a, the resolution becomes worse, an effect of the insufficient response of the BEMC and the BBE to hadrons. In contrast to the backward region, the deterioration in of the resolution in the very forward region becomes even worse, see fig. 6.2c. Here backscattering and the limited granularity of the calorimeter are the major reasons for the double peak structure observed.

The resolutions, as a function of η, are summarised in fig. 6.2d by using the values for σ from the Gaussian fits as the measure. A parabola can be fitted to the points, this has been superimposed in the figure. It is seen that the chosen bin width of 0.25 is larger than the resolution for the whole angular range, except for the very forward region, where the σ is not relevant since the distributions are not Gaussian. The bin-to-bin migrations have to be less than 40%, as required by the OPAL method [55], this is obeyed only when the bin width is greater than the resolution obtained. This requirement can be met in two ways, either by using pseudorapidity bins of width $\gtrsim 0.25$ (this gives less than the 40% migrations allowed) or, when using smaller bins, the angular region has to be restricted to the more central region. Since the very forward part is of high significance for the physics study, the first approach is favoured.

Although the resolution rapidly decreases in the very forward region, one has to bear in mind that a single pion study is not necessary relevant for the corrections concerning energy flows. In the kinematic range studied, the energy flows are rather flat distributions in the forward region, so migrations should not cause severe problems for the application of the bin-by-bin correction method.

6.2 Energy Flows in Different Lorentz Systems

Corrected transverse energy flows in the HERA laboratory system and in the hadronic CMS are compared to the predictions of the Monte Carlo generators, discussed in chapter 4. The flows are shown as function of the pseudorapidity, $\eta = -\ln\tan\frac{\theta}{2}$ and $\eta^* = -\ln\tan\frac{\theta^*}{2}$, respectively. The distributions are normalized to the number of events N.

Given the special situation at HERA, where many event features are a result of kinematics, one important variable is the polar angle, θ_q, of the current jet in the QPM. Neglecting QCD effects, the direction of the current quark in the lab frame is a simple function of the kinematic variables:

$$\cos\theta_q = 1 - \frac{2E_e y}{E_e y + Q^2(1-y)/(4E_e y)} . \tag{6.2}$$

The isolines of θ_q have already been shown in fig. 5.8. This picture is no longer valid when hard QCD radiation takes place, but helps to assess the features of energy flows.

6.2.1 The Transverse Energy Flow in the Laboratory System

In fig. 6.3a, the transverse energy flow in the laboratory system is presented for $x < 10^{-3}$. The current jet goes into the backward region of the detector, the peak is clearly seen at $\eta \approx -0.6$. The distribtution for $x < 10^{-3}$ exhibits a plateau at a height of 2 GeV/unit of pseudorapidity for $\eta > 0$. Overlaid are the predictions of the different Monte Carlo generators. Although the QPM model contains effects such as primordial $p_\perp$ and string fragmentation, the description of the current jet region is not satisfactory. This indicates the influence of hard and soft QCD radiation on the structure of the current jet. The CDM model is the only one to give a good overall description of the data. All the other models fail in the forward region, $\eta > 1$. MEPS and PSWQ predict far too little energy in this region and HERWIG overshoots the data. With the exception of the QPM, the other models give a good description of the left flank of the current jet region, which should only be sensitive to final state parton showers and hadronisation.

The differences between the models and the data in the higher x bin, shown in fig. 6.3b, are not as pronounced. Here the deviations concentrate solely on the forward part, the region between the current and the target jet. Since the current jet moves forward in the detector with increasing x and Q^2, studies of the region between the current and the remnant jets become more difficult.

Tracks can of course also be used to measure energy flows. The corrected energy flow for tracks is compared to the calorimetric measurements, in fig. 6.4a,b, for the two x bins. Comparing the track flow to the Monte Carlo leads to very much the same conclusions as above, but the differences between data and Monte Carlo are not as pronounced, since the forward region is not covered by the central tracker acceptance. So only the data are compared. From isospin invariance one would expect the total energy flow to be 1.5 times the energy flow of charged particles. The calorimeter/tracker ratio is presented in fig. 6.4c,d. The data is seen to lie systematically above the factor of 1.5 expected, however, this higher factor is also observed when using the Monte Carlo generator and comparing charged only to charged+neutral particle samples.

6.2.2 Comparison to Theoretical Calculations

Transverse energy flows in the forward region in the HERA laboratory system have been analytically calculated recently [54]. Since the data presented here form the basis of the comparison in [54] and the result of the calculations is given in the form of curves, only the main conclusions from [54] will be quoted, see fig. 6.5.

The calculations, based on the BFKL dynamics [2], result in a much higher transverse energy flow in the forward region than predicted by MEPS, which is based on Altarelli-Parisi evolution. Also the Altarelli-Parisi calculations carried out in [54] give an energy flow of the order of 0.8 GeV per unit of rapidity, i.e. much less than the BFKL prediction. At low x, terms in the QCD evolution which contain $\ln \frac{1}{x}$, become large and cannot be neglected anymore, as done in the Altarelli-Parisi evolution equations. As a consequence, BFKL evolution leads, on average, to higher $p_\perp$'s in the initial state parton shower, since the ordering in $p_\perp$ in the initial parton shower, which is inherent to Altarelli-Parisi evolution, is not present in the BFKL

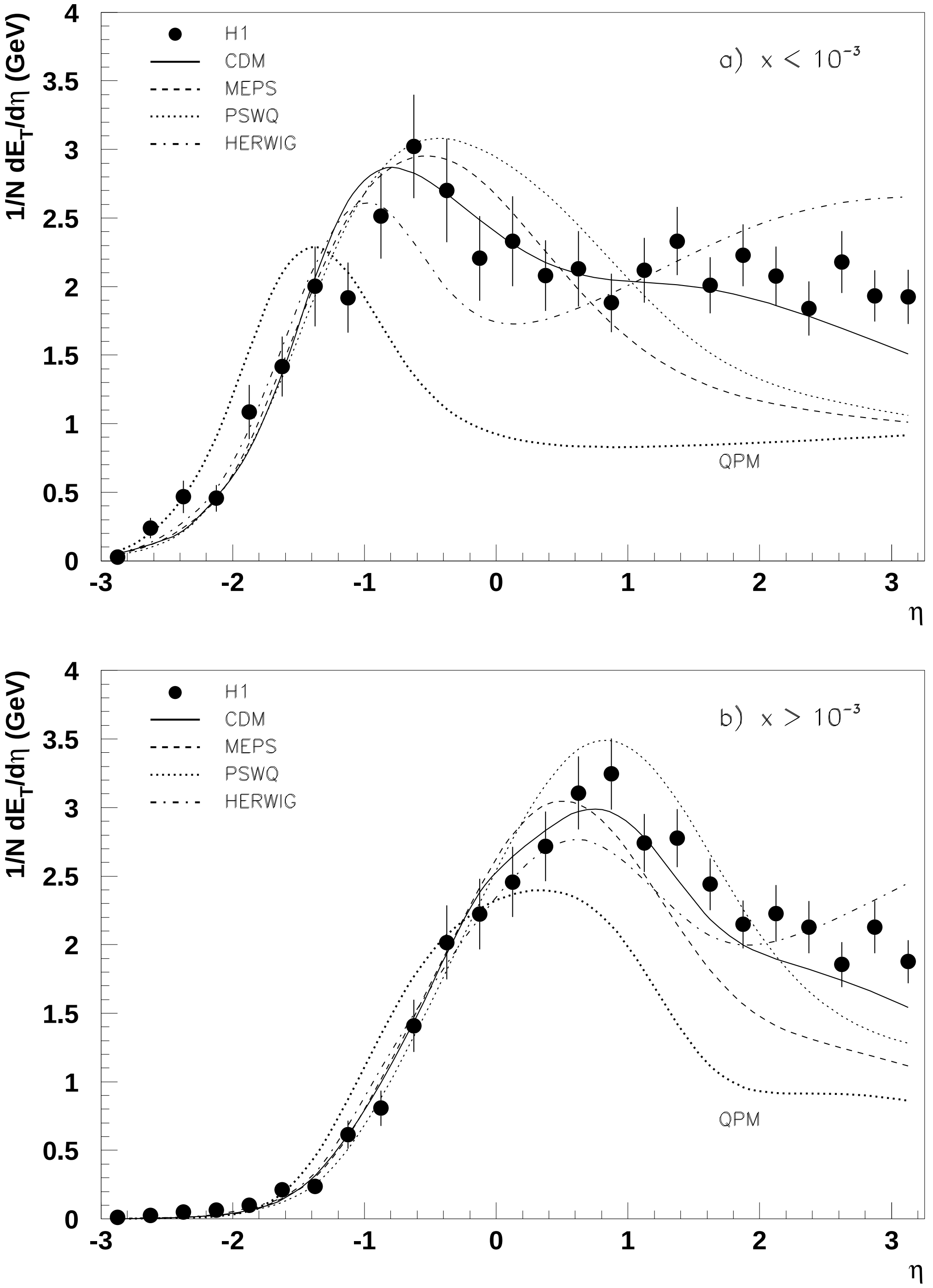


Figure 6.3: *Transverse energy flow $E_\perp$ in the laboratory system as a function of the pseudorapidity η for events with a) $x < 10^{-3}$ and b) $x > 10^{-3}$. The proton direction is to the right. The error bars contain the statistical and systematic errors added in quadrature, except for an overall 6% energy scale uncertainty.*

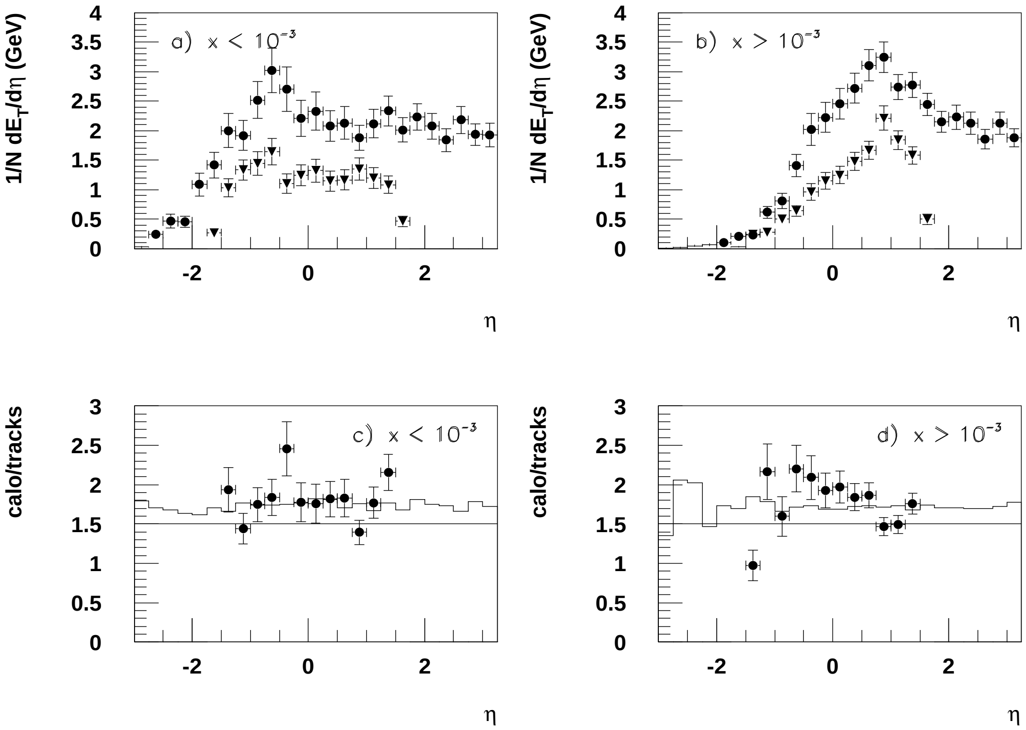


Figure 6.4: *Transverse energy flows $E_\perp$ in the laboratory system as a function of the pseudorapidity η for events with a) $x < 10^{-3}$ and b) $x > 10^{-3}$. Compared are calorimetric measurements (full circles) with tracks (triangles). The outermost triangles are lower since the binning of the track plots has been adjusted to the calorimeter, so the polar angle cut on the tracks does not correspond to the bin edges. In c),d) the ratio of the two flows is plotted. The histogram is the prediction of the CDM generator.*

dynamics [54]. The observation of this effect would be an indication of BFKL dynamics being important already in the kinematic region accessible at HERA.

The comparison with the calculation suffers two drawbacks. Firstly the calculation is carried out for a fixed x and Q^2, corresponding to the average values of the $x < 10^{-3}$ sample, to allow for the transformation to the HERA lab frame. Secondly, the size of the hadronisation corrections is highly uncertain. The estimate, derived from a modified MEPS model where the initial state parton shower has been switched off, can only be a very crude approximation.

One prediction of the theory is the dependence of the energy flow in the forward region on x. Altarelli-Parisi evolution leads to a decrease of the $E_\perp$ with x, whilst BFKL predicts an increase of the energy flow [54], for fixed Q^2. This crucial test of the signature of BFKL dynamics is studied in terms of the data in the next section.

