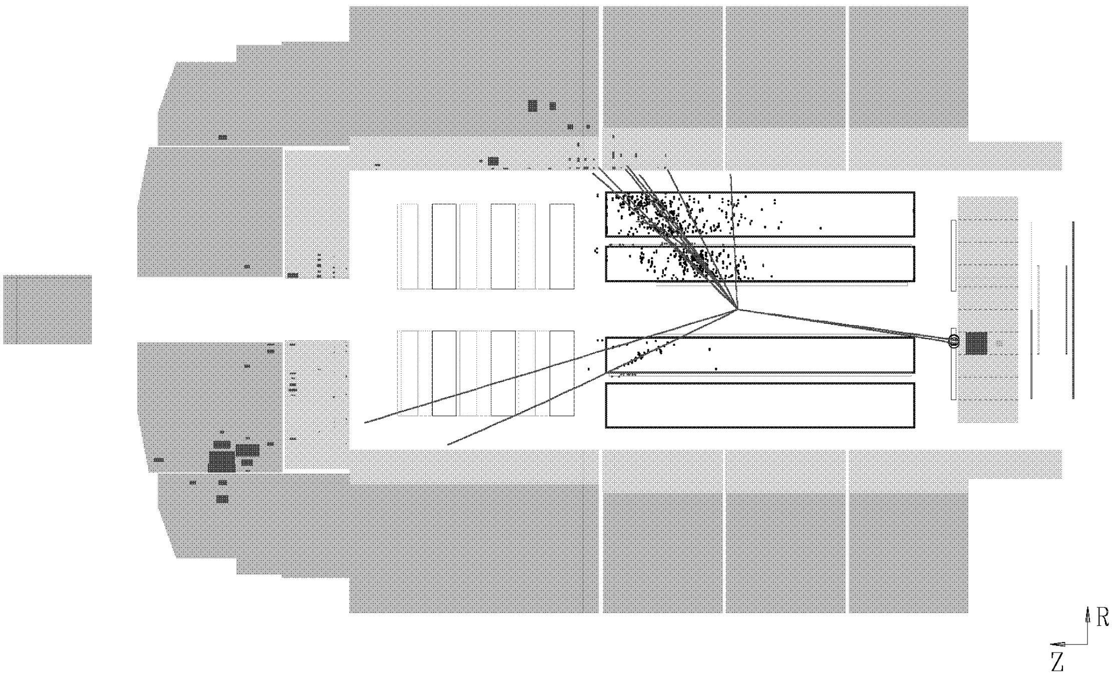
Universiteit Antwerpen
Universitaire Instelling Antwerpen
Departement Natuurkunde

Multiplicity structure of the hadronic final state in deep inelastic *ep* scattering

Proefschrift voorgelegd tot het behalen van de graad
Doctor in de Wetenschappen aan de Universitaire Instelling Antwerpen
te verdedigen door **Pierre Van Mechelen**



Promotor: Prof. Dr. E. De Wolf

Antwerpen, april 1998

Multiplicity structure of the hadronic final state
in deep inelastic ep scattering

Acknowledgements

First and foremost, I would like to express my gratitude to the promotor of this work, Prof. Eddi De Wolf. He always had time to answer my questions and to discuss problems I encountered with the experimental analysis. His guidance was invaluable for my understanding of the experimental results presented in this work. His profound knowledge and experience in high energy physics allowed me get an in depth account of HERA physics, without disregarding the link with other fields of science.

I also owe my thanks to Prof. Frans Verbeure for accepting me as a Ph. D. student and for encouraging me to participate in various schools, workshops and conferences.

In his capacity as Physics Coordinator for the H1 data taken in 1994, Albert De Roeck constantly encouraged me to produce new results. His help during the analysis and his support when satisfactory results were finally presented to the H1 Collaboration were invaluable.

During the first years of my work as a Ph. D. student Michael Kuhlen was the convener of the H1 Physics Working Group I was working in. His enthusiasm inspired me while I took my first steps in the exciting world of experimental particle physics. With the arrival of Frank Botterweck in the H1-Antwerpen group, my work boosted into a higher gear. His energy and tenacity were amazing and he certainly helped me getting to know my way around in the H1 Collaboration. Thank you Frank !

I appreciate all members of the IIHE who cheered up my stays at DESY. I particularly thank Pierre Marage and Robert Roosen for encouraging me and Patrick Van Esch for introducing me to the technical aspects of the multi-wire proportional chambers and the H1 data acquisition system.

I acknowledge all my H1 colleagues who contributed to the achievement of the results presented here. Especially for the analysis of diffractive deep inelastic scattering, I am grateful for the input of the Diffraction Working Group, in particular from Chris Cormack, Andrew Mehta and Julian Phillips.

I thank Sabine De Brabandere and the other colleagues of the UIA for the pleasant company and for creating a nice and stimulating atmosphere.

I acknowledge the IIKW for the financial support.

Finally, I also thank my girlfriend Sabine for proofreading the draft and for helping me focus on relevant subjects by asking critical questions.

Contents

Introduction

The scattering of electrons from hadrons has historically been an important tool for the understanding of the basic constituents of matter. Electrons are particularly useful probes, since they are point-like and stable and can be easily produced and accelerated. Moreover, their interactions can be described in the well understood and calculable electroweak sector of the Standard Model.

The resolution with which the target hadron is probed is measured by the square of the four-momentum transfer, Q^2. Early experiments at SLAC and DESY observed a scaling behaviour, such that, other than the $1/Q^4$ term arising from the exchange photon propagator, there is no discernible dependence of the cross section on Q^2 when the Bjorken scaling variable $x = Q^2/2m\nu$ is fixed, where ν is the energy transferred in the target rest frame, and m is the invariant mass of the target.

Such deep inelastic ep interactions have been instrumental in the development of the modern picture of hadrons, as being composed of point-like quarks with spin one half and fractional electric charge. These *quarks* interact with one another via the strong force, by the exchange of intermediate gauge bosons, the *gluons*.

In the theory of quarks and gluons, *Quantum Chromodynamics*, the coupling constant α_S plays a central role. The main difficulty is that this coupling constant is dependent on the scale (say, momentum transfer) of the process involved and that it becomes large for large scales (small Q^2). This means that a perturbative expansion in α_S will eventually break down and that the problem of confinement of quarks and gluons in colourless particles cannot be solved in perturbative quantum chromodynamics.

In deep inelastic ep scattering, the hadronic final state is the result of both a perturbative stage which describes what happens to the parton shortly after (and before) it is kicked out of the proton, and a non-perturbative stage where the 'partonic final state' is transformed in the hadronic final state. The problem is then to disentangle the dynamics of the hard subprocess and the soft hadronisation phase.

HERA is the world's first electron-proton collider. The large centre-of-mass energy ($\sqrt{s} \approx 300$ GeV) allows to explore new regimes: new particles with masses up to 300 GeV can be produced, the structure of the proton can be studied with a resolving power varying over 5 orders of magnitude down to dimensions of 10^{-18} m and partons with very small fractional proton momenta (Bjorken-x down to 10^{-6}) become experimentally accessible. The much larger centre-of-mass energy, with respect to previous fixed-target experiments, also offers a large phase space for production of hadrons in the final state. This allows the detailed characteristics of the hadronic final state to be probed in a previously unexplored kinematical domain. An advantage of HERA is that most measurements can be performed at a variable scale. The mean multiplicity of charged particles can e.g. be studied as a function of the hadronic centre-of-mass energy and also the dependence on the virtuality of the exchanged photon can be investigated.

The H1 experiment consists of a multipurpose detector, set up around the northern interaction point of HERA, that can measure both the scattered electron and a large part of the hadronic final state. Provided that the transition from partons to hadrons can be properly modelled, this opens up the possibility to study the underlying parton dynamics. Variables based on the hadronic final state can be defined for specific regions in phase space such that e.g. properties of hadrons ranging from the current region to the region of phase space close to the proton remnant can be investigated. By comparing the fragmentation of the current quark to hadronic final states in e^+e^- annihilations the universality of the quark fragmentation process can be tested.

Because no-one can anticipate all future theoretical developments, one of the purposes of this work is to record as much information as possible on the hadronic final state in deep inelastic ep scattering for future reference. One can however, already with present theories and models, interpret the data and draw some conclusions. The comparison with data from other types of processes will further help in this. Together with other results from the H1 and ZEUS experiments, a complete picture of the final state in deep inelastic scattering in the kinematical range available at HERA is slowly emerging.

This work is organised as follows. The first two chapters discuss the theoretical framework that is used to describe particle production in deep inelastic scattering and the statistical tools that will be used in the analysis of the hadronic final state. Chapter 3 gives a short overview of the experimental setup. Chapters 4 and 5 present the experimental results on the multiplicity structure of the hadronic final state in non-diffractive and diffractive deep inelastic scattering, respectively. These results are further analysed and compared to theoretical predictions and to Monte Carlo models. The comparison with experimental results from lower-energy lepton-nucleon interactions, as well as results from e^+e^- annihilations and hadron-hadron collisions, covering a wide energy range, will further help to understand the dynamical processes that are involved in the formation of the hadronic final state in deep inelastic scattering. To conclude, Chapter 6 gives a summary of the main results. For easy reference, all numerical results are listed in Appendix B. They can also be obtained in digital form from the Durham HEPDATA database (for more information, see Appendix B).

A note on units

In this work, a system of natural units is used, whereby $\hbar = c = 1$.

1

Theoretical framework

The multiparticle final state in deep inelastic ep scattering (DIS) is thought to be the result of the interplay of several mechanisms governing the hard ep interaction and the subsequent fragmentation into hadrons. This chapter introduces the various theoretical ideas that are used to describe DIS. Special attention is given to the diffractive photon dissociation process, which accounts for about 10% of the total DIS cross section. Finally, an overview is given of how these ideas are implemented in Monte Carlo models. In the subsequent chapters, these models are put to the test by comparisons to experimental data.

1.1 Deep inelastic scattering

1.1.1 Cross section and kinematical variables

The generic diagram for the DIS process $ep \to eX$ is shown in Fig. 1.1. At fixed ep centre-of-mass energy $\sqrt{s}$, two variables are sufficient to fully determine the kinematics of the scattering process. They are conventionally chosen out of a number of Lorentz-invariant candidates: Q^2, the four-momentum squared of the exchanged boson; Bjorken-x, which in the 'naive' Quark-Parton Model is the fraction of the proton momentum carried by the struck quark (measured in the proton infinite momentum frame); y, the fraction of the energy transferred from the initial electron to the hadronic system in the rest frame of the incoming proton; W, the invariant mass of the hadronic final state and ν, the absolute energy transferred from the electron to the hadronic system in the proton rest frame.

These variables can be expressed as a function of the four-momenta of the incoming and outgoing particles as follows [1]:

$$
\begin{aligned}
Q^2 &= -q^2 = -(k-k')^2, && (1.1)\\
x &= \frac{Q^2}{2P\cdot q}, && (1.2)\\
y &= \frac{P\cdot q}{P\cdot k}, && (1.3)\\
W &= (P+q)^2 = sy - Q^2 + m_p^2, && (1.4)\\
\nu &= \frac{q\cdot P}{m_p}, && (1.5)
\end{aligned}
$$

with m_p the proton mass, $s = (P+q)^2 \approx 4E_eE_p$ and the four-momenta as defined in

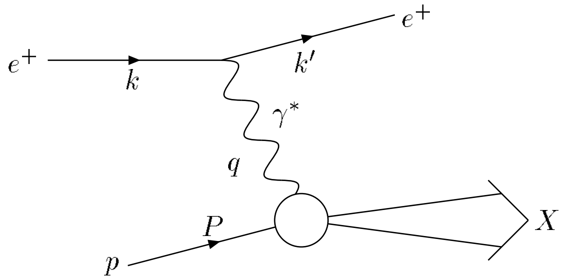


Figure 1.1 : Generic diagram for deep inelastic $ep \to eX$ scattering
In the squared momentum transfer (Q^2) range accessible by HERA, the scattering process is dominated by the exchange of a virtual photon, γ^*, which is radiated off the scattering positron and then interacts with the proton. The lepton vertex can be calculated exactly in Quantum Electrodynamics (QED). The proton vertex, however, has to be parameterised using structure functions, which can be interpreted as describing the parton content of the proton.

Fig. 1.1.

"Deep" and "inelastic" refer to $Q^2 \gg m_p^2$ and $W^2 \gg m_p^2$ respectively, which ensures that perturbative QCD is applicable to the hard scattering process and that nucleon resonance production can be neglected. With the present HERA beam energies, s reaches approximately 90 000 GeV^2. Using the relation $Q^2 = sxy$, one can conclude that x-values down to 10^{-4} can be reached in the deep inelastic scattering regime where $Q^2 > 10$ GeV^2.

Generally the momentum transfer Q^2 is still small compared to the mass of the Z^0 ($\sim$ 91.2 GeV). The scattering process is therefore dominated by the exchange of a virtual photon* and the differential cross section can be expressed in terms of two independent 'structure functions' F_1 and F_2 [1]:

$$\begin{aligned} \frac{\mathrm{d}^2\sigma}{\mathrm{d}x\mathrm{d}Q^2} &= \frac{4\pi\alpha_{EM}^2}{xQ^4}\left[y^2 xF_1(x,Q^2) + (1-y)F_2(x,Q^2)\right], && (1.6)\\ &= \frac{4\pi\alpha_{EM}^2}{xQ^4}\left[1 - y + \frac{y^2}{2}\frac{1}{1+R(x,Q^2)}\right]F_2(x,Q^2), && (1.7)\\ &= \frac{4\pi\alpha_{EM}^2}{xQ^4}\left[\left(1-y+\frac{y^2}{2}\right)F_2(x,Q^2) - \frac{y^2}{2}F_L(x,Q^2)\right]. && (1.8)\end{aligned}$$

α_{EM} is the electromagnetic coupling constant. In the above equations, the cross section is also expressed in terms of the longitudinal structure function F_L and the ratio R defined as:

$$\begin{aligned} F_L &= F_2 - 2xF_1, && (1.9)\\ R &= \frac{F_L}{F_2 - F_L} = \frac{F_2 - 2xF_1}{2xF_1} = \frac{\sigma_L}{\sigma_T}. && (1.10)\end{aligned}$$

R can be interpreted as the ratio of the cross sections σ_L and σ_T for the absorption of longitudinally and transversely polarised virtual photons on protons, respectively, with $\sigma_{tot}^{\gamma^* p} = \sigma_L + \sigma_T$. This ratio is however rather small for $Q^2 > 10$ GeV^2, as determined

*For $Q^2 \approx 1000$ GeV^2 the contribution from Z^0 exchange is about 1%.

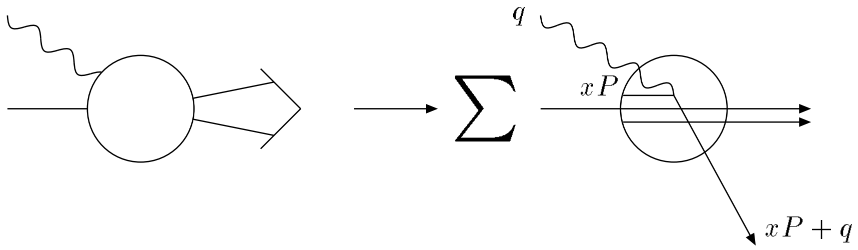


Figure 1.2 : The Quark-Parton Model approximation of deep inelastic scattering
The hadronic vertex of Fig. 1.1 is approximated as a sum over all the quark and antiquark constituents of the proton. The interaction between the virtual photon and a quark is calculable in QED.

in [2].

The structure functions F_i are arbitrary functions of the kinematical variables x and Q^2 describing the $\gamma^* p \to X$ transition. They have to be obtained from experiment and their determination is a major part of the experimental program at HERA [3, 4].

1.1.2 Views on deep inelastic scattering

The Quark-Parton Model

It is traditional to analyse DIS from the standpoint of the proton infinite momentum frame, since in that frame the current is essentially trivial and the physics of the DIS structure functions can be identified with the quark distributions inside the proton. This leads to the Quark-Parton Model (QPM) [1], which is an attempt to gain insight in what happens at the $\gamma^* p \to X$ vertex of Fig. 1.1.

The basic idea is that, at large Q^2, the $\gamma^* p$ interaction can be expressed as the sum of independent scatterings from point-like quark constituents, which behave as if they were free inside the proton during the interaction (see Fig. 1.2). The argument is that the scattering occurs over a short timescale of order $1/Q$ and that, due to relativistic time dilatation, the photon sees a frozen state of non-interacting quarks. The final hadronisation process, which is a direct consequence of the colour confinement property of QCD, occurs only much later.

In this case, the relation between the structure functions and the quark densities inside the proton is simple:

$$\frac{F_2}{x} = 2F_1 = \sum_q e_q^2 f_q(x), \tag{1.11}$$

where the sum runs over all quark flavours, e_q is the quark charge and $f_q(x)$ is the density of quarks with momentum fraction x inside the proton (the left equation is known as the Callan-Gross relation and only holds exactly in the QPM). Note that the structure functions exhibit so-called Bjorken scaling, i.e. they are independent of Q^2.

In addition to the three 'valence' quarks which carry its quantum numbers, the proton also contains an infinite 'sea' of $q\overline{q}$ pairs. When probed at a scale Q, all sea quark flavours

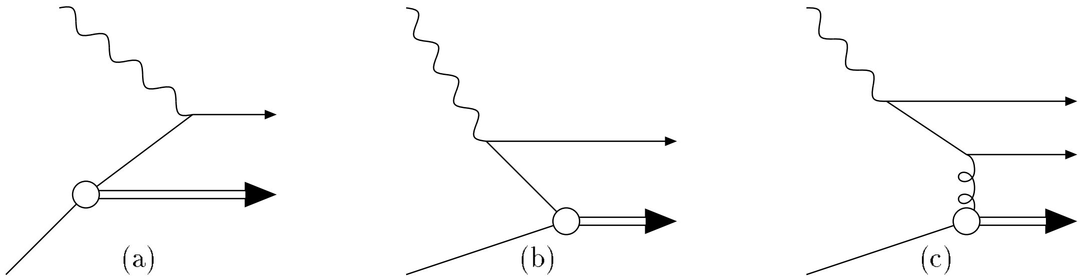


Figure 1.3 : Time-ordered contributions to DIS in the proton rest frame
Time runs from left to right. Diagram (a) shows how the photon can scatter on a parton already present in the proton's wave function. However, in the proton rest frame, the dominant time-ordered process is (b) where to photon first fluctuates in a $q\overline{q}$ pair and then interacts with the proton. (c) shows how this can be achieved through the exchange of a gluon.

with mass $m_q \lesssim Q$ are active. Experimental measurements of the structure functions for deep inelastic $\mu p, \mu D \to \mu X$ and $\nu N \to \mu X$ scattering were used to estimate the $f_q(x)$ distributions and early on it was found that the quarks carry only about half of the momentum of a nucleon [5]. The remainder can only be explained using Quantum Chromodynamics (QCD) and is attributed to the gluon constituents of the proton.

The photon fluctuation picture

When viewed in the proton rest frame, an alternative picture for DIS emerges [6]. In general, one must consider time-ordered processes where the virtual photon creates pairs (as in Fig. 1.3b) or scatters on the quark constituents already present in the proton wave function (Fig. 1.3a).

In the limit of large photon energy ν $(= E_e - E_e')$, the dominant time-ordered process in the proton rest frame is virtual $q\overline{q}$ pair production (as in Fig. 1.3b), where one of the quarks of the pair scatters, annihilates, or is captured in the proton. Thus, at high ν in the proton rest frame, the physics of DIS is dominated by the Fock-state structure of the virtual photon, rather than that of the proton.

In fact, it is natural to identify sea- and heavy-quark contributions to the DIS structure functions at small x, with pair-production processes such as shown in Fig. 1.3c. Alternatively, one can identify this type of pair production process with the photon-gluon fusion mechanism or with the evolution of the proton structure functions (see also Sec. 1.2). However, from the standpoint of physics in the proton rest frame, it is most natural to identify pair production at high ν as the materialisation of the photon's hadronic structure by the proton target.

An important concept in the analysis of DIS in the proton rest frame is the Ioffe time τ_I [7], which is defined as the effective distance between the production of the quark pair and its interaction with the proton. In [6] it is shown that:

$$\tau_I = \frac{C_I}{x m_p}, \tag{1.12}$$

where C_I is a dimensionless constant of order 1. Thus, at small x, the photon converts to a quark pair at a large distance before it interacts in the proton. At HERA, where one can study the structure functions down to $x \approx 10^{-4}$, the Ioffe distance can be as large as 10^4 fm. This means that there is even enough time for the $q\overline{q}$ pair to evolve into a more complex partonic system ($q\overline{q}$, $q\overline{q}g$, ...) before the actual interaction occurs.

The $\gamma^* p$ collision will thus bear some resemblance to hadron-hadron interactions. However, theoretically, the $\gamma^* p$ process is much more attractive than hadron-hadron collisions because, thanks to the high virtuality of the photon, the fluctuation into a partonic system is actually calculable in perturbative QCD [8].

1.1.3 Optical model formulation

The proton is not necessarily always destroyed in DIS. Sometimes the proton remains intact or is only excited to a low-mass system. These are processes of the type $ep \to eXY$, where Y is the proton or a low-mass hadronic system with the same quantum numbers as the incoming proton. This is called diffractive DIS in analogy to diffraction in hadron-hadron collisions, where one of the hadrons scatters elastically and the other hadron dissociates. For a fundamental understanding of diffraction in high energy physics, it is necessary to develop a formalism analogous to diffraction or 'shadow scattering' in optics.

As explained in the previous section, the photon can be regarded as a superposition of Fock-states:

$$|\gamma^*\rangle = \sum_k c_k |\psi_k\rangle, \tag{1.13}$$

where the sum over $|\psi_k\rangle = |q\overline{q}\rangle$, $|q\overline{q}g\rangle$, ... runs over states with definite numbers of partons with definite impact parameters and longitudinal momenta. The photon state $|\gamma^*\rangle$ is normalised, so that:

$$\langle\gamma^*|\gamma^*\rangle = \sum_k |c_k|^2 = 1. \tag{1.14}$$

The various states in this superposition are absorbed (into inelastic channels) in different amounts by the proton, so that the superposition of states which arises from shadow scattering is not simply proportional to the incident one. Shadow scattering thus leads not only to elastic scattering, but also to production of inelastic states which have the same internal quantum numbers as the incident photon.

In [9] it is argued that the states $|\psi_k\rangle$ are approximate eigenstates of the diffractive part of the S matrix, because the partons have no internal structure and all the variables on which the absorption of the state may depend have fixed values. One can thus write:

$$\mathrm{Im}T|\psi_k\rangle = t_k|\psi_k\rangle, \tag{1.15}$$

where $\mathrm{Im}T$ is the imaginary part of the scattering amplitude operator, $\mathrm{Im}T = 1 - \mathrm{Re}S$ (in [10] it is shown that the scattering amplitude is mainly imaginary, i.e. the scattering process is driven by a large absorption). The eigenvalue t_k is the absorption coefficient for the state $|\psi_k\rangle$. These eigenvalues vary, of course, with the number of partons and their impact parameters and longitudinal momenta.

The imaginary part of the elastic amplitude is:

$$\langle\gamma^*|\mathrm{Im}T|\gamma^*\rangle = \sum_k |c_k|^2 t_k = \langle t\rangle. \tag{1.16}$$

In other words, it is given by the average over absorption coefficients, which are weighted according to the probability of occurrence of the states $|\psi_k\rangle$ in the virtual photon state $|\gamma^*\rangle$. Using the optical theorem, the total cross section and the elastic cross section (ignoring any contribution from the real part) are given by:

$$\begin{aligned} \sigma_{tot} &= 2\langle t\rangle, & (1.17)\\ \sigma_{el} &= \langle t\rangle^2. & (1.18) \end{aligned}$$

The cross section for the production of diffractive states (i.e. with the same quantum numbers as the incoming virtual photon), with elastic scattering removed, is:

$$\begin{aligned} \sigma_{diff} &= \sum_k |\langle\psi_k|\mathrm{Im}T|\gamma^*\rangle|^2 - \sigma_{el} \\ &= \sum_k |c_k|^2 t_k^2 - \left(\sum_k |c_k|^2 t_k\right)^2 \\ &= \langle t^2\rangle - \langle t\rangle^2. \end{aligned} \tag{1.19}$$

From this, one can see that the cross section for inelastic diffraction is proportional to the dispersion $\langle(t-\langle t\rangle)^2\rangle$ in absorption coefficients for diagonal channels. Diffractive DIS can thus only arise if the parton states $|q\overline{q}\rangle$, $|q\overline{q}g\rangle$, ... are absorbed by different amounts.

Eqs. (1.17)-(1.19) imply $\frac{1}{2}\sigma_{tot} - \sigma_{el} - \sigma_{diff} = \langle t\rangle - \langle t^2\rangle$. The requirement $0 \leq t_k \leq 1$ for the absorption probabilities leads to the upper bound $\sigma_{diff} + \sigma_{el} \leq \frac{1}{2}\sigma_{tot}$ (Pumplin bound [9]). The bound is saturated if each $|\psi_k\rangle$ is either completely transparent or fully absorbed, so that $\langle t\rangle = \langle t^2\rangle$. The remainder of the cross section describes non-diffractive inelastic scattering.

In [9] a specific model to describe the absorption coefficients in a phenomenological manner is used. One of the results is that, for small momentum transfers t between the photon and the proton, the main contribution to the (large) dispersion comes from the variation in the number of partons.

1.2 Particle production in deep inelastic scattering

1.2.1 QCD extension of the parton model

In QCD it becomes clear that quarks cannot be regarded as completely free when struck, but that they couple to gluons and that the QPM is only the zeroth order approximation in a perturbative expansion in the strong coupling constant α_S.

Quarks may radiate (and absorb) gluons, which in turn may split into quark-antiquark or gluon pairs. With increasing Q^2, more and more of these fluctuations inside the proton can be resolved, which results in a depletion of quarks at large x and a corresponding accumulation at lower x. In fact, at small x and large Q^2, it is the gluon content that governs the interaction with the proton and gives rise to the DIS cross section via the

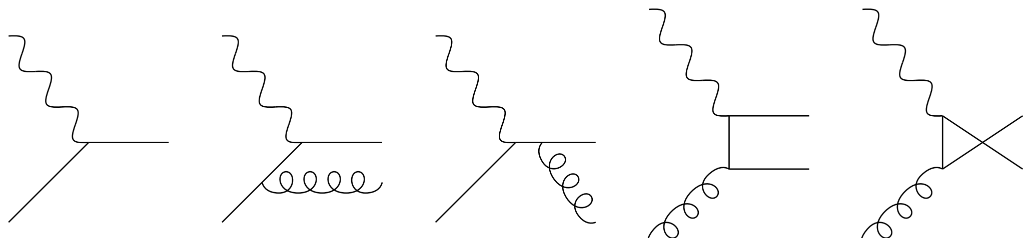


Figure 1.4 : Feynman diagrams for deep inelastic scattering up to $O(\alpha_S)$
In addition to the lowest order QPM diagram (a), the following $O(\alpha_S)$ corrections are shown: (b) initial state QCD Compton radiation; (c) final state QCD Compton radiation; (d) and (e) the Boson-Gluon Fusion (BGF) process.

creation of $q\overline{q}$ pairs. Such scale-violating Q^2 dependence has been observed in the data [3, 4, 11].

Higher order QCD corrections to the $\gamma^* q \rightarrow q$ process have been calculated to $O(\alpha_S^2)$. In most Monte Carlo models, however, only the matrix elements up to $O(\alpha_S)$ are used. Figure 1.4 shows the Feynman diagrams for some of the $O(\alpha_S)$ corrections to the QPM. These diagrams are divergent in the limit when the gluon energy or the opening angle vanishes (soft or collinear singularities). The regularisation of these divergences leads to the running of the coupling constant α_S. Furthermore, in order for the parton model to survive, remaining singularities have to be absorbed in the parton densities, with the consequence that these become Q^2 dependent. By far the most common way to accomplish this is the modified Minimal Subtraction scheme ($\overline{\mathrm{MS}}$) [12].

1.2.2 Parton showers

Leading-Log parton showers

A complete calculation of the matrix elements to orders beyond $O(\alpha_S^2)$ is at present not possible. Perturbative QCD can however go a long way in describing the parton densities inside the proton as well as after the hard interaction with the photon, down to a resolution scale Q_0 of a few GeV (the perturbative description must break down when α_S becomes too large). This is done by considering processes as illustrated in Fig. 1.5, where gluons are emitted and give rise to initial and final state parton showers.

While in the initial-state cascade a parton, typically close to the mass-shell, radiates a gluon and becomes further off the mass-shell, in the final state cascade its virtuality will decrease at each step in the cascade. The initial state cascade is constrained by the proton structure functions. In fact, the parton shower approach is equivalent to the DGLAP equations [13] which describe the evolution of the structure functions with Q^2.

In the Leading-Logarithmic Approximation (LLA) [14] the probability $\mathcal{P}$ for a branching $q \rightarrow qg$ to take place with a small change in virtuality $\mathrm{d}t$ is:

$$\frac{\mathrm{d}\mathcal{P}_{q\rightarrow qg}}{\mathrm{d}t} = \int \mathrm{d}z \frac{\alpha_S}{2\pi} P_{qg}(z). \tag{1.20}$$

$P_{qg}(z)$ are the Altarelli-Parisi splitting kernels [13] which describe the probability that a parton of four-momentum zp $(0 \leq z \leq 1)$ will be found in what started as a parton of

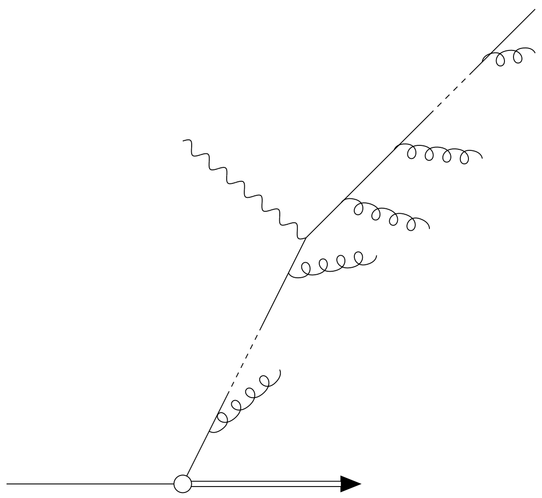


Figure 1.5 : Parton showers in DIS
Gluon emissions before and after the boson vertex give rise to the initial and final state parton showers, respectively.

initial momentum p before the branching. These splitting kernels can be calculated in perturbative QCD.

Because parton showers rely on emission probabilities, not amplitudes, coherence effects require special attention. In particular, the separation of the initial and final state parton shower implies that interference terms, which were calculated correctly in the matrix element approach, are neglected. In practice, interference effects can be realised by phase space restrictions for parton emission. For example, the effect of destructive interference between subsequent emissions in a parton shower can be approximated by imposing 'angular ordering': in the final state cascade, only emissions with a smaller opening angle than the previous one are allowed, so that $\theta_{i+1} < \theta_i$. The physical reason for this is that large wavelength quanta cannot be emitted from dipole sources with small transverse momenta [14].

The evaluation of a diagram such as in Fig. 1.5 involves integrations over the internal momenta exchanged between the partons [5]. The result will be something like:

$$\alpha_S \int \frac{\mathrm{d}k_{Ti}^2}{k_{Ti}^2} \cdots \int \frac{\mathrm{d}x_i}{x_i} \cdots, \tag{1.21}$$

where the dots represent functions which depend on the actual nature of the emitted partons and their dynamics and k_T and xP are the transverse and longitudinal momentum components of the partons, respectively (P is the proton momentum).

With strong k_T ordering, $Q_0^2 \ll \ldots k_{Ti-1}^2 \ll k_{Ti}^2 \ll \ldots Q^2$, the nested integration over all n branches can be carried out. The result is an expression $\propto (\alpha_S \ln(Q^2/Q_0^2)^n$. The k_T integration thus yields large logarithms when the k_T's are strongly ordered. They compensate for the smallness of α_S. Since α_S decreases only logarithmically with Q^2 and is compensated by a logarithmically growing term in Q^2, all graphs with branches up to $n = \infty$ need to be summed up. This is called a Leading-Log Approximation (LLA), since each power n in α_S is accompanied by the same (maximal) power of $\ln(Q^2/Q_0^2)$. Subleading terms would be $\propto \alpha_S^n \left(\ln(Q^2/Q_0^2)\right)^{n-1}$. The LLA is expected to be a good approximation when Q^2 is large, provided that x is not too small in order not to produce large $\ln 1/x$ terms also.

A special case for which the integration in Eq. 1.21 can still be carried out analytically, occurs when in addition to the above conditions also strong ordering in x is required, $x \ll \ldots x_{i+1} \ll x_i \ll x_{i-1} \ll \ldots x_0$. The large logarithmic terms arising from the integration are then of the form $\propto (\alpha_S(Q^2) \ln \frac{Q^2}{Q_0^2} \ln \frac{1}{x})^n$. This is the Double Leading-Log Approximation (DLLA) [5, 14]. It is expected to hold for large Q^2 and small x.

The DLLA turns out, however, to overestimate the gluon multiplicities in the parton cascade. One of the shortcomings is that the DLLA completely ignores the energy-momentum balance since the energy of the radiating parton remains unchanged after a gluon emission. To control the parton production, the Modified Leading-Log Approximation (MLLA) takes into account non-leading single logarithmic effects [14].

The Colour-Dipole Model

Another type of parton shower model is the Colour Dipole Model (CDM) [15]. The colour charges of the scattered quark and the proton remnant are assumed to form a colour dipole, from which gluons can be radiated. Subsequent gluon radiation emanates from dipoles spanned between the newly created colour charges and the others, and so on. To good approximation it can be assumed that these dipoles radiate independently. The CDM uses the leading order cross section for the emission of a gluon with transverse momentum p_T at rapidity $y^\dagger$ in the soft gluon approximation [1]:

$$\mathrm{d}\sigma = \frac{N_c \alpha_S}{2\pi} \frac{\mathrm{d}p_T^2}{p_T^2} \mathrm{d}y, \tag{1.22}$$

with N_c the number of colour charges. The cross section is uniform in rapidity and $\ln p_T^2$. Kinematically, the phase space is bounded by $|y| < \ln(E/p_T)$, where E is the total energy of the radiating system.