Finally, the reader may wonder, why the CDM is able to describe the data. Ac-

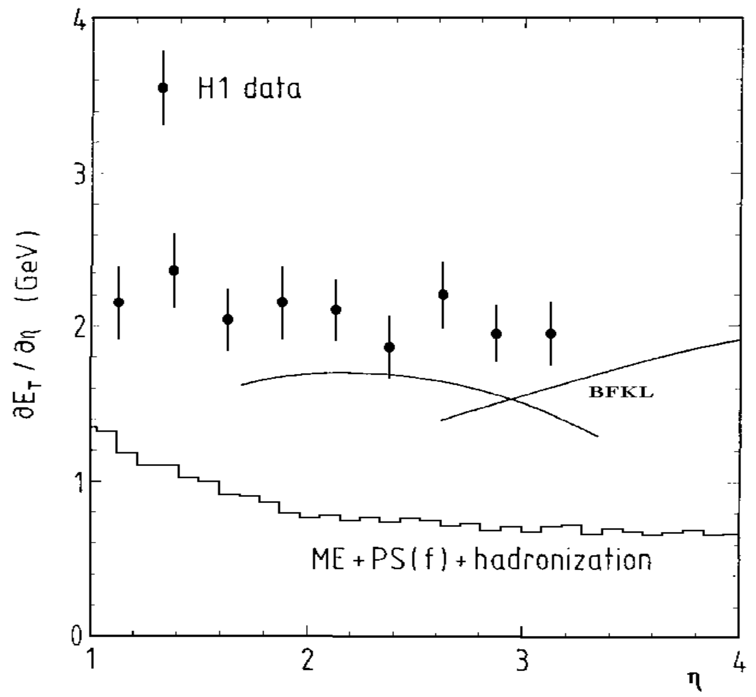


Figure 6.5: *Transverse energy flows in the laboratory system calculated using BFKL dynamics, from [54]. The low x data are compared to theoretical predictions and the ME+PS curves shows the estimated size of hadronisation effects, which in principle should be added to the BFKL lines.*

cording to [60], all the models implement Altarelli-Parisi evolution, but all modify it slightly. MEPS only modifies the first emission, whereas CDM takes into account coherence between the initial and the final state emission, as well as coherence effects, due to the extension of the proton remnant (see section 4.2). The authors of the colour dipole model claim, that the CDM sets an upper limit of possible QCD radiation, even when compared to BFKL dynamics [61], or, [25], that it is at least sufficiently close to BFKL dynamics to expect a good description in the low x domain.

6.2.3 Studying the Region Towards the Remnant

Following the path indicated by the studies of the transverse energy flow in the laboratory frame, it seems obvious to look for means to emphasize the region towards the remnant. A very nice solution to this problem is to employ the θ_q introduced in formula 6.2. Since the direction of the current jet is known purely from kinematics, it is straightforward to plot the energy flows as a function of $\eta - \eta_q$. The resulting distributions are shown in fig. 6.6.

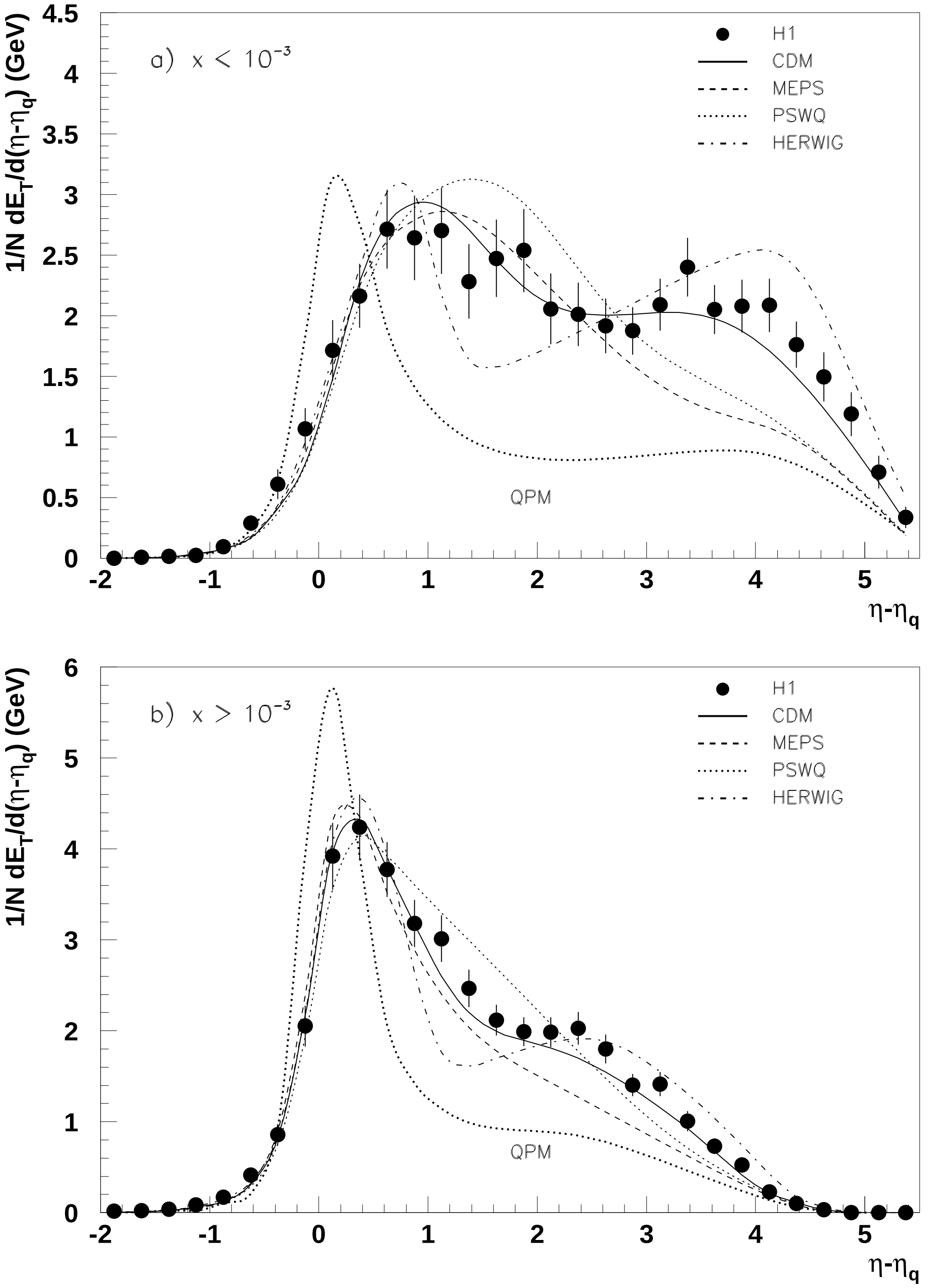


Figure 6.6: *Transverse energy flows $E_\perp$ in the laboratory system as a function of the pseudorapidity distance to the 'current quark' $\eta - \eta_q$ for events with a) $x < 10^{-3}$ and b) $x > 10^{-3}$. The proton direction is to the right. The error bars contain the statistical and systematic errors added in quadrature, except for an overall 6% energy scale uncertainty.*

It is striking that the position of the current jet is not at 0, where one would expect it in the QPM, but shifted towards the remnant. This is a kinematic effect due to initial state QCD radiation. The struck quark is not massless at the boson vertex, as expected in the QPM, and so the angle θ_q will be systematically overestimated. The small shift observed in the QPM is an effect of the string fragmentation, which of course also contributes in the models with QCD radiation.

The fall of the distributions at high values of $\eta - \eta_q$ is an effect of the limited detector acceptance. The correction procedure applied does not attempt to correct for the limited angular acceptance of the calorimeter, hence, the available range in $\eta - \eta_q$ becomes smaller with higher x, as the current jets moves forward.

Comparing the data to the QCD models it is confirmed that the CDM is the only model that is able to give an adequate description of the data. The MEPS and PSWQ model both have a similar shape, which is different from the data distribution. They underestimate the amount of transverse energy the more, the closer one gets towards the proton remnant. HERWIG also has a shape that is rather different from that of the data; the dip between the current and the remnant part, predicted by HERWIG, is not present in the data.

For the low x bin one can see a plateau in the transverse energy flow. For the range $1.5 < \eta - \eta_q < 4$ the data are almost constant at about 2 GeV per unit of rapidity. This plateau is reproduced by the CDM, and, at a much lower level, also by the QPM. In the latter case this can be explained by the hadronisation of the string spanned between the remnant and the current jet.

6.2.4 The x Dependence of the Forward Energy

From the theoretical knowledge of QCD effects at low x, one would expect the agreement between the data and all Monte Carlo generators to become better with increasing x. Special QCD effects, such as BFKL dynamics or shadowing, are expected only at $x \lesssim 10^{-3}$, thus the measured range covers both regions. To study this, one should concentrate, again, on the region towards the remnant, where the deviations are largest and also where new QCD effects are most likely to manifest themselves, as a predicted rise in energy with decreasing x. Therefore the transverse energy in the range $1.5 < \eta < 3.25$ is summed up for each event and this quantity is plotted as a function of $\log x$, as shown in fig. 6.7. The generators offer a wide range of predictions for this quantity; HERWIG shows a rise with decreasing x, CDM gives an almost flat curve and PSWQ and MEPS both give a falling distribution, but with very different slopes. The data clearly favour a flat distribution, slightly higher than that predicted by the CDM. It should be noted that the QPM also predicts a flat distribution, but at a level of 2.5 GeV less than the measured value of ≈ 4 GeV.

From the figure, one can hardly conclude that the agreement between data and the generators gets better at higher x. The agreement of HERWIG with the data seems to get better, but this could also be the effect of two crossing curves. The same is

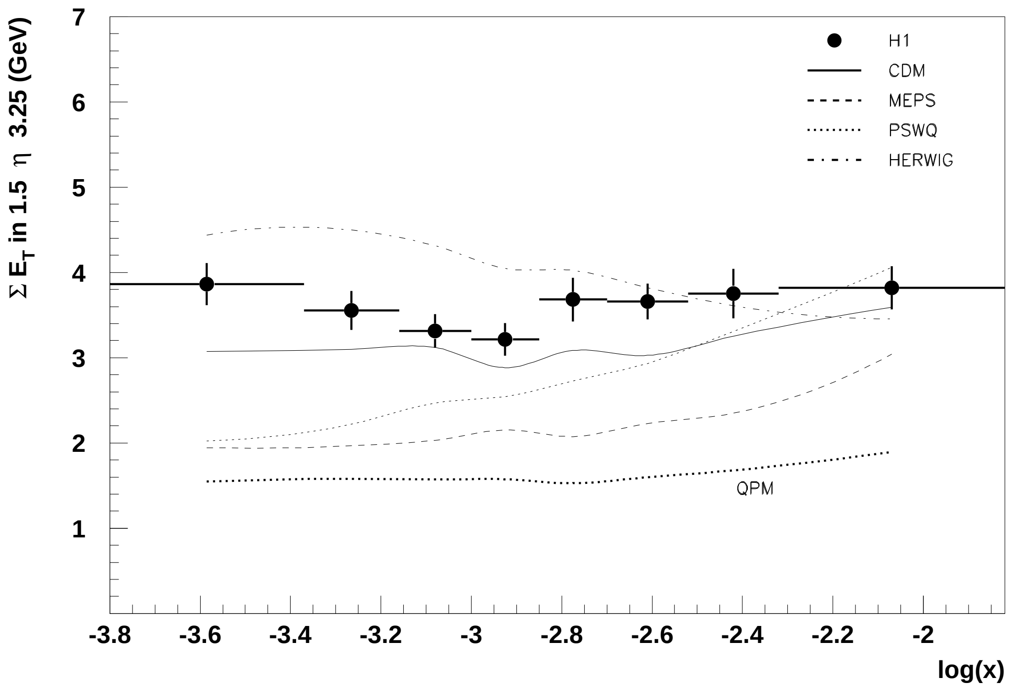


Figure 6.7: *Transverse energy in the range* $1.5 < \eta < 3.25$ *as a function of* $\log x$. *The error bars contain the statistical and systematic errors added in quadrature, except for an overall 6% energy scale uncertainty.*

probably true for PSWQ, and MEPS doesn't show any better agreement at higher x at all. A slight rise of the data is visible with decreasing x, which backs the theoretical predictions from BFKL calculations. The rise at high x is an effect of the current jet entering into the region where the $E_\perp$ is summed, thus obscuring the expected falling of the distribution. In order to draw firm conclusions from this plot on the observation of a BFKL-like behaviour, one has either to restrict the angular range, where the $E_\perp$ is summed, even further, to exclude the current jet at higher x, or, one should only use x values below 10^{-3}. This analysis can only be carried out when high statistics data become available, i.e. when the systematics for the 1993 data is sufficiently understood to allow corrections for detector effects.

6.2.5 Energy Flows in the Hadronic CMS

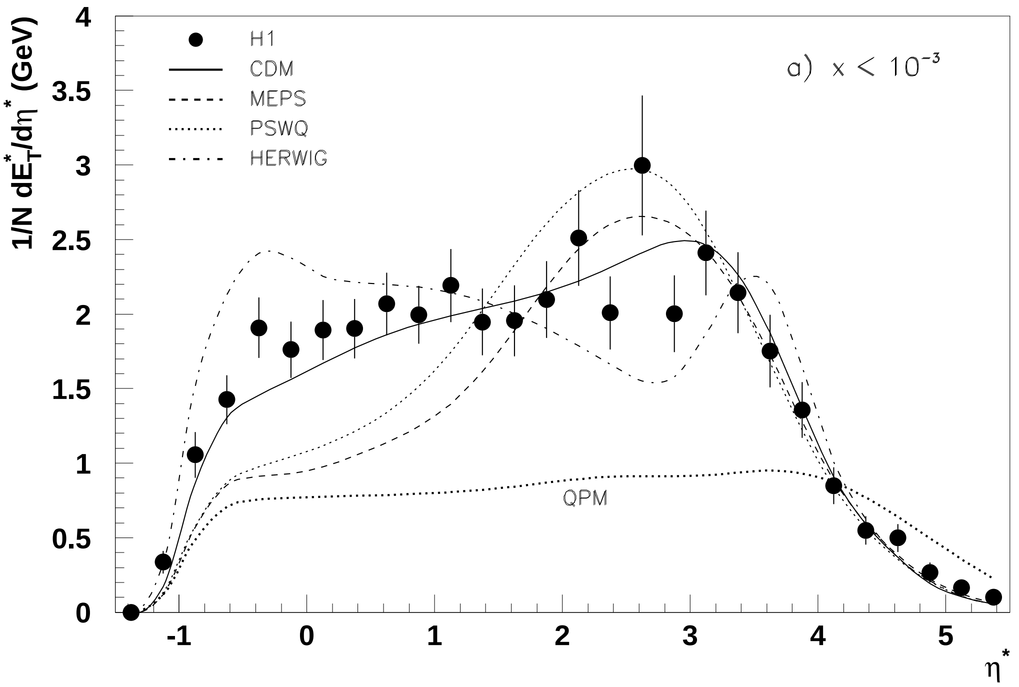
To get an estimate of the true amount of hard QCD radiation one has to study energy flows in the hadronic CMS. In the hadronic CMS, energies are not obscured by the boost to the laboratory frame and $p_\perp$'s balancing the $p_\perp$ of the electron; this allows a more direct observation of QCD radiation. As has been shown in section 5.5, the transformation of the event to the hadronic CMS has very little effect on the transverse energies. The overall gain in $E_\perp$ is negligible compared to the height of the plateau. The flow of transverse energy as function of the pseudorapidity in the hadronic CMS is presented in fig. 6.8, again for the two x-bins, $x < 10^{-3}$ and $x > 10^{-3}$.

With reference to the terms 'current' and 'target' hemisphere (see section 3.1.2), it is obvious that H1 sees in fact very little of the proton remnant. The plateau of the energy flow stops at $\eta^* < -0.4$, falling steeply towards smaller values of η^*, due to limited detector acceptance. Also, at $\eta^* < -0.4$ the influence of the boost becomes strong, see fig. 5.10. Although there is no clear distinction between the current and the target hemisphere, this distribution suggests that what has been termed above as 'towards the remnant' is in fact still part of the central region. Without the limited detector acceptance, the distribution would extend to -5 in pseudorapidity. The plateau exhibited for $\eta^* > 0$ is almost flat at $E_\perp^* \approx 2$ GeV per unit of rapidity, in both x-bins. It should also be noted that the plateau is slightly wider for $x < 10^{-3}$, an effect of the larger W, which governs the length of the plateau.

Comparing the data to the expectation of the QPM Monte Carlo shows that the larger part of the energy is generated by QCD radiation. The height of the QPM curve can be explained by $p_\perp$ from hadronisation, the intrinsic $p_\perp$ of the partons in the proton and gain due to the transformation to the CMS. All these effects are part of the other Monte Carlo predictions too, thus the difference is solely due to gluon radiation. The HERWIG curves show again a dip close to the right edge of the plateau and seem to 'oscillate' around the data. Both PSWQ and MEPS, give similar predictions, that don't match the data at all. The CDM Monte Carlo provides an overall good description of the measured distributions, though it falls slightly short in the higher x bin.

Finally, the mean transverse energy in the hadronic CMS is given as a function of W^2. This provides an overall measure of parton radiation activity, since W^2 sets the phase space limit for QCD radiation. A rise of the mean transverse energy with increasing W is therefore expected and confirmed by the data, as shown in fig. 6.9.

HERWIG gives an excellent description of the data for this distribution, which is a result of the tuning effort made for this generator by one of its authors. However, a simultaneous good agreement in this quantity and the energy flows could not be achieved. This is due to the lack of hard, large angle, QCD radiation that is effectively suppressed in the coherent parton shower approach of HERWIG [62]. A new version, which is due to be released, will match the matrix elements with parton showers, using a different strategy than the MEPS model. All the other models fall short for this distribution, but all show a similar rise in slope as the data. Although the QPM does not contain any QCD radiation, it too predicts a rise, though less steep. This rise is a result of the higher multiplicity at higher W.



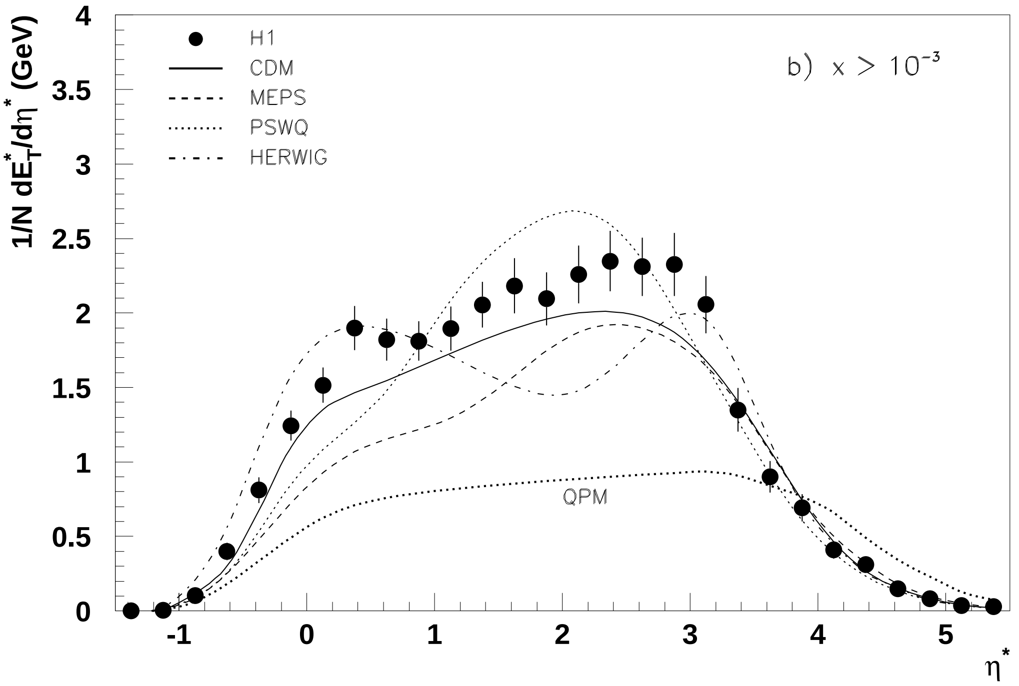


Figure 6.8: *Transverse energy flows E_T^* in the hadronic CMS as a function of the pseudorapidity η^* for events with a) $x < 10^{-3}$ and b) $x > 10^{-3}$. The proton direction is to the left. The error bars contain the statistical and systematic errors added in quadrature, except for an overall 6% energy scale uncertainty.*

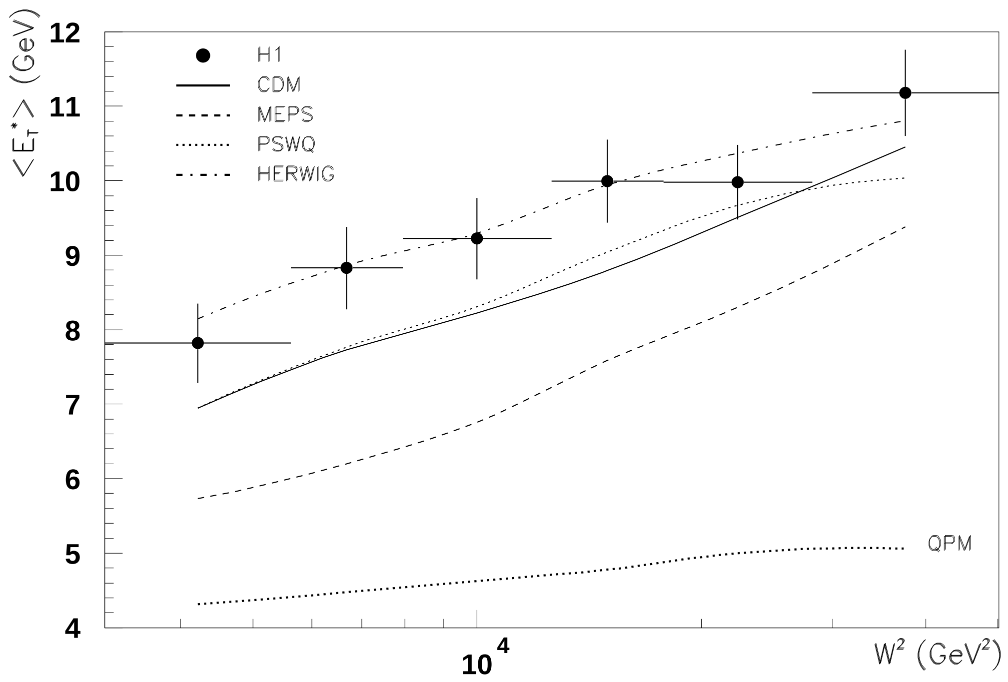


Figure 6.9: *Mean transverse energy in the hadronic CMS as a function of W^2. The error bars contain the statistical and systematic errors added in quadrature, except for an overall 6% energy scale uncertainty.*

6.3 Systematic Errors

The energy flows presented did include systematic errors which will be discussed subsequently. As will be shown later, the overall size of the systematic errors is much smaller than the statistical errors of the distributions, see fig. 6.12. The systematic studies of basic quantities such as the number of cells in a calorimeter energy cluster, energy, θ and ϕ distributions of clusters in section 5.3 should convince the reader that the Monte Carlo, after detector simulation, describes the data, in these aspects, very well. Studies have also shown that cluster multiplicities are well reproduced in the Monte Carlo and the relation between energies measured from charged tracks and the calorimeter are in agreement between data and MC. The electronic noise in the calorimeter is suppressed by appropriate cuts and is also included in the detector simulation by superimposing recorded H1 events with random triggers. All these factors give confidence that the detector simulation is adequate enough to correct for the calorimetric response in real data. It has also been verified that possible pile-up is not the source for the higher energies observed in the forward region of the detector. The level of pile-up is completely negligible in the 1992 data sample. The applied event selection cuts effectively exclude proton induced background (beam gas events). After the cuts accidental overlays of DIS events with proton gas events constitute well below 1% of the sample.

A study of systematic uncertainties will therefore concentrate on the more physics motivated sources of errors entering the correction procedure. The following list gives an overview of these errors, which are discussed later in detail:

- The overall hadronic calorimeter scale,
- the ad hoc factors for the hadronic energy measured in the BEMC and TC,
- the Monte Carlo model used to derive the corrections,
- the parton distribution functions used in the Monte Carlo generation,
- the hadron shower model used in the detector simulation,
- QED radiative effects,
- the remaining photoproduction background,
- the background of DIS diffractive events.

In sections 2.2 and 5.3, it has already been mentioned that the overall calibration of the LAr calorimeter is presently known to a precision of 6%. Varying the scale factor of 1.06, applied to the energy measurements in the data, by 6%, will give an estimate of the systematic error due to this uncertainty of the calibration. Since all the figures presented in this chapter are energy flows, this would result in shifting the data points up or down by 6%. Because the understanding of the calorimeter energy scale and calibration continuously improves, and because changes of the scale have very straightforward consequences on the energy flow, the scale uncertainty is not included in the total systematic error.

The BEMC and BBE are electromagnetic calorimeters and are therefore rather transparent to hadronic energy. This effect is clearly seen in the correction constants in fig. 6.1c. On the other hand the backward region is only of importance for very low x, when the current jet is expected in the BEMC. To account for the high degree of uncertainty of hadronic energy measurements in the BEMC, the factor 1.6, which is applied to correct for losses, has been varied by a factor of 2. This change will influence the energy flow in the laboratory frame exclusively for $\eta < -0.4$, but will affect the boost to the hadronic CMS and thus influence the whole distribution in the CMS. The size of the error, on the transverse energy flow in the CMS, is illustrated in fig 6.10a and is of the order of a few percent.

Using the MEPS model, instead of CDM, to derive the correction factors, gives a slightly different result, as shown in fig. 6.1c,d. This can be interpreted as a systematic error on the correction constants and is of the order of 5% in the very forward and backward regions, and less elsewhere. It is very satisfactory that this difference in the correction constants is much less than the difference of the input energy flows predicted by these two generators. This indicates that the corrections are well under control, even in the forward region, where they are large.