In DIS one of the colour charges, the proton remnant, is not pointlike. The suppression of radiation with short wavelength from an extended source is taken into account with a special parameter of $O(1\ \mathrm{fm})$ that describes the extent of the remnant.

In practice, these parton shower models are used in combination with the exact fixed order matrix elements to take care of the hard emissions that are not properly covered. The CDM for example covers the QCD-Compton graphs by dipole radiation, but for the BGF graph the matrix element is used.

In contrast to the leading logarithmic parton shower, in the CDM it is not possible to make the distinction between initial and final state radiation. Soft gluon interference is automatically taken into account.

1.2.3 Hadronisation

Eventually the perturbative calculation of parton showers has to be stopped because, at a scale Q_0 of the order of 1 GeV, the coupling constant α_S becomes too large. The subsequent hadronisation stage describes the non-perturbative process of transforming the

†Rapidity is defined in Sec. 1.2.4.

final, coloured partons into a jet of colour-neutral hadrons. Two different approaches exist to compare the parton level result of perturbative QCD calculations and the hadron-level measurements from real-world experiments.

Local Parton-Hadron Duality

A first way out is to assume that a parton jet resembles a hadron jet in some aspects. An initial argument for such a behaviour was the proof of 'pre-confinement', i.e., while still in the perturbative region, the parton cascade already prepares colour singlet clusters with finite mass independent of energy [16]. One can further distinguish two kinds of observables:

- **Infrared and collinear safe observables.** In this case the value of the observable does not change if a soft particle is added or if one particle is split into two collinear particles. Such observables do not depend on the cutoff scale Q_0 and are therefore likely to be independent of the final stage of the jet evolution. Quantities of this type are e.g. energy flows and -correlations and variables describing the global shape of the event in momentum space.

- **Infrared sensitive observables.** Here particles are counted. Examples are multiplicities, inclusive spectra and particle correlations. These observables depend on the cutoff Q_0. If one cannot 'fix' these observables by proper normalisation or rescaling so that they become infrared safe, one can interpret Q_0 as a hadron mass (say $Q_0 \sim m_\pi$) and compare the observables for a parton jet evolved down to hadronic scales directly to the experimental data. This procedure has been shown to work for momentum spectra and is called 'Local Parton-Hadron Duality' (LPHD) [17].

The connection between parton level and hadron level dynamics has been investigated by the Lund group for parton cascades treated in the CDM [18, 19]. For centre-of-mass energies above ~ 50 GeV, the fluctuations in the *hadron* multiplicity are to better than 90% determined by the hardest and second hardest gluons emitted. Further softer radiation and subsequent (Lund string) hadronisation, adds only small (sub-Poissonian) fluctuations to those induced in the initial stage of the shower development. These analytical results provide a quantitative realisation of the notion of LPHD derived directly from perturbative QCD.

The LPHD hypothesis has been experimentally investigated by the LEP and HERA experiments [20–22] and the data were found to be consistent with the predictions of the MLLA and LPHD.

Lund string fragmentation

Specific models have been developed to describe the hadronisation process. The Monte Carlo models described in Sec. 1.4 all use the Lund string model. However, alternatives do exist, e.g. the cluster model by Marchesini and Webber [23], implemented in the HERWIG Monte Carlo generator [24].

In the Lund string model [25], a quark and an antiquark moving apart stretch a colour field between them. The field is thought to be string-like with constant energy density per unit length of $O(1\ \mathrm{GeV/fm})$, and with transverse dimension of $O(1\ \mathrm{fm})$. When the stored energy becomes large enough, the string can break up and create a $q\overline{q}$ pair from the vacuum. The newly created q and $\overline{q}$ terminate the loose string ends, such that the new string pieces are themselves colour neutral. The process is iterated with the new strings until all the available energy is depleted. The string model gives rise to scaling fragmentation functions, a rapidity plateau with uniform particle density, and a logarithmic increase of mean multiplicity $\langle n \rangle = a + b \ln E$, where E is the string invariant mass.

1.2.4 Phase space for the hadronic final state

The properties of the hadronic final state will be studied in the hadronic centre-of-mass system (CMS), defined by the condition $\vec{P} + \vec{q} = 0$. The positive z-axis is usually defined by the direction of the virtual photon, $\vec{q}$. Particles moving forward are therefore said to belong to the current or photon hemisphere, while particles going backwards are assigned to the target or proton remnant hemisphere.

The rapidity y of a particle with momentum $p = (E, p_x, p_y, p_z)$ is defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \tag{1.23}$$

However, often the mass of a particle is not known and the particle is taken to be a pion with a mass equal to 0.140 GeV. Neglecting the mass altogether, one defines the pseudorapidity η:

$$\eta = \frac{1}{2} \ln \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} = -\ln \tan \frac{\theta}{2}, \tag{1.24}$$

with θ the polar angle of the particle momentum. Rapidity transforms under a Lorentz boost in the z-direction with velocity β as:

$$y \to y' = y - \coth \beta. \tag{1.25}$$

In the hadronic CMS, rapidity can take any value with $y_{min} < y < y_{max}$ (see Fig 1.6), with

$$y_{max} \approx \ln \frac{W}{m} \qquad \text{and} \qquad y_{min} = -y_{max}. \tag{1.26}$$

At HERA, with $W \lesssim 300$ GeV, pions can be produced with $-7.5 < y < 7.5$. The rapidity distribution thus has a total width of $2 \ln(W/m_\pi)$. According to a simple model of independent quark fragmentation [26], the rapidity distribution is expected to have a flat plateau, with a height that is energy independent. The total multiplicity would therefore rise approximately logarithmically with energy W.

Another useful frame of reference is the Breit frame. Here the virtual photon is entirely space-like, $q = (0, 0, 0, Q)$ and the incoming parton with momentum xP has its momentum reversed after being struck. If universality of quark fragmentation holds, the height of the rapidity plateau in the current hemisphere of the Breit frame is expected to be as in e^+e^- annihilation. This can easily be seen in the photon fluctuation picture where a photon with virtuality Q^2 fluctuates in a $q\overline{q}$ pair and one quark interacts with the proton, while the other fragments similarly as in $e^+e^- \to q\overline{q}$ interactions. The current hemisphere in the Breit frame covers a range in rapidity with a width $\sim \ln Q/m$.

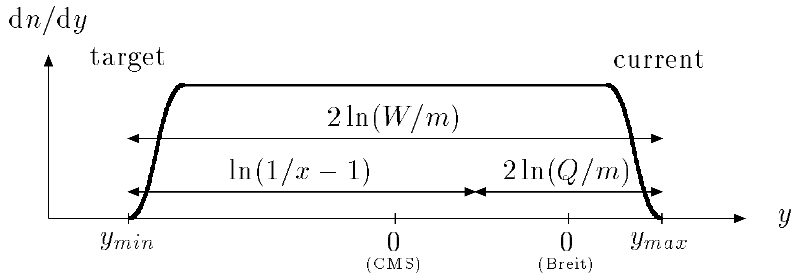


Figure 1.6 : The rapidity distribution in DIS
The sizes of the target and current hemispheres in the CMS and Breit frame are indicated. They can easily be expressed by using the relation $W^2 = Q^2(1/x-1)$ (neglecting the proton mass).

The remainder of the rapidity plateau is expected to have a height similar to the plateau height in hadron-hadron interactions. Therefore, the total multiplicity is approximately given by:

$$\langle n \rangle = C_{e^+e^-} \ln \frac{Q^2}{m^2} + C_{hh} \ln(\frac{1}{x} - 1). \tag{1.27}$$

If the heights $C_{e^+e^-}$ and C_{hh} are not too different, as suggested in [27], the multiplicity will mainly depend on the centre-of-mass energy W alone and not on $1/x$ or Q^2 separately.

In QCD it is expected that hard multi-jet production is less frequent in lepton-hadron collisions than in e^+e^- annihilations [14]. Radiation from the target remnant – a composite colour source – is expected to be strongly suppressed by the so-called 'antenna effect'. (Due to coherence only a certain fraction of the size of an antenna participates in the emission and consequently only the energy available in that fraction can be used for the emission. Short wavelength emission on the other hand corresponds to large energy momentum of the quanta and that is the origin of the damping.) For sufficiently small values of Bjorken-x it also suppresses hard gluon radiation in the current region. The probability per event for hard gluon radiation is consequently larger in e^+e^- annihilation. As shown in [19], the antenna effect leads to a smaller increase with energy W and to smaller multiplicity fluctuations in DIS in the central rapidity region, compared with e^+e^- annihilation.

1.3 Diffraction

1.3.1 Large Rapidity Gap events

At HERA, a class of DIS events has been detected of the type $ep \to eXY$ with a large gap in rapidity with no hadronic activity (see Fig. 1.7a) between the hadronic systems X and Y [28, 29]. In a normal fragmentation process, colour fields span the rapidity range from the struck quark to the proton remnant, leading to a relatively homogeneous distribution of hadrons. Large gaps are therefore infrequent and, moreover, exponentially suppressed. However, in processes where the colour flow is interrupted and where two colour-neutral objects fragment independently, large gaps can be produced. It is thus natural to interpret the 'Large Rapidity Gap' (LRG) events as due to the exchange of a colourless object as in Fig. 1.7b.

Some additional variables, beside the normal DIS variables (x, Q^2), are need to specify

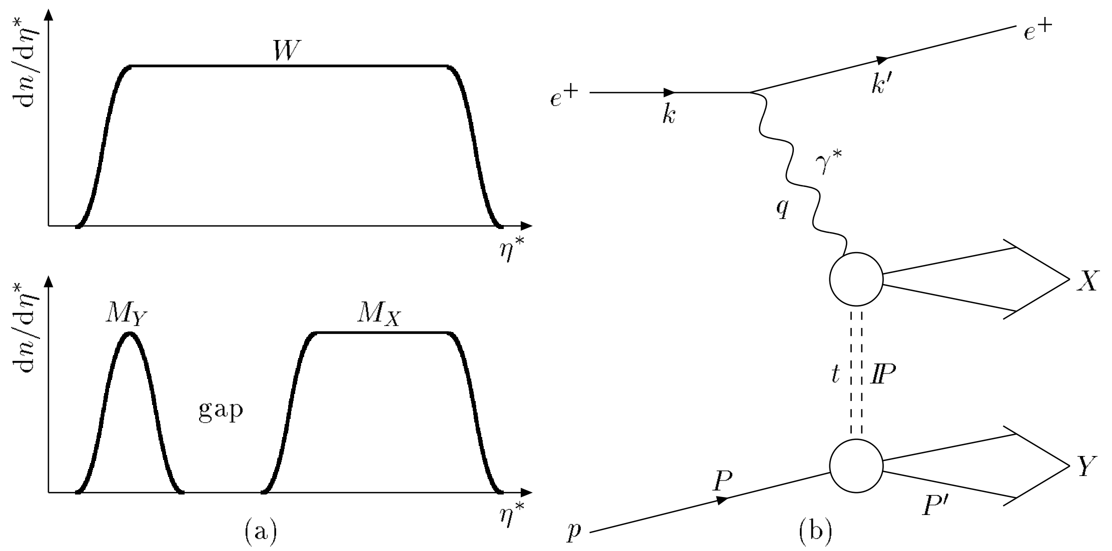


Figure 1.7 : Kinematics of Large Rapidity Gap events
(a) Rapidity distributions for normal DIS events (top) and for LRG events (bottom). The systems X and Y are separated by the largest rapidity gap in the event; Y contains the proton or its remnant.
(b) Generic diagram for LRG events. The virtual photon scatters off a colourless object – a pomeron in Regge phenomenology – inside the proton. Because there are no colour connections between the systems X and Y, a large rapidity gap can develop.

the kinematics of a LRG event [30, 31]. First, the hadronic final state is separated by looking for the largest rapidity gap in the event. The two resulting hadronic systems are labelled X and Y, the last one being the one containing the proton or proton remnant. Using the four-momenta as in Fig. 1.7b the following kinematical variables can be defined:

$$M_Y = P'^2, \tag{1.28}$$
$$M_X = (P - P' + q)^2, \tag{1.29}$$
$$t = (P - P')^2, \tag{1.30}$$
$$\beta = \frac{Q^2}{2(P - P')\cdot q}, \tag{1.31}$$
$$x_{I\!P} = \frac{x}{\beta}. \tag{1.32}$$

Here, the label $I\!P$ is used to indicate the kinematics of the pomeron, although the momentum exchange between the systems X and Y need not be interpreted as pomeron exchange. The above kinematical variables have the following meaning: M_X and M_Y are the invariant masses of the hadronic systems X and Y, respectively; t is the squared momentum transfer at the pY vertex and β and $x_{I\!P}$ can be interpreted in a QPM-like picture as the momentum fraction of the $I\!P$ exchange carried by the struck quark and the momentum fraction of the proton carried by the $I\!P$ exchange, respectively. It is also useful to note that M_X is the centre-of-mass energy of the $\gamma^* I\!P$ interaction and has thus a similar meaning as W in the $\gamma^* p$ CMS.

These new kinematical variables are not independent. The following relations hold:

$$x_{I\!P} = \frac{Q^2 + M_X^2 - t}{Q^2 + W^2 - m_p^2}, \tag{1.33}$$

$$\beta = \frac{Q^2}{Q^2 + M_X^2 - t}. \tag{1.34}$$

LRG events are mainly attributed to diffractive photon dissociation. In the following, the terms "LRG events" and "diffractive events" will be used synonymously, although, in practice, some contribution from non-diffractive and double-diffractive (where also the proton dissociates) processes is to be expected.

1.3.2 QCD-Regge approach to diffractive DIS

Regge phenomenology, which describes elastic hadron-hadron collisions at high energy, can be used to parameterise diffractive DIS. Just like in hadronic interactions, LRG events are thought to take place via the exchange of a pomeron ($I\!P$), carrying vacuum quantum numbers, or reggeons ($I\!R$) related to mesons. Using a QCD approach, the pomeron (and reggeons) is endowed with a partonic sub-structure and the concept of parton distributions in the pomeron and sub-leading reggeons is used to model diffractive DIS [32].

The cross section for diffractive events can then be expressed as follows :

$$\frac{\mathrm{d}\sigma_{ep\to eXY}}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}t\mathrm{d}\beta} = \frac{4\pi\alpha_{EM}^2}{\beta^2 Q^4}\left[1 - y + \frac{y^2}{2}\right] F_2^{D(4)}(x_{I\!P}, \beta, Q^2, t) \tag{1.35}$$

In this model, only single diffractive processes of the type $ep \to eXp$ are considered, such that the proton remains intact ($M_Y = m_p$). The structure function for diffractive events can then be parameterised as follows:

$$F_2^{D(4)}(x_{I\!P}, \beta, Q^2, t) = \frac{\mathrm{e}^{B_{I\!P}t}}{x_{I\!P}^{2\alpha_{I\!P}(t)-1}} F_2^{I\!P}(\beta, Q^2) + \frac{\mathrm{e}^{B_{I\!R}t}}{x_{I\!P}^{2\alpha_{I\!R}(t)-1}} F_2^{I\!R}(\beta, Q^2), \tag{1.36}$$

where the expressions in front of the pomeron and meson structure functions ($F_2^{I\!P}$ and $F_2^{I\!R}$) are the pomeron and meson fluxes in the proton, as motivated by Regge theory [33].

This approach, adopted in a recent study of the diffractive DIS cross section at HERA [31, 34], and assuming the DGLAP evolution equations, indicates that the pomeron must have a large hard gluon content at a resolution scale of a few GeV (see Fig. 1.8). A sub-leading meson exchange is however needed at larger values of $x_{I\!P}$ or M_X, if the hypothesis of factorisation into flux-factors and pomeron and reggeon structure functions is to be maintained.

1.3.3 Soft Colour Interactions

An alternative way to model LRG events is provided by the concept of Soft Colour Interactions (SCI) [35]. In this model the rapidity gaps are due to changes in the colour structure of the event *after* the hard interaction has taken place. At low x, most events are due to photon-gluon fusion, resulting in a colour octet-octet configuration in the final state. In the string model, there would be two fragmenting strings spanned (Fig. 1.9a). It is assumed that while the produced quark and antiquark travel to leave the proton, soft

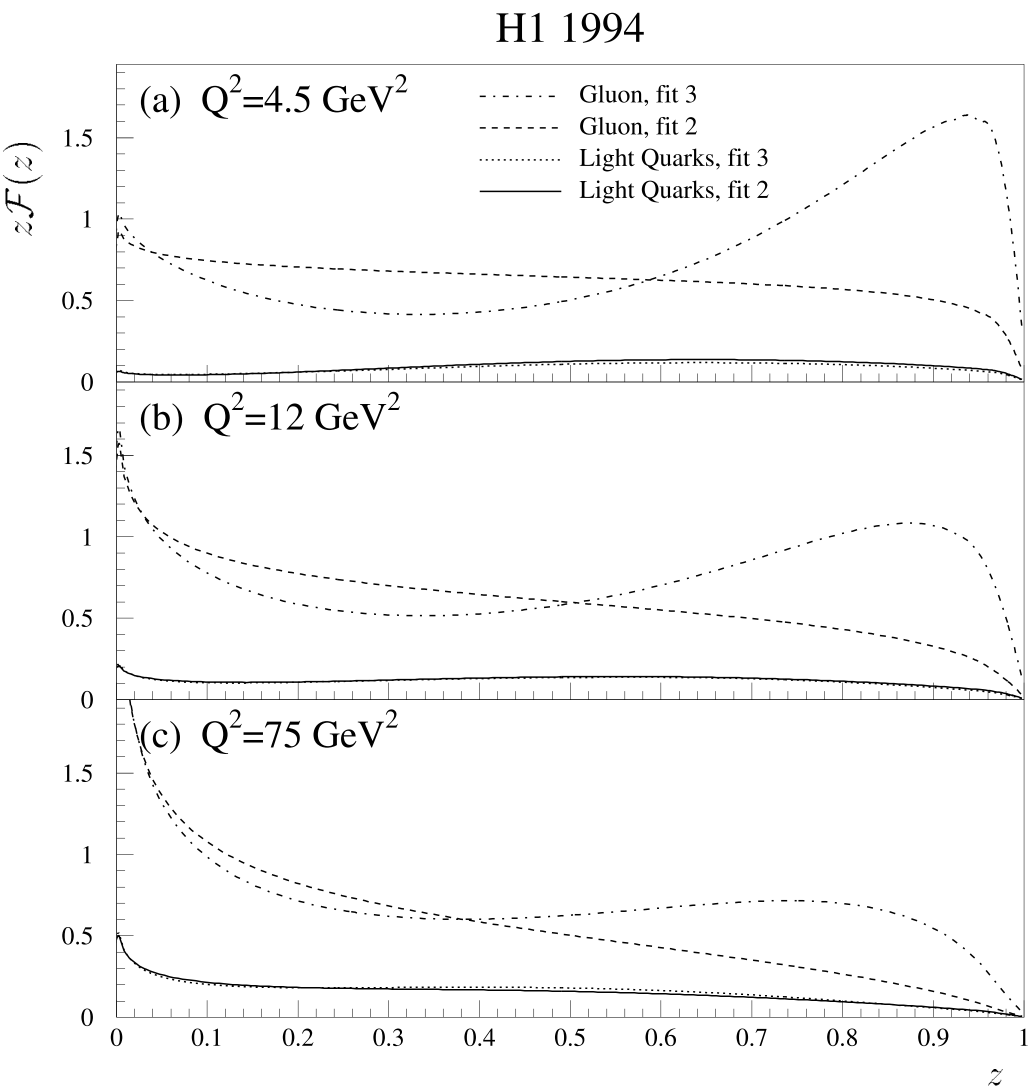


Figure 1.8 : Parton distributions inside the pomeron resulting from a QCD-analysis of the structure function [31]

The sum of the light quark and the gluon distributions are shown at different resolution scales Q^2. In obtaining these results, different parameterisations were used. In "fit 2" the gluon distribution is parameterised at the starting scale $Q_0^2 = 3$ GeV^2 with only the first term in a polynomial expansion, while in "fit 3" the first 3 terms are included. "Fit 1" (not shown) uses only quarks at the starting scale and does not describe the data well. The figures are normalised such that they represent the parton distributions multiplied by the flux factor at $x_{I\!P} = 0.003$. z is the fractional momentum of the pomeron carried by the struck parton.

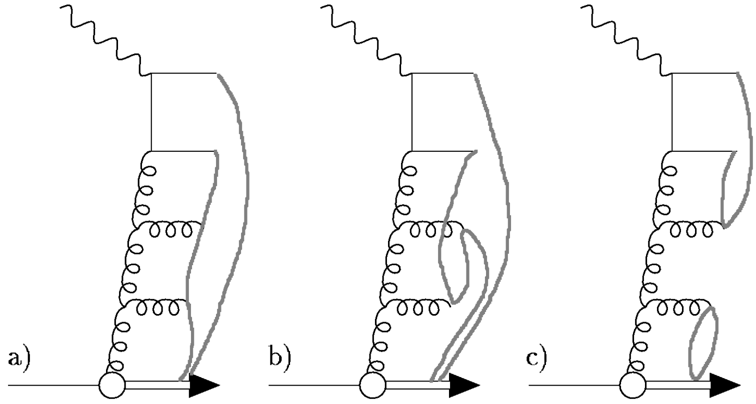


Figure 1.9 : String rearrangement due to Soft Colour Interactions
String configurations for the normal case (a), and after soft colour interactions, leading to either longer strings (b) or a rapidity gap (c).

gluons are exchanged between them and the proton remnant (see Fig. 1.9b). These gluons do not change the momenta significantly, but may rotate the colour configuration of the $q\overline{q}$ pair. The final configuration may turn out to be a colour singlet, in which case there would be no colour field connecting it with the remnant (Fig. 1.9c).

One of the attractive features of this model is an explanation of the absolute rate of LRG events. If soft gluon exchange results in a random colour configuration, the ratio of colour singlet to colour octet configurations would be $1/(1+8) \approx 10\%$, close to the experimental measurement.

1.4 Monte Carlo generators

1.4.1 The JETSET Monte Carlo generator for e^+e^- annihilations

In order to make comparisons with the hadronic final state produced in e^+e^- annihilations, the JETSET Monte Carlo generator [36] is used throughout this work. This generator has been tuned by several LEP experiments (see e.g. [37]) and describes data from e^+e^- experiments very well.

The production process $e^+e^- \rightarrow q\overline{q}$ is a pure electroweak process. Once the quarks are produced they can evolve into gluons and more quarks: a process that is described by Leading-Log parton showers. The hadronisation model implemented in JETSET is the *de facto* standard for Monte Carlo generators. It is based on the Lund string fragmentation model [18, 25].

1.4.2 Monte Carlo models for deep inelastic scattering

In the following chapters, several Monte Carlo models are used to compare the data to theoretical predictions for the hadronic final state of both non-diffractive and diffractive DIS. These models generate events using some experimentally determined parameterisation for the parton densities as input (the GRV parameterisation [38] e.g. describes the HERA F_2 results very well). QCD evolution is implemented using different approximations and phenomenological models are used for the non-perturbative hadronisation phase.

LEPTO

LEPTO [39] is based on the leading order QCD matrix elements with Leading-Log DGLAP parton showers for soft emissions. Therefore, the model is often called "MEPS" (Matrix Elements + Parton Showers). Angular ordering is imposed to model colour coherence. The rather sophisticated Lund string model as implemented in JETSET is used for hadronisation. LEPTO is capable of producing a fraction of events with a large rapidity gap by using the concept of 'Soft Colour Interactions' to rearrange the colour strings during the formation of the hadronic final state (see Sec. 1.3.3).

ARIADNE

In ARIADNE [40] perturbative QCD radiation is modelled with radiating colour dipoles according to the CDM. The same matrix elements as in the MEPS model are used. Therefore this model is often called "MEAR" (Matrix Elements + Ariadne). Hadronisation is performed with the string model as implemented in JETSET. ARIADNE allows to model rapidity gaps either by scattering on a pomeron, or by colour reconnections.

DJANGO

DJANGO is based on HERACLES [41] for the electroweak interaction and on the LEPTO program to model the hadronic final state. HERACLES includes first order radiative QED corrections (see also Sec. 4.1.3) and the simulation of real Bremsstrahlung photons. The CDM as implemented in ARIADNE is used to model QCD parton cascades. The hadronisation is again modelled as in JETSET and is based on string fragmentation.

RAPGAP

The RAPGAP generator models LRG events as deep inelastic scattering of a virtual photon off a pomeron or reggeon coupled to the initial state proton. The pomeron and reggeon are given a partonic content. The meson structure function is taken to be that of the pion [42]. To assess the sensitivity to the quark-gluon content of the pomeron, results are presented for two sets of parton densities (labelled "RG F_2^D (fit 3)" and "RG F_2^D (fit 1)"): i) a 'hard gluon' distribution ("fit 3" in [31]) where gluons carry $\geq 80\%$ of the momentum at the starting scale $Q_0^2 = 3$ GeV^2; ii) a '$q\overline{q}$ only' distribution where only quarks are present at Q_0^2 ("fit 1" in [31]). Different partonic sub-processes are implemented using Born term and first order perturbative QCD matrix elements: $eq \to eq$ scattering, QCD-Compton scattering ($eq \to eqg$) and BGF off a gluon in the colourless exchange ($eg \to eq\overline{q}$). Higher order effects in the QCD cascade are treated with the parton shower model, as implemented in LEPTO. Hadronisation is carried out with the Lund string fragmentation scheme, as in JETSET. QED radiative processes are included via an interface to HERACLES.

2

Statistical analysis of multiplicity distributions

The multiplicity distribution of hadrons produced in high energy interactions is one of the basic measures characterising multiparticle final states. The fluctuation pattern of the number of particles produced in a given domain of phase space reveals the nature of the correlations among hadrons and is, therefore, sensitive to the dynamics of the process. Whereas the total event multiplicity remains of considerable importance, interest has shifted with time towards studies of the multiplicity distribution in subdomains of phase space. In these restricted domains, global conservation constraints are minimised and dynamical correlation effects are better revealed [43, 44].

This chapter presents the statistical tools that will be used in the analysis of the multiplicity structure of the hadronic final state in deep inelastic ep scattering. First the link is made between the multiplicity distributions and its moments and the q-particle correlation functions. Scaling, a recurrent concept in the study of multiplicity distributions, is treated in Sec. 2.2. Finally, some parametric distributions which naturally lead to scaling will be discussed in Sec. 2.3.

2.1 The multiplicity distribution and its moments

The set of probabilities P_n to produce n charged hadrons in a given region of phase space is known as the multiplicity distribution. P_n is simply the ratio σ_n/σ of the inclusive cross section σ_n for n particles in the final state to the total cross section $\sigma = \sum_n \sigma_n$. Because of this, it is a very robust observable with respect to systematic uncertainties which have a global effect on the measurement of the absolute event rates ('luminosity', see Ch. 3).

When studying a multiparticle final state, one can ask whether the production of particles was independent or whether the particles were correlated. Two particles are positively (negatively) correlated when the presence of the first particle increases (decreases) the probability to find the second. It is evident that the fluctuation of the total number of particles will be larger if the particles are positively correlated, while the average multiplicity remains unaffected. In high energy physics, particle correlations were first exploited by Mueller in the formulation of the concept of short range order [45]. Since then, correlations have been extensively used as probes of the interaction and hadronisation dynamics [44]. In the following some variables are defined that can help to quantify these results.

2.1.1 Particle correlations

Consider some inclusive reaction $a + b \to c + X$, $c_1 + c_2 + X$, ... , where the position in phase space of some particles c is fixed and X can be anything. The q-particle density function ρ_q is equal to the inclusive differential cross section:

$$\rho_q(y_1, y_2, \dots y_q) = \frac{1}{\sigma}\frac{\mathrm{d}^q\sigma}{\mathrm{d}y_1\mathrm{d}y_2\dots\mathrm{d}y_q}. \tag{2.1}$$

Here, y_i is a variable needed to specify the position in phase space of particle i (e.g. rapidity).

Generally, these particle densities contain 'trivial' contributions from lower order correlations. Therefore it is advantageous to consider the cumulant correlation functions κ_q, known from statistical physics, which vanish whenever one of their arguments becomes statistically independent of the others. The relation between the ρ_q and κ_q is:

$$\begin{aligned}
\kappa_1(y_1) &= \rho_1(y_1), && (2.2)\\
\kappa_2(y_1, y_2) &= \rho_2(y_1, y_2) - \rho_1(y_1)\rho_1(y_2), && (2.3)\\
\kappa_3(y_1, y_2, y_3) &= \rho_3(y_1, y_2, y_3) - \sum \rho_1(y_1)\rho_2(y_2, y_3) + 2\rho_1(y_1)\rho_1(y_2)\rho_1(y_3), && (2.4)\\
\kappa_4(y_1, y_2, y_3, y_4) &= \rho_4(y_1, y_2, y_3, y_4) - \sum \rho_1(y_1)\rho_3(y_1, y_2, y_3) \\
&\quad - \sum \rho_2(y_1, y_2)\rho_2(y_3, y_4) + 2\sum \rho_1(y_1)\rho_1(y_2)\rho_2(y_3, y_4) \\
&\quad -6\rho_1(y_1)\rho_1(y_2)\rho_1(y_3)\rho_1(y_4), && (2.5)
\end{aligned}$$

and so on (see [46] for the extension to higher order correlations). The summation runs over all possible permutations of particles 1, 2, $\dots q$.

2.1.2 Multiplicity moments

Various types of moments and their combinations play a central role in the analysis of probability distributions. The (plain) multiplicity moments are defined as:

$$\mu_p = \langle n^p \rangle = \sum_{n=0}^{\infty} n^p P_n. \tag{2.6}$$

These moments are frequently normalised to the mean multiplicity as follows:

$$C_p = \frac{\langle n^p \rangle}{\langle n \rangle^p}. \tag{2.7}$$

Additionally, central moments (taken with respect to the mean) are also important:

$$m_p = \langle (n - \langle n \rangle)^p \rangle = \sum_{n=0}^{\infty} (n - \langle n \rangle)^p P_n. \tag{2.8}$$

The most common measure of the width of the multiplicity distribution is the dispersion, which is just the square root of the second central moment:

$$D = \sqrt{m_2} = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}. \tag{2.9}$$

The normalised factorial moments, $R_q = \tilde{R}_q / \langle n \rangle^q$, are defined as:

$$R_q = \frac{\langle n(n-1)\ldots(n-q+1)\rangle}{\langle n \rangle^q} = \frac{1}{\langle n \rangle^q} \sum_{n=0}^{\infty} n(n-1)\ldots(n-q+1)P_n. \tag{2.10}$$

Finally, the normalised factorial cumulants, $K_q = \tilde{K}_q / \langle n \rangle^q$, can be expressed as a function of the normalised factorial moments:

$$\begin{aligned}
K_1 &= R_1, & (2.11)\\
K_2 &= R_2 - 1, & (2.12)\\
K_3 &= R_3 - 3R_2 + 2, & (2.13)\\
K_4 &= R_4 - 4R_3R_1 + 12R_2R_1^2 - 3R_2 - 6, & (2.14)
\end{aligned}$$

and so on. The extension of these formulae to higher orders can be found in [46].

The connection between the multiplicity distribution and particle correlations is made explicit in the relations between the factorial moment (cumulant) of order q, $\tilde{R}_q$ ($\tilde{K}_q$), in a phase space domain (say, rapidity Δ) and the q-particle densities ρ_q and correlation functions κ_q:

$$\begin{aligned}
\tilde{R}_q &= \int_\Delta \ldots \int_\Delta \rho_q(y_1, y_2, \ldots y_q) \mathrm{d}y_1 \mathrm{d}y_2 \ldots \mathrm{d}y_q, & (2.15)\\
\tilde{K}_q &= \int_\Delta \ldots \int_\Delta \kappa_q(y_1, y_2, \ldots y_q) \mathrm{d}y_1 \mathrm{d}y_2 \ldots \mathrm{d}y_q. & (2.16)
\end{aligned}$$

2.1.3 Forward-backward correlations

An alternative method to study the multiplicity structure of the final state is to classify the final state particles according to the z-component of their momentum. Particles with $p_z > 0$ ($p_z < 0$) are assigned to the 'forward' ('backward') event hemisphere. One can then look at the correlation between the number of particles produced in these forward and backward hemispheres. Forward-backward correlations are useful if the statistics is too low to investigate in detail the differential correlation function, while it may already allow to discriminate between various models of multiple production and different hadronisation schemes.

The forward-backward correlation can be studied [47] by plotting the average number of particles produced in one hemisphere against the number of particles in the opposite hemisphere. Experimentally, one finds that this relation is well parameterised by a linear dependency:

$$\begin{aligned}
\langle n_F \rangle &= a_F + b_F\, n_B, & (2.17)\\
\langle n_B \rangle &= a_B + b_B\, n_B. & (2.18)
\end{aligned}$$

Here n_F and n_B are the number of particles produced in the forward and backward hemispheres, respectively. The parameters $a_{F,B}$ and $b_{F,B}$ can be obtained from a fit. The slope b_F can also be calculated using:

$$b_F = \frac{\langle n_F\, n_B \rangle - \langle n_F \rangle \langle n_B \rangle}{\langle n_B^2 \rangle - \langle n_B \rangle^2}. \tag{2.19}$$

A similar equation holds for b_B.

The correction of the two-dimensional probability distribution $P(n_F, n_B)$ for detector effects can however be troublesome, so that it is sometimes easier to calculate the correlation parameter ρ:

$$\rho = \frac{D^2 - D_F^2 - D_B^2}{2 D_F D_B}, \tag{2.20}$$

with D, D_F and D_B the dispersions of the multiplicity distribution in full phase space, and in the forward and backward hemispheres, respectively. This correlation parameter is equal to the slope parameter b in Eq. (2.18) for forward-backward symmetric processes.

In the case that particles are randomly (binomially) distributed between the forward and backward hemispheres, the correlation parameter ρ is given by:

$$\rho = \frac{D^2 - \langle n \rangle}{D^2 + \langle n \rangle}. \tag{2.21}$$

Finally, it is useful to note that, by using particles in opposite hemispheres, ρ_{FB} explicitly measures long-range correlations.