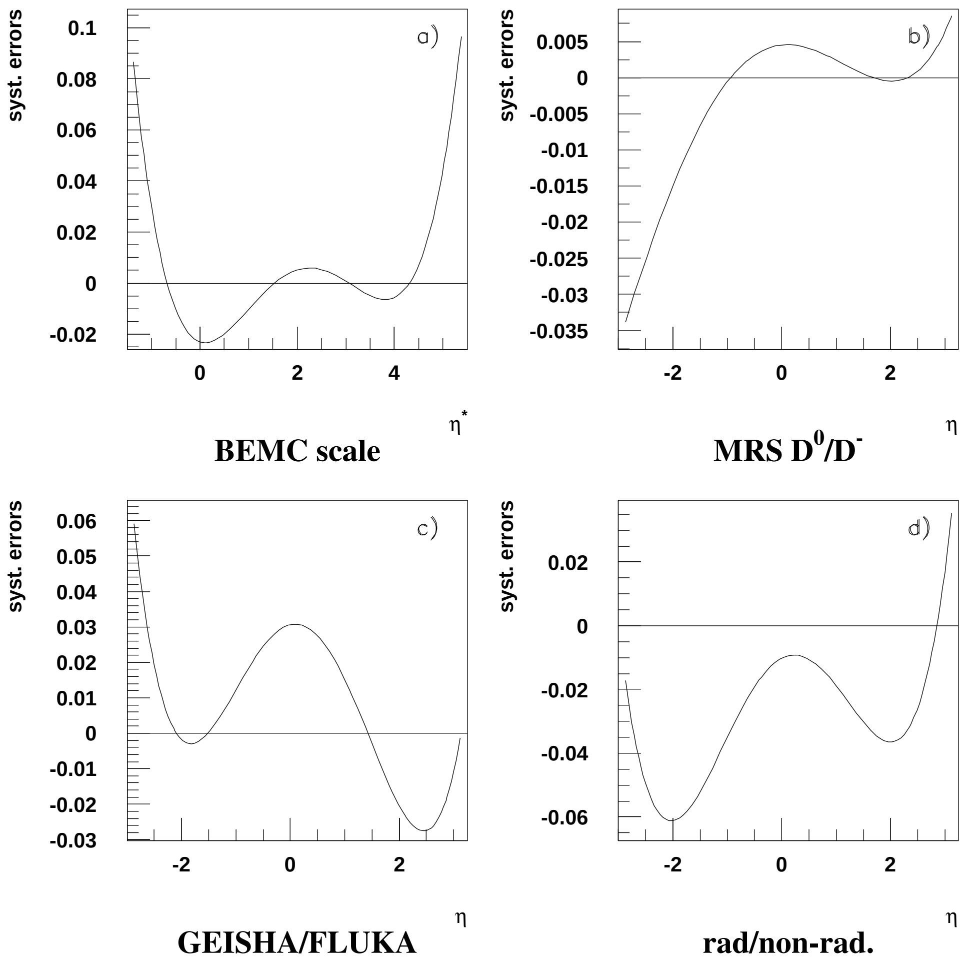


Figure 6.10: *Smoothed systematic errors for $x < 10^{-3}$. The distributions give the systematic error of a given source for the energy flow in the lab (except for a) in parts of a hundred. a) the error introduced by the BEMC scale factor uncertainty, for the energy flow in the hadronic CMS, b) structure function dependence, c) influence of the hadron shower model, d) effect of QED radiative corrections.*

Choosing different parton density functions for the Monte Carlo generation will result in different predictions for the energy flows, since the relative number of events in the two x bins will be different. This can also be accounted for by studying the change of the correction constants when using the MRS D^0 parametrisation instead of MRS D^-. The effect is very small, generally below 3%, and is shown in fig. 6.10b, for the low x bin.

At the moment there are two programs widely used to simulate hadronic showers in calorimeters. GHEISHA [63] is the default option of the H1 detector simulation package and FLUKA [64] is the alternative. Both programs give different predictions, especially for low energy showers [57]. From what has been said about the increase of energy in the forward region of the detector, it is obvious that the correct description of secondary interactions in dead material and the scattering of the resulting particles into the calorimeter is crucial to account for the effect. The other main ingredient is a correct description of the materials and the geometry of the components in the detector simulation. Studies have shown [59] that the latter point is not of major importance. Varying the size and material of the beam collimator or the beam pipe thickness, within reasonable limits, will not result in significant changes of the forward energy flow. Here only the effect of using FLUKA, instead of GHEISHA, is presented. Although the difference between GHEISHA and FLUKA is significant when the energy flow resulting from particles that are generated below $3.7°$ is studied [59], the overall effect on the energy flow is rather small. The error on the $E_\perp$-flow in the laboratory system is shown in fig. 6.10c and is less than 5%.

QED radiative corrections have an influence on the energy flows too, and have not been corrected for in the correction procedure. However, the effects are expected to be small, due to the restricted range in y. The effect has been investigated with a version of CDM interfaced to the HERACLES [65] generator. HERACLES generates photon radiation of the incoming and outgoing electron according to the first order matrix elements. Comparing the output of this combined generator with CDM, one can obtain an estimate for the systematic error introduced; the result for the energy flow in the lab is shown in fig. 6.10d. The systematic error is of the order of 5%. For future studies one would directly use a radiative Monte Carlo as DJANGO6 [66] to compare with the data, instead of including this in the systematic error.

Photoproduction background can survive the event selection cuts for this analysis if a high energy π^0 fakes a scattered electron in the BEMC and some other charged hadron makes a hit in the BPC within the required distance of the cluster. These are very rare cases, but the photoproduction cross section is by orders of magnitude higher than the DIS one. Using the PYTHIA [67] Monte Carlo, after detector simulation and normalising to the luminosity of the 1992 data, 3 events survive all selection cuts. These events have a very distinct energy flow in the lab, as shown in fig. 6.11a,b. The energy flow of these background events can either be subtracted from the data distribution or, as done here, be included in the systematic error.

The cut on the energy in the forward region of the detector will remove events with large rapidity gaps (see section 5.1). But from a physics point of view it will not

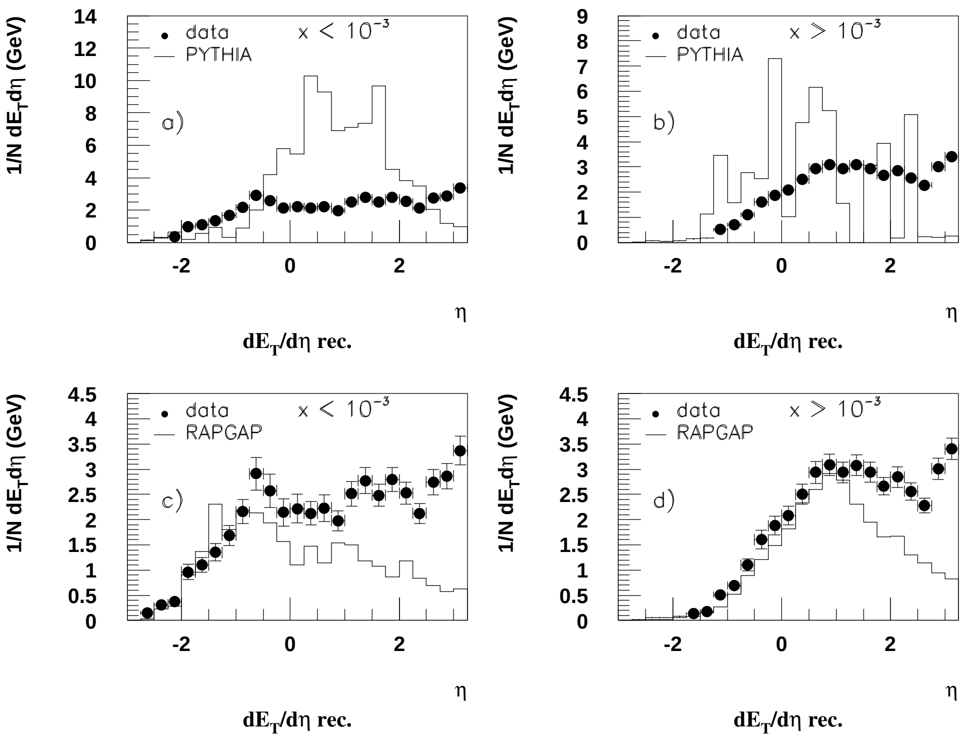


Figure 6.11: *Energy flow in the laboratory frame for photoproduction background events a),b) and diffractive like events c),d). Both flows are uncorrected Monte Carlo after detector simulation. Note, that a) contains 19 events and b) only 2 events, so the errors on the PYTHIA points are large. These 21 events after normalizing to luminosity correspond to 3 events.*

remove all diffractive type events from the sample. Studies using the RAPGAP [68] generator have shown that up to 10% of the events in the selected sample could be of a diffractive nature. The real percentage of diffractive like events is very difficult to establish, since the cross section for this sort of diffractive production is not well known. 10% is a conservative estimate when treating this as a systematic error. Fortunately RAPGAP predicts energy flows that are not very different from the MEPS prediction, as can be seen in fig. 6.11c,d, so the corrections are no more affected than in the case when MEPS is used instead of CDM. The influence of diffractive events on the correction constants is therefore already accounted for in the systematic error. Nevertheless the data will contain diffractive events, so it is the task of the Monte Carlos to include this process.

Summing up the different contributions of systematic errors one ends up with the energy flows presented in fig. 6.12. The error bars contain both the statistical and the systematic error, distinguished by the horizontal bars. The hatched histogram illustrates the systematic error alone. As can be seen the error in 1992 is dominated by statistics, this will change drastically with the 20fold statistics of the 1993 data.

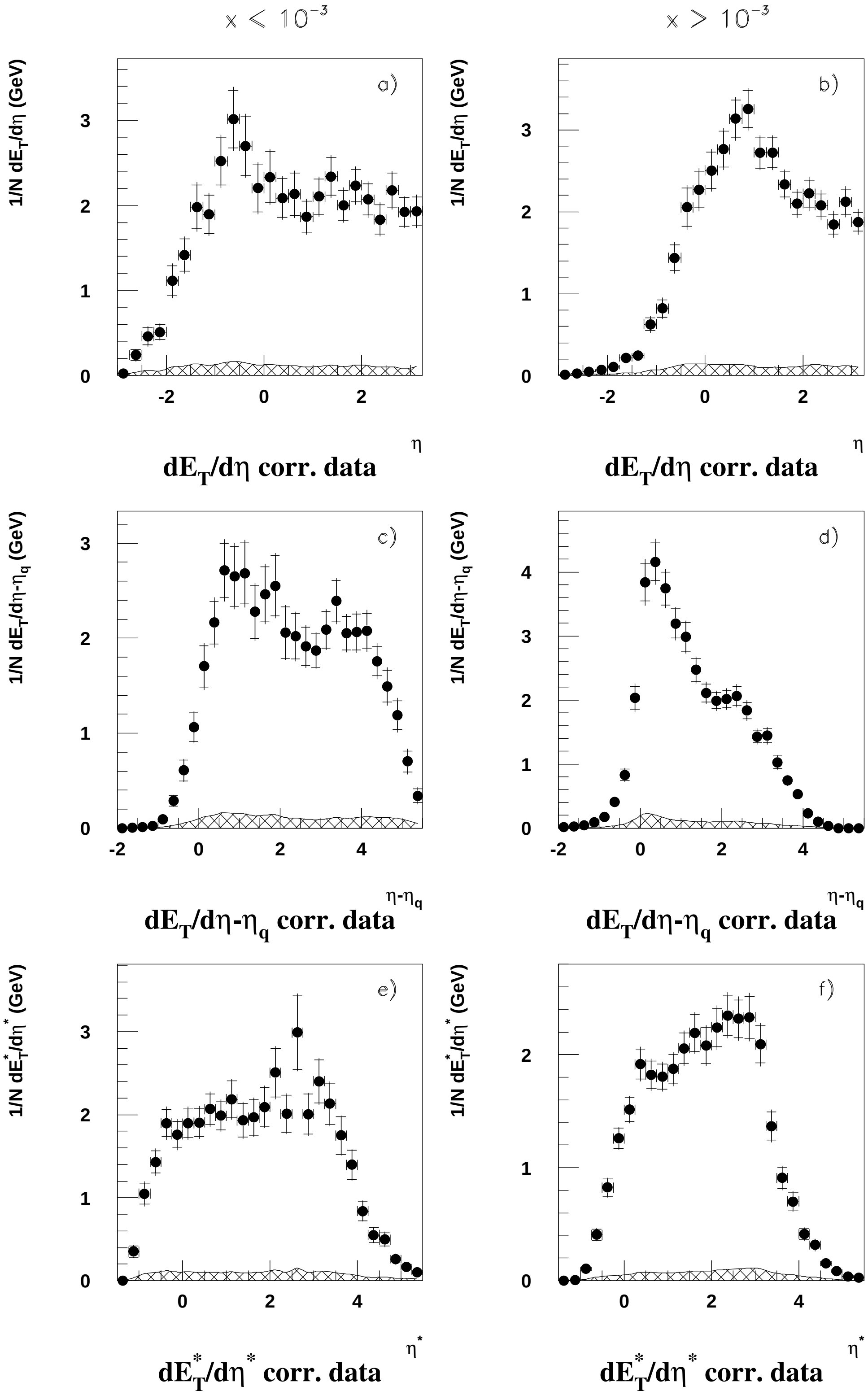


Figure 6.12: *The energy flows in the lab and the hadronic CMS with separated statistical error (bars up to the little horizontal lines) and systematic error (hatched histogram). Adding both errors in quadrature results in the outer error bars.*

6.4 Variations of Monte Carlo Parameters

Although no attempt has been made to systematically tune the Monte Carlo parameters, based on the energy flows measured, some parameter variations have been studied. Concerning HERWIG, extensive tuning was tried by one of the authors, resulting in the parameters presented in section 4.3. Therefore, no further studies with HERWIG will be discussed here, only the main results are quoted. For all generators, only parameters that are specific to ep-scattering have been varied. Hadronisation parameters and parameters concerning QCD features, not unique to HERA, have already been tuned to LEP data. These parameters should then be valid at HERA too.