2.1.4 Superposition of multiplicity distributions

In the above discussion the link is made between the width of the multiplicity distribution and particle correlations. Typically, one thinks of short range correlations between particles which are created during the hadronisation process (due to e.g. hadronic resonance decays). There is, however, also another way to interpret broad multiplicity distributions.

Consider the situation where the observed multiplicity distribution is the result of the superposition of processes with a different topology (different average multiplicity). Let's assume that these individual processes produce particles without any correlation. In that case they yield a Poissonian and the total observed multiplicity distribution is equal to:

$$P_n = \int f(\lambda) \mathrm{e}^{-\lambda \overline{n}} \frac{(\lambda \overline{n})^n}{n!} \mathrm{d}\lambda. \tag{2.22}$$

$f(\lambda)$ is a function that describes the distribution of the average multiplicity of the different processes. (Eq. (2.22) is called a Poisson transform.) The function $f(\lambda)$ is constrained by the following normalisation conditions:

$$\int f(\lambda) \mathrm{d}\lambda = \int \lambda f(\lambda) \mathrm{d}\lambda = 1. \tag{2.23}$$

This has the result that the total average multiplicity is simply $\langle n \rangle = \overline{n}$. Furthermore, it is easy to show that the dispersion is equal to:

$$D^2 = \overline{n} + \left(\int \lambda^2 f(\lambda) \mathrm{d}\lambda - 1 \right)^2 \overline{n}^2. \tag{2.24}$$

Thus, the dispersion of the total multiplicity distribution is simply proportional to the spread in λ.

The conclusion is that broad multiplicity distributions can be interpreted as the result of the superposition of distinct event topologies with different average multiplicities. In reality, short range correlations will always be present, so that, if phase space effects can be neglected, the multiplicity distribution will be broader than a Poissonian. Differences in the multiplicity structure of different processes (e^+e^-, lepton-hadron, hadron-hadron, ...) that can not be explained by the correlation introduced in the hadronisation process, should be attributed to fundamental differences in the event topology.

2.2 Scaling

2.2.1 Koba-Nielsen-Oleson scaling

To introduce the scaling laws for the hadronic final state, it is convenient to define the dimensionless variable Feynman-x:

$$x_F = \frac{p_L}{p_{max}}, \tag{2.25}$$

which measures, in the hadronic centre-of-mass frame, the ratio of the longitudinal momentum component p_L of a detected particle to the maximally allowed momentum p_{max}. In the case of deep inelastic scattering $p_{max} \approx W/2$. Clearly x_F varies from -1 to 1.

When plotting a distribution of x_F, one finds that the spectrum decreases rapidly in the range from $|x_F| = 0$ to 1, like $\propto (1 - |x_F|)^n$ for $x \to 1$. At low energy, the shape of the x_F distribution is found to be energy independent (to good approximation), an effect known as 'Feynman scaling' [48]. In the meantime, violation of Feynman scaling has been observed in e^+e^-, lepton-hadron and hadron-hadron collisions.

As a result of Feynman scaling, Koba, Nielsen and Oleson derived a scaling law for the multiplicity distribution [49]. They studied the limit for $n \to \infty$ and $\langle n \rangle \to \infty$ with $z = n/\langle n \rangle$ fixed. In this case one obtains the KNO form:

$$\langle n \rangle P_n = \psi(z). \tag{2.26}$$

The function $\psi(z)$ was shown to become asymptotically independent of the total energy. $\psi(z)$ is normalised by the following conditions:

$$\int_0^\infty \psi(z) \mathrm{d}z = \int_0^\infty z\psi(z) \mathrm{d}z = 1 \tag{2.27}$$

Exact KNO scaling implies that, besides $\psi(z)$, the normalised factorial moments R_q as well as the normalised moments C_q are energy independent. Furthermore, the generalised dispersions $D_q = (m_q)^{1/q}$ are proportional to the mean multiplicity $\langle n \rangle$ (Wróblewski relations [50]).

2.2.2 Golokhvastov generalisation of KNO scaling

A successful generalisation of KNO scaling was proposed by Golokhvastov [51] (called KNO-G scaling). KNO-G scaling needs one assumption, namely that there exists a probability distribution $P(\tilde{n})$ and that this distribution is a continuous function of a continuous parameter $\tilde{n}$. The discrete probability P_n measured in an experiment is then defined in the following way:

$$P_n = \int_n^{n+1} P(\tilde{n})\mathrm{d}\tilde{n} = \int_{n/\langle \tilde{n} \rangle}^{(n+1)/\langle \tilde{n} \rangle} \psi(\tilde{z})\mathrm{d}\tilde{z}, \tag{2.28}$$

$$P(\tilde{n}) = \frac{1}{\langle \tilde{n} \rangle}\psi\left(\frac{\tilde{n}}{\langle \tilde{n} \rangle}\right), \qquad \left(\tilde{z} = \frac{\tilde{n}}{\langle \tilde{n} \rangle}, \quad \langle \tilde{n} \rangle = \int_0^\infty \tilde{n}P(\tilde{n})\mathrm{d}\tilde{n}\right), \tag{2.29}$$

where $\psi(\tilde{z})$ is a scaling function normalised by the equations:

$$\int_0^\infty \psi(\tilde{z})\mathrm{d}\tilde{z} = \int_0^\infty \tilde{z}\psi(\tilde{z})\mathrm{d}\tilde{z} = 1. \tag{2.30}$$

It is worth noting that the normalisation relations for the KNO-G scaling can always be fulfilled (even for small $\langle n \rangle$), which is not the case for standard KNO scaling. Furthermore, it is easy to show that, for $\langle n \rangle \geq 1$, $\langle \tilde{n} \rangle = \langle n \rangle + 0.5$, with very good precision [52].

Graphically, KNO-G scaling can be tested by plotting the cumulative sums $S_n = \sum_{i=n}^{\infty} P_i$ versus $\tilde{z}$.

Scaling, either in KNO or in KNO-G form, is experimentally well established for the full phase space and single hemisphere multiplicity distributions in e^+e^- annihilations, in DIS lepton-hadron interactions and in hadron-hadron collisions [52], except at the highest SPS collider energies [53, 54]. This is remarkable since Feynman scaling is strongly violated in all these processes. However, already in 1970, Polyakov [55] derived the KNO scaling law for e^+e^- hadronic final states within a broad class of field theories. This pre-QCD model can be reformulated in terms of a scale-invariant stochastic branching process. An energy independent coupling constant is assumed and the mean multiplicity rises as a power of energy. The Polyakov derivation, and subsequent work [55], demonstrates that KNO scaling should hold at least approximately in any model based on similar principles.

2.3 Parametric models

In the following chapters, parametric forms will be used to describe the reconstructed multiplicity distributions. These are the results from mainly phenomenological considerations, but are also useful in their own right as an economic representation of the data. In the correction of the observed data distributions, parametric models are used to obtain a correction method that is independent of the Monte Carlo generator models (see App. A).

2.3.1 Negative Binomial Distribution

A much studied distribution is the Negative Binomial Distribution (NBD) defined as:

$$P_n(\overline{n}, k) = \frac{k(k+1)\dots(k+n-1)}{n!} \left(\frac{\overline{n}}{\overline{n}+k}\right)^n \left(\frac{k}{\overline{n}+k}\right)^k, \tag{2.31}$$

with parameters k (or $1/k$) and $\overline{n}$. The average $\langle n \rangle$ and the dispersion D of the NBD are related to these two parameters by:

$$\langle n \rangle = \overline{n} \qquad \text{and} \qquad \frac{D^2}{\langle n \rangle^2} = \frac{1}{\overline{n}} + \frac{1}{k}. \tag{2.32}$$

For $1/k \to 0$ Eq. (2.31) reduces to a Poissonian. The parameter k is equal to the second order factorial cumulant K_2. The NBD is well-behaved in the KNO scaling limit; the scaling function is given by:

$$\psi_k(z) = \frac{k^k}{(k-1)!} z^{k-1} \mathrm{e}^{-kz}. \tag{2.33}$$

The NBD can be regarded as a generalised Bose-Einstein distribution for the ways to distribute n particles in k Bose-Einstein cells of equal average occupancy $\overline{n}/k$. Many phenomenological models for hadroproduction predict multiplicity distributions of a Negative Binomial form. They are reviewed in [56, 57]. In QCD, the NBD is obtained as a solution in the LLA for the gluon multiplicity distribution in a quark jet [58]. Within the framework of the MLLA and LPHD, the lower order factorial moments behave approximately as those of the NBD [16, 17, 59].

It is well-established experimentally in a large variety of collision processes that multiplicity distributions, in full phase space as well as in restricted phase space domains, are approximately of Negative Binomial form. However, deviations are observed in high statistics e^+e^- experiments for final states with hard jets [60–63] and in hadron-hadron interactions [54]. In fixed target DIS lepton-hadron and lepton-nucleus collisions the NBD adequately describes the multiplicity distributions in the full phase space, in limited (pseudo-) rapidity intervals, as well as in the full current and target hemispheres [64, 65]. NBD fits to the multiplicity distribution observed by the H1 Collaboration are discussed in Ch. 4.

2.3.2 Lognormal Distribution

A multiplicity distribution exhibiting explicit KNO-G scaling has been derived by assuming a scale-invariant multiplicative branching mechanism as the basis of the multihadron production [66–68]. Application of the central limit theorem leads to a scaling function ψ of Lognormal form. In this model the multiplicity distribution P_n is related to a continuous probability density $f(\tilde{n})$ and defined as $P_n = \int_n^{n+1} f(\tilde{n})\mathrm{d}\tilde{n}$, where $f(\tilde{n})$ is the Lognormal distribution (LND). The mean continuous multiplicity $\langle \tilde{n} \rangle$ is again approximately given by $\langle \tilde{n} \rangle = \langle n \rangle + 0.5$. Following [66] one finds:

$$P_n = \int_{n/\langle \tilde{n} \rangle}^{(n+1)/\langle \tilde{n} \rangle} \frac{N}{\sqrt{2\pi}\sigma} \frac{1}{\tilde{z}+c} \exp\left(-\frac{[\ln(\tilde{z}+c)-\mu]^2}{2\sigma^2}\right) \mathrm{d}\tilde{z}. \tag{2.34}$$

The integrant in Eq. (2.34) defines the Lognormal scaling function in KNO-G form. It depends on two parameters. Here, $\tilde{z} = \tilde{n}/\langle \tilde{n} \rangle$ is the scaled continuous multiplicity; N, σ, μ

and c are parameters of which only two are independent due to normalisation conditions. In fits to data, correlations between the parameters are reduced if d and c are used as free parameters. $sigma$ and μ can be expressed in terms of d and c as follows:

$$\sigma = \sqrt{\ln\left[\left(\frac{d}{1+c}\right)^2 + 1\right]} \qquad \text{and} \qquad \mu = \ln(c+1) - \frac{\sigma^2}{2}. \tag{2.35}$$

The parameter d is equal to the dispersion of the scaling function. The expression for P_n also depends on $\langle \tilde{n} \rangle$ or $\langle n \rangle$. The latter fit-parameter is denoted by m in the tables in App. B. Exact scaling implies that d and c are energy independent. Finally, the mean continuous multiplicity is found to grow as a power of the energy. Comparisons of the LND with data from e^+e^- annihilation, $\nu(\overline{\nu})p$ and $p\overline{p}$ collisions can be found in [63, 66–71]. Fits to the deep inelastic data observed by the H1 Collaboration can be found in Ch. 4.

3

HERA and H1

The experimental results presented in this work are based on data taken by the H1 experiment at the HERA collider facility. This chapter discusses the experimental setup and the acquisition and reconstruction of the data produced by collision events.

3.1 The DESY site

DESY (*Deutsches Elektronen-Synchrotron*) is a German research centre based in Hamburg and Zeuthen. Its purpose is to conduct basic research in the natural sciences, with activity focused on the investigation of the fundamental properties of matter in particle physics and on the use of synchrotron radiation in surface physics, materials science, chemistry, molecular biology, geophysics and medicine.

Besides the 365 scientific members of the local staff, some 2600 scientists from 280 universities or research institutes and 35 countries are participating in research at DESY.

The institute was founded on December 18, 1959 in Hamburg. The construction of the electron synchrotron "DESY" began in 1960 and operations started with the commissioning of the electron linear accelerator LINAC I in 1964. In 1969 LINAC I was joined by the positron linear accelerator LINAC II. The double storage rings DORIS and PETRA were built from 1969 to 1974 and from 1976 to 1978, respectively. An important milestone was reached in 1979 with the discovery of the gluon at the e^+e^- ring accelerator PETRA [72]. PIA, a positron intensity accumulator, was built in 1979.

In 1980 the synchrotron radiation laboratory HASYLAB was established. This laboratory continues the synchrotron radiation research with DORIS until today. A pre-accelerator for DORIS (and later also for HERA), DESY II, started operations in 1987.

The construction of the hadron-electron ring HERA began in 1984. Because protons would be needed for HERA, a H^- linear accelerator (LINAC III) and a proton synchrotron (DESY III) were added to the existing (pre-) accelerators. The assembly of the detectors H1 and ZEUS in the North and South areas of HERA lasted from 1988 to 1992. The first electron-proton collisions in HERA were already achieved in 1991 and the first events were recorded by the particle physics experiments in June 1992. In 1995 and 1997 the H1 and ZEUS experiments were joined by two fixed-target experiments (HERMES and HERA-B) in the East and West areas of HERA.

In the near future DESY plans a test facility for a linear collider named TESLA, which will collide electrons on positrons with a centre-of-mass energy up to 1 TeV. Such an accelerator would provide new opportunities to study $t\bar{t}$ quark production and a chance for the discovery of a Higgs particle with a mass up to 350 GeV.

year	e^+/e^-	$\langle z \rangle$ (cm)	$\langle I_e \rangle$ (mA)	$\langle I_p \rangle$ (mA)	$\langle L_{sp} \rangle$ (cm^{-2} s^{-1} mA^{-2})	$\int L_{HERA}$ (pb^{-1})	$\int L_{H1}$ (pb^{-1})
1992	e^-		1.40	1.07	$1.87{\cdot}10^{29}$	0.055	0.030
1993	e^-	-5.7	7.72	10.79	$3.15{\cdot}10^{29}$	0.989	0.565
	e^-	$+72.0$	8.33	11.39	$1.68{\cdot}10^{29}$	0.010	0.003
1994	e^-	-0.2	10.49	28.55	$3.37{\cdot}10^{29}$	0.938	0.532
	e^+	$+5.0$	17.00	41.02	$3.62{\cdot}10^{29}$	4.892	3.442
	e^+	$+67.9$	16.78	41.17	$2.77{\cdot}10^{29}$	0.086	0.067
1995	e^+	-1.7	18.40	54.00	$4.06{\cdot}10^{29}$	10.673	5.978
	e^+	$+70.1$	17.91	57.50	$2.85{\cdot}10^{29}$	0.260	0.153
	e^+	-72.7	17.84	49.44	$2.45{\cdot}10^{29}$	0.122	0.064
1996	e^+	-1.5	20.60	60.26	$4.05{\cdot}10^{29}$	14.366	8.919
1997	e^+	-2.1	28.18	73.51	$4.82{\cdot}10^{29}$	34.084	27.345

Table 3.1 : Summary of HERA operations from 1992 to 1997
The lepton ring can be filled with electrons or positrons, the latter having better beam characteristics because they repel remaining anions in the beam pipe. In order to cover a wider kinematical range, the interaction point was shifted along the beam line several times ($\langle z \rangle$ is the mean distance along the beam line of the actual to the nominal interaction point; positive shifts mean that the interaction point was shifted in the proton beam direction). The mean electron and proton currents ($\langle I_e \rangle$ and $\langle I_p \rangle$) increased over the years and, together with a better average specific luminosity ($\langle L_{sp} \rangle$), this resulted in an increase in integrated luminosity from year to year ($\int L_{HERA}$ is the integrated luminosity produced by HERA at final beam energies, $\int L_{H1}$ is the integrated luminosity recorded on tape by the H1 collaboration). The numbers for the year 1997 are still preliminary.

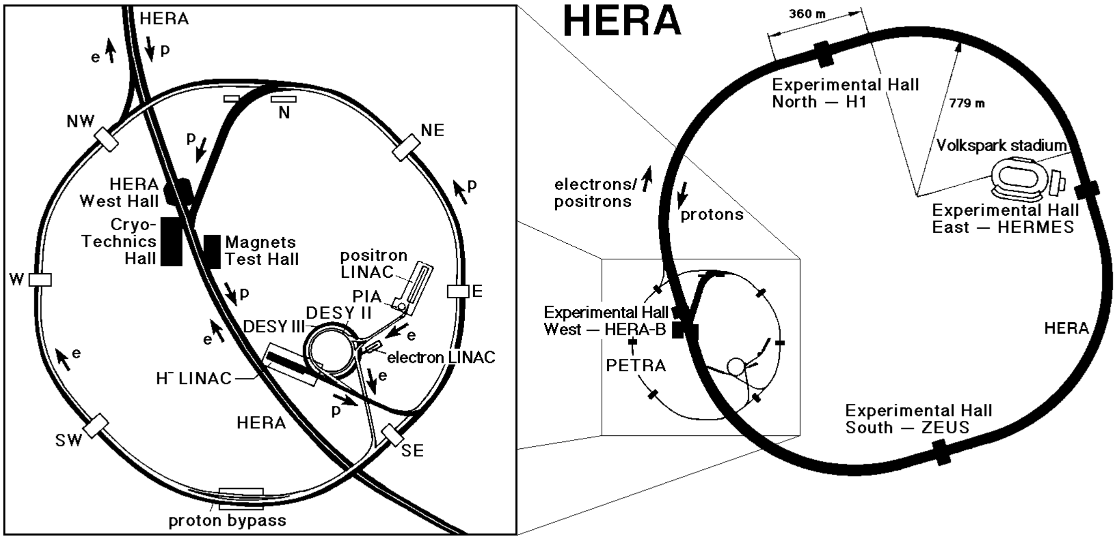


Figure 3.1 : Lay-out of HERA
HERA is a nearly circular ep collider with a diameter of approximately 2 km. The four experimental halls in the North, East, South and West areas are equipped with the H1, HERMES, ZEUS and HERA-B detectors, respectively. The enlarged box shows the system of pre-accelerators with the storage rings DESY and PETRA.

3.2 The HERA machine

The layout of HERA (*Hadron-Elektron-Ring-Anlage*) is shown in Fig. 3.1. HERA collides electrons or positrons* on protons and consists of two separate accelerator systems for the electron and the proton beam. The rings have a circumference of about 6300 m. The beams cross each other in four interaction points in the North, East, South and West areas.

HERA was the first storage ring that collided leptons on nucleons, thereby increasing the centre-of-mass energy of lepton-nucleon collisions by two orders of magnitude with respect to conventional fixed-target experiments.

The protons can reach an energy of 820 GeV, the limiting factor being the magnetic bending field. Superconducting magnets have to be used to achieve a field of 4.86 T. The electrons circulate at the much lower energy of 30 GeV (design value). The limiting factor here is the energy loss due to synchrotron radiation. A continuous power source of 13 MW is needed to keep the particles at their maximum energy.

The DESY and PETRA storage rings are now used as pre-accelerators for HERA (when not needed for HERA, they still serve other purposes as well). The DESY II and DESY III rings accept and accumulate 450 MeV electrons and 50 MeV protons from LINAC II and LINAC III, respectively. The particles are accelerated to 8 GeV and injected in PETRA. There, they are accelerated further to 12 GeV for the electron beam and 40 GeV for the proton beam. The particles enter HERA in groups of 70 bunches. Six injection cycles are needed to fill HERA with 210 proton and 210 electron bunches. This corresponds

*In the following the generic name "electron" is used for electrons as well as for positrons.

to $0.8 \cdot 10^{13}$ electrons and $2.1 \cdot 10^{13}$ protons. The bunch buckets cross each other every 96 ns near the interaction points, but not all the bunches have a collision partner. These unpaired 'pilot' bunches are used for the study of beam induced background.

The design value for the luminosity is $1.5 \cdot 10^{31}$ cm^{-2} s^{-1}. In the first years of operation however, not all the bunch buckets were filled, yielding a lower luminosity. Table 3.1 shows the evolution of the integrated luminosity, delivered by HERA and observed by H1, over the years. Thanks to the continuous efforts of the HERA machine group the integrated luminosity grows steadily with time and at the end of 1997 a total integrated luminosity, delivered by HERA, of 66.5 pb^{-1} was reached.

The results discussed in this work are based on data taken in 1994. That year the HERA machine group switched from electrons to positrons and the beam energies were 27.55 GeV for electrons and positrons and 820 GeV for protons, yielding a centre-of-mass energy of 300 GeV. Collisions were achieved for only 153 bunches and the luminosity was a factor 10 lower than the design value. The integrated luminosity accumulated on tape by the H1 experiment that year was 4.04 pb^{-1}. This covers three distinct running periods: a first one where electrons were used (0.53 pb^{-1}), a second one where positrons were used (3.44 pb^{-1}) and a last one where again positrons were used but where the interaction vertex was shifted by 70 cm in the proton beam direction (0.07 pb^{-1}).

It is planned to upgrade HERA in about 2 years to increase the integrated luminosity by an order of magnitude by going to the low β mode. This will require the installation of additional deflection and focusing magnets inside the particle detectors [73].

3.3 The H1 detector

The following describes the H1 detector with emphasis on the components most relevant to the analysis presented in this work. A much more extensive account can be found in [74].

The description given here reflects the status of the H1 detector at the end of 1994. During the HERA winter shutdown of 1994-95, the detector was upgraded with a new backward calorimeter and backward tracking chamber and an improved time-of-flight system. A proton spectrometer was installed in the HERA tunnel, downstream the proton beam, to measure protons that are deflected under very low angles. In 1995 and 1996 a silicon tracker was added to measure tracks close to the interaction point.

Figure 3.2 shows a schematic overview of the H1 detector. The H1 coordinate system is chosen in such a way that the positive z-axis follows the proton beam direction. The x-axis points to the centre of the HERA ring and the y-axis completes the right-handed system.

The large variety of reactions produced by HERA results in a widely differing energy flow in the ep centre-of-mass system. In the laboratory however, the large momentum imbalance of the incoming beams causes most particles to be boosted into a narrow cone around the proton direction. Therefore, the H1 detector is asymmetric and much more instrumented on the forward (i.e. outgoing proton) side. The detector was designed for good identification of electrons and high granularity and resolution for jets. The solenoid encloses the calorimeter and trackers, thereby minimising the amount of dead material encountered by particles before their energy is measured. To be able to recognise missing transverse energy in charged current interactions $ep \to \nu X$, where the neutrino remains undetected, the detector is also nearly hermetic.

Background conditions at HERA present another challenge. The fraction of genuine ep collisions is tiny (10^{-3} – 10^{-5}) when compared to background events caused by collisions

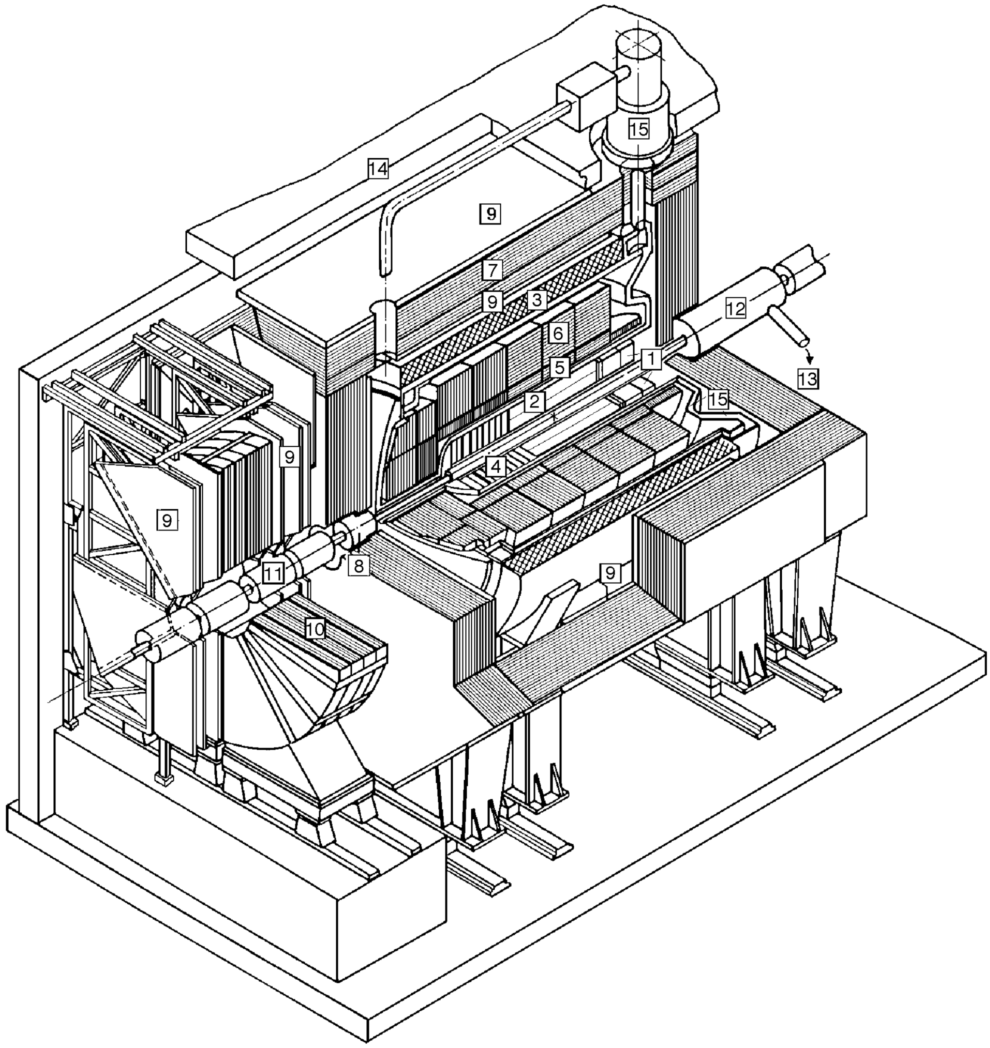


Figure 3.2 : Schematic view of the H1 detector
The Backward Electromagnetic Calorimeter (BEMC) $\boxed{1}$ *detects energy depositions from backward scattered particles (like the electron in DIS events). The hadronic final state is measured by tracking detectors surrounded by calorimeters. The Central Tracker* $\boxed{2}$ *is mounted concentrically around the beam line, inside a homogeneous magnetic field of 1.15 T generated by a superconducting cylindrical coil* $\boxed{3}$*. The momentum of forward going particles is measured by the Forward Tracker* $\boxed{4}$*. The Liquid Argon (LAr) Calorimeter surrounds the trackers and includes an electromagnetic* $\boxed{5}$ *and a hadronic section* $\boxed{6}$*. The central detectors are enclosed by an iron return yoke (IRON)* $\boxed{7}$ *which is instrumented for the detection of hadronic energy leakage from the LAr Calorimeter and for the measurement of muon tracks. A warm silicon-copper calorimeter (PLUG)* $\boxed{8}$ *measures forward energy very close to the beam pipe. Finally, the outside of the detector is equipped with muon chambers* $\boxed{9}$*. On the forward side a toroidal magnet* $\boxed{10}$ *is installed to measure the momentum of forward going muons. Not shown are the luminosity detectors (Electron Tagger and Photon Detector) which are installed in the HERA tunnel along the beam line. Other marked items are the beam pipe and beam magnets* $\boxed{11}$*, the compensating magnet* $\boxed{12}$*, the helium cryogenic system* $\boxed{13}$*, the concrete shielding* $\boxed{14}$ *and the LAr cryostat* $\boxed{15}$*.*

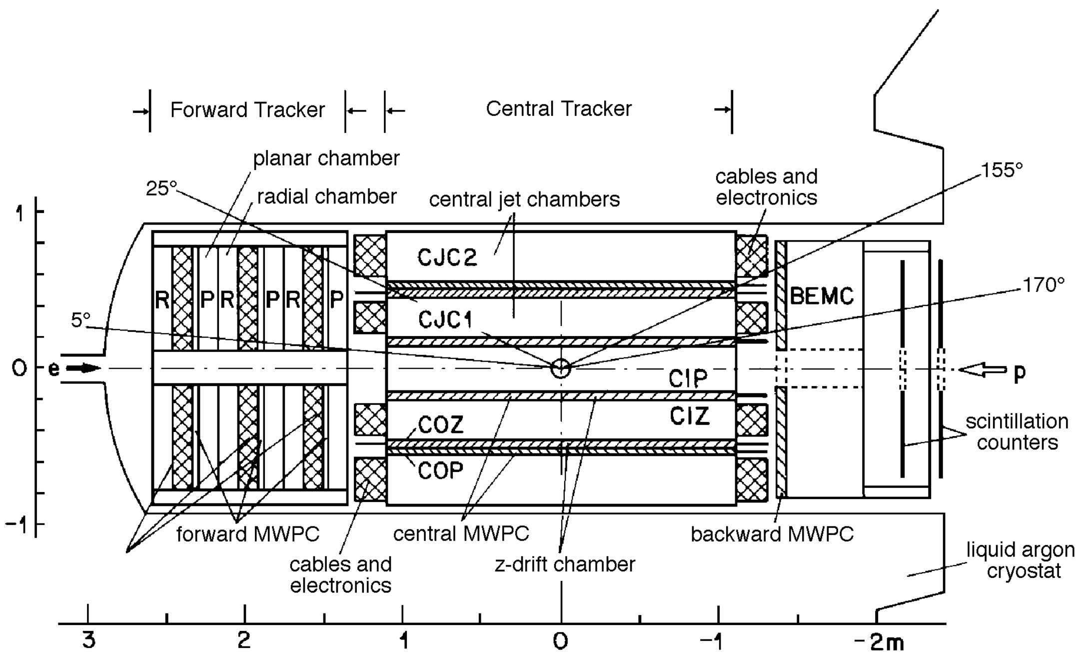


Figure 3.3 : The H1 Tracking Detectors
Both the Central Tracker and the Forward Tracker contain different subdetectors. The Central Tracker consists of inner and outer cylindrical jet chambers, z-drift chambers and Multi-Wire Proportional Chambers (MWPC) and the Forward Tracker is a hybrid of wire chambers, proportional chambers and transition radiation detectors. Also shown are the BEMC and backward MWPC (BPC) and the scintillation counters of the Time-of-Flight hodoscope (ToF).

between beam protons and the beam pipe wall or the residual gas. Together with the high bunch crossing frequency and the large number of readout channels (270 000), this requires an efficient trigger and a fast data acquisition system.

3.3.1 Tracking detectors

Tracking detectors are chambers in which a number of anode and cathode wires are strung to create an electrostatic field. The chambers are filled with a gas mixture that ionises when energetic charged particles traverse the detector. The electrons from ionised gas atoms drift towards the anode wires and create a signal. Tracking detectors are non-destructive, i.e. the amount of energy lost by the particles is minimised and the particles will leave the chamber on a nearly unaltered course.

The tracking detectors from the inner part of the H1 detector (see Fig. 3.3) provide track triggering, reconstruction and particle identification in the polar angular range $8^\circ \lesssim \theta \lesssim 165^\circ$. The tracking system can be divided in two distinct components : the Central Tracker and the Forward Tracker. They have a different orientation and are optimised for their angular acceptance region.

Figure 3.4 shows a section in the $r-\phi$ plane of the Central Tracker. A jet chamber cell consists of a plane of anode sense wires parallel to the beam line and of two planes of cathode wires, providing a nearly uniform drift field. The cells are tilted by about 30°

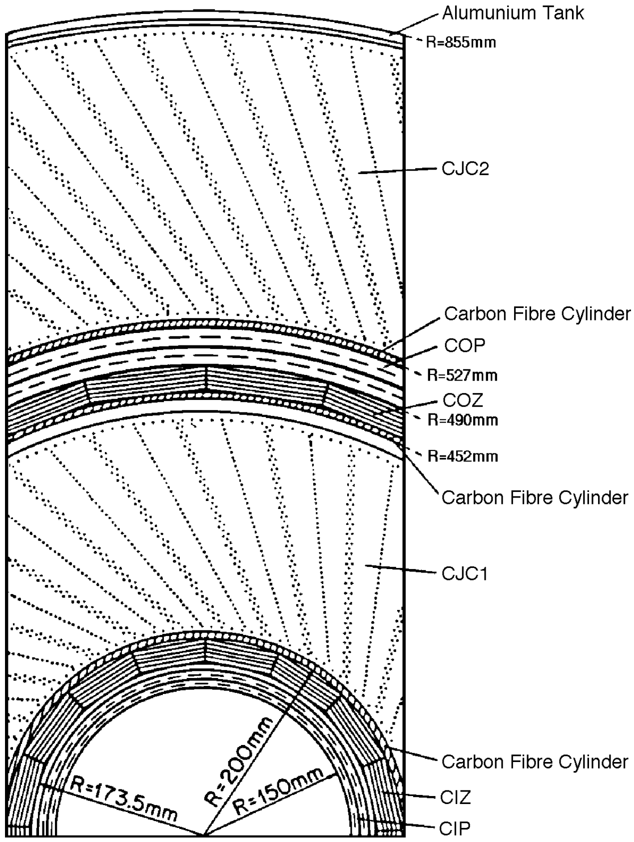


Figure 3.4 : The Central Tracker in $r - \phi$ view
Particles produced at the interaction point first encounter the Central Inner Proportional chamber (CIP), consisting of 2 layers of 480 wires each, and the Central Inner z-chamber (CIZ), which has 15 cells with 4 wires each. The 30 cells of the first Central Jet Chamber (CJC1) are each equipped with 24 sense wires. This setup is repeated for the outer part of the Central Tracker with 24×4 wires for the Central Outer z-chamber (COZ), a layer of 1574 and a layer of 1615 wires for the Central Outer Proportional chamber (COP) and 60 cells with each 32 sense wires in CJC2. The whole structure is enclosed in an aluminium tank.

so that, in the presence of the magnetic field, the ionisation electrons produced by high momentum particles drift approximately perpendicular to the track, yielding an optimal track resolution. Every particle travelling through both Central Jet Chambers, CJC1 and CJC2, will cross at least one sense wire plane in each chamber, no matter its momentum. The exact crossing time can be determined with an accuracy of 0.5 ns and the drifting electrons arrive at neighbouring sense wires with a time shift of 100 ns, thereby minimising the amount of cross-talk between bunch crossings. The sense wires are read out at both ends and provide a z-measurement through charge division. Up to 56 space points can be measured for tracks with sufficiently large momentum perpendicular to the beam line.