There are several parameters that are common to all the generators used, for instance the parton density functions used, or the intrinsic $p_\perp$ of the partons in the proton. These are also left at the (default) values given in chapter 4, since they are not of major importance for this analysis. The emphasis here lies on parameters, which alter perturbative QCD features. The main parameter of perturbative QCD is Λ_{QCD}, but very little happens to the energy flows when Λ_{QCD} is varied within reasonable limits. In the following each model is discussed individually in terms of its behaviour on undergoing parameter variations.

- PSWQ

  The most important parameter of LEPTO parton showers is the variable, setting the maximum of the evolution variable, LST(9). $W\sqrt{Q^2}$ is chosen here, but other scales will result in significant changes of the energy flows predicted, as shown in fig. 6.13. From the available choices, the $W\sqrt{Q^2}$ scale clearly provides the best description of the data.

- MEPS

  The value of the cut-off parameter, y_{cut}, is crucial in this model. This once again demonstrates that LEPTO uses an adding procedure for the parton showers, not a matching of the PS to the ME; in such a case a strong dependence on the cut-off is not expected. How the cut-off value influences the amount of QCD radiation, and thus the energy flows, has already been explained in section 4.1.4. To illustrate this, fig. 6.14 shows what happens to the energy flow in the laboratory frame when y_{cut} is varied between 0.0025 and 0.05. Choosing 0.05 results in an energy flow which does not describe the data at all, whereas 0.0025 produces too little $E_\perp$, especially for the low x bin.

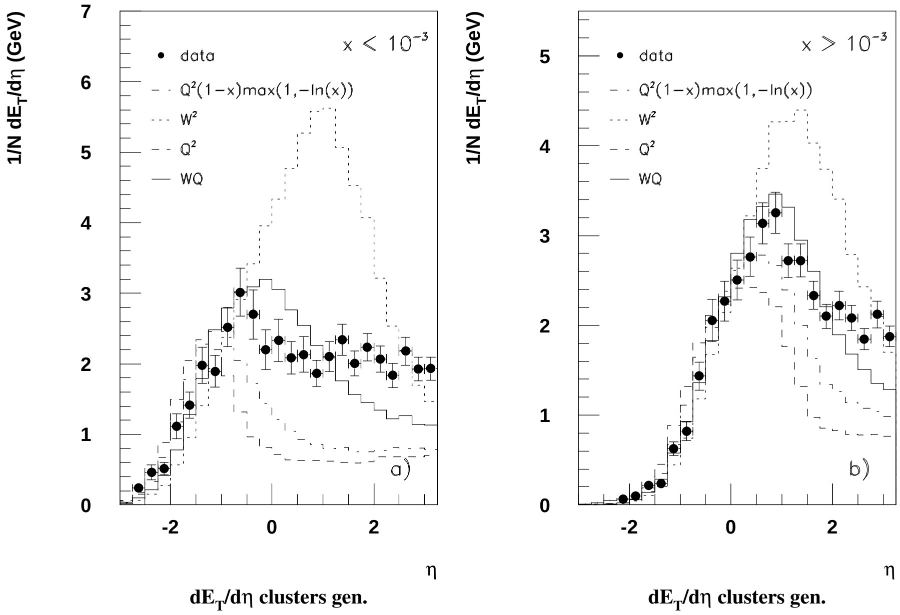


Figure 6.13: *Energy flow in the laboratory frame with LEPTO PS, varying the parton shower scales. The* $Q^2(1-x)\max(1,\ln\frac{1}{x})$ *scale is the default in LEPTO, WQ was used to compare to the data in this thesis.*

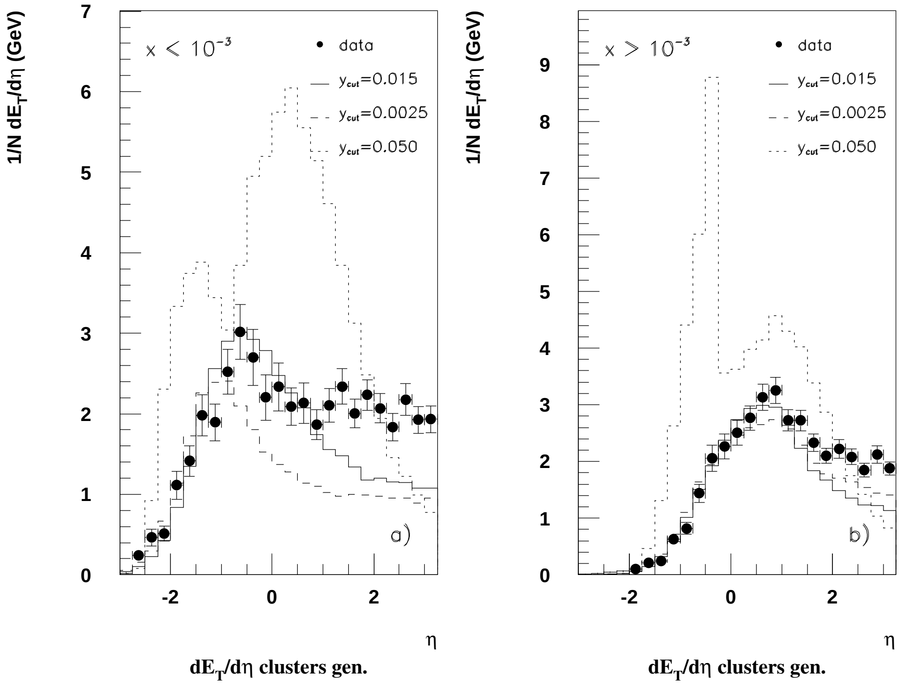


Figure 6.14: *Energy flow in the laboratory frame with MEPS, varying the cut-off for the matrix elements,* y_{cut}. *The value of 0.015 is the default in LEPTO.*

- CDM

  Although one might think that the y_{cut} parameter will have a similar influence on the CDM, it turns out that varying y_{cut} doesn't change the energy flows in the CDM greatly. Two 'protection mechanisms' inhibit a strong influence: QCD Compton events are abandoned by the CDM and in case of BGF the cut-off $\frac{p_\perp^2}{Q^2} > \mathrm{PARA}(20) = 0.5$ does a pre-selection independent of the cut-off on the matrix elements. Varying PARA(20) from 0.5 to 1.0 or 0.25 also has little influence on the flows. It turns out that μ (PARA(11)), the inverse radius of the proton remnant, is the parameter that the CDM is most sensitive to, when studying energy flows. The result of a variation of μ by a factor of 2.5 and 0.5 respectively, is shown in fig. 6.15. The energy flow in the remnant direction is changed by about 15%.

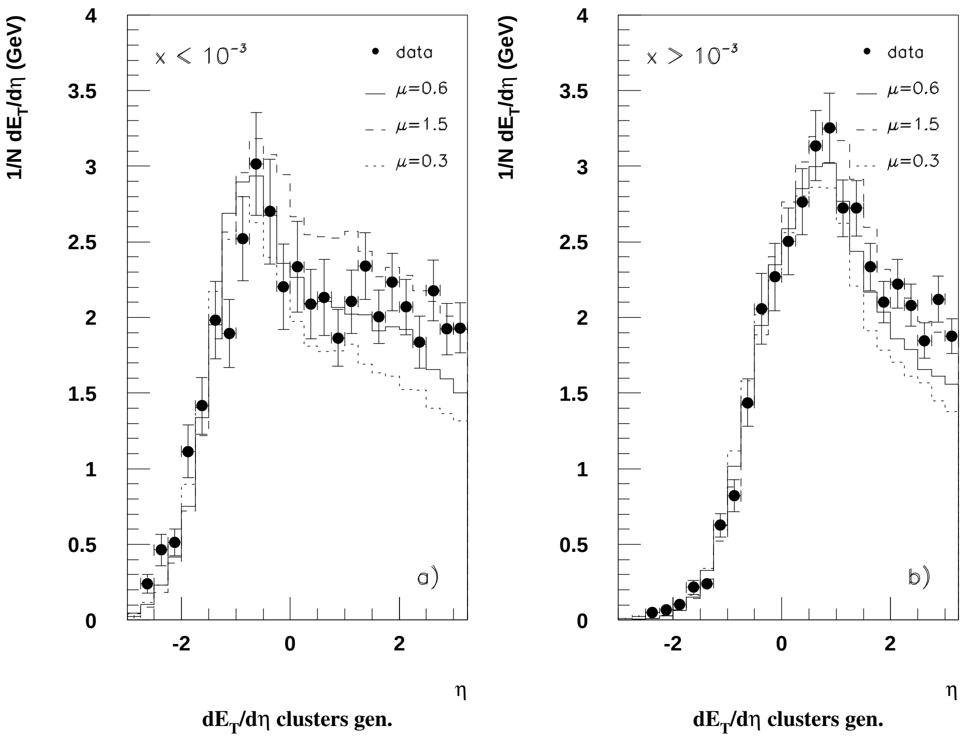


Figure 6.15: *Energy flow in the laboratory frame with CDM, varying the inverse size of the remnant, μ.*

- HERWIG

  One result of the tuning is that it is necessary to have the percentage of events with a soft underlying event close to 100%, i.e. the PRSOF parameter close to 1. HERWIG with too few soft underlying events will not be able to describe the data. The predictions for the energy flows depend heavily on the PRSOF parameter, low values result in far too little energy in the forward region and a deeper dip in fig. 6.8. Hence, the soft collision of the beam remnant with part of the initial state partons results not only in an enhanced energy close to the remnant, but also fills the gap between the current and the target regions. The other parameters such as PTRMS, PSPLT, CLMAX and QSPAC are of less importance for this area of study.

Chapter 7

Charged Particle Spectra

Nothing travels faster than the speed of light
with the possible exception of bad news,
which obeys its own special laws.

– Douglas Adams

In this chapter the charged particle spectra as observed in the hadronic CMS and in the so-called Breit frame are presented. The scaling behaviour of the longitudinal momentum fraction, $x_F = 2p_z^*/W$, and the transverse momentum as function of x_F are investigated in the hadronic CMS and compared to other experiments. A first look, using the 1993 data, at the energy spectra in the Breit frame is presented. Though lacking a determination of systematic errors, these spectra allow preliminary comparisons with theoretical calculations. The Breit frame differs from the CMS by an x dependent longitudinal boost. The correction procedure for the charged particle spectra and the systematic errors are treated in analogy to the energy flows, as described in section 6.1, however, the following modifications are introduced:

- The data are corrected to the full $+z^*$ hemisphere for the plots in the hadronic CMS, taking into account acceptance losses due to the limited angular coverage of the central tracker.

- For the Breit frame plots no acceptance cuts are applied to particles on the generator level, as it is required to make comparisons with analytical calculations possible.

- The central and forward tracker momentum and angular resolutions are much superior to those of the calorimeter, so, the arguments concerning bin width and resolutions (see section 6.1.3) are less important. The resolution in x_F, due to the boost, has already been discussed in section 5.5 (see also fig. 5.11).

- The track finding efficiency, as discussed in section 5.2.1, must also be taken into account. It is not possible to account for these defficiencies in the correction procedure, unless the detector simulation attemps to incorporate malfunctioning electronics, which to-date it does not. The central track finding efficiency was determined for the 1992 data, via an eye scan, to be 92±3%. There is no efficiency determination for the 93 data yet, but the ϕ distribution, fig. 5.3, suggests a similar or higher value.

7.1 Charged Particle Spectra in the Hadronic CMS

In the following charged particle spectra in the hadronic CMS are presented, for longitudinal (p_z^*) and transverse ($p_\perp{}^*$) momentum components with respect to the virtual boson direction. For these studies central tracks from the 1992 running period have been used exclusively. Possible scaling violations due to QCD are investigated by comparing the spectra in different ranges of W. The longitudinal momentum is expected to scale with W [70], if QCD effects are neglected.

The scaled longitudinal momentum is defined by $x_F = 2p_z^*/W$, here $\frac{W}{2}$ is the maximum energy of a particle in the hadronic CMS, resulting in $0 < x_F < 1$. The x_F spectra, for three different W ranges between 50 and 200 GeV are shown in fig. 7.1a. The acceptance corrections are generally 20-30%, 15% of which is due to the dead sector in CJC1, see fig. 5.3. However, for the W range $50-100$ GeV, the corrections rise to 60% at $x_F = 0.03$. This increase is due to forward going particles leaving the central tracker acceptance in the $+z$ direction. The systematic error, due to model dependent corrections, rises from around 5% in the lowest x_F bin to around 25% in the largest x_F bin. The determination of the track finding efficiency introduces a small systematic error of 3%. Due to the large x_F bins the momentum resolution does not contribute to the systematic error.

The spectra are seen to fall steeply with increasing x_F. Within the experimental errors the distributions for the different W ranges are in agreement, this is in accordance with the scaling hypothesis; so scaling violations, due to QCD radiation, are presently not observable in the data. All the models describe the data to within the experimental errors, and for this reason only the CDM is shown in fig. 7.1a.

In fig. 7.1b, the data, averaged over all values of W, are compared to EMC lepton-nucleon scattering data, at much lower $< W >= 14$ GeV [71]. The slope describing the H1 data appears steeper than that for EMC, this difference is an indication of scaling violations. At smaller x_F a proliferation of particles is observed at HERA energies comparing to that of EMC. This is correlated with the increase of the available phase space for the fragmentation process, starting with a higher energy primary quark, more particles can be produced in the fragmentation chain.

Data from e^+e^- annihilation at LEP [72], at a similar W, is also compared to the data[1]. The DELPHI data agree well with the H1 measurements when $x_F > 0.15$. However, at smaller x_F, the H1 data lie below the LEP measurements.

[1]This is possible as, due to the large W and the dominantly small $p_\perp{}^*$, the H1 spectra in terms of $x_F = 2p_z^*/W$, $x_E = 2E^*/W$ and $x_p = 2p^*/W$ differ by less than the size of the symbols, and so can be compared directly to the x_p spectrum from DELPHI.

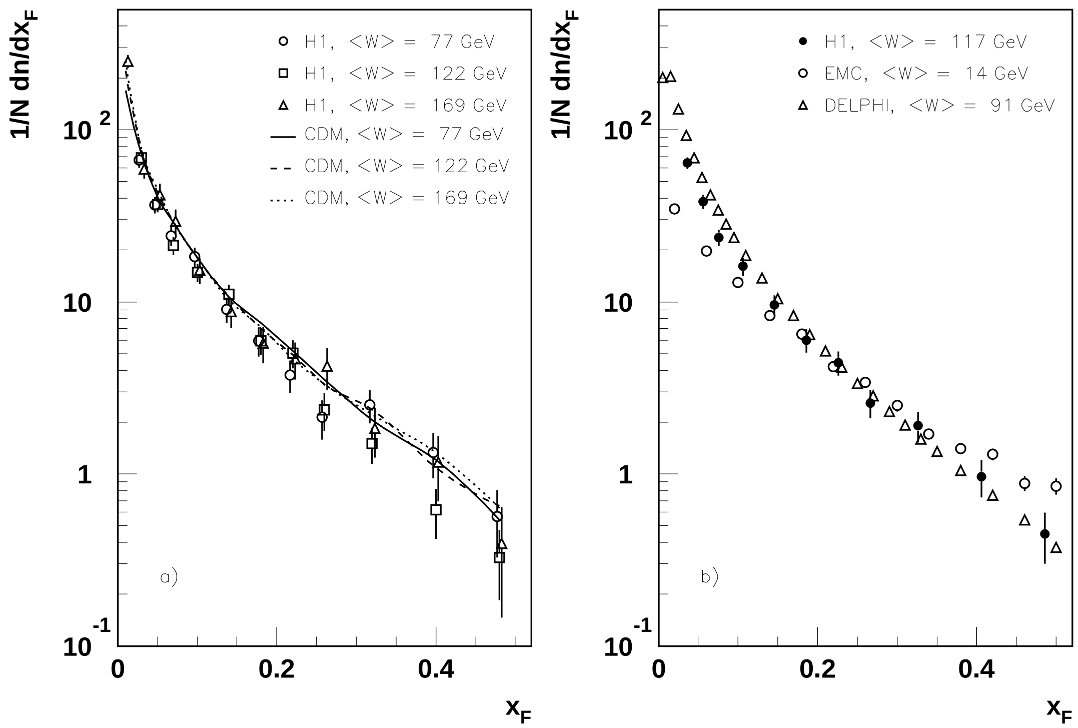


Figure 7.1: *Scaled charged particle spectrum ($x_F = 2p_z^*/W$) in the hadronic CMS. The distributions are normalized to the number of events N, and n refers to the number of charged particles in a given bin. Statistical and systematic errors are added in quadrature. In a) the data for three different W bins are shown, together with the CDM prediction. In b) the data are compared with data from EMC and DELPHI. The DELPHI data are divided by two to account for the two jets in e^+e^- annihilation.*

At HERA, the transverse momentum of the observed particles is dominated by gluon radiation, the contributions to this due to the intrinsic $p_\perp$ in the proton and from fragmentation, are expected to be small. As a sensitive measure of these effects the so-called seagull plot, in fig. 7.2, gives the average transverse momentum squared, $<\mathrm{p}_\perp^{*\,2}>$, as a function of x_F [71]. Corrections applied vary from 2% at small x_F to 27% at large x_F. The systematic errors take into account the model dependence of the corrections (3-7%, rising with x_F), the calorimeter calibration (similar magnitude) and the dependence of the results on the choice of the $p_\perp <$ 20 GeV cut. This latter cut gives rise to a systematic error of up to 10% in the highest x_F bin.

With increasing x_F, the average transverse momentum squared increases by approximately an order of magnitude. Comparing H1 data, with an average W of 117 GeV, to EMC data [71], with $<W>=$ 14 GeV, much larger transverse momenta are observed for H1. As intrinsic $p_\perp$ and fragmentation $p_\perp$ are considered constant, this increase has to be attributed to the larger phase space available for QCD radiation.

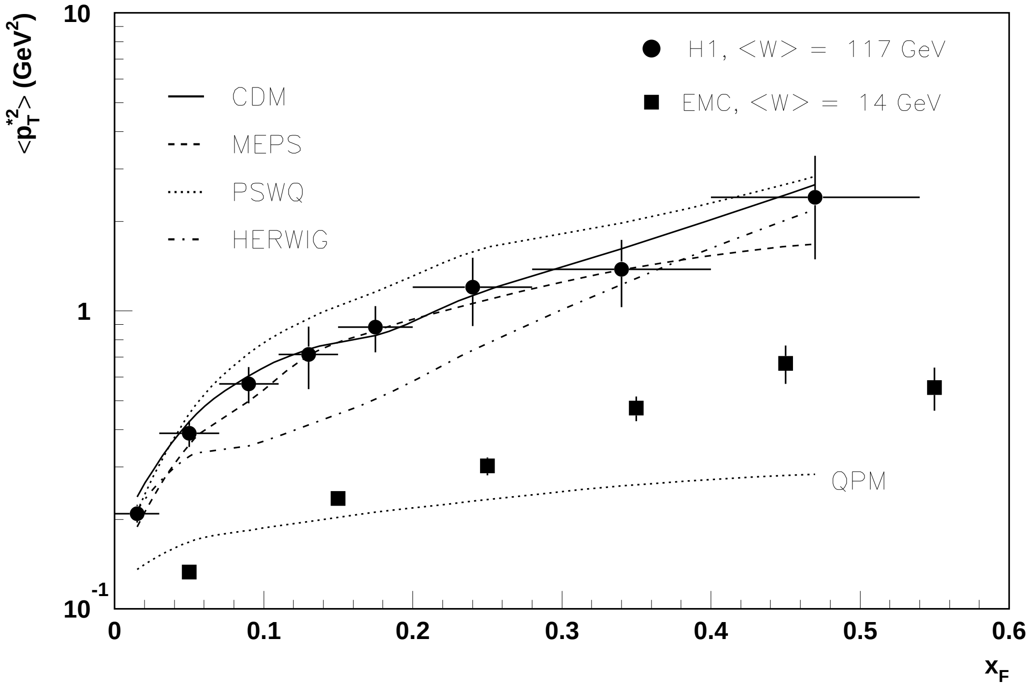


Figure 7.2: *Seagull plot: average $p_\perp^2$ as a function of $x_F = 2p_z^*/W$. The error bars for the data contain statistical and systematic errors added in quadrature. Compared to the data are the model predictions and EMC data at lower W.*

The QPM, which contains only these two non-perturbative effects, gives a factor of about 5-10 less $<p_\perp^{*\,2}>$.

As to the strength of the increase in $<p_\perp^{*\,2}>$, the models differ in their prediction. MEPS and CDM describe the data adequately, while PSWQ is only marginally compatible with the data. HERWIG differs significantly from the data, not only in magnitude, but also in the shape of the spectrum.

7.2 Colour Coherence and the MLLA

Final state properties of charged particles in DIS can be investigated in the so-called Breit frame[2], which is defined by $q = (0, 0, 0, -2xP)$. In this frame the current quark is backscattered from the boson, i.e. its momentum sign is flipped. In the approximation of the QPM the current quark is purely longitudinal along the z-axis, as is the exchanged boson. From this it is clear that the Breit frame is related to the hadronic CMS by a longitudinal boost. The boost vector, i.e. the 4-vector that will be at rest in the Breit frame, is given by $2xP + q$. The transformation to the Breit frame is performed using the method described in section 5.5.

[2]4-vectors are given as (E, p_x, p_y, p_z), so the boson carries only momentum in the Breit frame.

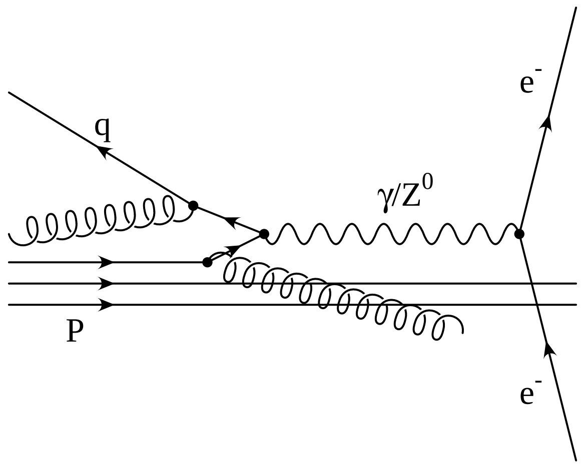


Figure 7.3: *A diagrammatic representation of an event in the Breit frame. The interacting quark emits initial state radiation and so does not meet the boson head on in the z-direction.*

One of the claims about the Breit frame is that jets produced in the current region, $p_z < 0$, should be identical to jets produced in e^+e^- annihilations, at an e^+e^- CMS energy of $\frac{Q}{2}$ [73]. This means that the development of the energy spectra of particles in e^+e^- experiments, as a function of the experiment's CMS energy, can be compared to the Q^2-dependence of energy spectra produced in DIS, when observed in the current region of the Breit frame [27, 74].

Theoretical calculations are carried out in the Modified Leading Logarithmic Approximation (MLLA, [75]), which takes into account next-to-leading logarithms. The aim of the MLLA is a correct treatment of coherent emission of soft gluons inside a jet, leading to angular ordering and an effective transverse momentum cut-off for the partons. To derive predictions for hadrons, the idea of LPHD (see section 3.5.1) is employed. The theoretical MLLA approach is similar to building parton shower models obeying QCD coherence.

The MLLA allows predictions to be made for quantities, such as multiplicities, which otherwise cannot be calculated easily. The infrared singularities of QCD give rise to an infinite number of soft gluons, so a resummation of these singularities is crucial to obtain predictions for multiplicities etc.. The MLLA provides a way to renormalise the infrared singularities and can therefore give quantitative predictions for parton distributions.

Since the main emphasis of the MLLA is on colour coherence, studies at LEP have focussed on this issue. However, the fact that a Monte Carlo (a modified version of JETSET) without coherent parton showers, also describes charged particle spectra data very well, leads to the conclusion that coherence has not yet been demonstrated on the basis of hadron momentum spectra [76]. At the same time, the LEP results back the idea of LPHD.

DIS opens a new domain for charged particle spectra, namely the target region, defined by $p_z^{BF} > 0$, in the Breit frame[3]. It is important to note, that what is referred to as the target region, in the Breit frame, is partly called the current region in the hadronic CMS. The impossibility to get a real distinction between the current and target regions has already been mentioned in section 3.1.2.