To achieve a constant resolution, a large number of calibration constants, like the event timing, average drift velocity, average Lorentz angle and wire-dependent constants, have to be determined and monitored through time. (The Lorentz angle defines the average direction which the drifting electrons will follow under influence of the electric and magnetic fields, while making collisions with the atoms present in the gas mixture.)

The resolutions in $r\phi$ and z are 170 μm and 2.2 (3.3) cm for protons (pions) respectively.

The z-drift chambers, CIZ and COZ, serve to improve the z-resolution of the jet chambers. They deliver track segments with a typical resolution of 300 μm in z, which can be linked to the track segments of the jet chambers. They cover the polar angular ranges of $16° < \theta < 169°$ (CIZ) and $25° < \theta < 156°$ (COZ).

The Forward Tracker contains three nearly identical supermodules. Each supermodule consists of wire chambers, proportional chambers and transition radiation detectors. Closest to the interaction point in one supermodule are three planar wire chambers, each 4 wires deep and rotated 60° with respect to each other in azimuth. Here the wires are parallel and strung perpendicular to the beam direction. The resolution of the planar chambers is typically 150–170 μm. Next, when going further away from the interaction point, is a proportional chamber, followed by a transition radiator, consisting of an array of polypropylene layers. An ultra-relativistic charged particle, passing through these radiators, emits soft X-rays, which are subsequently detected in the radial wire chambers as an enhanced track ionisation. Because the amount of radiation is dependent on the Lorentz factor γ, it is possible to discriminate between pions and electrons. The radial chambers, mounted on the back of each supermodule, are 12 wires deep. They have wires radiating outwards from the beam pipe, also strung perpendicular to the beam direction. The resolution achieved is 150–200 μm.

The H1 detector is equipped with a number of Multi-Wire Proportional Chambers (MWPC) which together cover the continuous polar angular range $5° < \theta < 175°$. Their prime purpose is to deliver a fast signal – better than the separation between to HERA bunch crossings – which can be used to trigger events at the first level trigger. Hits in the different chambers are combined so that an event vertex can be reconstructed. At least three space points are requested per track to minimise the effect of spurious background hits. If these tracks point to the interaction region, the z-vertex trigger element is set. The forward and central MWPCs also provide moderately accurate space points for the reconstruction of central and forward going tracks. The Backward Proportional Chamber (BPC), mounted in front of the Backward Electromagnetic Calorimeter (BEMC), serves to discriminate electrons and photons and to provide an accurate impact point of the scattered electron. It has four anode wire planes which can be combined for triggering purposes. The BPC has a resolution of 0.5 mrad in θ and when requiring three out of four hits in coincidence with the BEMC energy cluster, an efficiency of 98% is obtained.

COP and the front end electronics and read out systems of all the MWPCs were designed and constructed by the Belgian team in the H1 collaboration.

3.3.2 Calorimetry

A calorimeter measures the energy and direction of both charged and neutral particles. Due to interactions with the calorimeter material, a shower of secondary particles is generated until the particles' energies are too low and the particles are absorbed. The length of the shower is proportional to the energy of the primary particle.

The main calorimeter of the H1 detector is the Liquid Argon (LAr) Calorimeter (shown in Fig. 3.5). It measures energy depositions in the range $4° < \theta < 153°$ and has an electromagnetic and a hadronic section. The electromagnetic part of the LAr Calorimeter is made of 2.4 mm thick lead absorber plates alternated with comparable thick layers of liquid argon. The total thickness varies between 20 and 30 radiation lengths. Liquid argon is chosen as sampling medium because of its large density, stability, homogeneous response and ease of calibration. Because hadrons need more material to develop showers of secondary particles, the hadronic section uses stainless steel plates, each 16 mm thick.

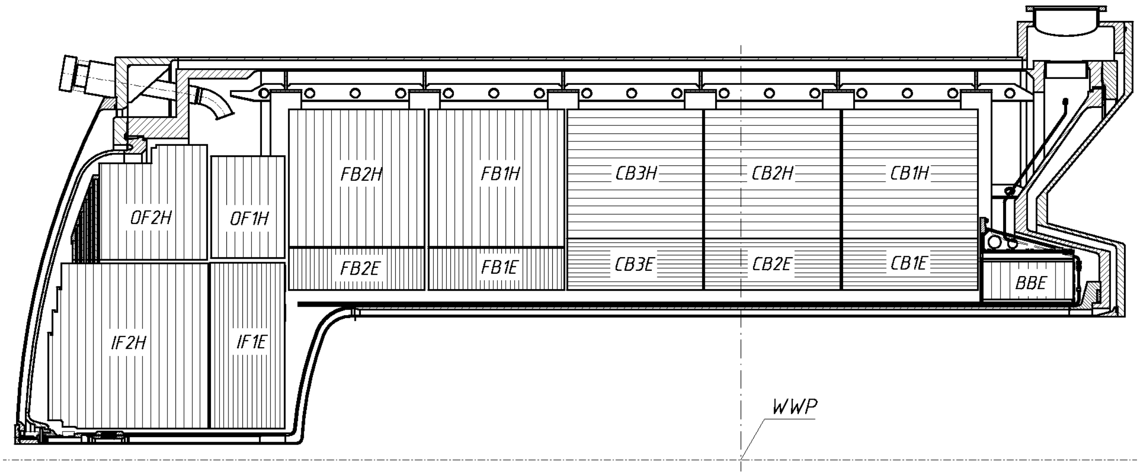


Figure 3.5 : The Liquid Argon Calorimeter
The LAr Calorimeter is segmented along the z-axis in eight 'wheels', each divided in eight identical octants. Lead and stainless steel plates are used as absorber material. The orientation of the plates is perpendicular to the beam line in the forward part and parallel to the beam line in the central part, such that the angle of incidence of particles coming from the interaction region is always less than 45°. A single cryostat keeps the calorimeter at the temperature of liquid argon (~ 87 K).

The total thickness of the LAr Calorimeter varies between 5 and 8 interaction lengths. The charges deposited in the liquid argon are collected in 45 000 cells. These cells consist of high voltage planes and read out pads on the ground side.

The LAr Calorimeter energy resolutions for leptons and hadrons have been determined in test beams [75] and are $\sigma_E/E \approx 0.12/\sqrt{E} \oplus 0.01$ and $\sigma_E/E \approx 0.50/\sqrt{E} \oplus 0.02$, respectively (in these formulae and the ensuing ones, the energy is expressed in units of GeV). From the comparison of the energy of the scattered electron in DIS events measured by the LAr Calorimeter and by the tracking detectors, the absolute electromagnetic energy scale uncertainty is estimated to be 3%. A study of the transverse momentum balance between the hadronic final state and the scattered electron in DIS events has shown that the absolute hadronic energy scale is known to an accuracy of 4%.

The backward region of the H1 detector is equipped with a lead-scintillator sandwich calorimeter (shown in Fig. 3.6). The primary purpose of this Backward Electromagnetic Calorimeter (BEMC) is to trigger on and measure the energy of electrons scattered under small angles in DIS events in the range $5 < Q^2 < 100$ GeV^2.

The BEMC consists of 88 lead-scintillator stacks and has a total depth of 21.7 radiation lengths. The calibration is done by using the scattered electron energy spectrum in the region close to the beam energy, which is known to be dominated by a pronounced peak (commonly referred to as the kinematic peak). The overall energy scale uncertainty for electrons is 1%. Test beams results [76] show that the electron energy resolution is $\sigma_E/E \approx 0.39/E \oplus 0.10/\sqrt{E} \oplus 0.017$. The BEMC can also detect hadronic energy, but the response is poor because the depth is only about 1 interaction length. The energy resolution for hadrons is of the order of $\sigma_E/E = 1.0/\sqrt{E}$; the scale uncertainty is 20% [76].

To cover the region between the forward edge of the LAr Calorimeter (3.5°) and the beam pipe (0.6°) a copper-silicon calorimeter (PLUG) is installed inside a cylindrical hole left around the beam pipe by the return yoke. It uses nine 7.5 cm layers of copper as

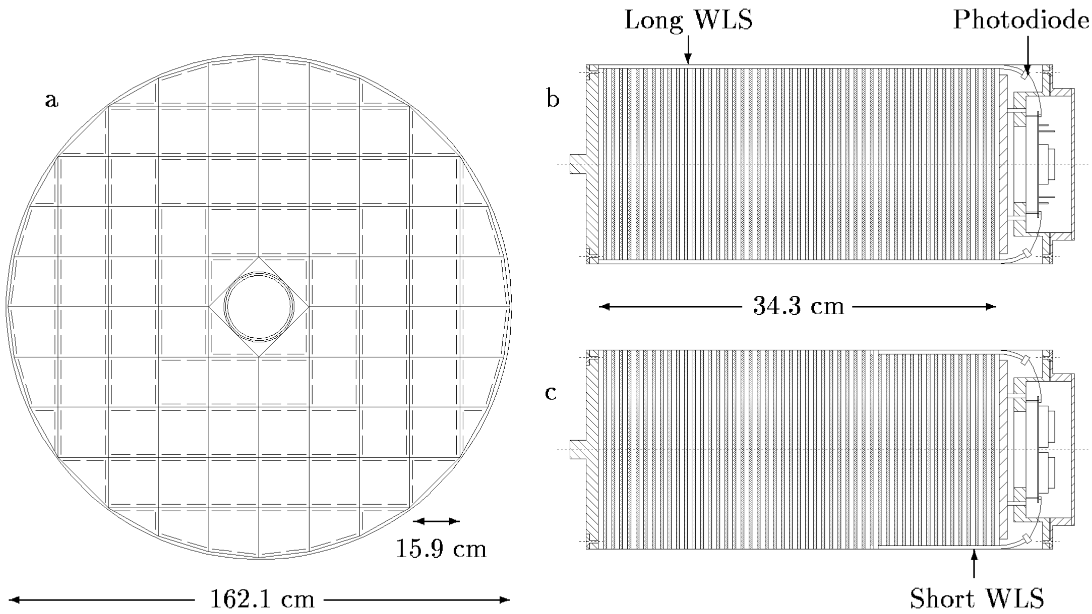


Figure 3.6 : The Backward Electromagnetic Calorimeter
(a) Transverse view of the BEMC. (b) and (c) Longitudinal view of a BEMC stack. It is a sandwich of 50 layers of 4 mm scintillating material and 2.5 mm lead absorber. The ultraviolet radiation produced in the scintillator is guided through wavelength shifters (WLS) to photodiodes. Short and long wavelength shifters are used to provide a twofold segmentation in depth.

absorber and eight 40 μm layers of silicon as sampling material. Its energy resolution is designed to be $\sigma_E/E \approx 1.5/\sqrt{E}$ [74], but this value has never been reached due to radiation damage to the silicon detectors.

3.3.3 Muon detection

Penetrating muons are detected by the outermost parts of the H1 detector. The iron return yoke (IRON), which surrounds the calorimeters and trackers, is used as absorber material and has been instrumented with limited streamer tubes, a kind of wire chambers. The most backward part of the IRON (the 'Tail Catcher') also detects hadronic showers that escaped the LAr Calorimeter and BEMC.

The momentum resolution of the main tracking detector for muons travelling under low polar angles is rather poor because the transverse motion with respect to the field lines produced by the solenoid is small. Therefore the Forward Muon Spectrometer is installed outside the IRON on the outgoing proton side of the detector. A toroidal magnet generates a magnetic field of approximately 1.5 T and muon tracks are measured by drift chambers in front of and behind the magnet. Three pairs of drift chambers are installed on either side of the toroidal magnet, two of which have a wire orientation that is optimal for the measurement of the polar angle θ and one is optimised for a good resolution in the azimuth ϕ.

3.3.4 Time-of-flight systems

Time-of-flight detectors are essential to reject background collisions at the first trigger level. They signal particles from outside the H1 detector, travelling along with the particle beams. These particles can produce significant signals in the central detectors, but arrive at a time $|z|/c$ before the bunches collide (z is the position of the time-of-flight detector along the beam line). Particles originating from genuine ep interactions arrive at a time $|z|/c$ after the collision time.

The Time-of-Flight hodoscope (ToF) is situated at $z \approx -2$ m, close behind the BEMC, and has a similar angular coverage as the BEMC. It consists of two walls of plastic scintillator and has a time resolution of 4 ns. The ToF signals background events due to interactions between the proton beam and the beam pipe wall or the residual gas, occurring upstream the proton beam. Similar devices are the inner and outer Vetowalls, located outside the iron return yoke ($z = -6.5$ m and $z = -8.1$ m). Their purpose is to detect 'halo' muons travelling along with the proton beam. The Proton Remnant Tagger, at $z = +26$ m, detects particles originating from proton fragmentation and is thus very useful in diffraction studies.

3.3.5 Luminosity measurement

The main task of the luminosity (LUMI) detectors is to provide an accurate estimate of the instantaneous luminosity of the HERA machine. This luminosity is determined from the rate of Bethe-Heitler events $ep \to ep\gamma$ [77], where the electron and photon are scattered at very low angles along the electron beam. This is a well-known QED process with a high cross section (70.38 mb for the kinematical region considered), so that a statistically accurate measurement is possible. The main source of background is Bremsstrahlung from the residual gas in the beam pipe $eA \to eA\gamma$. This background is subtracted using data from the electron pilot bunches.

The LUMI system also provides electron beam monitoring for the HERA machine, tagging of photoproduction events and energy measurement for electrons scattered under small angles and for photons from initial state QED radiation.

Two LUMI detectors are installed in the HERA tunnel: a Photon Detector at $z = -107$ m and an Electron Tagger at $z = -37$ m. Both are electromagnetic calorimeters consisting of KRS-15 crystals. The electrons or photons entering the crystals cause a shower of secondary electrons which is detected by photomultipliers through Čerenkov radiation. The thickness of each detector is 22 radiation lengths.

3.4 Data acquisition

Because the HERA bunch crossing frequency is about 10 MHz, the event rates observed by the H1 detector are very high. However, most events are due to background such as beam-gas and beam-wall interactions (about 10000 events per second) or cosmic muons (about 100 events per second). Genuine ep collisions are rare: only a few deep inelastic scattering events occur each second. Moreover, the time needed for the readout of all 270 000 channels from the central detector is typically of the order of 1–2 ms. During this time the H1 detector is unable to accumulate new data and is thus blind for new interactions.

Therefore, the readout of the detector is delayed until the triggering system has made a decision whether the event is worth keeping. In the meantime all data are stored in electronic pipelines.

The H1 trigger system consists of four on-line trigger levels, each making a compromise between the time needed to make a decision and the complexity of that decision. Only when an event is accepted as a possibly interesting interaction, it is passed on to the next level, where more time is made available and more detailed event characteristics can be investigated. If the event is rejected, the data simply fall off the end of the pipelines and the data taking continues.

3.4.1 Trigger level 1

Most of the subdetectors have trigger systems, each of which typically sends eight bits of information, called 'trigger elements', to the first level of the Central Trigger (CTL1) at each bunch crossing. All trigger elements must arrive at CTL1 within 22 bunch crossings. While waiting for the decision returned from CTL1, the subdetectors store their data in local pipelines so that the data taking can continue without interruption. Pipelining makes the first level trigger dead-time free, and it is therefore desirable that as much background rejection as possible is performed by CTL1.

CTL1 combines these trigger elements into subtriggers using programmable logic. Each subtrigger can produce a L1keep signal to stop the pipelines and initiate the event readout. Subtriggers can optionally be prescaled, which means that they will only fire once in a specified number of times.

Two kinds of subtriggers exist. Physics triggers are designed for obtaining a sample of events for physics analysis and are not prescaled, if at all possible. Either because of high physics rates, like e.g. in case of the photoproduction triggers, or because of high background rates, small prescale factors have to be applied. Monitoring triggers are designed for efficiency studies or to measure the rate of particular sources of background. These have very high prescales so that they do not significantly contribute to the dead-time.

The BEMC Single Electron Trigger (BSET) is the most important trigger for deep inelastic scattering events. It looks for localised energy depositions in the BEMC. The analogue signals coming from the long wavelength shifters in a BEMC stack are summed and when this sum exceeds a fixed threshold the stack is taken as initiator for a cluster which includes signals from neighbouring stacks. A number of different thresholds are used; in case of the so-called 'CL2' threshold the efficiency of the BSET trigger is about 100% for electrons with an energy E'_e above 10 GeV (see Fig. 3.7).

Beam-wall and beam-gas interactions produce particles that cause a signal in the scintillator walls behind the BEMC. A background (BG) and interaction (IA) timing window define whether the hits belong to particles arriving from upstream the proton beam, or from the nominal interaction point. The ToF-BG signal is the most effective background rejection criterium and is therefore applied as a veto condition in most subtriggers.

Other triggers elements typically look for activity in the trackers and calorimeters and try to decide whether the signals are compatible with a genuine ep collision (e.g. the z-vertex trigger which uses the fast signals from the MWPCs).

After the L1keep signal, each trigger system provides more detailed information to the Central Trigger. This can be used by two further levels that can make the decision to abort the readout of an event if it is identified as background. The Level 2 trigger provides a signal after 20 μs, and Level 3 can reject the event after 800 μs. During the 1994 running, however, these last two levels were not used (all events were accepted).

Once the L1keep signal is given, all subdetectors transfer their data asynchronously, with an event tag, to the Central Event Builder. When this is done, dead-time ends and the pipelines are freed again.

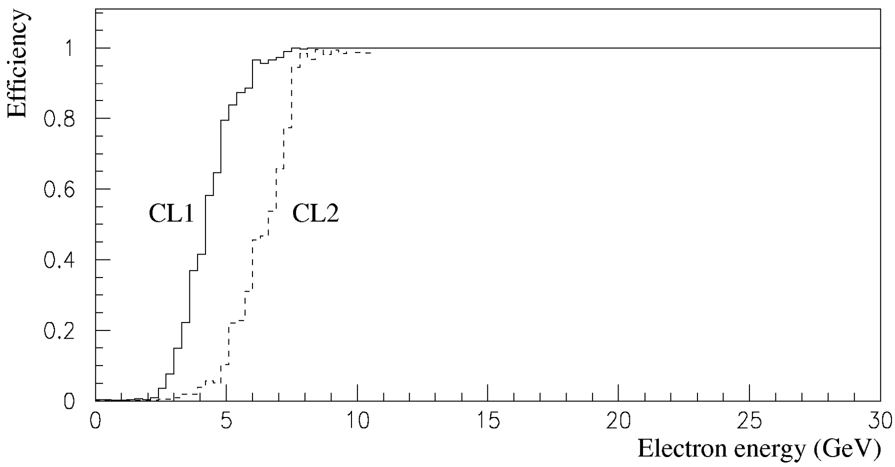


Figure 3.7 : The BEMC Single Electron Trigger efficiency as a function of electron energy
The BSET efficiency for the thresholds CL1 and CL2 is shown as a function of the scattered electron energy E'_e. For $E'_e \gtrsim 10$ GeV this trigger element is 100% efficient.

3.4.2 The level 4 filter farm

The level 4 filter farm is a fast, on-line software event filter, consisting of 32 (in 1994) MIPS R3000 processor boards, running in parallel. It is the last filtering step that rejects events with no tape backup. Each board processes one event until a decision is reached, based on the L1 trigger information, raw event data and event quantities that are reconstructed on-line.

The L4 trigger rejects the bulk of remaining beam-gas, beam-wall and cosmic induced events as well as those resulting from trigger noise. The L4 decision can e.g. be based on explicit cuts applied on the reconstructed z-coordinate of the event vertex or on the d_{ca} (distance of closest approach to the z-axis) of a track. Beam-gas background is effectively suppressed with a cut in the $(y_h, \sum p_z / \sum p)$ variable plane[†]. Trigger noise is removed by explicit trigger verification using the reconstructed quantities. The BSET trigger e.g. suffers from a large background due to particles hitting a single photo diode (mainly synchrotron radiation related). Events with a disproportionate fraction of the total stack energy in only one diode are rejected.

Events accepted by the L4 filter farm are written permanently to tape at a rate of 5–10 events per second. They are approximately 120 kbytes in size. A small fraction ($\sim 1\%$) of the rejected events are written to L4-reject tapes for monitoring purposes.

3.5 Event reconstruction

The off-line event reconstruction (also called level 5) uses the raw detector information to produce variables that can directly be used in physics analyses. It also tags events for different physics analysis classes. As an example, the following gives a brief overview of the reconstruction of charged particle central tracks within the programme H1REC (for a detailed discussion see [74]).

[†]For a definition of y_h, see Ch. 4.

Tracks are parameterised using the five helix parameters: the signed curvature (κ, positive if the direction of a track coincides with a counter-clockwise propagation along the circle projected on the xy-plane), the signed distance of closest approach to the z-axis (d_{ca}, positive if the vector to the point of closest approach and the trajectory direction form a right-handed system), azimuth and polar angle (ϕ and θ) and the z-position (z_0) at the point of closest approach. The first three parameters are determined by a circle fit to the data in the xy-projection. The circle equation, in polar coordinates, is given by:

$$\frac{1}{2}\kappa(r^2 + d_{ca}^2) + (1 - \kappa d_{ca}) r \sin(\phi - \varphi) - d_{ca} = 0. \tag{3.1}$$

The last two parameters are obtained from a least-squares fit of:

$$z = z_0 + S^{xy} \frac{\mathrm{d}z}{\mathrm{d}S}, \tag{3.2}$$

where S is the track length and S^{xy} is the same length projected on the xy-plane. The polar angle θ can be determined from the slope $\mathrm{d}z/\mathrm{d}S$ by:

$$\theta = \arctan \frac{1}{\mathrm{d}z/\mathrm{d}S}. \tag{3.3}$$

Sometimes it happens that, due to multiple scattering between the two CJCs, one parameterisation is not sufficient. In that case two separate sets of parameters are determined with the constraint that they have to join in a point between the two chambers under an angle compatible with the mean multiple scattering angle.

Tracks are found in two passes. A fast reconstruction (also used by the L4 filter farm) first looks for hit triplets and calculates for each triplet the κ and ϕ parameters. Triplets belonging to the same track tend to cluster in the $\kappa\phi$-plane, so that track candidates can be identified. Track candidates with the largest number of triplets are checked first; the remaining triplets are used to look for other tracks. Drift sign ambiguities are normally resolved at this stage because the correct solution has a much larger number of hit triplets. The standard track reconstruction uses the information from the fast track finding to improve the time reference value for an event by fits to the drift length values of long tracks. Tracks are then fitted a second time and, if successful, they are kept and their hits are not used in the following steps. The search for track elements then starts again, this time producing track segments from chains of hit triplets that have one hit in common. These segments are then merged into larger ones, allowing for possible gaps due to limited double track resolution and inefficiencies.

The code then fits the primary event vertex and the secondary vertices from pairs of tracks. Because the z-resolution of non-vertex-fitted central tracks is poor, it is only after this vertex fit that the information from the proportional and z-drift chambers is added and that central and forward tracks are combined.

3.6 Detector simulation

The simulation of the detector response is necessary to study the observed signatures of certain types of interaction and to correct the reconstructed physics data for limited acceptance and detector inefficiencies. The simulation programme H1SIM takes the infor-

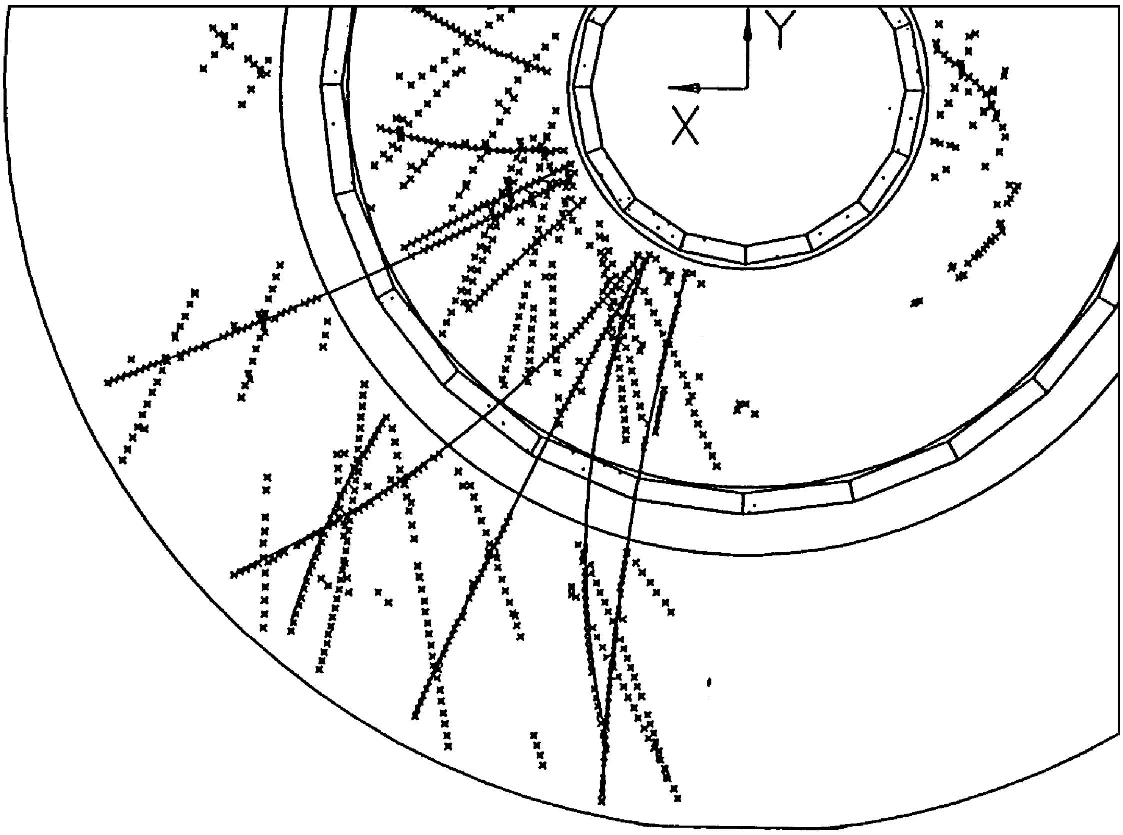


Figure 3.8 : Track reconstruction in the Central Jet Chambers
Electron-proton scattering event as seen in the CJC, showing tracks found by the pattern recognition programme and mirror tracks.

mation produced by an event generator, under the form of a set of particles with matching four-momenta, as input. It will then simulate the propagation of these particles through the detector and produce the same raw data banks as for real data. These events are then reconstructed and analysed in exactly the same way as the real data.

The GEANT [78] framework is used to describe the geometry of the H1 detector and the beam-line within ± 100 m around the interaction region. Two different levels of detail, called fine and coarse granularities, are implemented. For the tracking detectors e.g., separate volumes are used for each cell and wire in the detailed geometry, while for the coarse granularity an entire gas chamber is treated as one volume. The simulation with the fine granularity can take a lot of time (up to a few minutes per event on a SGI Challenge processor).

4

Multiplicity structure of the final state in non-diffractive DIS

This chapter discusses the multiplicity structure of the hadronic final state in non-diffractive deep inelastic ep collisions, observed by the H1 experiment at HERA*. Section 4.1 describes the experimental procedure. The selection of non-diffractive deep inelastic events and charged particle tracks, the reconstruction of the kinematical variables and Lorentz boost to the hadronic centre-of-mass system, the correction of the data for detector inefficiencies and limited acceptance and the investigation of possible sources of systematic uncertainties are reviewed. In Sec. 4.2 corrected results from the H1 experiment on the multiplicity distribution, in the full current hemisphere and in pseudorapidity bins of varying size, are shown and compared with data from e^+e^-, lepton-hadron and hadron-hadron collisions and to theoretical expectations and Monte Carlo models. A summary of the results concludes the chapter.

4.1 Experimental procedure

4.1.1 Event selection

During the part of the positron running period in 1994 that is considered in this analysis, the H1 detector recorded approximately 1.3 pb^{-1} of useful data on tape. However, only a small fraction of the observed events are the result of a neutral current deep inelastic scattering. The aim of the event selection is, therefore, to extract the genuine DIS events and to suppress background events due to e.g. beam-gas, cosmic ray and photoproduction interactions. The event selection also limits the acceptance to a region where the reconstructed variables have a good resolution and accuracy.

The event selection is done in a number of distinct steps. First, a very basic selection is made to define the event sample that will be the basis of the analysis. The scattered positron is then subjected to a number of selection criteria to ensure that the positron variables are well reconstructed and to suppress background events. The last step defines the kinematical region and excludes the so-called 'Large Rapidity Gap' events, which are attributed to diffractive dissociation of the virtual photon.

*The results of this chapter are published in [79].

Basic event selection

During the positron running period in 1994 a problem was detected with the readout electronics of the Central Tracker. This resulted in displaced hits, affecting the track reconstruction efficiency. This defect was not reproduced in the Monte Carlo simulation and, because the charged particle tracks are essential to this analysis, the results are based on events which were recorded after this readout problem was fixed.

Furthermore, runs where the HERA machine operated in shifted vertex mode are excluded.

The basic event sample is defined as follows:

- **Run quality.** Only runs with 'good' or 'medium' quality are considered. This means that major detector components (trackers, calorimeters, ...) were operating without problems. The medium quality label is assigned to runs where some non-essential subsystems are experiencing problems and are probably not participating in the data acquisition (during HERA luminosity runs the data acquisition is always kept running with as many readout branches as possible). Furthermore, the high voltage status of all subdetectors used in this analysis is checked.

- **Event classification.** The goal of the L5 event classification is to tag candidates for different physics analysis classes after the full reconstruction. Out of a total of nearly 20 distinct classes, two classes are considered for the analysis of DIS events:

  Neutral current low Q^2 (NCLQSQ class): These events have a positron candidate in the BEMC, defined as the most energetic cluster with an energy larger than 8 GeV and associated with a hit in the BPC.

  Neutral current high Q^2 (NCHQSQ class): These events have a positron candidate in the LAr Calorimeter. The candidate energy cluster must have an electromagnetic fraction larger than 50%. The event is further required to balance in p_T (the missing transverse momentum must be smaller than 40 GeV), at least one central or forward track has to be detected and the polar angle θ_e and transverse energy E'_{Te} of the scattered positron candidate must be contained in one of the following regions : $10° < \theta_e < 45°$ and $E'_{Te} > 8$ GeV, or, $45° < \theta_e < 160°$ and $E'_{Te} > 5$ GeV.

- **ToF and Vetowall.** Events that are tagged as background by the time-of-flight scintillators are excluded.

- **Coherent noise in the LAr Calorimeter.** Events with known patterns of coincident noise in the LAr Calorimeter are also rejected.

Selection of the scattered positron

The selection criteria for the scattered positron are of course different for positrons detected in the BEMC and positrons detected in the LAr Calorimeter. Because of the geometrical acceptance, events with a scattered positron in the BEMC are limited to a range in Q^2 from roughly 10 to 80 GeV^2. They are subjected to the following selection criteria:

- **BEMC cluster.** The radius of the positron candidate cluster must be smaller than 4 cm. This selection excludes the major part of wrongly identified energy depositions due to hadrons entering the BEMC.

- **BPC hit.** A hit in the BPC must be associated with the positron candidate cluster to ensure the particle entering the BEMC was charged. The distance between the BPC hit and the BEMC cluster, projected on the xy-plane must be smaller than 3.5 cm. The BPC hit must lie within 60 cm of the z-axis.

- **Primary vertex.** The primary vertex must be reconstructed from tracks and the z-coordinate must lie within 30 cm of the nominal interaction point. This primary interaction point is used to reconstruct the polar angle of the scattered positron.

- **Energy and polar angle.** The scattered positron energy E'_e must be larger than 12 GeV. Above this energy the L1 trigger that identifies the scattered positron in the BEMC has an efficiency larger than 99%. This requirement also ensures that photoproduction background is less than 1% of the final DIS sample [3]. In these events a quasi-real photon is exchanged and the positron is scattered at a low angle. Because of the absence of a high energetic positron in the BEMC, it can happen that some low energy deposits are mistaken for the scattered positron.

  The polar angle is required to be in the range $157° < \theta_e < 173°$. This ensures that the positron is well contained in the BEMC and BPC and excludes the inner triangular stacks of the BEMC. It also avoids the region of large θ_e were the resolution of the positron-only method for the reconstruction of the kinematical variables becomes problematic.

Events with a scattered positron detected in the LAr Calorimeter have a Q^2-value larger than 200 GeV^2. A positron candidate is found by looking for the electromagnetic cluster with the highest transverse energy. This candidate cluster is then subjected to the following criteria:

- **Hadronic energy.** The maximally allowed hadronic energy behind the electromagnetic cluster is 0.5 GeV.

- **Isolation.** The energy in the second road around the candidate cluster must be less than 1.2 GeV. (The first road is defined as a cylinder around the positron cluster with axis pointing to the event vertex and a radius of 15 cm; the second road is the volume outside the first road and within a cylinder with a radius of 30 cm.)

- **Primary vertex.** As for the low Q^2 sample, the z-coordinate of the primary event vertex is required to lie within 30 cm of the nominal interaction point.

- **Polar angle.** The positron candidate cluster is restricted to the range $10° < \theta_e < 150°$, thus avoiding the intermediate region between the BEMC and the LAr Calorimeter where energy measurements are problematic.

Selection of non-diffractive DIS events

Some further selections are applied to obtain the final sample of non-diffractive DIS events:

- **Large Rapidity Gap events.** Events with a Large Rapidity Gap adjacent to the proton beam are removed from the sample by requiring that the total energy deposited in the LAr Calorimeter in the range $4.4° < \theta < 15°$ is larger than 0.5 GeV. This excludes the diffractive DIS events which are studied separately in Ch. 5.

- **Kinematical selection.** The events are restricted to the kinematical region defined by:

$$0.05 < y < 0.6, \tag{4.1}$$
$$80 < W < 220 \text{ GeV}, \tag{4.2}$$
$$10 < Q^2 < 80 \text{ GeV}^2 \quad \text{or} \quad 200 < Q^2 < 100 \text{ GeV}^2. \tag{4.3}$$

  This ensures that the kinematical variables can be reconstructed with good resolution and that there is sufficient hadronic activity in the central detector. It also avoids the problematic low and high y regions where the resolution of the reconstructed kinematical variables is bad and QED radiative corrections are large.

4.1.2 Track selection

The reconstruction of charged particle tracks happens at several levels (initial track finding, vertex fitting and linking of track segments from different subdetectors) and each step involves increasingly stricter constraints. Spurious hits have, therefore, not much chance to survive and give rise to so-called 'ghost' tracks. In fact, a visual scan of the data and Monte Carlo events shows that essentially no fake tracks exist after the full reconstruction. This means that all reconstructed tracks are genuine and that no complicated selection is needed. The situation is somewhat different if one wants to produce a sample of high quality tracks, with good resolution and accuracy in θ and p_T. For a *multiplicity* study, however, the quality of the reconstructed track momenta is of less importance. Constraints on the quality of tracks will only be applied during the study of systematic uncertainties.

There are in principle two types of tracks. Primary tracks are fitted to the primary interaction point and arise from particles produced in the ep collision itself. Secondary tracks do not successfully fit to the primary vertex and are the result of decays or secondary interactions of particles with the detector material. Some contamination between the two samples is of course expected. The fraction of secondary particles that are still reconstructed as primary central tracks is of the order of 10%. The results presented in this and the next chapter are corrected for this and thus do not include the charged decay products of K^0_S, Λ and $\overline{\Lambda}$ and from weakly decaying particles with a lifetime longer than $8 \cdot 10^{-9}$ s.

The primary track finding efficiency in the Central Tracker has been determined from a visual scan and is found to be $94.1 \pm 0.8\%$ for real data and $94.3 \pm 1.3\%$ for Monte Carlo events. The reconstruction of forward tracks is considerably more complicated because of the large amount of dead material in front of the Forward Tracker which gives rise to a lot of secondary particles. Forward tracks were not included in the observed multiplicity distributions. The Monte Carlo acceptance correction covers this loss.

Table 4.1 lists the efficiency and purity of tracks as determined from a Monte Carlo study (based on the DJANGO 6.2 generator, H1SIM version 3.06.41 and H1REC version 6.04.01). Figures 4.1, 4.2 and 4.3 show some control plots where central track variables in data and from the same Monte Carlo simulation are compared.

Central tracks

Because the quality of tracks reconstructed in the Central Tracker is high, the selection is straightforward:

a) Efficiency

$8° < \theta < 15°$ $p > 0.5$ GeV	$15° < \theta < 165°$ $p_T > 0.15$ GeV	$40° < \theta < 140°$ $p_T > 0.15$ GeV
48.9%	90.9%	93.3%

b) Purity

Forward Tracker	Central Tracker
72.1%	90.8%

Table 4.1 : Track finding efficiency and primary track purity obtained from a Monte Carlo simulation study
a) The track finding efficiency for primary particles is shown for the indicated ranges in θ and p or p_T. The Forward Tracker roughly covers the range $8° < \theta < 15°$, while the Central Tracker stretches from $15°$ up to $165°$. The track finding efficiency for particles that fully traverse both CJCs (with $40° < \theta < 140°$) is slightly better than for particles that leave the Central Tracker on the forward or backward side.
b) Some secondary particles, produced in decays or secondary interactions with the detector material, are produced so close to the primary interaction point that they are reconstructed as primary tracks. The table lists the purity (i.e. the amount of genuine primary tracks) for primary tracks reconstructed in the Forward and Central Tracker.

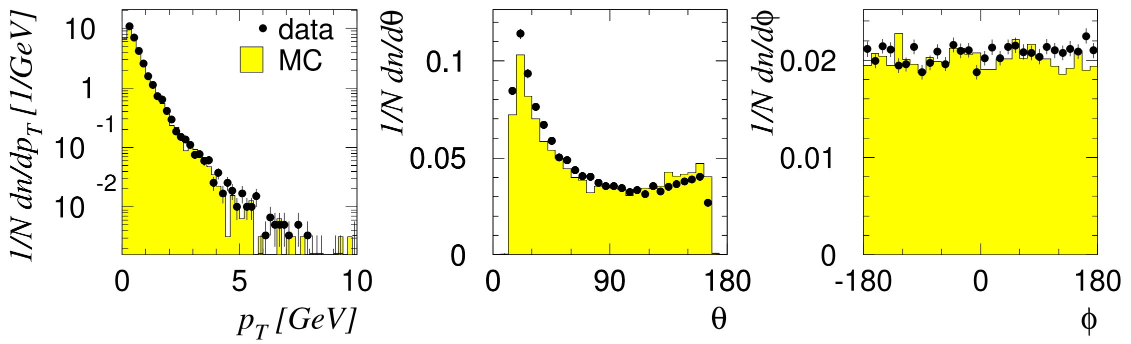


Figure 4.1 : Comparison of the reconstructed p_T, θ and ϕ distributions of central tracks in data and Monte Carlo events
Event-normalised distributions are shown of the track transverse momentum p_T, polar angle θ and azimuth ϕ. The real data are plotted as black dots and the Monte Carlo distributions are represented by the shaded histograms. N is the number of events; n is the number of tracks.

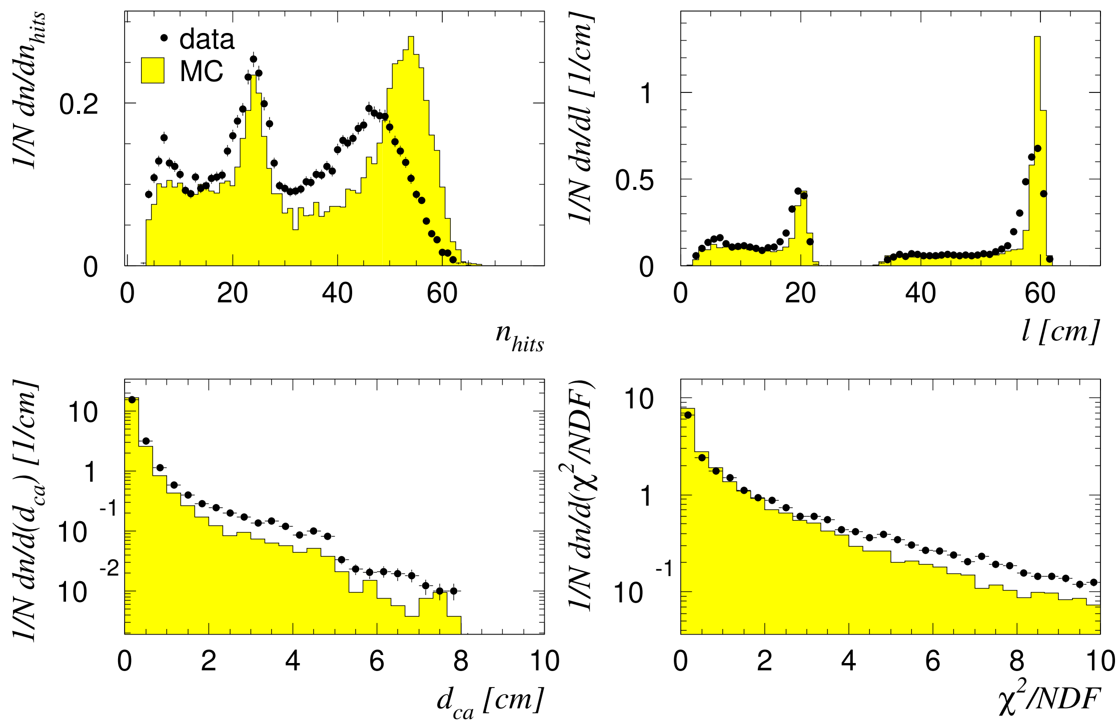


Figure 4.2 : Comparison of technical track variables of central tracks in data and Monte Carlo events
The event-normalised distributions of the number of hits per track n_{hits}, the track length l, the distance of closest approach d_{ca} and the χ^2 per degree of freedom of the fit to the primary vertex are shown and the comparison between data (black dots) and Monte Carlo events (shaded histograms) is made (here, the d_{ca} is calculated in the xy-plane and with respect to the reconstructed event vertex). It is clear that the hit efficiency in data is somewhat lower than in the Monte Carlo simulation, especially for tracks with many hits.

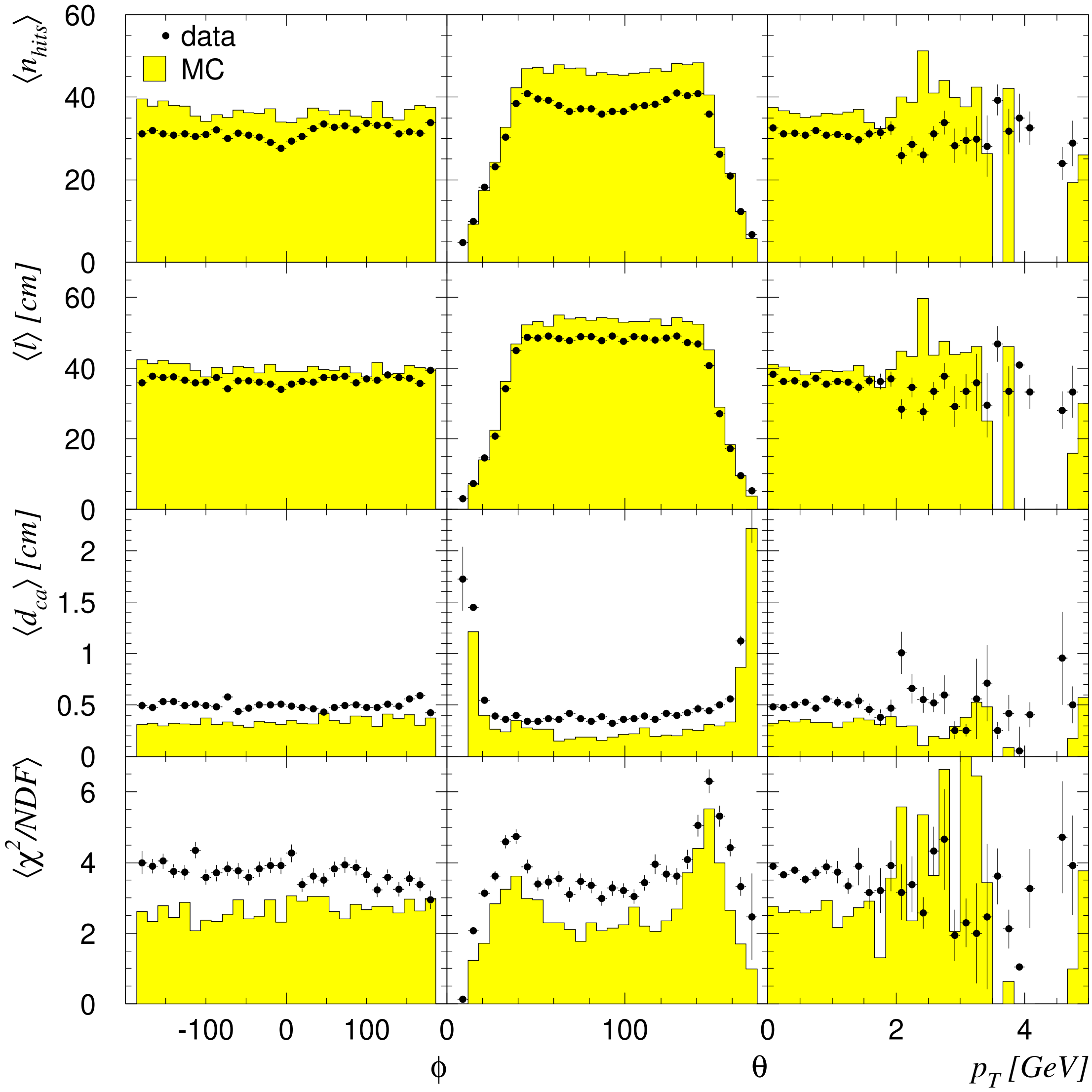


Figure 4.3 : Dependence of technical track variables on ϕ, θ and p_T of the track

The average number of hits $\langle n_{hits} \rangle$, average track length $\langle l \rangle$, average distance of closest approach $\langle d_{ca} \rangle$ and $\langle \chi^2/NDF \rangle$ of the vertex fit are plotted as function of ϕ, θ and p_T for data (black dots) and Monte Carlo events (shaded histograms). From this figure it is clear that the smaller hit efficiency in data is not restricted to a particular region in phase space and thus is not due to some dead or inefficient areas in the Central Tracker. The smaller number of hits leads to shorter tracks that are less well fitted to the primary vertex.

- **Primary vertex fitted tracks.** Tracks must successfully fit to the primary event vertex. This removes most of the secondary particles and all ghost tracks.

- **No split tracks.** Sometimes it happens that the two track segments from CJC1 and CJC2 cannot be linked, while they are produced by the same particle. These split tracks are removed so that no double counting occurs. This problem was solved in a later version of the reconstruction software.

To study the systematic uncertainty of the track selection some cuts on the track quality are introduced. This is done by requiring at least 10 hits, a track length of at least 10 cm and a starting point in CJC1 (thus avoiding the split track problem altogether).

Forward tracks

Forward tracks are only used for the study of systematic uncertainties. They are selected using the following criteria:

- **Track momentum.** The track must be contained in the region defined by $\theta > 8°$ and $p > 0.5$ GeV. The uncertainty in momentum $\delta p/p$ must be smaller than 1.

- **Track segments.** The track must be made up of at least one planar segment.

- **Track and vertex fit.** The resulting χ^2 from the track fit must be smaller than 25 and the χ^2 per degree of freedom from the vertex fit must be smaller than 10. Furthermore, when extrapolated to the event vertex, the distance in the xy-plane between the track and primary vertex must be smaller than 10 cm.

4.1.3 Event kinematics

Reconstruction of the DIS kinematical variables

Because the H1 detector measures both the scattered positron and a large part of the hadronic final state, the determination of the event kinematics is an over-constrained problem and several reconstruction methods exist [80]. These methods have to be judged by their resolution over the kinematical plane and their sensitivity to QED radiative corrections and the limited hadronic measurement due to the beam pipe holes. In the following the positron and proton beam energies are labelled E_e and E_p, respectively.

Positron-only method: This method makes only use of the measurement of the scattered positron energy E'_e and polar angle θ_e:

$$y_e = 1 - \frac{E'_e}{E_e}\sin^2\frac{\theta_e}{2}, \tag{4.4}$$

$$Q^2_e = 4E_eE'_e\cos^2\frac{\theta_e}{2}. \tag{4.5}$$

The effect of the finite resolution of the positron measurement can be studied by taking the partial derivatives of the above expressions:

$$\frac{\Delta y_e}{y_e} = \frac{1-y_e}{y_e}\frac{\Delta E'_e}{E'_e} \oplus \frac{1-y_e}{y_e}\cot\frac{\theta_e}{2}\cdot\Delta\theta_e, \tag{4.6}$$

$$\frac{\Delta Q^2_e}{Q^2_e} = \frac{\Delta E'_e}{E'_e} \oplus \tan\frac{\theta_e}{2}\cdot\Delta\theta_e. \tag{4.7}$$

From this it can be seen that the method fails at low y_e or θ_e close to 180°. The selections $y > 0.05$ and $\theta_e < 173°$ are thus necessary to ensure a good resolution of the reconstructed kinematical variables. Because no measurement of the hadronic final state is needed, the method is insensitive to leakage through the beam pipe holes. Furthermore, the Lorentz boost to the hadronic centre-of-mass system can be calculated independently of the hadronic system. The biggest disadvantage of the positron-only method is its sensitivity to QED radiation. This can be remedied by excluding events with low or high y, or with a large difference between y_e and y_h (see below).

Jacquet-Blondel method: The event kinematics can also be determined using hadrons only. The Jacquet-Blondel method [81] uses the total four-momentum of all reconstructed hadrons $(E_h, p_{xh}, p_{yh}, p_{zh})$:

$$y_h = \frac{(E_h - p_{zh})}{2E_e}, \tag{4.8}$$

$$Q^2_h = \frac{p^2_{xh} + p^2_{yh}}{1-y_h}. \tag{4.9}$$

Note that the loss due to leakage is minimised because forward going hadrons have small transverse momentum and $E - p_z \approx 0$. Because of the asymmetric beam energies at HERA, the leakage in the backward direction is small. The Jacquet-Blondel method is especially useful for the reconstruction of the kinematics of charged current events where there is no scattered positron to be measured.

Double angle method: Using the polar angles of both the scattered positron and the hadronic recoil jet one can reconstruct y and Q^2 as follows:

$$y_{DA} = \frac{(1-\cos\theta_h)\sin\theta_e}{\sin\theta_h + \sin\theta_e - \sin(\theta_e+\theta_h)}, \tag{4.10}$$

$$Q^2_{DA} = 4E^2_e\frac{(1+\cos\theta_e)\sin\theta_h}{\sin\theta_h + \sin\theta_e - \sin(\theta_e+\theta_h)}, \tag{4.11}$$

with the polar angle of the hadronic recoil jet, θ_h, calculated from the Jacquet-Blondel variables:

$$\cos\theta_h = \frac{Q^2_h(1-y_h) - 4E^2_e y^2_h}{Q^2_h(1-y_h) + 4E^2_e y^2_h}. \tag{4.12}$$

The positron-only method was found to be adequate for the analysis of the multiplicity structure in non-diffractive DIS. Figure 4.4 shows the resolution and accuracy of the reconstructed kinematical variables Q^2, y, x and W. Bjorken-x and the hadronic centre-of-mass energy W are calculated by:

$$x = \frac{Q^2}{sy}, \tag{4.13}$$

$$W^2 = sy - Q^2 + m_p^2, \tag{4.14}$$

with s the squared ep centre-of-mass energy and m_p the proton mass. The other methods are reviewed here because of their use in the analysis of diffractive DIS.

QED radiation

The kinematical variables determined with the positron-only method suffer from the effect of QED radiation. The problem is that the idealised diagram shown in Fig. 4.5a, for the lepton vertex in deep inelastic ep scattering, does not describe the real world where higher orders in α_{EM} are significant. These higher order diagrams are shown in Fig. 4.5b,c and d.

The result of initial state radiation is that the effective energy of the incoming positron is decreased by an amount equal to the energy of the radiated photon :

$$E_e^{eff} = E_e - E_\gamma. \tag{4.15}$$

If the photon is intercepted by the photon tagger in the beam pipe, one can correct for this. However, in most events the photon disappears undetected in the beam pipe and one has to cope with an additional error in the measurement. Using a Monte Carlo generator like HERACLES [41] which describes QED radiation one can calculate the true y and Q^2 on the generator level and thus correct for the migration in the kinematical plane.

Final state radiation occurs when the positron radiates a photon after the hard interaction with the proton. However, this photon is highly collinear with the positron and will end up in the same calorimeter energy cluster. This means that no correction has to be made, although one has to be careful when defining the correct kinematical variables in the Monte Carlo generator (by using the scattered positron momentum before the photon is radiated).

A more subtle complication is the problem of vertex corrections. Although the four-momentum of the photon that is exchanged with the proton can now correctly be calculated from the incoming and outgoing positron, the interpretation of the cross section changes. Both a pure 'Born' cross section and a 'QED' cross section, with higher order QED effects switched on, can be defined. However, for $y < 0.6$ the difference between both is smaller than 4% [82]. Furthermore, if the occurrences at the lepton and proton vertex are independent of each other, these vertex corrections are not important in this analysis.

Lorentz boost to the hadronic centre-of-mass system

Because of the very asymmetric beam energies, the laboratory frame is a very peculiar system to study the hadronic final state. It is much better to boost all detected particles to the hadronic centre-of-mass system (CMS), so that comparison with other experiments and to theoretical models can easily be made. This Lorentz transformation involves both a boost and a rotation. The Lorentz boost is given by:

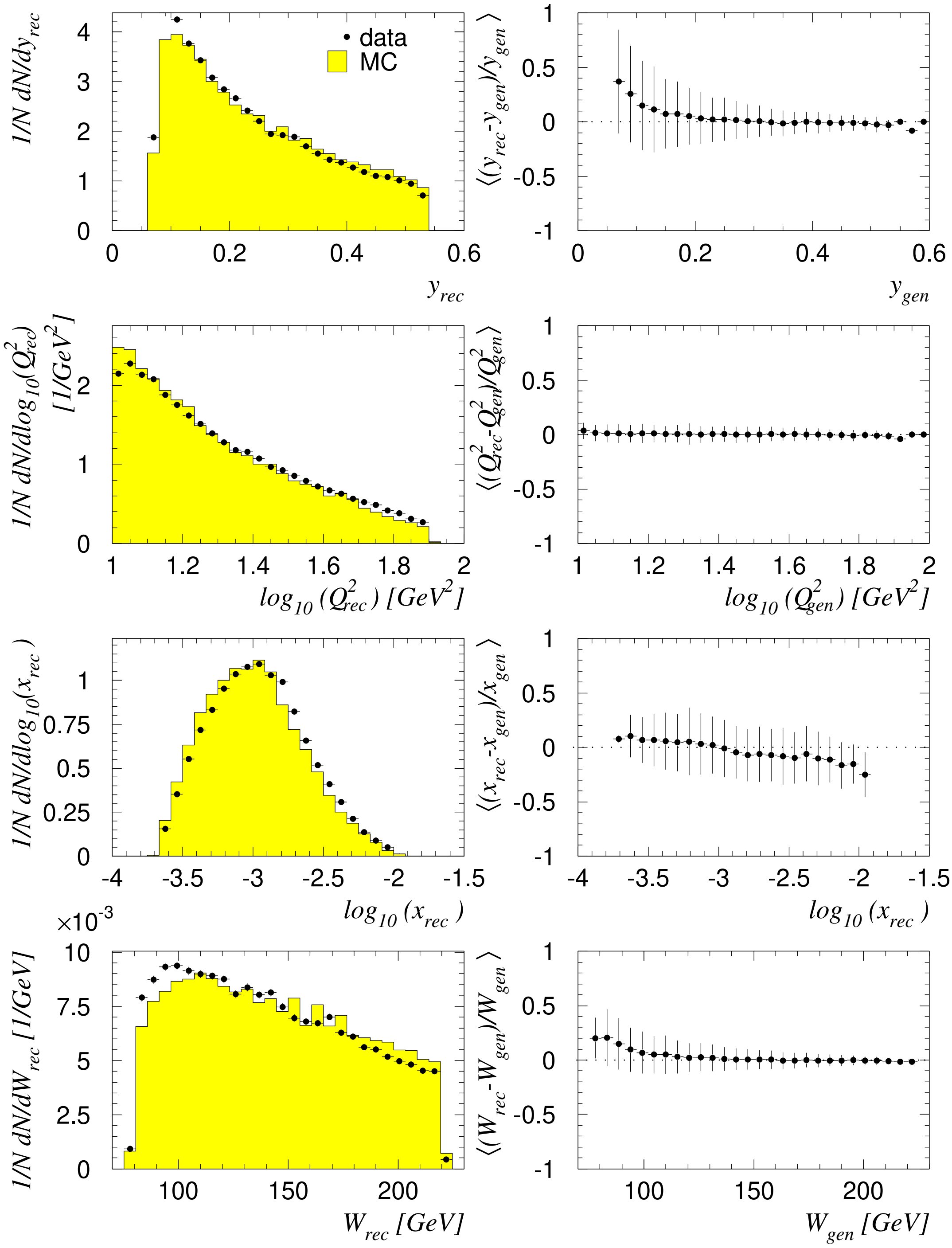


Figure 4.4 : Kinematical variables y, Q^2, x and W reconstructed with the positron-only method

The event-normalised distributions of y, Q^2, x and W are shown on the left. The data are represented by the black dots, while the Monte Carlo distributions are plotted as shaded histograms. The result of a Monte Carlo simulation study of the resolution and accuracy is shown on the right. For each kinematical variable the evolution of the mean relative difference between the reconstructed and generated variable is plotted as a function of the true (generated) variable. The error bars reflect the spread of the relative difference.

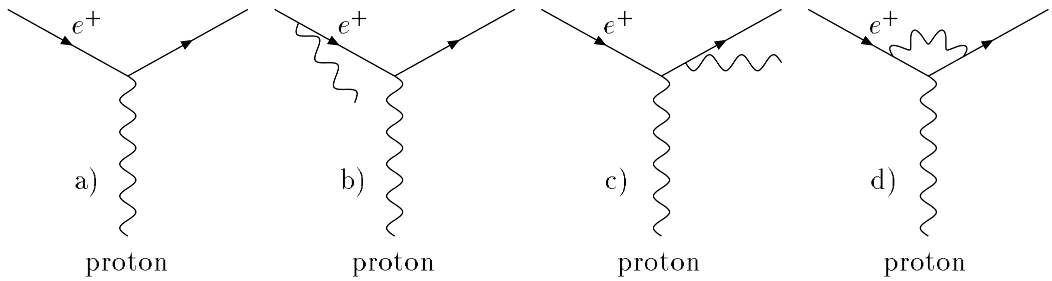


Figure 4.5 : Diagrams for QED radiation effects
(a) The Born diagram with no radiation; (b) initial state radiation; (c) final state radiation; (d) vertex correction due to the exchange of a virtual photon

$$E^{cms} = \gamma(E^{lab} - \vec{\beta}\cdot\vec{p}^{\;lab}), \tag{4.16}$$
$$p_{\|}^{cms} = \gamma(p_{\|}^{lab} - \beta E^{lab}), \tag{4.17}$$
$$\vec{p}_{\perp}^{\;cms} = \vec{p}_{\perp}^{\;lab}, \tag{4.18}$$

where p is the four-momentum $(E, \vec{p})$ of an arbitrary particle and $p_{\|} = \vec{\beta}\cdot\vec{p}/\beta$ and $\vec{p}_{\perp} = \vec{p} - p_{\|}\vec{\beta}/\beta$. The Lorentz-$\gamma$ and β factors are calculated from the four-momentum of the hadronic system P' :

$$\vec{\beta} = \frac{\vec{p}_{P'}}{E_{P'}}, \tag{4.19}$$
$$\gamma = \sqrt{1-\beta^2} = \frac{E_{P'}}{\sqrt{\vec{p}_{P'}^{\,2}}}, \tag{4.20}$$
$$P' = P + q = P + k - k', \tag{4.21}$$

where P, q, k and k' are the four-momenta of the incoming proton, exchanged photon, incoming and outgoing positron, respectively, measured in the laboratory frame.

The subsequent rotation aligns, by convention, the positive z-axis along the direction of the virtual photon. This has the effect that a particle that was travelling backwards in the laboratory will suddenly go forward in the hadronic CMS, but is consistent with the definition of the hadronic CMS in fixed-target DIS experiments. In this chapter the superscript * is used in the notation of variables measured in the hadronic CMS.

Because of the boost, the Central Tracker acceptance window shifts back and forth in the hadronic CMS. This means that the high efficiency is smeared out over a larger rapidity range, yielding a lower probability to detect a particle when averaged over events. Figure 4.6 shows the acceptance and resolution in pseudo-rapidity in the hadronic CMS.

4.1.4 Data correction

The raw multiplicity distribution in a given region of (W, Q^2) and pseudorapidity needs to be corrected for several effects. These include loss of events and particles due to limited geometrical acceptance and resolution of the tracking detectors, limited track finding efficiency, contamination by tracks from particle decays and interactions in the

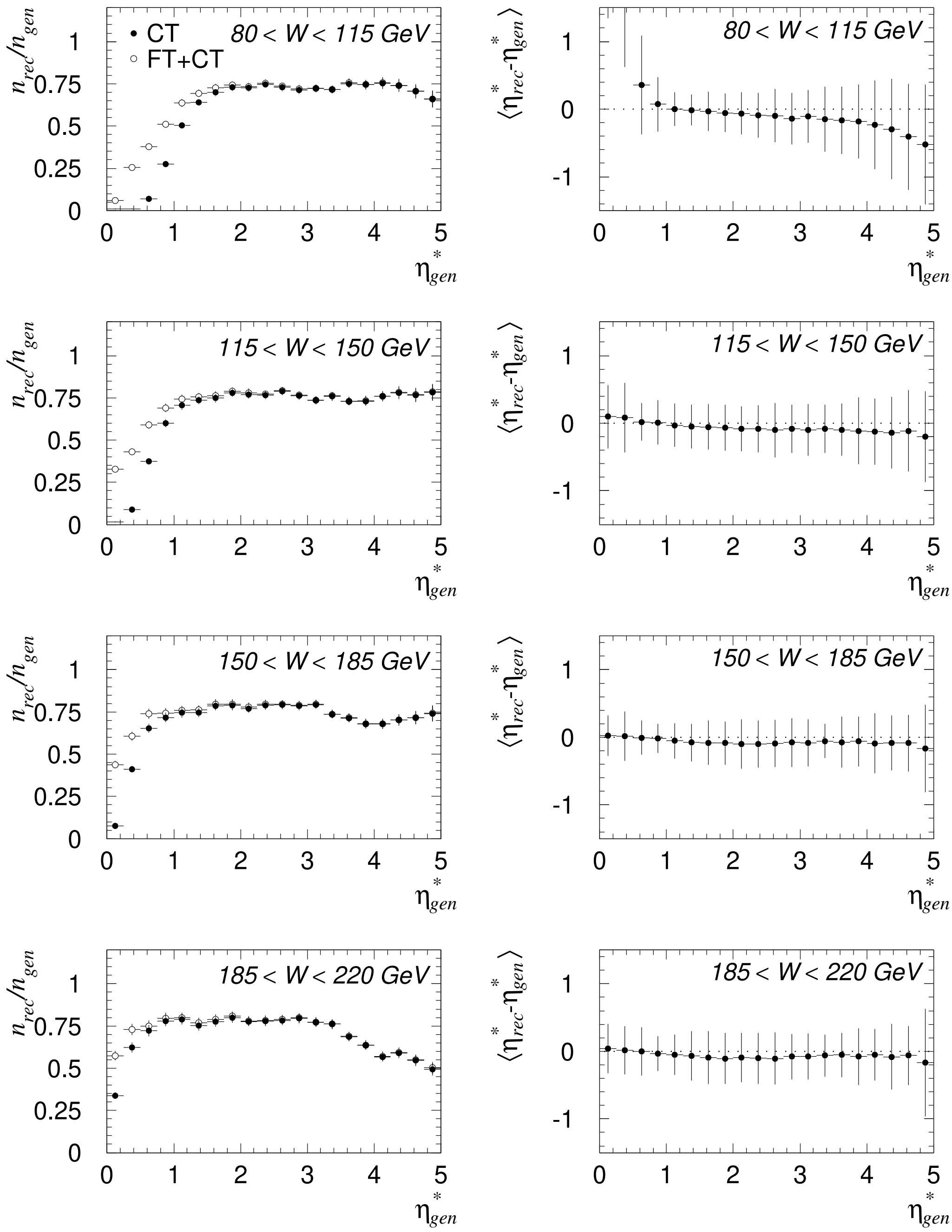


Figure 4.6 : Acceptance and resolution in pseudorapidity in the hadronic CMS as a function of pseudorapidity

The figures on the left show the fraction of generated particles that are detected in the trackers as a function of pseudorapidity in the hadronic CMS for different ranges in W. Because the exact position of the Central Tracker acceptance window in rapidity space is different for each event, the high efficiency in the laboratory is smeared out over a larger area. The black points were obtained using the Central Tracker only, while the white dots include forward tracks. The plots on the right show the mean difference between reconstructed and generated pseudorapidity in the hadronic CMS for the same ranges in W. Only central tracks were used and the error bars reflect the spread of the difference.

detector material which are assigned to the primary vertex, and also QED initial state radiation which affects the event kinematics. The results presented in Sec. 4.2 are corrected for all the above-mentioned effects using the DJANGO 6.0 generator.

Correction factors are obtained from Monte Carlo simulation by comparing the 'true' generated distributions before the detector simulation with the 'observed' distributions after this simulation followed by the same reconstruction, selection and analysis as the real data. The true distributions do not include the charged decay products of K^0_S, Λ, $\overline{\Lambda}$ and from weakly decaying particles with lifetime larger than $8 \cdot 10^{-9}$ s.

The multiplicity unfolding method used here is discussed in detail in App. A. In short, a multiplicity migration matrix is determined from the Monte Carlo simulation, which contains the probabilities μ^{mc}_{ij} for an event with observed multiplicity j to have originated from an event with true multiplicity i. This migration matrix is then used, together with the bin-by-bin correction factors ϵ_i for event losses and migration in the kinematical plane, to unfold the observed multiplicity distribution. One of the underlying assumptions of the method is that the probabilities occurring in the migration matrix derived from the Monte Carlo generator are the same as in the real data. Otherwise, the unfolded multiplicity distribution will be biassed towards the generator input distribution. To reduce this bias it has been found necessary to use an iterative procedure whereby the predicted multiplicity distribution at the generator level is reweighted using a previous approximation of the unfolded multiplicity data, until convergence is reached. This is illustrated in Fig. 4.7.

As a cross-check of the measurement of the means and dispersions of the distributions obtained by the matrix unfolding method, and to allow the study of certain multiplicity distributions given in parametric form, a second method is used. Instead of correcting the observed multiplicity distribution, a theoretical distribution $f^{th}_i(\vec{a})$ which depends on the parameter set $\vec{a}$, is transformed into a raw reconstructed distribution taking all detector effects into account. The function $f^{th}_i(\vec{a})$ represents the true multiplicity distribution for ideal non-radiative deep inelastic collisions. The transformed distribution $g^{th}_j(\vec{a})$ is then compared to the measured one and the parameters $\vec{a}$ are determined with a standard minimum χ^2 procedure. In Sec. 4.2 this method is used to make comparisons to the Negative Binomial and Lognormal distributions.

4.1.5 Systematic errors

This section discusses the various possible sources of systematic uncertainties. Each possible systematic effect is independently varied and the analysis is repeated as outlined in the previous sections. The difference in the final result, relative to the quoted result based on the correction with the DJANGO 6.0 Monte Carlo generator, is taken as one contribution to the systematic uncertainty. The errors from all sources are added in quadrature.