What do the actual predictions of the MLLA for DIS look like? Two distinct cases have been studied in [73], the Q^2 development of the spectra at fixed x and the x development at fixed Q^2. The predictions from [73] are shown in fig. 7.4a,b; clearly

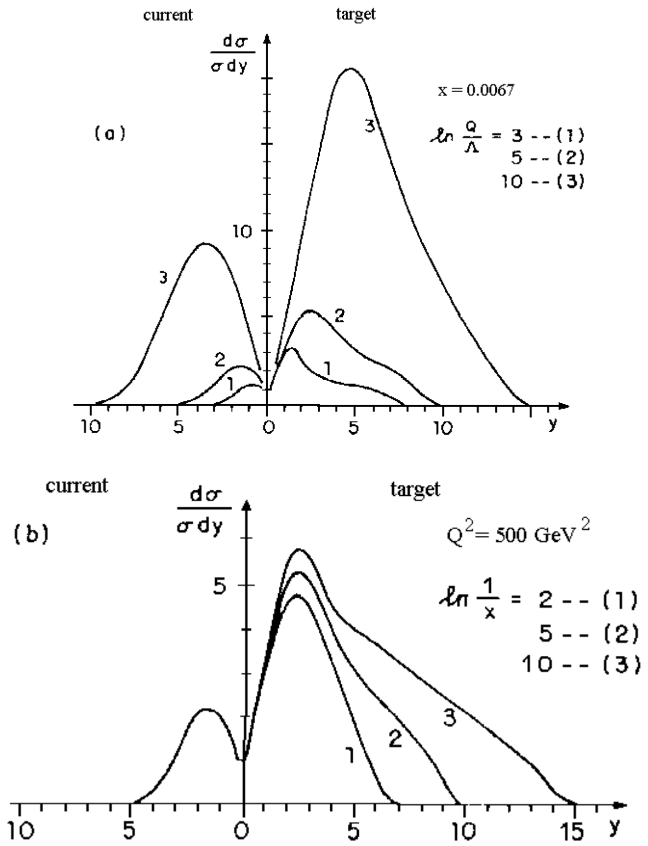


Figure 7.4: *The MLLA predictions for charged parton (or particle) flows in the Breit frame. a) fixed x, three different $Q^2 = 9, 500, 11000$ GeV^2 and b) at fixed Q^2 for three different $x = 0.14, 0.0067, 0.00005$. y is defined as $\ln \frac{E^{BF}}{\Lambda_{eff}}$.*

the x and Q^2 dependencies have very different effects on the current region (to the left) and the target region. Whilst both sides of the event are affected by changes in Q^2, only the target region is predicted to show an x dependence. The shift of the peak position in the current region is expected to follow a linear Q^2 dependence (if Q^2 is the fundamental scale in DIS). Such a shift can be directly compared to

[3]Note that in the hadronic CMS the target region is defined as $p_z^* < 0$, so in the Breit frame the two sides are flipped.

similar LEP studies, if the energies are scaled with $\frac{Q}{2}$ (the maximum energy in the current region). This issue is discussed in depth in [74]. The target region is unique to DIS, so no previous results exist.

How does one read these plots? y is defined as $\ln E^{BF}/\Lambda_{eff}$, the energy of the particle in the Breit frame divided by the effective QCD scale. The scale Λ_{eff} of the MLLA is not well defined, i.e. it can not be related to the known values of $\Lambda_{\overline{MS}}$, due to the lack of some higher order calculations [77]. Hence the value favoured by the LEP experiments, $\Lambda_{eff} = 150$ MeV, is used here.

The left part of the abscissa in fig. 7.4 *does not* indicate a negative logarithm. After a particle has been identified as belonging to either the current or the target region, it will enter either the left hand side or the right hand side of the plot, respectively. So for both regions $\ln E^{BF}/\Lambda_{eff}$ is plotted separately, but into the same plot.

7.3 Charged Particle Spectra in the Breit Frame

As has already been mentioned at the beginning of this chapter, the corrections for the charged particle spectra in the Breit frame follow a slightly different procedure than that for the energy flow analysis. No acceptance cuts are applied to the charged hadrons on the generator level, and so the corrections also account for acceptance losses. One of the arguments which is put forward against this procedure for the energy flows (see section 6.1), is not valid for investigations using track information, as secondary tracks can be (at least partly) elliminated. Only primary tracks, i.e. tracks from the main event vertex, are selected and the effect of additional forward energies is strongly suppressed. However, the argument concerning the possibly unjustified extrapolation using a Monte Carlo model, is still valid and one has to bear in mind this effect when evaluating the measurements presented below. The corrections are illustrated in fig. 7.5. Although the spectra extend beyond 4 in the target region, the high correction factors make an analysis for $y > 4$ unreasonable. As can been seen in fig. 7.5a, the forward tracker contributes exclusively in the target region. The dip at $y = 0$ is due to the energy cut, $E^{BF} > \Lambda_{eff}$, that has been applied to the data.

This study presents corrected results for the 1993 data, however, systematic errors are not included for the distributions shown. Hence, the results have to be treated with caution, as the statistical errors are very small and the overall error is dominated by systematics. Due to the limited Q^2 range of the event selection, the data are binned in just 3 Q^2 regions. However, the wider x range allows for 6 bins in x to be made. For the Q^2 dependence no restrictions on x are applied, this results in the following values:

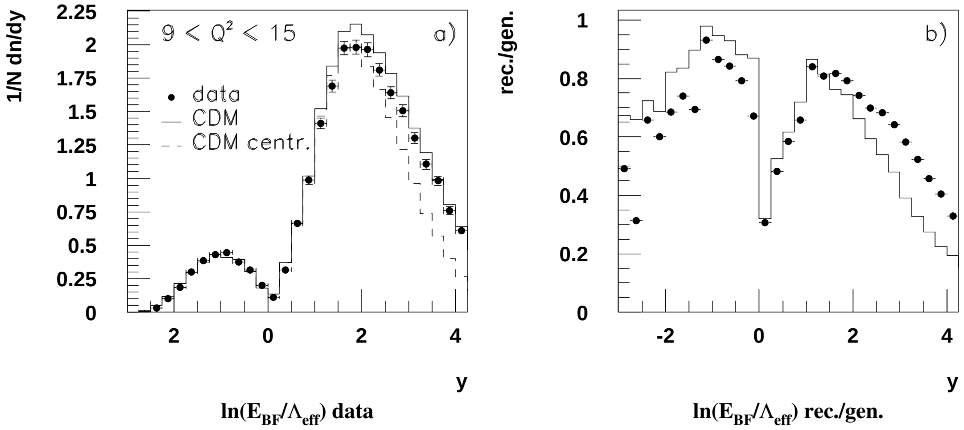


Figure 7.5: *a) Charged particle spectrum in the Breit frame, uncorrected, with statistical errors only. Compared are data and CDM Monte Carlo for the lowest Q^2 bin. The agreement is equally good in the other bins. The dashed line gives the spectrum when only central tracks are used. b) The correction constants using the CDM, for the lowest (dots) and highest (full line) Q^2 bins.*

Q^2 bin (GeV^2)	Average Q^2 (GeV^2)	Average $\log x$
9-15	12.1	-3.0
15-26	19.6	-2.8
26-100	43.8	-2.5

Accordingly, for the x bins, Q^2 has been restricted to the range 10 GeV^2 $< Q^2 <$ 20 GeV^2; this makes for a fixed Q^2, as required in the theoretical calculations. The resulting values are summarised below:

x bin	Average $\log x$	Average Q^2 (GeV^2)
0.0001-0.0006	-3.39	13.0
0.0006-0.0009	-3.13	14.0
0.0009-0.0013	-2.97	14.3
0.0013-0.0020	-2.80	14.5
0.0020-0.0034	-2.59	14.8
0.0034-0.0100	-2.38	16.1

7.3.1 The Q^2 Dependence

Using the corrected charged particle spectra from the 1993 data in the three Q^2 bins listed above, the peak position in the target and in the current regions are studied; for this purpose Gaussian curves are fitted to the peaks. In the current region no restriction on the fit range is applied, while in the target region the fit is restricted to $1.0 < y < 3.5$. The latter restriction is attributed to the non-Gaussian shape of the spectrum in the target region. The distributions and the respective fit results are shown in fig. 7.6.

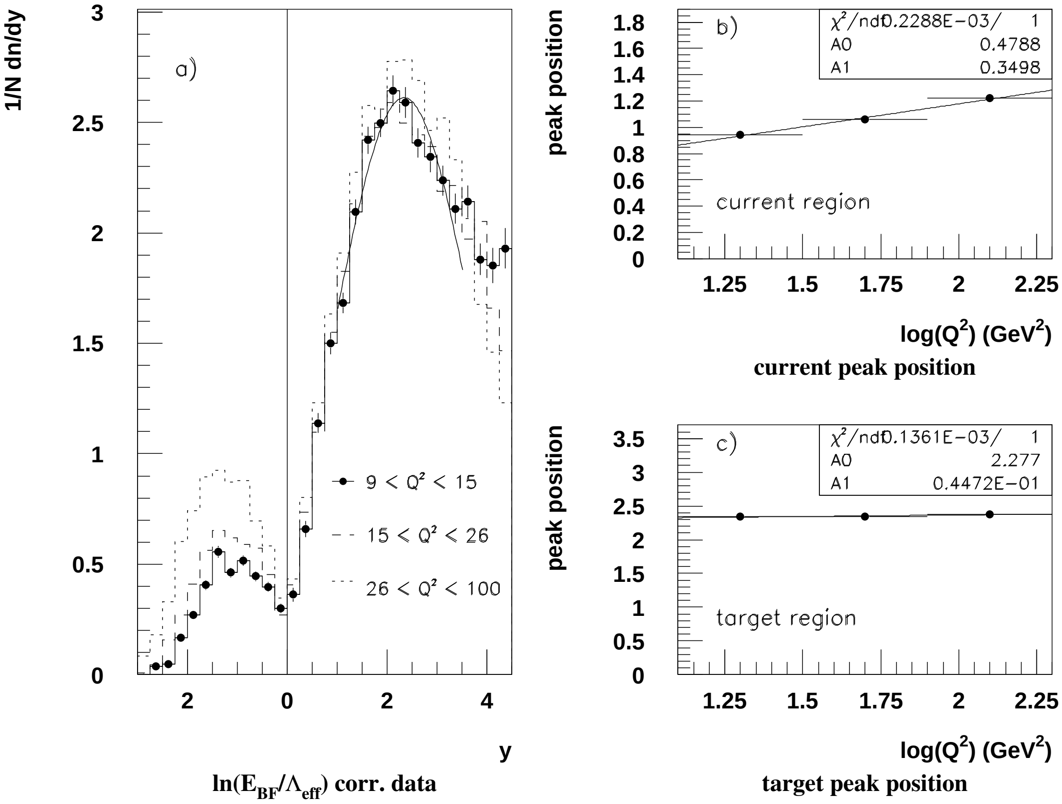


Figure 7.6: *Charged particle spectra and the peak positions as a function of Q^2, using the full range in x. a) Development of the spectra with Q^2. b),c) Plotted are the mean values of the Gaussians fitted independently to both sides of a spectrum. b) current region, c) target region. The straight lines are results of a linear fit.*

Compared to the theory, the small Q^2 dependence seen in the target region may be a surprise. However, the theoretical curves are for a $Q^2 = 9,\ 500$ and 11000 GeV^2 respectively, and so the range analysed is only a small part of the full range covered by the predictions. In fig. 7.6b, seeing a linear rise with Q^2 in the current region is in nice agreement with the predictions.

7.3.2 The x Dependence

In analogy to the previous section, the x dependence of the peak positions has been explored for both the current and the target region. The very strict constraint placed on the Q^2 range ensures that the results are not obscured by Q^2 variations, however, this does result in reduced statistics. The findings are shown in fig. 7.7.

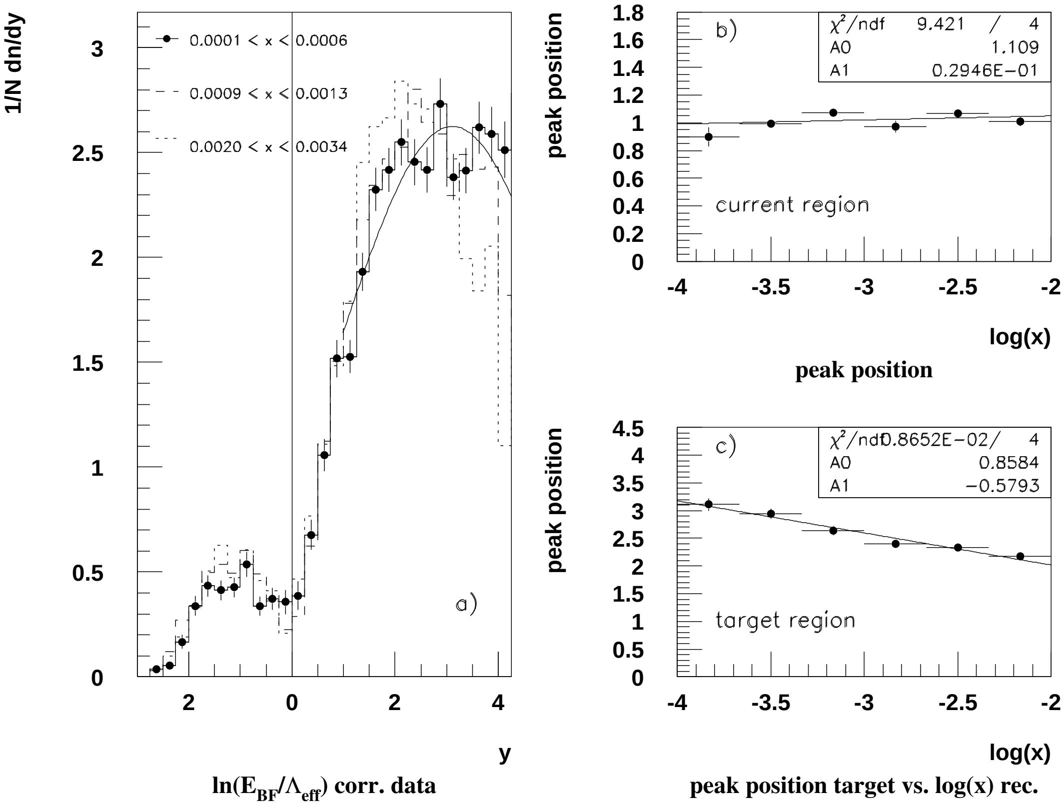


Figure 7.7: *Charged particle spectra and peak positions as a function of x, restricting Q^2 to 10-20 GeV^2. a) Development of the spectra with x. b),c) Plotted are the mean values of the Gaussians fitted independently to both sides of a spectrum. b) current region, c) target region. The straight lines are results of a linear fit.*

For the interpretation of the results in the target region, the main obstacles are the limited detector acceptance and the large correction factors. One could in principle extend the measurements up to $y = 5.5$, but at a cost of obtaining correction factors of greater than 5. This violates the applicability of the bin-by-bin correction method, where the corrections are required to be small. So it is difficult to draw firm conclusions on the x dependence of the spectra in the target region. The MLLA (at $Q^2 = 500$ GeV^2) does not predict the shift in the peak position seen in the data ($Q^2 = 14$ GeV^2), only an extension of the distribution towards higher y, see fig. 7.4b. Given the difference in the Q^2 and the lack of systematics, this cannot be verified or dismissed with the results at present.