For illustration, the change in the mean charged multiplicity due to various effects are summarised in Tab. 4.2. Results are given for both the full current hemisphere ($\eta^* > 0$) where the corrections are largest and the pseudorapidity range $1 < \eta^* < 3$, where the detector acceptance is high.

- Photoproduction events that survive the DIS selection will have a fake positron detected in the BEMC. Above a reconstructed scattered positron energy E'_e of 14 GeV the contribution of photoproduction events is practically zero. For $12 < E'_e <$ 14 GeV the background contribution is estimated to be 5% [3]. This means that photoproduction events only contribute to the highest W interval. To suppress photoproduction events in this interval, the E'_e cut is raised from 12 to 14 GeV, a smaller

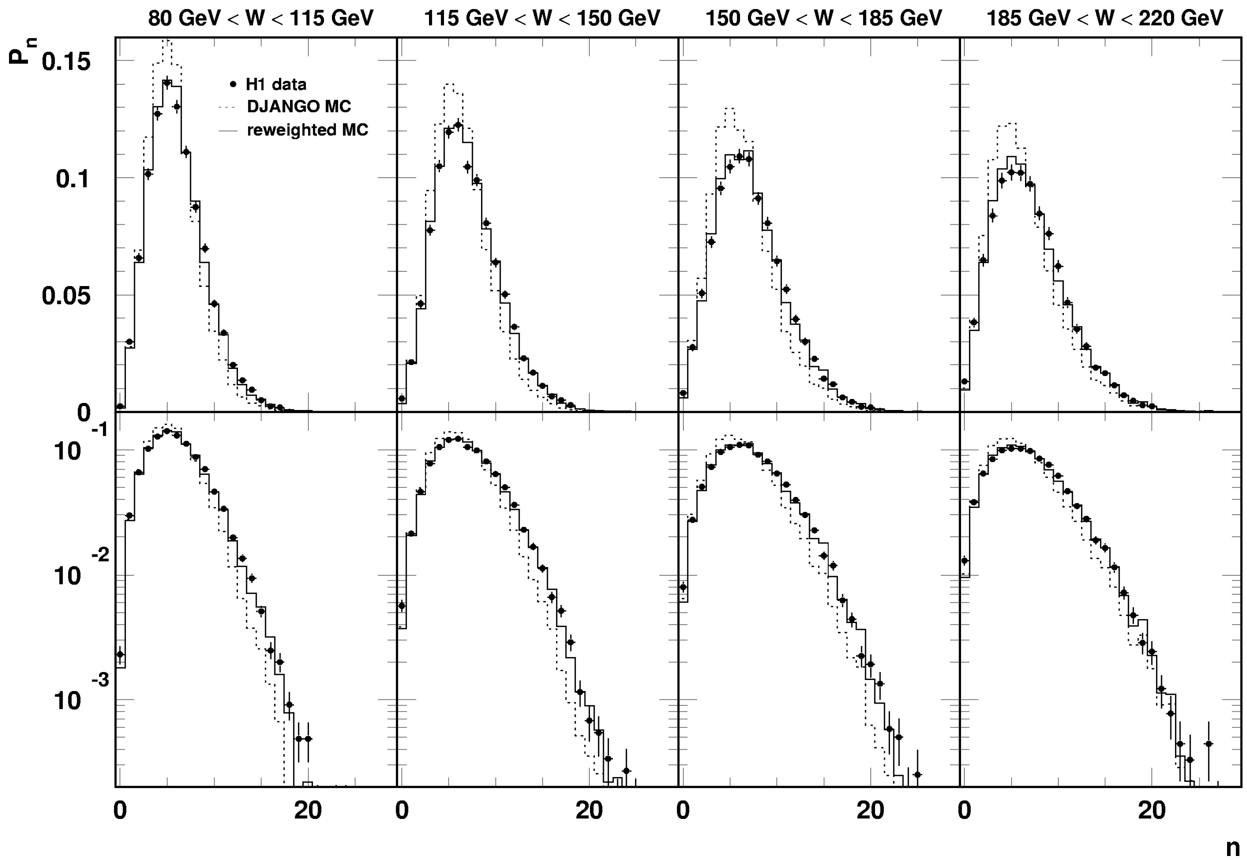


Figure 4.7 : Reconstructed multiplicity distributions in non-diffractive DIS for data, Monte Carlo and reweighted Monte Carlo events

Raw data multiplicity distributions in the pseudorapidity range $1 < \eta^* < 5$ *are compared to DJANGO 6.0 and reweighted DJANGO 6.0 distributions, after detector simulation and reconstruction, in four* W *regions and on a linear (upper figures) and logarithmic scale (lower figures). The dotted histograms represent the result from the DJANGO 6.0 generator and are narrower and have a lower average multiplicity than the data. After the reweighting procedure, however, the Monte Carlo distributions are in good agreement with the observed data.*

	$\eta^* > 0$				$1 < \eta^* < 3$			
W (GeV)	$80 \div 115$	$115 \div 150$	$150 \div 185$	$185 \div 220$	$80 \div 115$	$115 \div 150$	$150 \div 185$	$185 \div 220$
photo-production	−0.7% (1.3%)	−0.8% (1.5%)	−0.8% (1.6%)	−0.0% (2.2%)	−0.9% (1.7%)	−1.2% (2.4%)	−0.5% (2.5%)	+0.3% (3.3%)
BEMC energy scale	−2.1% (1.2%)	−1.7% (1.5%)	−0.8% (1.5%)	+0.2% (1.7%)	−0.4% (1.4%)	−1.1% (2.0%)	+2.7% (2.2%)	+2.4% (2.8%)
track selection	−4.0% (1.2%)	−3.7% (1.5%)	−3.9% (1.6%)	−3.1% (1.5%)	−3.0% (2.1%)	−3.5% (1.9%)	−4.3% (2.0%)	−4.5% (2.4%)
unfolding method	+0.1% (1.2%)	−0.8% (1.4%)	+0.9% (1.7%)	+1.8% (1.8%)	+0.6% (1.7%)	−1.2% (1.9%)	+1.4% (2.4%)	+1.1% (2.7%)
LEPTO 6.1	−4.9% (1.6%)	−4.5% (0.9%)	−4.6% (1.2%)	−3.8% (1.2%)	−0.7% (1.1%)	−2.3% (1.0%)	−2.5% (1.2%)	−2.1% (1.2%)
TOTAL	6.8%	6.6%	6.6%	5.9%	3.7%	5.3%	6.4%	6.5%

Table 4.2 : Summary of systematic effects in the analysis of the multiplicity structure in non-diffractive DIS
The change in average multiplicity is given for the full current hemisphere and for a restricted pseudorapidity interval. For comparison, the statistical error on $\langle n \rangle$ for the given data sample is given between brackets.

radial extension of the BEMC energy cluster is demanded and events with a positron detected in the electron tagger are rejected. This results, however, in a negligible change of the mean multiplicity in the highest W interval.

- The BEMC energy scale is known to an accuracy of 1%. Decreasing the energy scale by its error reduces the mean multiplicity by a maximum of $\sim 3\%$.

- The track selection criteria ensure a high efficiency (above 95% inside the acceptance region). The contribution to the total systematic error is estimated by applying additional selections on the track quality: requiring more than 10 hits, a track length of more than 10 cm and a starting point in CJC1. One should bear in mind, however, that by selecting long tracks with many hits, tracks with small and large θ are disfavoured, thus enhancing the bias due to the Monte Carlo acceptance correction.

- The multiplicity unfolding is performed iteratively to reduce the dependence on the generated Monte Carlo multiplicity distribution. To estimate the importance of any residual systematic bias on the mean and dispersion of the multiplicity distribution, a Negative Binomial distribution, smeared for detector effects, is fitted to the reconstructed data multiplicity distribution. The effect on the determination of the mean multiplicity is well below 2%. In addition, $\langle n \rangle$ has been estimated from the fully corrected single-inclusive charged particle η^* spectrum in each W interval, using the standard bin-by-bin correction method. The (generally small) difference between extremes is taken as one contribution to the overall systematic error.

- The predictions for $\langle n \rangle$ of the Monte Carlo generators used here differ by 17%, at most, in the pseudorapidity domain $0 < \eta^* < 1$. This region lies partly outside the Central Tracker acceptance at low W. Since a change in generator multiplicity outside the acceptance region does not affect the reconstructed distribution, a generator dependent bias cannot be removed by reweighting the input distribution. A systematic uncertainty has been assigned to the results pertaining to the full current hemisphere ($\eta^* > 0$) by taking the maximum difference between results derived from different event generators. Table 4.2 illustrates the maximum size of the effect obtained from the generator LEPTO 6.1

The contribution from decay products of K^0_S, Λ, $\overline{\Lambda}$ etc. is subtracted via the unfolding procedure. However, recent studies at HERA indicate that the K^0_S production rate may be overestimated by about 10% [83, 84] in the models used for correction. Such a difference leads to a systematic underestimation of the mean multiplicity of at most 0.4%.

From event simulation it is estimated that about 0.1 tracks per event assigned to the primary vertex are, in fact, due to photon conversions in the detector material. To account for possible differences between the event simulation and the real detector response an additional 1% systematic error on the mean multiplicity is assumed.

Diffractive events are removed from the H1 DIS sample by rejecting events with a Large Rapidity Gap. *Not* removing these events would result in a 3% decrease of the mean multiplicity. This effect is not included in the quoted systematic errors.

4.2 Results

Charged particle multiplicity distributions, in different pseudorapidity domains, have been measured in the kinematical regions listed in Tab. 4.3. The data are integrated over particle

W range (GeV)	$\langle W \rangle$ (GeV)	Q^2 range (GeV^2)	$\langle Q^2 \rangle$ (GeV^2)	no. of events
80 – 115	96.9	10 – 20	13.9	9150
		20 – 40	27.6	5021
		40 – 80	55.0	2509
		200 – 1000	385.3	377
		10 – 80	22.9	16680
115 – 150	132.0	10 – 20	13.9	8202
		20 – 40	27.5	4360
		40 – 80	55.1	2421
		200 – 1000	372.8	411
		10 – 80	23.2	14983
150 – 185	166.8	10 – 20	13.9	6778
		20 – 40	27.6	3662
		40 – 80	55.1	1751
		200 – 1000	378.4	439
		10 – 80	23.3	12191
185 – 220	201.9	10 – 20	13.9	5299
		20 – 40	27.6	2919
		40 – 80	54.9	1037
		200 – 1000	374.2	349
		10 – 80	23.5	9255

Table 4.3 : Kinematical bins used in the study of the multiplicity structure of the final state in non-diffractive DIS
The number of events, average Q^2 and average W are listed for the different kinematical regions studied. Apart from the full current hemisphere (defined as the domain $\eta^* > 0$), several domains in pseudorapidity are considered: $1 \leq \eta^* \leq \eta^*_c$ with $\eta^*_c = 2, 3, 4, 5$, as well as intervals of unit pseudorapidity centred at $\eta^* = 2.5, 3.5, 4.5$.

transverse momentum to allow comparison with other experiments. All data can be found in numerical form in App. B. There, tables list the lower moments of the multiplicity distributions, the fit-parameters resulting from negative binomial and lognormal fits and the multiplicity distributions themselves, together with their statistical and systematic errors.

4.2.1 Current hemisphere multiplicity moments

Because of the multitude of existing single hemisphere data from fixed target lepton-nucleon collisions and from e^+e^- annihilations, which can be used as comparative material, the moments of the charged particle multiplicity distribution in deep-inelastic ep scattering at HERA have been measured in the full current hemisphere $\eta^* > 0$. As discussed in Sec. 4.1.2, the limited experimental acceptance in the interval $0 < \eta^* < 1$, in particular in the range $80 < W < 115$ GeV, renders the H1 data in this region more sensitive to the Monte Carlo generators used in the correction procedure than in other pseudorapidity domains. The quoted systematic errors include this additional uncertainty.

Data on multiplicity distributions in a single hemisphere and in restricted domains of rapidity are available from several e^+e^- experiments [60, 61, 63, 69, 70, 85–89]. In the following, the JETSET Monte Carlo generator, with parameter settings as used by the DELPHI Collaboration [37], is used to make comparisons between DIS data and results from e^+e^- annihilations. This model reproduces in detail the multiplicity distributions of the final state in e^+e^- annihilations, in the full phase space as well as in restricted intervals of rapidity. Because the production of charm and bottom quarks is suppressed in DIS, the JETSET e^+e^- predictions presented below (labelled as "JETSET e^+e^-" in the figures) are obtained for a mixture of 'primary' light quarks pairs only. The contribution to $\langle n \rangle$ from $c\bar{c}$ ($b\bar{b}$) events is estimated at LEP to be a factor 1.04 ± 0.03 (1.14 ± 0.01) larger than for light quarks [90]. Decays of K^0_S, Λ and $\overline{\Lambda}$ are excluded, as in the DIS data and simulations, thereby avoiding experiment and energy dependent corrections to published data. The event axis, which is not a priori known in e^+e^-, is taken to be the thrust axis. The JETSET predictions are also used in domains of pseudorapidity where no direct e^+e^- measurements exist. For these, the conclusions should be treated with caution.

Mean charged multiplicity

Figure 4.8a shows the mean charged multiplicity for $\eta^* > 0$, measured by this and other lepton-nucleon experiments [65, 71, 91, 92], as a function of the centre-of-mass energy W.

In the W range covered by HERA, $\langle n \rangle$ is compatible with a linear increase with $\ln W$. Combined with the data at lower energy, however, it is evident that the mean multiplicity increases faster than $\ln W$. The HERA data confirm, for the first time in DIS lepton-proton scattering, the faster-than-linear growth of $\langle n \rangle$ with $\ln W$, a feature already well-known from e^+e^- annihilations and hadron-hadron collisions, and expected in perturbative QCD.

Various models predict the evolution of the mean multiplicity with energy and several parameterisations have been fitted to the data plotted in Fig. 4.8a. Because of the systematic difference between the EMC and E665 results, discussed in [65], fits are performed *without* the E665 data points. The errors on best-fit parameters quoted in this section include systematic uncertainties.

From [66] an appropriate power-law form, compatible with the KNO-G prescription and following from a scale-invariant branching mechanism, is

$$\langle n \rangle = a \cdot (W/W_0)^{2b'} - c, \tag{4.22}$$

where the constant c may be regarded as a 'discreteness correction' with the value $c = 0.5$ and $W_0 = 1$ GeV. The fit yields $a = 1.40 \pm 0.04$, $b' = 0.20 \pm 0.01$ with $\chi^2/NDF = 39/23$. The value for b' agrees with that obtained in [67] ($b' = 0.221$) in a comprehensive analysis of the *full* phase space multiplicity distribution in e^+e^- annihilations.

The DIS data are further compared to the Modified Leading-Log (MLLA+LPHD) prediction in the form proposed by [14, 93] and valid for running QCD coupling α_S:

$$\langle n \rangle = c_1 \frac{4}{9} \mathcal{N}_{LS} + c_2, \tag{4.23}$$

with

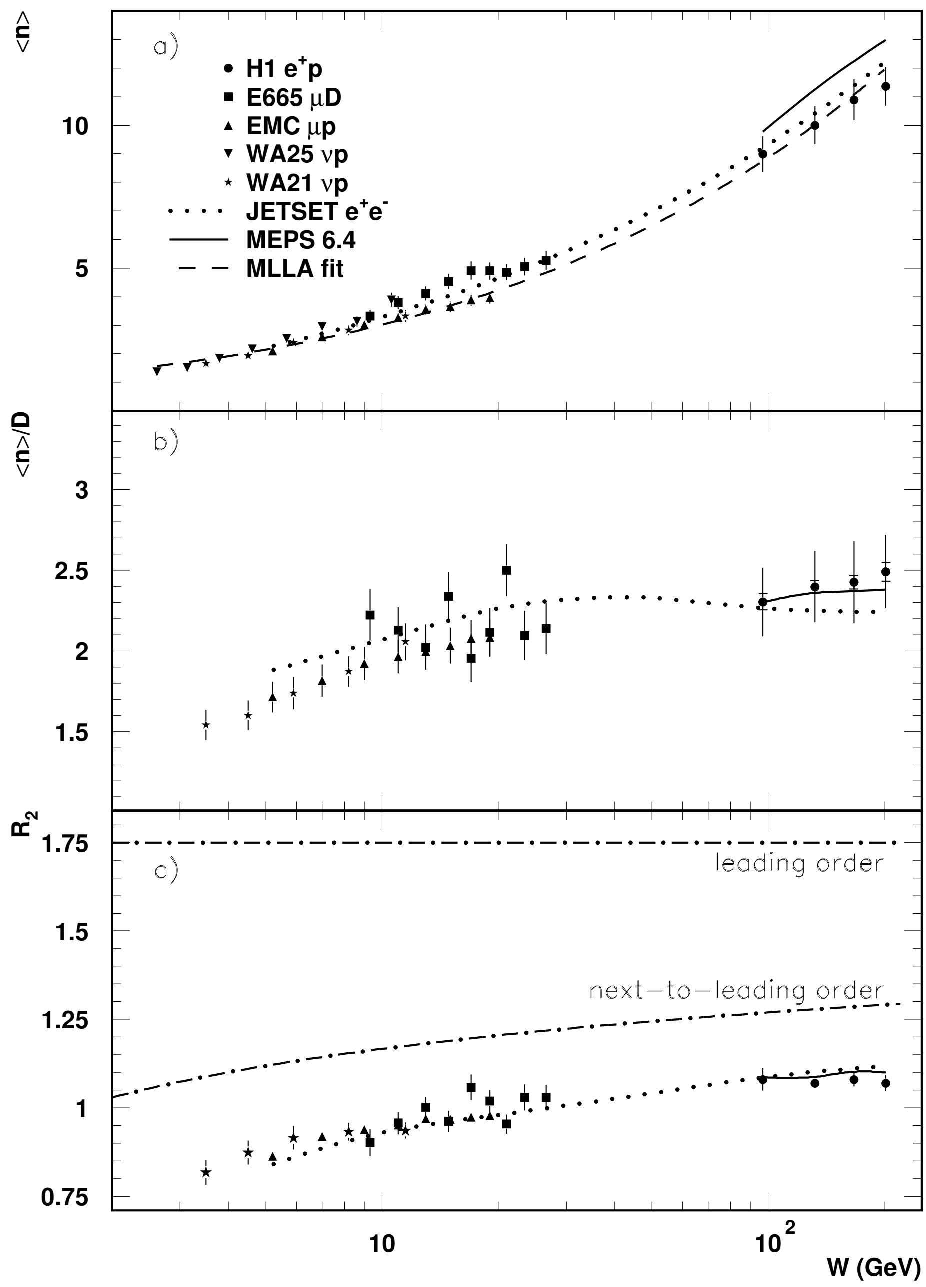


Figure 4.8 : The lower moments of the full current hemisphere charged hadron multiplicity distribution in non-diffractive DIS as a function of W

The average full current hemisphere charged hadron multiplicity $\langle n \rangle$, the ratio $\langle n \rangle / D$ and the normalised second order factorial moment R_2 measured in non-diffractive deep inelastic ep scattering are compared with data from other lepton-nucleon experiments, with the JETSET model which represents the data from e^+e^- annihilations to high accuracy, with the LEPTO 6.4 model (MEPS) and to theoretical predictions and parameterisations. For clarity, only a representative sample of νp data is plotted. The total errors on the H1 data are the overall uncertainties computed by adding the statistical and systematic errors in quadrature. When two error bars are displayed the inner error bar is the statistical error and the outer one is the total error. For data from other experiments the published systematic errors are used, whenever available. Otherwise a systematic error of 5% is assumed.

$$\mathcal{N}_{LS}(Y) = \Gamma(B)\left(\frac{z}{2}\right)^{1-B} I_{1+B}(z), \tag{4.24}$$

where $z \equiv \sqrt{\frac{48}{\beta_0}Y}$, with $Y = \ln(W/2Q_0)$, $a = 11 + \frac{2}{27}N_f$, $\beta_0 = 11 - \frac{2}{3}N_f$, $B = a/\beta_0$ and N_f the number of active flavours; I_ν is a modified Bessel function of order ν and Γ is the Gamma function. In [93] Eq. (4.23) was shown to describe the mean charged multiplicity in e^+e^- annihilations from LEP energies down to centre-of-mass energies of 3 GeV. The factor $\frac{4}{9}$ accounts for the multiplicity difference in a quark and gluon jet; c_1 is a (non-perturbative) normalisation parameter and c_2 is the 'leading parton' contribution, not included in the theoretical calculation. Using the same shower cut-off value $Q_0 =$ 270 MeV and $N_f = 3$ as in [93], the fit-parameters $c_1 = 1.21 \pm 0.05$ and $c_2 = 0.81 \pm 0.08$ are obtained with $\chi^2/NDF = 45/23$. The best-fit curve is shown in Fig. 4.8a (dashed line) and describes the data over a wide W range.

The mean multiplicity has also been computed as a function of $\alpha_S(W)$ including the resummation of leading and next-to-leading corrections [94] with the result

$$\langle n \rangle = a\ \alpha_S^b\ \exp\left(c/\sqrt{\alpha_S}\ \right)\left[1 + d\cdot\sqrt{\alpha_S}\ \right], \tag{4.25}$$

where the parameter a cannot be calculated from QCD. The constants b and c are predicted by theory. Note that the formula (4.23) reduces to (4.25) at large z. This last QCD prediction has been successfully tested over a wide energy range in several analyses of the mean charged multiplicity in e^+e^- annihilations, including the recent LEP measurement at 130 GeV [60, 69, 70, 89]. The running coupling constant is calculated with the two-loop expression

$$\frac{\alpha_S\left(W^2\right)}{4\pi} = \frac{1}{\beta_0 \ln\left(W^2/\Lambda^2\right)} - \frac{\beta_1 \ln\ln\left(W^2/\Lambda^2\right)}{\beta_0^3 \ln^2\left(W^2/\Lambda^2\right)}, \tag{4.26}$$

with the constant β_0 as defined before, $\beta_1 = 102 - \frac{38}{3}N_f$, $b = \frac{1}{4} + \frac{10N_f}{27\beta_0}$ and $c = \sqrt{96\pi}/\beta_0$. According to [94] Λ needs not to be identical to $\Lambda_{\overline{\mathrm{MS}}}$, although both are expected to be rather similar, in particular if the $O\left(\sqrt{\alpha_S}\right)$ correction turns out to be small. A fit of the data to the form (4.25) with a and d as free parameters, $\Lambda = 263$ MeV [95] and $N_f = 3$ yields $a = 0.041 \pm 0.006$ and $d = 0.2 \pm 0.3$ with $\chi^2/NDF = 27/23$. Neglecting the $O\left(\sqrt{\alpha_S}\right)$ correction in (4.25) and treating Λ as a free parameter, results in $\Lambda = 190 \pm 60$ MeV and $a = 0.034 \pm 0.005$ with $\chi^2/NDF = 29.9/23$.

The functions (4.22), (4.23) and (4.25) are fairly similar. The corresponding fitted curves practically coincide and only (4.23) is shown in Fig. 4.8a (dashed line) for reasons of clarity. Clear differences between them will only become visible at much larger energies.

The energy evolution of the mean multiplicity of partons emitted from a primary parton has been calculated for running as well as for fixed α_S [93]. In the latter case the multiplicity rises as a power of the energy as in Eq. (4.22). For running coupling constant, the growth is slower than any power of W, but faster than any power of $\ln W$, as in Eqs. (4.23) and (4.25). The various parameterisations discussed here show that, up to present energies, distinction between fixed and running α_S, based on measurements of $\langle n \rangle$, is still not possible.

The data are sensitive, however, to soft gluon interference in QCD. Neglecting interference would increase the multiplicity anomalous dimension [14], $\gamma = \mathrm{d}\ln\langle n\rangle/\mathrm{d}\ln W^2$, by a factor $\sqrt{2}$. This has been shown to be inconsistent with data for any reasonable value of Λ [96] and confirmed for DIS in [21,22].

The above comparisons confirm and extend earlier indications from μp interactions [91] that the rate of increase with energy of $\langle n\rangle$ in the current fragmentation region of DIS lepton-nucleon interactions is roughly similar to that observed in e^+e^- annihilation in the presently covered energy range. Figure 4.8a compares the DIS multiplicity with single hemisphere results as expected from the JETSET e^+e^- generator for light quark-antiquark pairs (dotted line). The H1 results are consistent with the presence of a small multiplicity excess in e^+e^- annihilations, relative to DIS, above $W = 10$ GeV, which has been noted before [97, 98]. In measurements of inclusive charged particle spectra at HERA [99, 100] a similar excess is seen near zero longitudinal momentum in the hadronic centre-of-mass. It is usually attributed to a more prolific gluon emission in e^+e^- (cf. the 'antenna effect').

The MEPS 6.4 generator for DIS at HERA energies (solid line) overestimates the mean charged multiplicity substantially.

The ratio $\langle n\rangle/D$ and R_2

Figure 4.8b shows the ratio of $\langle n\rangle$ to D in a comparison with fixed target DIS data. This ratio is expected to be energy independent if KNO scaling holds. The data above 10 GeV are indeed constant, within large errors, after a clear rise at lower W. The JETSET e^+e^- prediction (dotted curve) exhibits rather similar energy dependence and, moreover, illustrates that KNO scaling is only approximately valid. The MEPS 6.4 generator (solid line) agrees well with the HERA data.

The normalised second order factorial moment $R_2 = \langle n(n-1)\rangle/\langle n\rangle^2$ is plotted in Fig. 4.8c. This quantity is equal to the integrated two-particle inclusive density and is, therefore, a direct measure of the strength of the hadron-hadron correlations. It shows little, if any, energy dependence over the HERA range but rises steadily at lower W. The behaviour is very similar for e^+e^- final states. The experimental values of R_2 are further compared to a QCD calculation which, for a quark jet, predicts R_2 to behave as

$$R_2 = \frac{7}{4}\,[1 - \kappa\sqrt{\alpha_S}\,]\,, \tag{4.27}$$

with $\kappa = 0.88$ for three flavours [101]. Leading and next-to-leading order predictions are plotted, with α_S calculated according to the two-loop formula (4.26) and the scale parameter set to $\Lambda = 263$ MeV. Both the curves are significantly above the data. Nevertheless, it is interesting that the next-to-leading order calculation comes closer to the data, although the disagreement remains considerable. The data are rather well reproduced by the JETSET model in the case of e^+e^- annihilations, a fact confirmed by differential measurements of the two-particle correlation function by OPAL [102]. The prediction of MEPS 6.4 for DIS (solid line) is in agreement with the HERA result.

The results presented in this and the previous section indicate similarities of the low order moments of the single hemisphere multiplicity distribution in DIS and e^+e^- annihilations, in conformity with the hypothesis of approximate universality of quark and gluon fragmentation.

QCD predictions for the energy dependence of the mean *parton* multiplicity, derived from analytical solutions of the evolution equations, are in agreement with that observed for *hadrons* and thus add to the existing support for the LPHD ansatz at the single-inclusive level. The large disagreement between data and QCD next-to-leading order calculations of [101] for R_2 suggests that extension of the LPHD hypothesis to higher order inclusive correlations may not be justified and that calculations beyond next-to-leading order are necessary.

4.2.2 Q^2 dependence of the multiplicity moments

In this and the following sections the multiplicity distribution and its moments are studied in domains of pseudorapidity, limited to the current fragmentation region $\eta^* > 1$. The Q^2 and W dependence is examined and compared with data from other types of interactions and with predictions from the MEPS 6.4 generator.

In the simple quark-parton model the properties of the total hadronic system produced in a deep inelastic lepton-hadron collision depend on the lepton kinematical variables Bjorken-x and Q^2 only through the invariant mass W of the hadronic system. In QCD, scaling violations of the quark fragmentation functions and of the parton distributions introduce an explicit Q^2 dependence even at fixed W. Fixed target DIS electron, muon and (anti-) neutrino experiments at low energies confirm that the global characteristics of the hadronic final states and the average number of produced hadrons in particular, vary most significantly with W. At fixed W, only weak dependencies on Q^2 are observed [64, 103–105]. Such results are in accord with the Bjorken-Kogut correspondence principle [27] and imply that the densities in the hadron plateaus spanning the current, central and target regions are quite similar [27, 106]. Only recently, a statistically significant Q^2 dependence of the mean charged hadron multiplicity has been established in $\mu^+ p$ and $\nu(\overline{\nu})p$ interactions [105, 107]. The effect is limited to a restricted region in Feynman-x: $-0.15 < x_F < 0.15$, where x_F is the fractional longitudinal momentum of a hadron in the hadronic CMS frame. There are no published results on a possible variation with Q^2 of the *shape* of the multiplicity distribution.

The H1 experiment allows us to investigate the charged particle multiplicity distribution over a wide range in Q^2 in a novel energy domain. In Fig. 4.9 the mean charged multiplicity and the dispersion of the multiplicity distribution are plotted in four intervals of W, and in the pseudorapidity domain $1 < \eta^* < 5$, covering part of the current fragmentation hemisphere. Within errors, no significant variation with Q^2, in a fixed interval of W, is observed for $\langle n \rangle$ and D in the range $10 < Q^2 < 1000$ GeV^2.

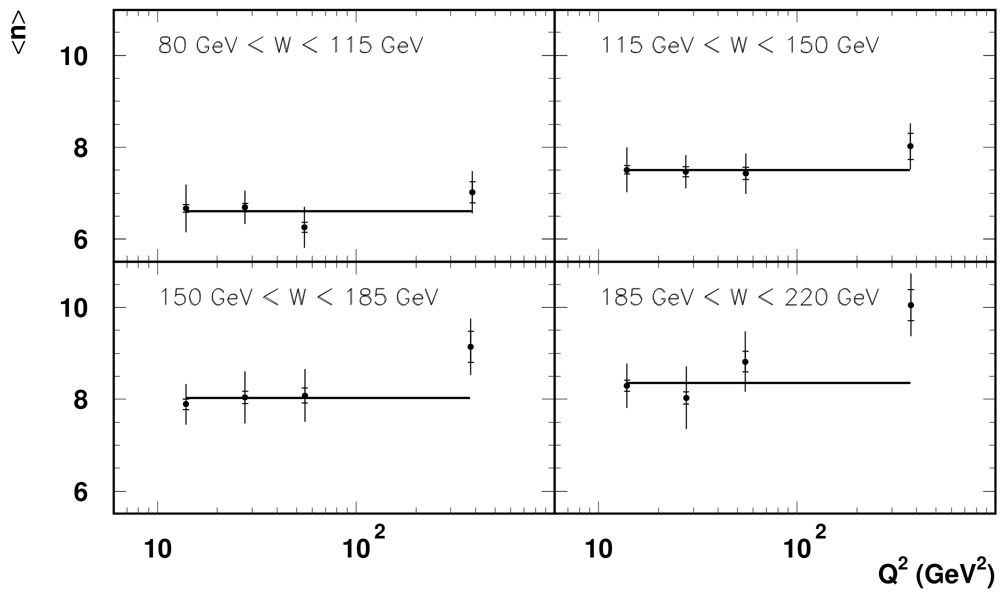
To ascertain the evolution with Q^2 seen at much lower energy, it will be necessary to study the multiplicity distribution in the central and proton remnant pseudorapidity regions which are not covered in this analysis. A similar study with quasi-real photons at HERA would be of evident interest.

The data presented in subsequent sections have been obtained from data samples integrated over the full Q^2 region ($Q^2 > 10$ GeV^2) covered by this experiment.

4.2.3 The shape of the multiplicity distribution in pseudorapidity domains

Fig. 4.10 shows, as a representative example, the multiplicity distribution in the interval $115 < W < 150$ GeV, measured in various η^* domains.

The figure illustrates that the multiplicity distribution, at fixed W, becomes narrower as the size of the η^* interval is reduced. However, the same distributions plotted in KNO form (not shown) widen under the same conditions. The latter property is in part related to the diminishing influence of global conservation constraints and was first predicted in [43].



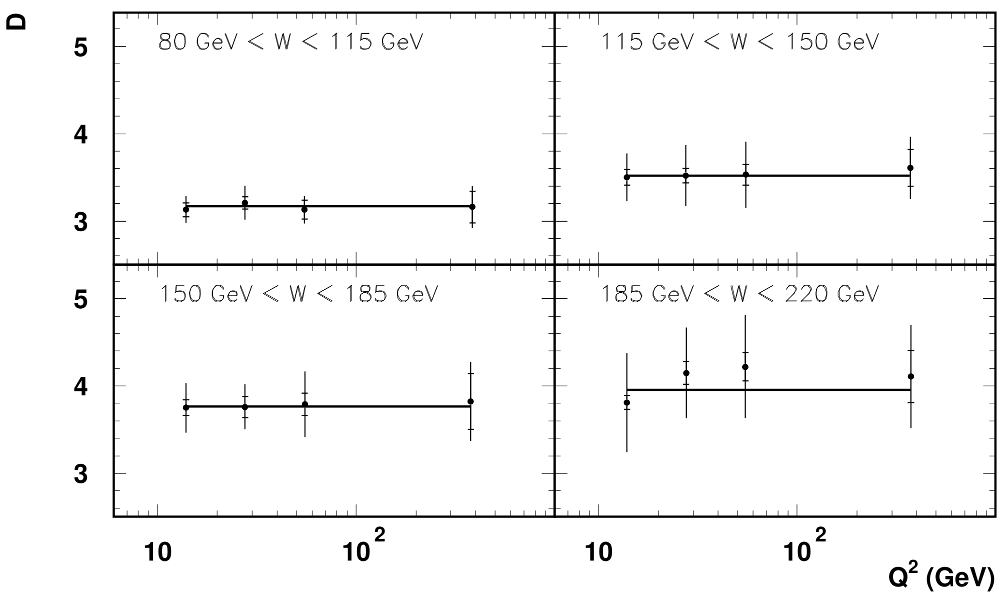


Figure 4.9 : Q^2 dependence of the lower moments of the charged hadron multiplicity distribution in non-diffractive DIS

The Q^2 dependence of the mean charged particle multiplicity (top) and the dispersion (bottom) is plotted in intervals of W and in the domain $1 < \eta^* < 5$. The inner error bars represent the statistical error, the outer error bars represent the total (quadratic sum of statistical and systematic) errors. The solid lines are fits to a constant.

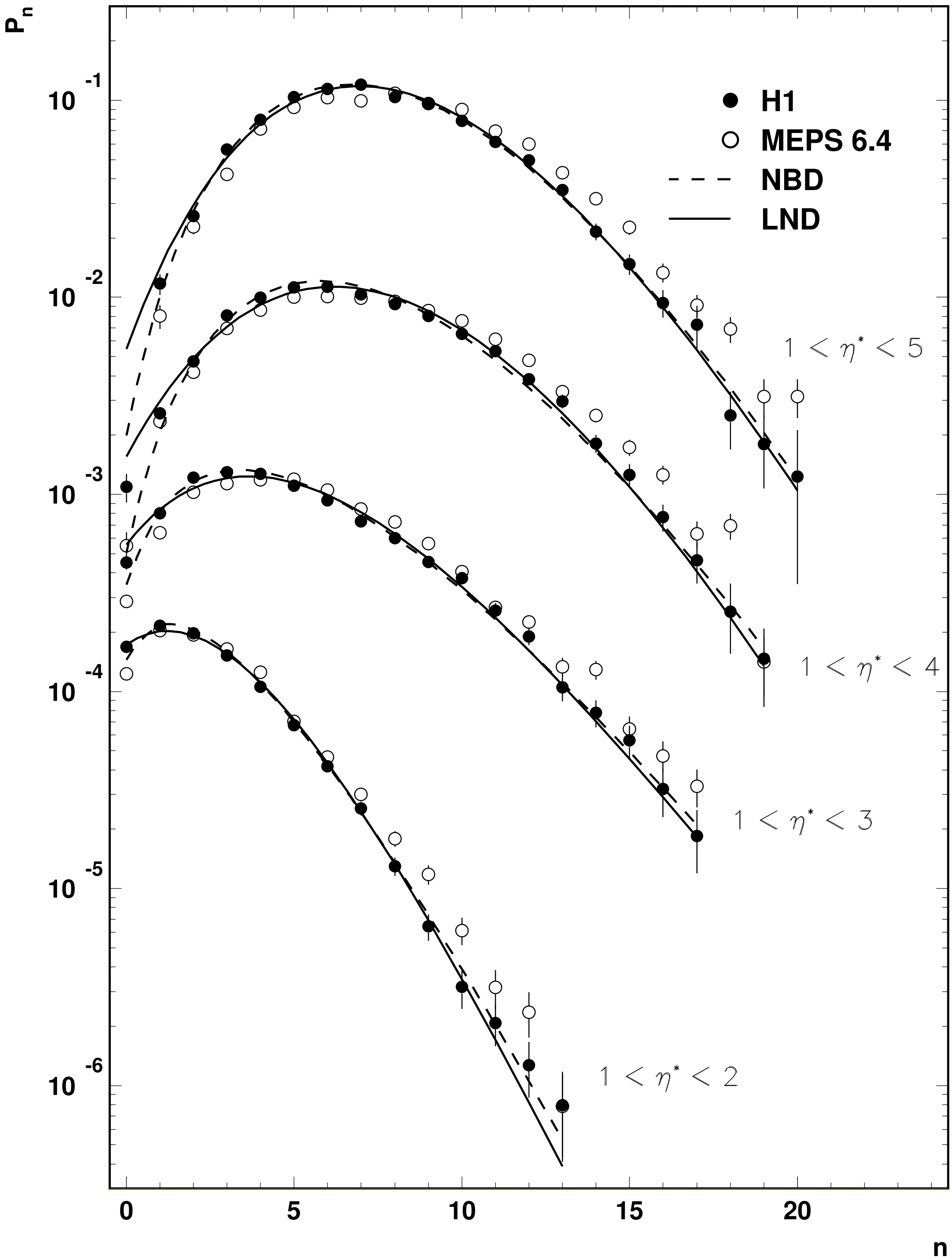


Figure 4.10 : The fully unfolded multiplicity distributions in non-diffractive DIS events with $115 < W < 150$ GeV

The distribution for $1 < \eta^* < 5$ is plotted at its true scale; each consecutive distribution is shifted down by a factor of 10. The H1 data points (solid symbols) are compared with MEPS 6.4 (open symbols) and to fits to the Lognormal (full) and Negative Binomial (dashed) distribution. Statistical errors only are plotted.

Figure 4.10 also shows the predictions from MEPS 6.4 (open symbols). Significant deviations are noted. The model overestimates the mean multiplicity (cf. Fig. 4.8a) and, consequently, does not reproduce the small and large n tail of the distribution. These defects are seen in all W and η^* intervals examined.

Numerous parameterisations for the shape of the multiplicity distribution are used in the literature. Here, the Negative Binomial distribution (NBD) and the Lognormal distribution (LND) are used. The phenomenological arguments leading to these forms are discussed in Ch. 2. The parameters of these parametric models are obtained from a least χ^2 fit to the uncorrected multiplicity distributions. The best-fit parameters for all studied pseudorapidity and W intervals are summarised in App. B. The quoted errors are the quadratic sum of the statistical error and the systematic uncertainties.

The solid (dashed) line in Fig. 4.10 shows how the LND (NBD) compares to the measurements. Inspection of this figure and of Tab. B.3 indicates that the LND gives a reasonably accurate description of the data, in particular in the smallest η^* domain. However, the quality of the fits deteriorates in larger domains. Likewise, the NBD fits are acceptable in the smallest η^* domain but become progressively worse for larger intervals. The two distributions are seen to differ most for low multiplicities. Nevertheless, it may be verified from Tabs. B.2 and B.3 that the estimates for the mean and dispersion of the multiplicity distribution, derived from these parameterisations, agree very well with those derived from the fully unfolded distribution. Parametric forms, such as the NBD and LND, therefore remain useful for phenomenology.

Several extensions of the NBD have been introduced. In [108] the H-function extension of the NBD (known as the HNBD) was used to fit the H1 data. The average multiplicity was fixed at its observed value, k was set to unity (this yields a Weibull distribution) and only the parameter μ was varied in the χ^2 minimisation procedure. This HNBD yields a substantially better χ^2/NDF and fixes the discrepancy at small multiplicities. In [109] another extension, the Modified NBD (MNBD) is proposed. Again, the improvement in χ^2/NDF is substantial.

In intermediate size rapidity intervals in e^+e^- annihilations at LEP, both the LND and the NBD are unable to describe the multiplicity distribution which exhibits a prominent shoulder at intermediate n values. The shoulder is most prominent in the single hemisphere distributions [61–63]. This structure results from a superposition of two-jet and three- or four-jet events. The effect demonstrates that the fluctuations in the number of hadrons, and therefore the multiplicity distribution, carries information on the hard partonic phase of the multihadron production process, even after soft hadronisation. The H1 measurements show no evidence for a shoulder structure of the type seen in e^+e^-. Neither is such a structure present in the multiplicity distribution predicted by MEPS 6.4.

An excess of high-multiplicity events would also be expected if a significant proportion of DIS events were induced by QCD instantons [110]. In order to determine an upper limit for the cross section of such events from our data, an analysis method recently applied by H1 in a study of strange particle production [83] is closely followed. The observed multiplicity distribution is assumed to be a superposition of two distributions, one associated with instanton-induced events, another with 'standard' DIS events. The former is calculated from the instanton generator described in [111]. For the latter a Negative Binomial form is adopted. Using a χ^2 minimisation procedure to determine the relative proportion of instanton-induced and standard DIS events, a 95% confidence level upper limit of 0.3 nb is derived for the cross section of instanton production in the pseudorapidity domain $1 < \eta^* < 5$ and $80 < W < 115$ GeV.

4.2.4 KNO scaling and correlations

To demonstrate the energy scaling of the multiplicity distribution in DIS in the energy range opened up by HERA, Fig. 4.11 shows the KNO distributions $\psi(z)$ in the domain $1 < \eta^* < 5$ for four intervals in W.

Exact KNO scaling implies that the function $\psi(z)$, and hence the moments C_q, are independent of W. Figure 4.12 displays the variation of the C moments with W for various η^* intervals (see also Tab. B.2). At fixed W the moments increase as the η^* domain decreases in size, reflecting the widening of the multiplicity distribution in KNO form.

The moments in the smallest and largest pseudorapidity domain show, within errors, little W dependence and thus exhibit approximate KNO scaling. However, violation of KNO scaling is seen in intermediate size intervals. This observation is consistent with the clear KNO scaling violations observed at HERA in DIS data on multiplicity distributions measured in the current region of the Breit frame of reference [21, 22].

The MEPS 6.4 generator (solid line) describes the data well in the largest η^* domains but tends to underestimate C_2 and C_3 in the smaller ones.

The dotted lines in Fig. 4.12 are expectations from JETSET for e^+e^- in the same pseudorapidity domains as covered by H1. Single hemisphere e^+e^- data are known to exhibit KNO scaling above $\approx$ 20 GeV [60, 69, 88]. Large differences between DIS and e^+e^- annihilations are predicted, in particular for higher order moments in small η^* domains. However, in the largest domains the e^+e^- results join smoothly with the DIS data. This is again in accord with the QCD expectation that hard multi-jet production is less frequent in lepton-hadron collisions. The H1 data cover larger W values than presently available in e^+e^- annihilations and it will be of interest to confirm these predictions with future measurements at LEP at centre-of-mass energy in the range studied here.

Also shown is a measurement for non-diffractive $p\overline{p}$ collisions at $\sqrt{s} = 200$ GeV in the interval $1 < |\eta^*| < 2$ from UA5 [54]. (The UA5 measurement at $\sqrt{s} = 900$ GeV (not shown) yields $C_2 = 1.84 \pm 0.02$, $C_3 = 4.0 \pm 0.1$ and $C_4 = 14.5 \pm 0.8$, the same, within errors, as the results at 200 GeV.) Here the fluctuations in particle density number near the central plateau are significantly larger than in DIS. It will, however, be shown in Sect. 4.2.5 that the particle density itself is quite similar in the two processes.

The values of the cumulants K_q in a given domain of phase space are a direct measure of the strength of 'genuine' correlations among hadrons. Inspection of Tab. B.2 shows that the three-particle correlation function is significantly different from zero only in the smallest interval $1 < \eta^* < 2$, in accord with measurements for other types of interactions [44].

To study the two-particle correlation function more directly, $\langle n \rangle$ and the second order factorial moment $R_2 = 1 + K_2$ are presented in Fig. 4.13 in η^* domains. The mean charged particle multiplicity increases slowly for $1 < \eta^* < 2$, i.e. near the central region, but faster in larger domains. The W dependence of R_2 for $1 < \eta^* < 2$ is less clear, in view of the errors, but compatible with the slow (logarithmic) rise well established in hadron-hadron and e^+e^- interactions. The MEPS 6.4 generator reproduces reasonably well the behaviour of R_2 but systematically overestimates the mean multiplicity in all η^* domains.

The moments of the multiplicity distribution have also been calculated in the framework of the CKMT model for small-x DIS [112]. In [113], a unified approach to hadron-hadron, real photon-hadron and virtual photon-hadron interactions is proposed, where the pomeron pole is assumed to be universal, i.e. the same in soft and hard processes. The observed differences between the different types of interactions arise from the difference in the size of unitarity corrections (multipomeron exchanges). These corrections are evalu-

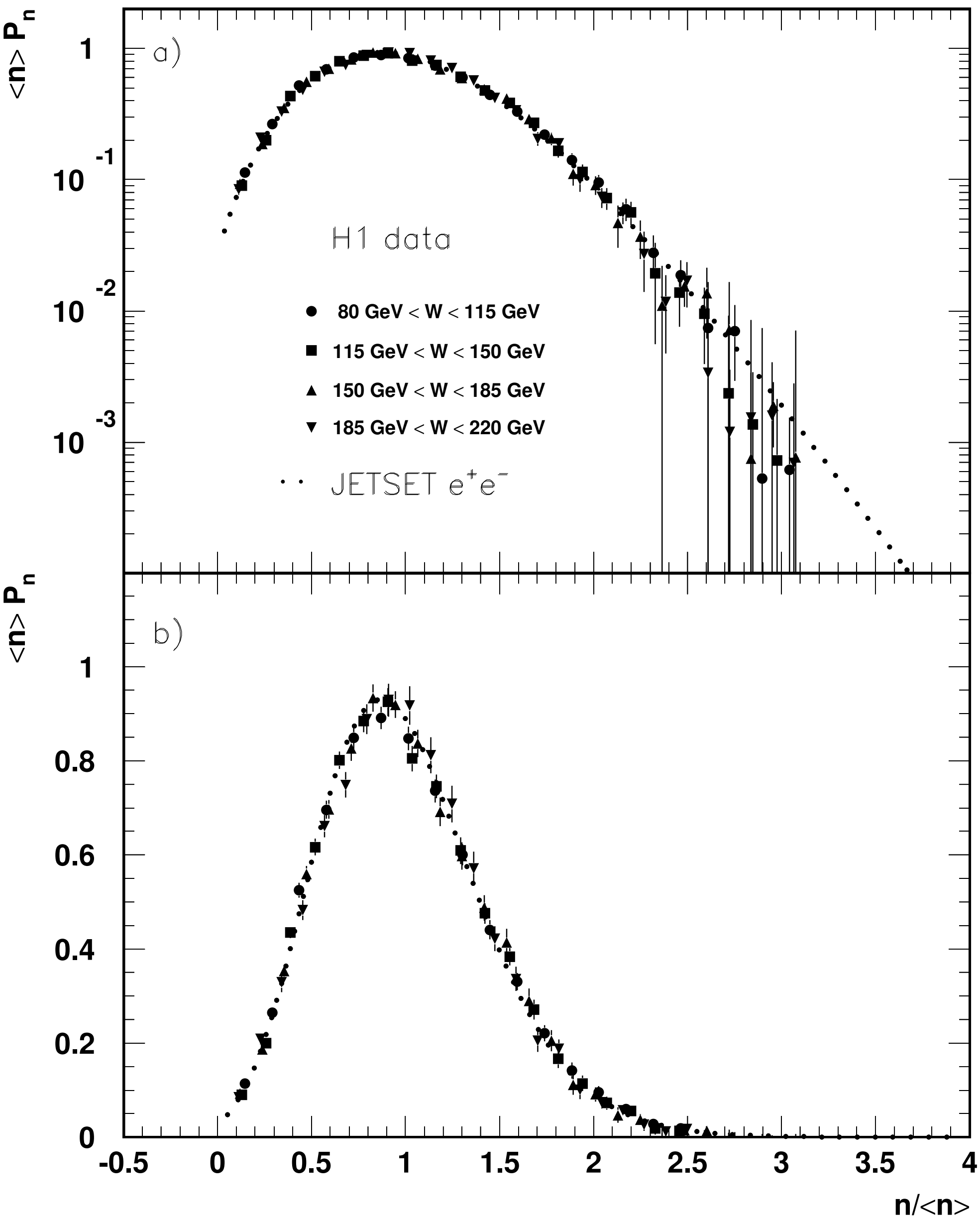


Figure 4.11 : Multiplicity distributions in non-diffractive DIS in KNO form
Unfolded multiplicity distributions in the pseudorapidity domain $1 < \eta^* < 5$ measured in different W intervals are plotted in KNO form, on a logarithmic (top) and linear scale (bottom). Statistical errors only are plotted. The dotted curve represents the KNO function for e^+e^- and was obtained with the JETSET Monte Carlo generator at a centre-of-mass energy of 91.2 GeV and for the same η^* range as the H1 data.

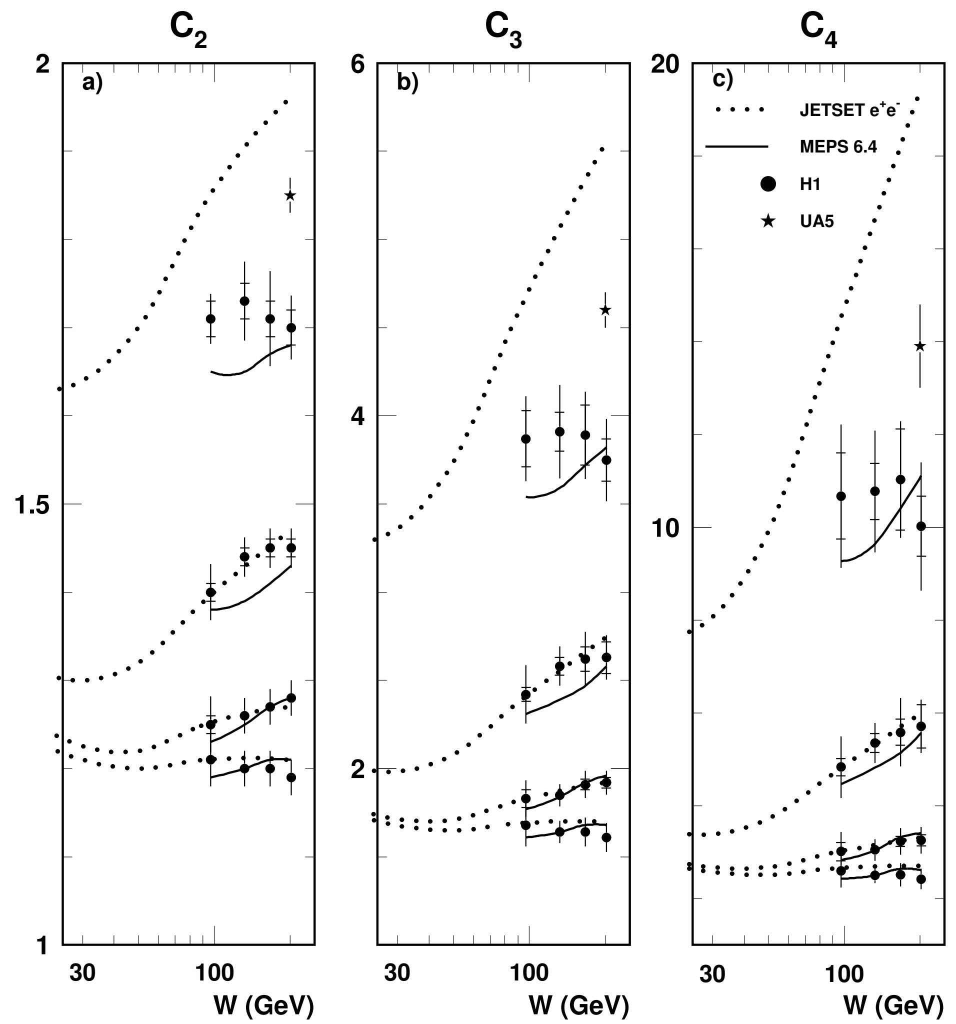


Figure 4.12 : W dependence of the multiplicity moments C_q in various η^* domains in non-diffractive DIS

From top to bottom, the η^* domains are $1 < \eta^* < 2$, $1 < \eta^* < 3$, $1 < \eta^* < 4$ and $1 < \eta^* < 5$. The curves are described in the text. The inner error bars represent the statistical errors, the outer error bars represent the total (quadratic sum of statistical and systematic) errors. Data from UA5 [54] in the interval $1 < |\eta^*| < 2$ are also shown.

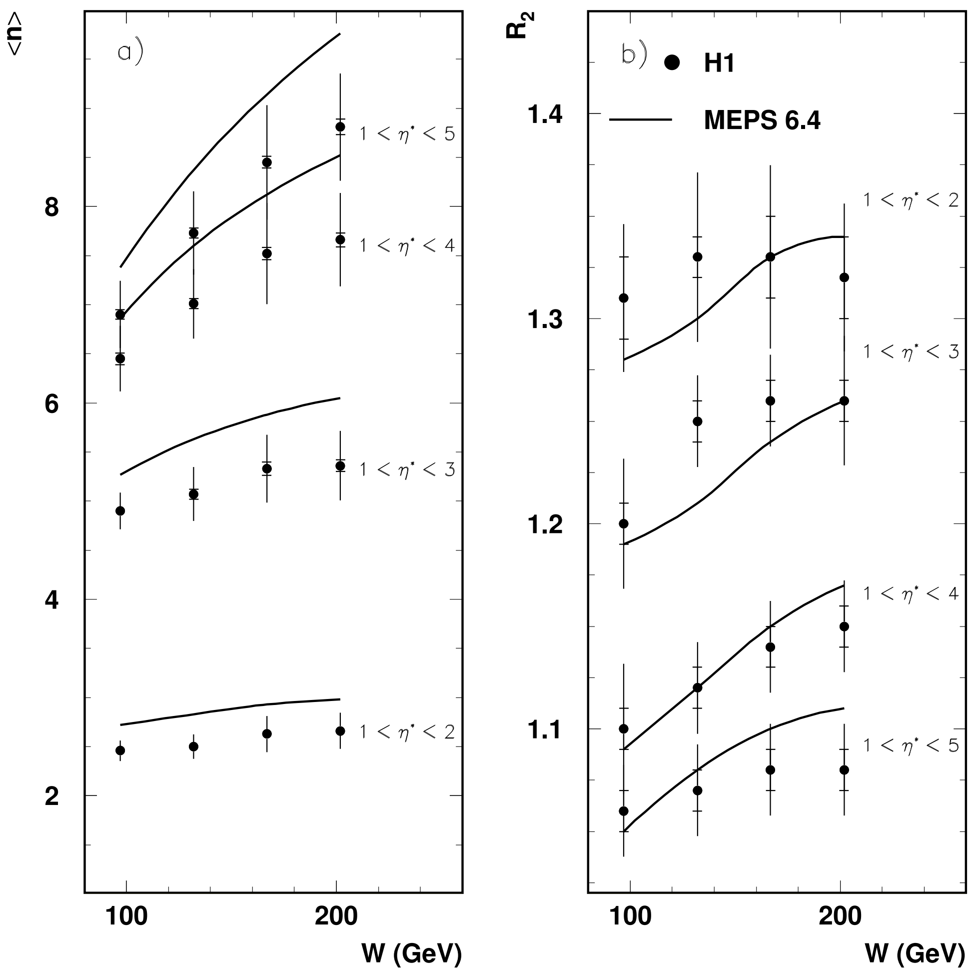


Figure 4.13 : Average charged hadron multiplicity and normalised second order factorial moment in non-diffractive DIS as a function *W*

The W dependence of $\langle n \rangle$ and R_2 in indicated pseudorapidity domains is compared with MEPS 6.4 predictions. The inner error bars represent the statistical errors, the outer error bars represent the total (quadratic sum of statistical and systematic) errors.

ated in DIS, via the two pomeron exchange diagram, from the observed ratio of diffractive to non-diffractive events. The multiplicity distributions are computed assuming a Poisson distribution in clusters for individual chains. The resulting distributions and moments describe the measured H1 data remarkably well [113].

4.2.5 Particle density and E_T flow in the central rapidity plateau

The flow of transverse energy, E_T, in multiparticle final states at high energy is studied intensively at hadron colliders and by the HERA experiments. The E_T distribution in a given phase space domain, and its moments, are convoluted observables. They depend not only on the particle density and the p_T structure of the collisions but also on multiparticle correlations, and, therefore, on the moments of the multiplicity distribution in that domain [114]. Here the evolution with W of the particle density is compared with that of the average transverse energy for DIS and hadron-hadron interactions.

Figure 4.14 shows a compilation of measurements in DIS [65, 91] and in non-diffractive hadron-hadron collisions [54, 115–118] of mean charged multiplicity per unit of rapidity or pseudorapidity (solid symbols). The H1 data are the same as those shown in Fig. 4.13a. The open symbols show measurements of the mean E_T in the region $-0.5 < \eta^* < 0.5$, presented and discussed in [119]. The solid line shows a parameterisation used by UA1 [118] with the form $\langle n\rangle = 0.35 + 0.74(W^2)^{0.105}$.

In spite of the differences on the rapidity region covered, the known differences between pseudorapidity and rapidity density, different experimental procedures and systematics, it remains of interest to note that the charged particle density for DIS at HERA interpolates quite smoothly with DIS and hadron-hadron data at lower and much higher energy. This observation is consistent with the analogy between virtual photon-hadron, real photon-hadron and hadron-hadron interactions originally advocated by Gribov and Feynman†, which suggests universality of dynamics in the central plateau [119]. However, the difference in correlation strength, noted in the previous section, suggests that such 'universality', while applicable to single-inclusive spectra [119], may not hold for higher order correlations.

Several mechanisms have been suggested to explain the dynamics of the large multiplicity fluctuations in soft hadron-hadron collisions: impact parameter averaged Poisson-like fluctuations; multiple soft parton interactions in the same event leading to mini-jets [120]; multi-pomeron exchange as in the Dual Parton Model [121] and multi-string configurations as in the Lund FRITIOF model [122]. These mechanisms have no direct analogues in DIS, except for the boson-gluon fusion QCD process which can lead to two-string colour topologies and could, therefore, mimic multi-string properties in hadron-hadron models.

In Fig. 4.14 the energy evolution of $\langle n\rangle$ in the central region in DIS is further compared with that in e^+e^- annihilation. The dotted curve is the prediction from the JETSET generator for the interval $1 < \eta^* < 2$. It shows that the hadron density evolves much faster with W than in the DIS and hadron-hadron data.

The energy evolution of $\langle n\rangle$ in perturbative QCD is controlled by the anomalous multiplicity dimension $\gamma = \mathrm{d}\ln\langle n\rangle/\mathrm{d}\ln W^2$. For e^+e^- annihilation, and in restricted (pseudo-) rapidity intervals, it is given by $\gamma_{ee}(\eta^*) \sim [3\alpha_S(k^2_{Tmax})/2\pi]^{(1/2)}$, where k_{Tmax} is the maximum possible transverse momentum at a given η^* [19]. From the JETSET e^+e^- predictions above $W = 20$ GeV it follows that γ_{ee} is constant with a value of 0.16. This is somewhat smaller than expected from the analytic result ($\sim$ 0.22 at 200 GeV) but

†For further discussion and references see [14], Ch. 4.

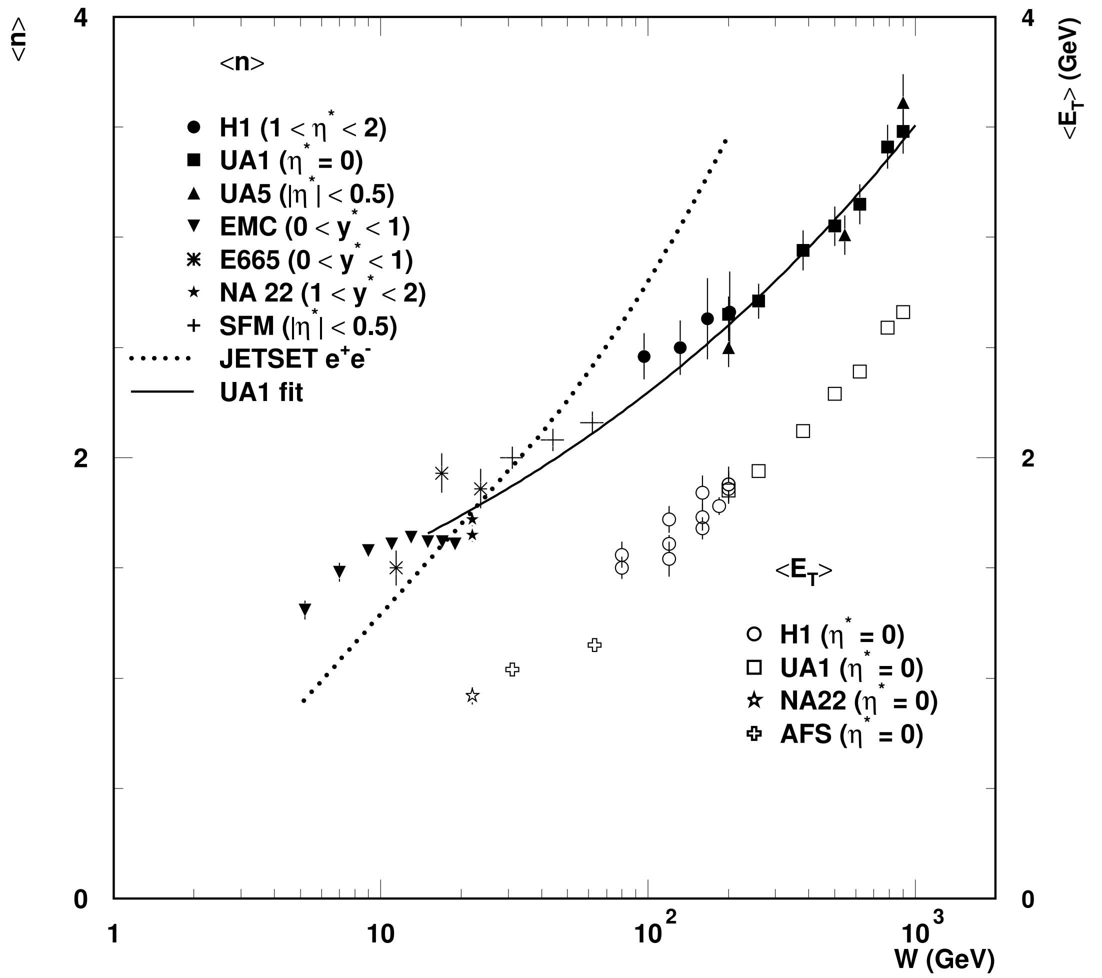


Figure 4.14 : W dependence of the average charged hadron density and the average transverse energy flow in non-diffractive hadron-hadron collisions and DIS

The W dependence of the charged particle density in DIS and hadron-hadron collisions (solid symbols, left scale) is plotted together with the mean transverse momentum flow per unit of (pseudo-) rapidity (open symbols, right scale). The curves are described in the text. η^* is pseudorapidity, while y^* is rapidity in the hadronic CMS.

agrees with that quoted in [19]. In leptoproduction, where gluon emission from the nucleon remnant is thought to be suppressed, $\gamma_{DIS}(\eta^*)$ depends more strongly on η^* with $\gamma_{DIS}(\eta^*) \sim \frac{1}{2}\gamma_{ee}(\eta^*)$ for not too large positive η^* [19]. From the DIS data in Fig. 4.14, using the UA1 parameterisation, it follows that $\gamma_{DIS} = 0.08$–0.09, consistent with expectations. This is the first semi-quantitative experimental confirmation of the 'antenna suppression' effect in DIS. However, better data are needed, also at lower W, to exploit these perturbative QCD predictions in a fully quantitative manner. This result also implies that the rate of increase with energy of $\langle n \rangle$ in the full current hemisphere in DIS and in single hemispheres for e^+e^- annihilation, discussed in Sect. 4.2.1, are expected to differ at higher centre-of-mass energies.

Non-asymptotic analytical QCD predictions for higher order multiplicity moments are at present not available. It is known, however, that the energy dependence of the moments is asymptotically controlled by $\gamma_{DIS}(\eta^*)$ or $\gamma_{ee}(\eta^*)$ [18]. It is therefore likely that the differences between the C moments in DIS and e^+e^-, described in Sect. 4.2.4, are a further reflection of suppressed gluon emission in DIS.

Finally, the comparison between $\langle n \rangle$ and $\langle E_T \rangle$ in Fig. 4.14 demonstrates that both evolve with energy rather similarly. This suggests that the increase of mean transverse energy with increasing W (decreasing Bjorken-x) at fixed Q^2, previously observed in this experiment [119], follows mainly from an increase of the hadron multiplicity and less so from a rise of the mean transverse momentum of individual hadrons.

4.3 Summary

Data are presented on the evolution with W and Q^2 of the charged particle multiplicity distribution and its statistical moments, over the ranges $80 < W < 220$ GeV and $10 < Q^2 < 1000$ GeV^2 in sub-domains of pseudorapidity space, including the full current hemisphere. The main results can be summarised as follows:

- The mean charged hadron multiplicity and the dispersion, measured in fixed intervals of W and in the domain $1 < \eta^* < 5$, show, within errors, no dependence on the virtuality of the exchanged boson over the Q^2 range covered by H1.

- The low order moments of the multiplicity distribution in the full current hemisphere show noteworthy similarities with single hemisphere data in e^+e^- annihilations, in conformity with the hypothesis of approximate environmental independence of quark hadronisation. In particular, the mean charged hadron multiplicity in DIS shows a similar rate of increase with W to that measured in e^+e^- annihilation up to 130 GeV. Analytical predictions from perturbative QCD on the mean parton multiplicity in jets, which are proven to be valid for hadrons in e^+e^- annihilation, are therefore confirmed for the first time in DIS at higher energies than presently available at LEP. Data on the second order factorial moment show that higher order QCD corrections and non-perturbative effects remain significant at present energies.

- The analysis of the multiplicity distribution in pseudorapidity domains of varying size proves that the well-documented property of KNO scaling, a general characteristic of scale-invariant stochastic branching processes, remains valid in DIS at HERA for small and for large pseudorapidity intervals, but not in intermediate size domains. The KNO phenomenon, also predicted in QCD at asymptotic energies, results from an intricate interplay of correlations, different in different types of collisions, and changing rapidly over phase space. The KNO function in the region $1 < \eta^* < 5$ is strikingly similar to that expected for e^+e^- annihilation under the same kinematical conditions.

- The charged particle density near the central region of the $\gamma^* p$ CMS grows significantly more slowly in DIS than in e^+e^- annihilation. The strength of particle correlations, as reflected in the C moments, is much larger in the latter process for small pseudorapidity domains. Such differences can be understood within perturbative QCD from calculations of the local anomalous multiplicity dimensions within the Lund dipole formalism. The comparison between e^+e^- annihilation and DIS data provides direct evidence for the 'antenna suppression' effect in current fragmentation, a characteristic of deep inelastic lepton-hadron dynamics at small Bjorken-x. The same mechanism offers a qualitative explanation for the absence of a multi-jet induced shoulder structure in the DIS multiplicity distributions in intermediate size pseudorapidity domains, now well established in e^+e^- annihilation data at LEP.

- The charged particle density near the central plateau in the deep inelastic process is of the same magnitude as that for minimum bias non-diffractive hadron-hadron interactions at the same value of the centre-of-mass energy. Its evolution with W is also comparable to that in hadron-hadron collisions measured up to 900 GeV. However, hadron-hadron collisions are characterised by substantially stronger correlations in small pseudorapidity domains. These features remain to be understood within the Gribov-Feynman pictures of DIS.