In the current region the data are compatible with no shift in the peak position with x being observed, this is in agreement with MLLA predictions. The correction factors are reasonable for the current region, and so it would be the best starting point for further studies. An in depth analysis of the target region needs a better understanding of the forward tracks and the extension to higher values of Q^2.

Chapter 8

Conclusions

I had to get up in the morning at ten o'clock at night,
half an hour before I went to bed, eat a lump
of cold poison, work twenty-nine hours a day down mill,
and pay mill owner for permission to come to work...
But you try and tell the young people today that
and they won't believe ya'.

– Monty Python

In this thesis properties of the hadronic final state in deep inelastic scattering have been studied with the H1 detector at HERA. The basis of the investigations was data obtained during the 1992 running period, meaning an integrated luminosity of ≈ 22.5 nb^{-1}. The high level of understanding of detector effects for tracks and calorimeter clusters has allowed the data to be corrected, enabling comparisons with Monte Carlo generators. The data are limited to the kinematical region 10 GeV^2 < $Q^2 < 100$ GeV^2 and $10^{-4} < x < 10^{-2}$, the average Q^2 being about 25 GeV^2.

Overall properties such as energy flows in the laboratory frame and in the hadronic CMS are presented, concentrating on possible new QCD effects in the low x region. Further investigation of the region between the current and the target jet has revealed differences between the data and the predictions of the various Monte Carlo generators. Only a generator based on the colour dipole model gives a satisfactory description of the data. The data are in good agreement with calculations based on the BFKL dynamics at small x. Although there are indications that Altarelli-Parisi based QCD evolution is not able to describe the measurements, the current size of the statistical errors and the uncerntainty of the size of hadronisation effects prevents firm conclusions to be drawn. A full analysis of the 1993 data with its 20fold statistics will allow to establish or abandon the existence of small x effects in the HERA kinematical range.

Comparing different models implemented in the form of Monte Carlo generators to the data, the studies revealed the importance of certain concepts. Firstly, in coherent parton shower models (HERWIG) the lack of hard, large angle radiation manifests itself in strong deviations from the data in the description of the energy flows. Secondly, for LEPTO parton showers the importance of the maximum of the evolution variable (also called the parton shower scale) is crucial and all the choices investigated failed to describe the data. The approach to add parton shower to $O(\alpha_s)$ matrix elements is far from the success of similar models in e^+e^-. Last, but not least, the colour dipole model, although it describes the data quite well, has the uncertainty in the treatment of boson-gluon fusion events and the freedom that arises from treating the remnant as an extended coloured object.

Charged particle spectra in the hadronic CMS and in the Breit frame allow studies of specific properties, in both the current or the target regions. The seagull plot backs results obtained from the studies of energy flows, and longitudinal momentum spectra in the hadronic CMS are in concordance with the scaling hypothesis. Momentum spectra of charged particles in the Breit frame, give more insight into the until now unexplored target region. They show an x dependence, which, if confirmed by further studies obtaining sytematic errors for the distributions, would contradict theoretical calculations based on the MLLA.

Intense collaboration has occured with the authors of the Monte Carlo generators discussed in this thesis and the results of this analysis have stimulated work into improving these generators. New versions of LEPTO, ARIADNE and HERWIG are being prepared or have already been released. The present studies have also initiated the recent theoretical calculations of energy flows at HERA. A similar analysis of the 1993 data, with its higher statistics, will allow a deeper insight into QCD at low x, as much finer binning in x and Q^2 will be possible, along with a significant reduction of the statistical errors.

Bibliography

[1] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438 and 675.
G. Altarelli and G. Parisi, Nucl. Phys. 126 (1977) 297.

[2] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Sov. Phys. JETP 45 (1972) 199.
Y.Y. Balitsky and L.N. Lipatov, Sov. J. Nucl. Phys. 28 (1978) 282.

[3] A.D. Martin, University of Durham preprint DTP/94/32, May 1994.

[4] A.H. Mueller, Nucl. Phys. B (Proc. Suppl.) 18C (1990) 125.

[5] H1 Collab., I. Abt et al., DESY preprint 93-103 (1993), to be submitted to Nucl. Instr. and Meth..

[6] H1 Calorimeter Group, B. Andrieu et al., Nucl. Instr. and Meth. A336 (1993) 460.

[7] H1 Calorimeter Group, B. Andrieu et al., Nucl. Instr. and Meth. A336 (1993) 499.

[8] H.P. Wellisch et al., Max-Planck-Institut für Physik preprint MPI-PhE/94-03, Jan. 1994.

[9] J. Gayler, private communication.

[10] H1 Collab., I. Abt et al., Nucl. Phys. B407 (1993) 515.

[11] V. Gates, E. Kangaroo, M. Roachcock, W.C. Gall, Physica 15D (1985) 289.

[12] D. Graudenz, Phys. Lett. B 256 (1991) 518, preprint LBL-34147 (1993);
T. Brodkorb, J.G. Körner, Z. Phys. C54 (1992) 519, T. Brodkorb, J.G. Körner, E. Mirkes and G.A. Schuler, Z. Phys. C44 (1989) 415;
F.A. Berends, W.T. Giele and H. Kuijf, Nucl. Phys. B321 (1989) 39;
K. Hagiwara and D. Zeppenfeld, Nucl. Phys. B313 (1989) 560.

[13] V.V. Sudakov, Sov. Phys. JETP 3 (1956) 65.

[14] G. Marchesini, B.R. Webber, Nucl. Phys. B238 (1984) 1.

[15] G. Marchesini, B.R. Webber, Nucl. Phys. B310 (1988) 461.

[16] T. Sjöstrand, Phys. Lett. B157 (1985) 321.

[17] G. Ingelman, *in* Proc. HERA workshop, Eds. W. Buchmüller and G. Ingelman, Hamburg (1991) Vol. 3, 1366.

[18] N. Brook, G. Ingelman, L. Lönnblad, *in* Proc. HERA workshop, Eds. W. Buchmüller and G. Ingelman, Hamburg (1991) Vol. 1, 275.

[19] G. Gustafson, Ulf Pettersson, Nucl. Phys. B306 (1988) 746.
G. Gustafson, Phys. Lett. B175 (1986) 453.

[20] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Phys. Lett. B165 (1985) 147.

[21] B. Andersson, G. Gustafson, L. Lönnblad, Ulf Pettersson, Z. Phys. C43 (1989) 625.

[22] B. Andersson, private communication.

[23] B. Andersson, L. Lönnblad, Phys. Lett. B277 (1992) 359.

[24] B. Andersson, L. Lönnblad, private communication.

[25] L. Lönnblad, CERN preprint CERN-TH.7307/94, June 1994.

[26] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, Z. Phys. C27 (1985) 65.

[27] Yu.L. Dokshitzer et al., Z. Phys. C55 (1992) 107.

[28] OPAL Collab., M.Z. Akrawy et al., Phys. Lett. 247B (1990) 617.
OPAL Collab., G. Alexander et al., Phys. Lett. 264B (1991) 467.

[29] L3 Collab., B. Adeva et al., Phys. Lett. 259B (1991) 199.

[30] R.P. Feynman, R.D. Field, Nucl. Phys. B136 (1978) 1.

[31] L3 Collab., B. Adeva et al., Z. Phys. C55 (1992) 39.

[32] T. Sjöstrand, Computer Phys. Comm. 39 (1986) 347;
T. Sjöstrand and M. Bengtsson, Computer Phys. Comm. 43 (1987) 367, and for JETSET 7.3, T. Sjöstrand, CERN-TH-6488-92 (1992).

[33] B. Andersson, G. Gustafson, C. Sjögren, Nucl. Phys. B380 (1992) 391.

[34] B.R. Webber, Nucl. Phys. B238 (1984) 492.

[35] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, Computer Phys. Comm. 67 (1992) 465.

[36] R. Nisius, Ph.D. thesis, RWTH-Aachen, July 1994.

[37] M. Nyberg, private communication.

[38] G. Ingelman, private communication.

[39] A.D. Martin, W.J. Stirling, R.G. Roberts, Phys. Lett. 306B (1993) 145, ibid. 309B (1993) 492.

[40] T. Sjöstrand et al., *in* Z Physics at LEP 1, eds. G. Altarelli, R. Kleiss and C. Verzegnassi, CERN 89-08, 1989, Vol. 3, p. 143

[41] Review of Particle Properties, Phys. Rev. D45, 1992.

[42] L. Lönnblad, Computer Phys. Comm. 71 (1992) 15.

[43] G. Gustafson, private communication.

[44] G. Marchesini et al., `herwig57.doc`, a computer readable manual available on *asis01.cern.ch* in `/pub/asis/share/cern/doc/txt`.

[45] J. Chýla, H1 internal note H1-05/93-291, Jan. 1994.

[46] V. Blobel, DESY internal report R1-88-01, Jan. 1988.

[47] Z. Feng et al., H1 internal note H1-06/90-139, June 1990.

[48] H1 Software Notes 12 and 13, Nov. 1990.

[49] ZEUS Collab., M. Derrick, et al., Phys. Lett. B315 (1993) 481.

[50] H. Lohmander, H1 internal note H1-03/94-347, March 1994.

[51] P. Reimer, private communication.

[52] F. Jacquet, A. Blondel *in* Proceedings of the study of an *ep* facility for Europe, DESY 79/48 (1979) 391.

[53] S. Bentvelsen et al., *in* Proc. HERA workshop, Eds. W. Buchmüller and G. Ingelman, Hamburg (1991) Vol. 1, 23.

[54] K. Golec-Biernat et al., University of Durham preprint DTP/94/30, May 1994.

[55] OPAL Collab., M.Z. Akrawy et al., Z. Phys. C47 (1990) 505.

[56] R. Brun et al., GEANT3 User's Guide, CERN–DD/EE 84–1, Genf (1987)

[57] M. Rudowicz, Ph.D. thesis, Max-Planck-Institut für Physik preprint MPI-PhE/92-14, September 1992.
S. Peters, Ph.D. thesis, Max-Planck-Institut für Physik preprint MPI-PhE/92-13, September 1992.

[58] H1 Collab., I. Abt et al., DESY preprint DESY-94-055, March 1994.

[59] A. Wegner, Ph.D. thesis, DESY internal report, July 1994.

[60] L. Lönnblad, private communication.

[61] B. Andersson, private communication.

[62] M.H. Seymour, private communication.

[63] H. Fesefeldt, RWTH Aachen preprint Pitha 85/02 (1985).

[64] P.A. Aarnio et al., CERN Fluka89 User Guide (1990).

[65] A. Kwiatkowski, H. Spiesberger, and H.-J. Möhring, Computer Phys. Comm. 69 (1992) 155.

[66] K. Charchula et al., CERN preprint CERN-TH-7133-94, Jan 1994.

[67] T. Sjöstrand, CERN preprint CERN-TH-6488-92, May 1992.

[68] H. Jung, DESY preprint DESY-93-182, Dec 1993.

[69] H1 Collab., T. Ahmed et al., Phys. Lett. B298 (1993) 469.

[70] R.P. Feynman, Phys. Rev. Lett. 23 (1969) 1415.

[71] EMC Collab., M. Arneodo et al., Z. Phys. C35 (1987) 417.

[72] DELPHI Collab., P. Abreu et al., Phys. Lett. B311 (1993) 408.

[73] L.V. Gribov et al., Sov. Phys. JETP 68(1) (1988) 1303.

[74] D. Kant et al., H1 internal note, in preparation.

[75] A.H. Mueller, Nucl. Phys. B213 (1983) 85 and B241 (1983) 141(E).
Yu. L. Dokshitzer, S. Troyan, Leningrad Physics Institute preprint 922 (1984).

[76] E.R. Boudinov et al., Phys. Lett. 309B (1993) 210.

[77] N.C. Brümmer, NIKHEF-H/94-09, subm. to Phys. Lett. B.