- A comparison of the evolution with W of mean transverse energy flow and charged particle density near the central region in DIS shows that the two phenomena are strongly correlated. The striking similarity with the behaviour seen in non-diffractive hadron-hadron collisions implies that a theoretical explanation within QCD should simultaneously address the dynamics of both types of processes.

- The multiplicity distributions in the smallest η^* domain examined can be well parameterised with Lognormal or Negative Binomial functions. However, the quality of the fits deteriorates in larger domains. The large n tail of the experimental distributions is well described but deviations occur for small multiplicities. Nevertheless, as economic representations of the data, these parametric forms continue to be useful for phenomenology.

- Among a variety of Monte Carlo generators presently being developed for DIS, predictions from MEPS 6.4 are shown. This model overestimates the mean charged multiplicity in the full current hemisphere, as well as in sub-domains of pseudorapidity space. Higher order moments of the multiplicity distributions are reasonably well described.

5

Multiplicity structure of the final state in diffractive DIS

The multiplicity structure of the so-called 'Large Rapidity Gap' (LRG) events, which were excluded from the analysis in the previous chapter and which are mainly attributed to diffractive dissociation of the virtual photon, is studied in the following sections*. Many details of the experimental procedure have already been explained in Chapter 4 and are not repeated here; only the aspects specific to diffractive DIS are reviewed. Section 5.1 discusses the selection of diffractive DIS interactions, the reconstruction of the kinematical variables of diffractive events, the boost to the $\gamma^* I\!P$ centre-of-mass frame and the study of systematic uncertainties. The results on diffractive DIS presented in Sec. 5.2 cover multiplicity moments and distributions in full phase space and single hemispheres, rapidity spectra and forward-backward correlations. The main results are summarised at the end of the chapter.

5.1 Experimental procedure

5.1.1 Event selection

The LRG events comprise about 10% of all DIS events at HERA. Because they form a subset of the DIS event sample, the basic event selection and the selection of the scattered positron is the same as described in Chapter 4. However, in view of the small number of diffractive events available and to increase the statistics, the cut on the scattered positron energy, E'_e was lowered to 10.5 GeV.

This analysis is based on the same event sample as used in Chapter 4 and uses data accumulated during the 1994 nominal vertex positron running period, after the tracker readout repair.

Selection of diffractive DIS events

After the basic event selection and the selection of the scattered positron, a clean sample of diffractive DIS events is obtained as follows:

- **y-balance.** To ensure agreement between the different methods for the reconstruction of the event kinematics, the following cut on the y-balance is applied:

*The results of this chapter are published in [123].

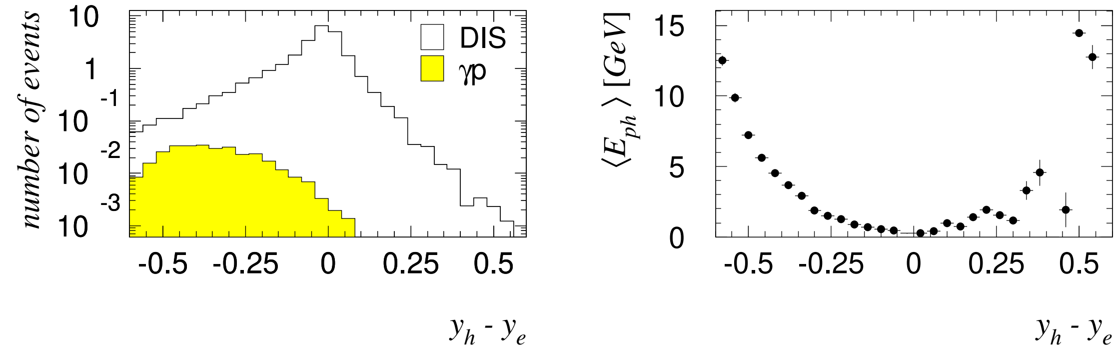


Figure 5.1 : Effect on the y-balance due to photoproduction and QED radiative background in DIS
(left) Distribution of the number of events in DIS and photoproduction (γp) interactions as a function of $y_h - y_e$, as obtained from the RAPGAP and PHOJET Monte Carlo generators, after the basic selection and the selection of the scattered positron. The histograms are normalised to the total number of events, $N_{DIS} + N_{\gamma p}$, and were both obtained for a luminosity of 1 pb^{-1}.
(right) The average energy of the initial state radiated photon in RAPGAP is plotted as a function of $y_h - y_e$.

$$|y_h - y_e| < 0.25. \tag{5.1}$$

Figure 5.1 shows that this selection effectively suppresses photoproduction interactions and events in which a high-energetic photon is radiated.

- **LRG selection.** LRG events are selected by requiring no activity above noise levels in any of the forward detectors or in the most forward part of the LAr Calorimeter [31]. Particles reach these detectors both directly from the interaction point and as a result of secondary scattering with the beam pipe wall or adjacent material like collimators. The forward detectors are thus sensitive to energetic particles going into directions beyond their geometrical acceptances. The effective ranges of sensitivity to energy flow are $\eta \lesssim 5.5$ for the LAr Calorimeter, $3.5 \lesssim \eta \lesssim 5.5$ for the PLUG Calorimeter, $4.5 \lesssim \eta \lesssim 6.5$ for the Forward Muon Spectrometer and $6.0 \lesssim \eta \lesssim 7.5$ for the Proton Remnant Tagger.

  An event is considered to have a LRG if the pseudorapidity η_{max} of the most forward cluster with an energy larger than 400 MeV in the LAr Calorimeter is smaller than 3.3 (see Fig. 5.2), the energy detected in the PLUG Calorimeter is smaller than 5 GeV, maximally 1 hit pair is detected in the Forward Muon Spectrometer and if there are no hits in the Proton Remnant Tagger.

- **Kinematical selection.** The events are restricted to the kinematical region defined by:

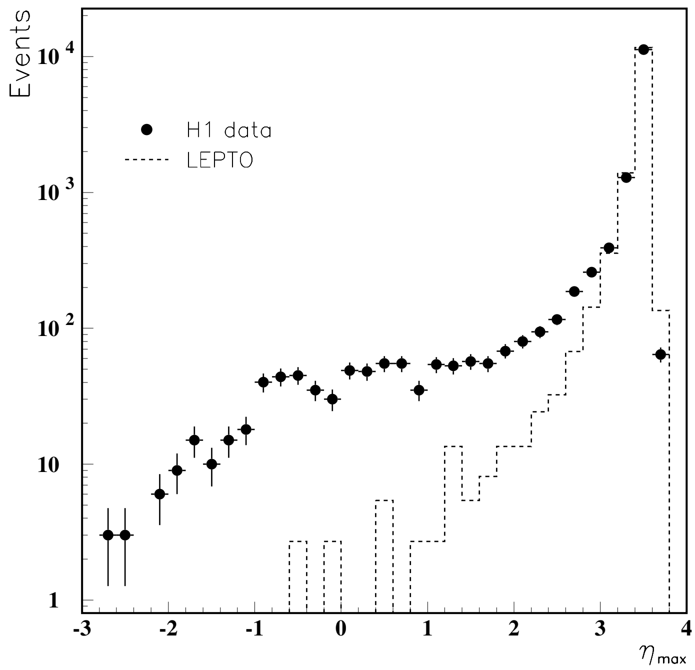


Figure 5.2 : Distribution of η_{max} for H1 DIS data and LEPTO without SCI
The η_{max} distribution is plotted for data (diffractive and non-diffractive DIS) and for events from the LEPTO Monte Carlo generator without 'soft colour interactions'. The data distribution shows an excess of events where the most forward energy cluster (with $E > 400$ MeV) detected in the LAr Calorimeter has a pseudorapidity η_{max} in the central or even backward region. These events are dubbed 'Large Rapidity Gap' events because they show an absence of hadronic activity in a large region in rapidity adjacent to the proton beam direction.

$$7.5\ \mathrm{GeV}^2 < Q^2 < 100\ \mathrm{GeV}^2 \tag{5.2}$$
$$0.05 < y < 0.6 \tag{5.3}$$
$$x_{I\!P} < 0.05 \tag{5.4}$$
$$3\ \mathrm{GeV} < M_X < 36\ \mathrm{GeV} \tag{5.5}$$
$$|t| < 1\ \mathrm{GeV}^2 \tag{5.6}$$
$$M_Y < 1.6\ \mathrm{GeV} \tag{5.7}$$

The restrictions on M_Y and $|t|$ cannot be applied directly on the reconstructed event sample because the system Y is not measured. However, they are approximately imposed by the requirement of an absence of activity in the forward detectors and the above list defines the kinematical region to which the data are corrected.

5.1.2 Event kinematics

Reconstruction of the hadronic final state

In view of the fact that an accurate measurement of the final state system X is needed to reconstruct the kinematical variables in diffractive DIS, the reconstruction of the total four-momentum of the hadronic final state has been optimised.

This is done by the 'FSCOMB' algorithm [30] which combines tracking and calorimeter information in a way that avoids double counting. Primary vertex constrained tracks are extrapolated into the calorimeter and energy clusters increasingly distant from the extrapolated track are discounted. The procedure continues until the total energy of the excluded clusters exceeds that of the track, under assumption that the particle yielding the track is a pion, or until all electromagnetic (hadronic) clusters within a cylinder of radius 30 cm (50 cm) around the extrapolated track have been removed. Detector noise is potentially problematic, particularly at small values of M_X. To minimise its effect, techniques have been developed that reject low energy isolated clusters in the LAr Calorimeter. Any calorimeter deposit that is not associated with a track or rejected as noise is accepted. This method yields the hadronic final state vector $(E_h, p_{xh}, p_{yh}, p_{zh})$ with a much improved resolution and accuracy.

Reconstruction of the diffractive DIS kinematical variables

The improved reconstruction of the hadronic final state vector makes it possible for the different kinematical reconstruction schemes based on hadrons (see Sec. 4.1.3), to compete with the positron-only method. In fact, it was found that a combination of the positron-only and the double angle method yields the highest quality for the reconstruction of y and Q^2:

$$y = y_e^2 + y_{DA}(1 - y_{DA}), \tag{5.8}$$
$$Q^2 = \frac{4E_e^2(1-y)}{\tan^2(\theta_e/2)}. \tag{5.9}$$

Here, y is obtained from the weighted average of y_e and y_{DA} in such a way that y_e dominates at high y and y_{DA} at low y, where the resolution of the respective variables is best.

The invariant mass M_X is reconstructed with a method that utilises the optimised reconstruction of y and corrects a possible mismatch in the energy scale of the hadronic final state:

$$M_X^2 = (E_h^2 + p_{xh}^2 + p_{yh}^2 + p_{zh}^2)\frac{y}{y_h}. \tag{5.10}$$

The quantities $x_{I\!P}$ and β are obtained from:

$$x_{I\!P} = \frac{Q^2 + M_X^2}{ys}, \tag{5.11}$$

$$\beta = \frac{x}{x_{I\!P}}, \tag{5.12}$$

using the approximation $|t| \ll M_X^2 + Q^2$.

Figure 5.3 shows the resolution and accuracy of the reconstructed kinematical variables y, Q^2, $x_{I\!P}$ and M_X.

Lorentz boost to the $\gamma^* I\!P$ centre-of-mass system

The hadronic final state is studied in the rest frame of the system X. Therefore all reconstructed tracks have to be boosted to the $\gamma^* I\!P$ CMS.

The pomeron direction cannot be unambiguously determined because the outgoing system Y is not measured. Since its transverse momentum is small, it has been assumed that the $I\!P$ direction is collinear with the incident proton in the rest frame of the system X, so that the four-momentum of the $\gamma^* I\!P$ CMS, measured in the laboratory, is given by:

$$x_{I\!P} \cdot P + q = x_{I\!P} \cdot P + k - k' \tag{5.13}$$

P, q, k and k' are again the four-momenta of the incoming proton, exchanged photon, incoming and outgoing positron, respectively, measured in the laboratory frame.

As for the hadronic CMS, the positive z-axis is chosen to point in the direction of the exchanged photon. In this chapter the superscript * is used to denote variables measured in the $\gamma^* I\!P$ CMS.

Figure 5.4 shows the acceptance and resolution in rapidity in three M_X-bins. Rapidity is calculated by assuming the pion mass for all particles.

The rapidity spectra have also been recomputed with rapidity defined along the thrust axis in the CMS of the system X. Except for the lowest M_X interval, no significant difference is seen with the results in Fig. 5.7. This is consistent with the observation in [124] that the thrust axis in LRG events is strongly aligned with the γ^* direction in the $\gamma^* I\!P$ system.

5.1.3 Systematic errors

Several sources of possible systematic errors are investigated. The analysis is repeated for each source and the changes to the results are added in quadrature. For illustration, the typical systematic error on the mean total charged multiplicity and on the central rapidity ($-0.5 < y < 0.5$) particle density are given in square brackets for each source separately.

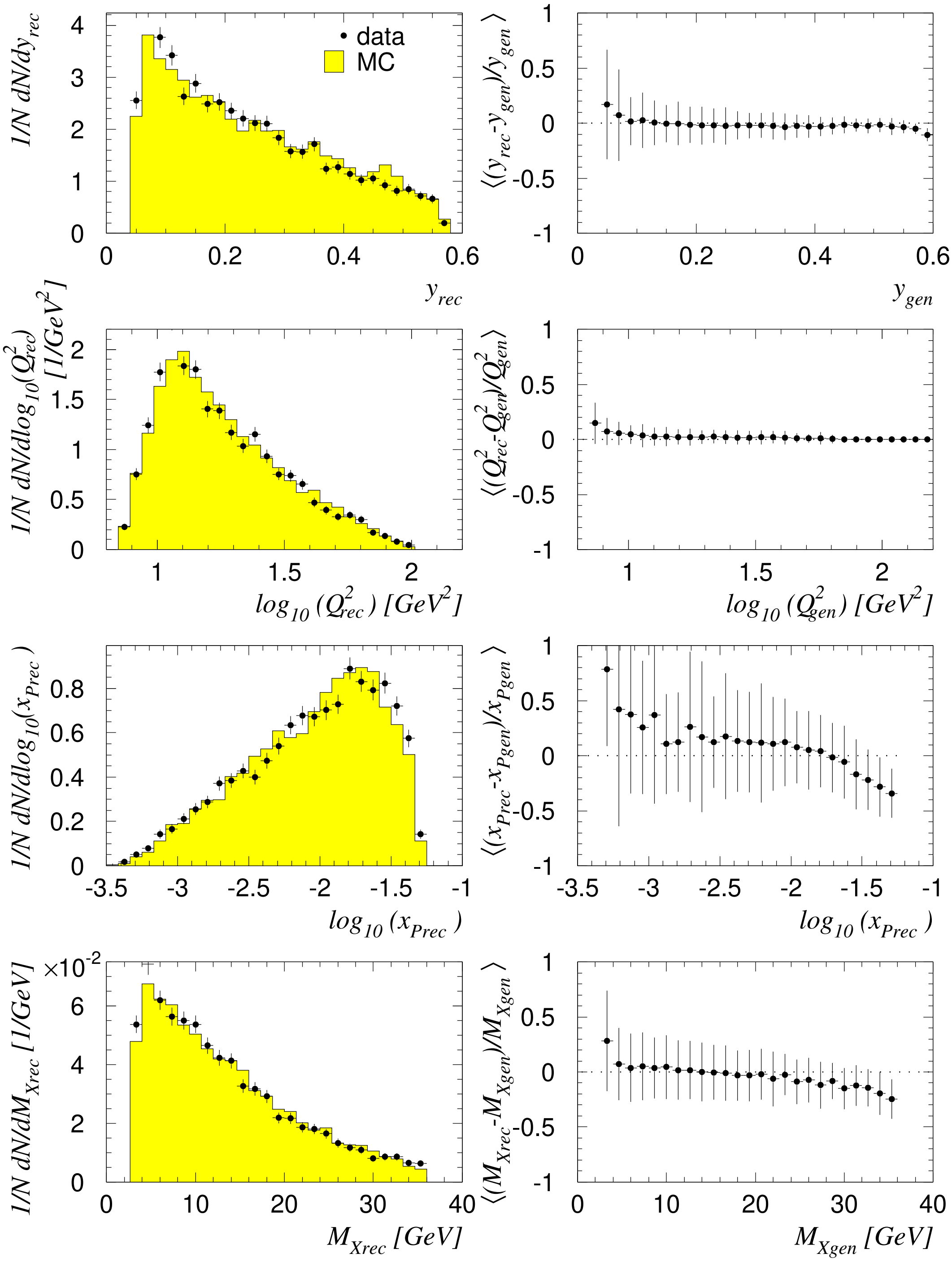


Figure 5.3 : Kinematical variables y, Q^2, $x_{I\!P}$ and M_X obtained using the improved reconstruction of the hadronic final state.
Event normalised distributions for y, Q^2, $x_{I\!P}$ and M_X are shown on the left. H1 data are plotted as black dots, while the Monte Carlo distributions are represented by the shaded histograms. The resolution and accuracy, obtained from a Monte Carlo simulation based on the RAPGAP generator, are shown on the right. For each variable the evolution of the mean relative difference is plotted as a function of the true (generated) variable. The error bars reflect the spread of the relative difference.

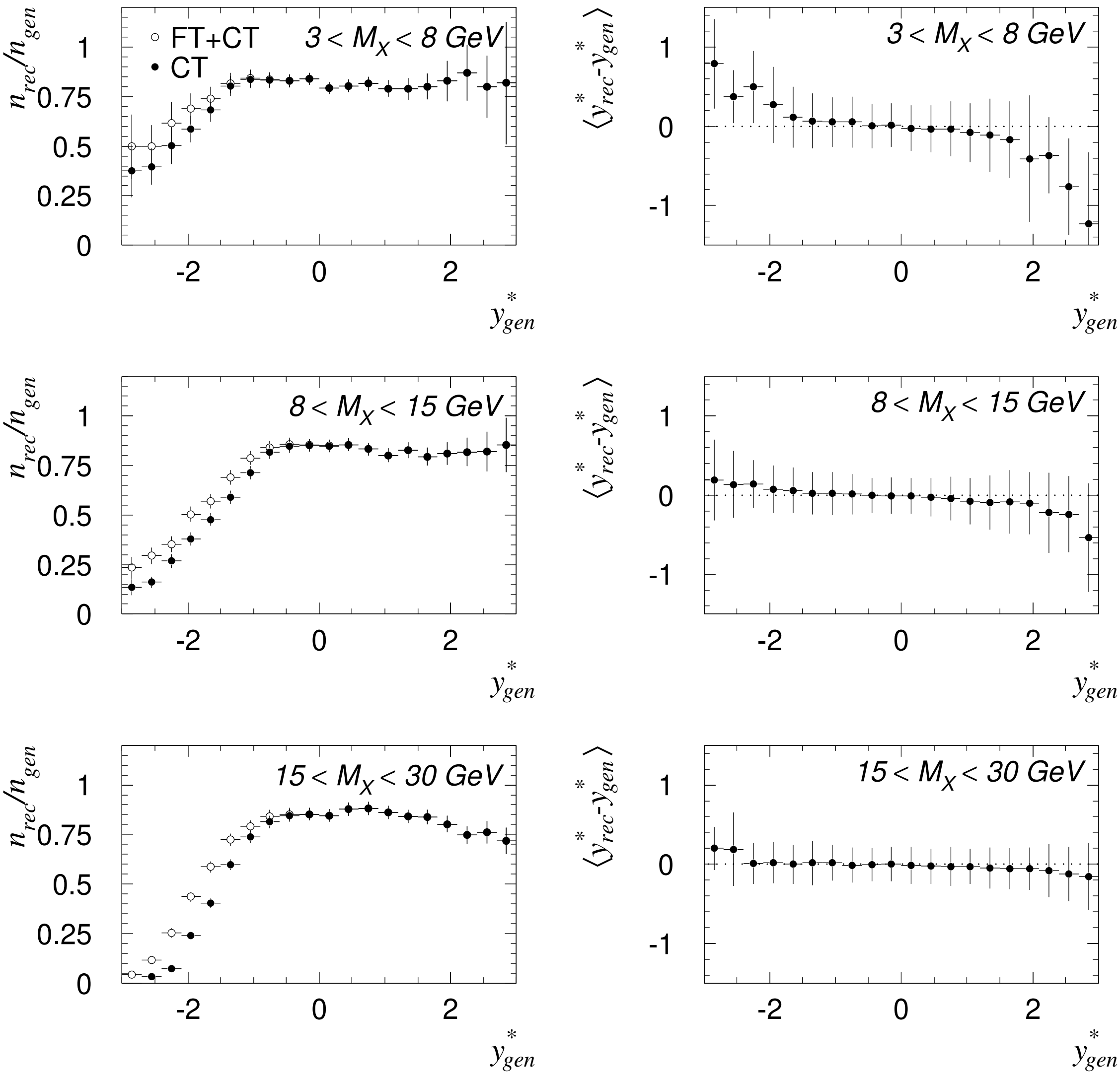


Figure 5.4 : Acceptance and resolution in rapidity in the $\gamma^* I\!P$ CMS as a function of rapidity

The figures on the left show the fraction of generated particles that are detected in the Central (CT) and Forward (FT) Trackers as function of rapidity in the $\gamma^* I\!P$ CMS for different ranges in M_X. The requirement of a forward rapidity gap ensures that the hadronic final state of the system X is well contained in the central detectors. Only central tracks are used for the results shown in Sec. 5.2. The plots on the right show the difference between reconstructed and generated rapidity in the $\gamma^* I\!P$ CMS. The error bars reflect the spread of the difference.

- The error due to the uncertainty of the energy scale of the hadronic final state is estimated by scaling the LAr Calorimeter, BEMC and Central Tracker energies by $\pm 4\%$, $\pm 20\%$ and $\pm 3\%$ respectively [0.4%, 0.4%].

  The systematic uncertainty in the reconstruction of the scattered positron is studied by varying the energy E_e' and polar angle θ_e by $\pm 1\%$ and ± 1 mrad, respectively [0.3%, 0.9%].

- The influence of the pomeron (reggeon) flux and pomeron (reggeon) structure function used in the Monte Carlo generator for correction is investigated by reweighting the β, $x_{\mathbb{P}}$ and t distributions for Monte Carlo events [β: 1.3%, 2.2%; $x_{\mathbb{P}}$: 0.3%, 1.6%; t:1.1%, 3.5%].

- Both the colour dipole model, as in ARIADNE [40], and the parton shower model are used to evaluate the influence of these event generation schemes on the corrections. The full difference is taken as the systematic error [0.6%, 0.6%].

- The strangeness-suppression parameter (`PARJ(2)` in JETSET [36]), affecting the rate of strange particle production in the simulation of the fragmentation process, has been varied in the range 0.2–0.3 according to recent results on strange particle production [125, 126] [0.1%, 0.1%].

- Track quality criteria (such as track length and the number of hits) are varied to estimate systematic errors related to an imperfect description of the acceptance and efficiency of the Central Tracker in the Monte Carlo simulation [1.3%, 3.1%].

  An uncertainty of 30% is assumed on the Monte Carlo correction outside the tracker acceptance (θ outside the range $8°$–$165°$ or $p_T < 0.1$ GeV; the range between $8°$–$15°$ has been cross-checked with data from the Forward Tracker) [3.6%, 1.5%].

- Background events are suppressed by the event selection criteria. Remaining background contamination is estimated by including events simulated with the PHOJET and DIFFVM generators in the Monte Carlo event sample [PHOJET: $< 0.1\%$, $< 0.1\%$; DIFFVM: 0.7%, 1.0%].

  The number of events with initial and final state QED radiation in the Monte Carlo is changed by $\pm 50\%$ [0.4%, 0.1%].

- A fit to a smeared Negative Binomial distribution is used as a cross-check on the unfolding results [error on central rapidity particle density: 2.8%].

5.2 Results

All data presented below can be found in numerical form in App. B where they are listed together with their statistical and systematic errors. The data presented in this chapter are corrected for the effects of acceptance and resolution of the H1 detector in the kinematical region specified in Sec. 5.1.1. The correction methods are described in Sec. 4.1.4 and App. A. The data span the M_X range from 3 to 36 GeV, distributed over the intervals as listed in Table 5.1. Statistical and systematic errors on the data points shown in the figures are combined in quadrature. When two error bars are displayed, the inner one is the statistical error, the outer one the total error. In comparing H1 LRG data at a given M_X with data from other processes, the corresponding centre-of-mass energy scale is chosen to be W for fixed-target DIS data, $\sqrt{s}$ for e^+e^- and non-diffractive hadron collisions and M_X for hadro-produced diffractive states.

M_X range (GeV)	$\langle M_X \rangle$ (GeV)	$\langle \beta \rangle$	$\langle Q^2 \rangle$ (GeV^2)	no. of events
3 – 8	5.4	0.41	21	1492
8 – 15	11.4	0.17	26	1515
15 – 30	21.1	0.06	27	1359
4 – 6	5.0	0.43	22	638
6 – 8	7.0	0.30	23	530
8 – 11	9.5	0.21	26	737
11 – 15	13.0	0.13	26	778
15 – 19	16.9	0.08	27	543
19 – 24	21.3	0.06	27	468
24 – 36	29.1	0.03	27	562

Table 5.1 : Kinematical bins used in the study of the multiplicity structure of the final state in diffractive DIS
Average M_X, β and Q^2 and the number of observed events for the different intervals in M_X considered in this analysis. Both a coarse and a fine grained binning in M_X are used.

5.2.1 Multiplicity moments

Fig. 5.5a shows the dependence of the mean charged particle multiplicity on M_X in full phase space. The H1 LRG data can be parameterised by a form $\langle n \rangle = a_1 + a_2 \log M_X^2 + a_3 \log^2 M_X^2$, with $a_1 = 2.2 \pm 0.4$, $a_2 = 0.08 \pm 0.17$ and $a_3 = 0.21 \pm 0.02$ (χ^2 per degree of freedom $= 0.4$), indicating that $\langle n \rangle$ increases faster than the logarithm of the centre-of-mass energy. In non-diffractive DIS at HERA, a similar rate of increase but with respect to W is observed (see Chapter 4). Also shown are data on $\langle n \rangle$ for the diffractively produced system X in the reactions $\pi^\pm p \to X^\pm p$ and $K^+ p \to X^+ p$ [127, 128]. Although the two data sets agree well for $M_X \lesssim 10$ GeV, $\langle n \rangle$ in LRG events exceeds that in meson diffraction at larger masses. There are no meson diffraction results with $M_X \gtrsim 15$ GeV. The meson diffraction data are close to the e^+e^- annihilation results, represented here by the predictions from the JETSET Monte Carlo model (dotted line) which is known to reproduce the e^+e^- multiplicity data very well over a wide energy range.

Results on the dispersion and the correlation parameter R_2, (Figs. 5.5b,c) confirm that also the second order moments in LRG data are quite similar within errors to meson diffraction and to e^+e^- for $M_X \lesssim 10$ GeV. Stronger multiplicity fluctuations and correlations than in e^+e^- are observed at larger M_X. The rise of R_2 with M_X demonstrates that KNO scaling [49] does not hold in the M_X range studied here.

These similarities seen in Figs 5.5a–c have lead to the view [127, 129] that in meson diffraction the (mainly) longitudinal momentum exchange with the target leads to an excited meson state which can be pictured as a colour-string of invariant mass M_X stretched between the valence q and $\overline{q}$ of the meson. This string subsequently hadronises in a similar way as a quark pair in $e^+e^- \to q\overline{q}$ at the corresponding centre-of-mass energy $\sqrt{s} = M_X$. A comparative study of the thrust distribution, energy and quantum number flow in the rest frame of the system X with that measured in e^+e^- added further support for this interpretation [130]. For $M_X \lesssim 10$ GeV the same idea has been successfully applied to proton dissociation assuming that the baryonic system X now results from the fragmen-

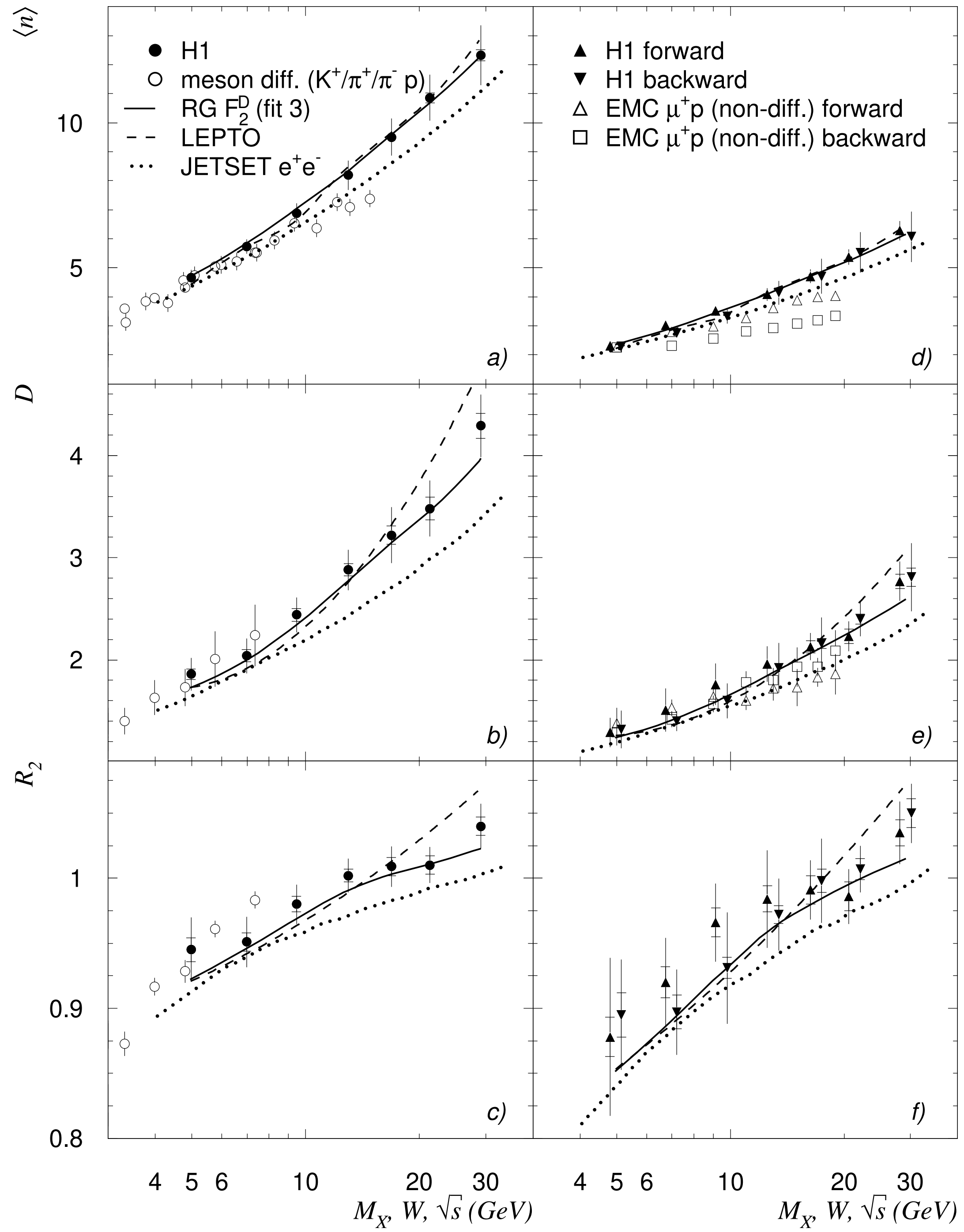


Figure 5.5 : The lower moments of the multiplicity distribution in diffractive DIS as a function of M_X, in full phase space and in single hemispheres

The multiplicity moments $\langle n \rangle$, D and R_2 in full phase space (a–c) and in single hemispheres (d–f) for charged hadrons in the system X, as a function of M_X (H1 and meson diffraction), W (EMC) and $\sqrt{s}$ (e^+e^-). For clarity, H1 data points in single hemispheres are slightly shifted in the horizontal direction with respect to their true positions. Also shown are predictions of several Monte Carlo models (see text). The Monte Carlo curves in forward and backward hemispheres are symmetric.

tation of a (valence) quark-diquark string, thus explaining observed similarities with DIS lepton-nucleon data at values of W comparable to M_X [127, 131]. Due to the larger values of Bjorken-x involved, the latter reaction is dominated by quark-diquark fragmentation.

Combining these experimental results, it follows that low-mass diffraction (above the resonance region) in hadron collisions and photon dissociation in DIS at HERA could be interpreted as the hadronisation of a single string, or colour dipole, with colour triplet-antitriplet endpoints. The larger multiplicity moments seen in LRG events for $M_X \gtrsim 10$ GeV relative to the other processes suggest, however, that the above interpretation is incomplete and that high-mass diffraction involves additional mechanisms.

In high-mass proton diffraction at $p\overline{p}$ collider energies, measurements show that, there too, the multiplicity structure of the system X deviates from the expectations for quark-diquark fragmentation and becomes, in fact, similar to that of soft non-diffractive interactions with $\sqrt{s} = M_X$ [132]. Within the framework of the Dual Parton Model (DPM), which is phenomenologically very successful [133], soft non-diffractive collisions are described by the fragmentation of two or more strings corresponding to single or multiple pomeron exchange in the elastic channel. The similarity between non-diffractive and diffractive processes is explained in DPM [134] by assuming that the colourless exchange in the latter becomes resolved in a $q\overline{q}$ pair at large M_X and subsequently interacts with the dissociating hadron. The diffractive state is then described by two colour strings, one stretched between a valence quark of the excited hadron and a quark in the exchange, the other between a diquark and the remaining quark of the $q\overline{q}$ pair. Multi-string systems are known to lead to a faster than logarithmic increase of $\langle n \rangle$ with energy, to wider multiplicity distributions and stronger long-range particle correlations than single-string fragmentation [133]. It is also important to note that only colour triplet-antitriplet strings are considered in the DPM.

Present models for diffractive DIS also include additional mechanisms besides simple $q\overline{q}$ fragmentation. Fig. 5.5a–c show model calculations with RAPGAP (fit 3) (solid line) which describe the data well. The difference between RAPGAP and JETSET e^+e^- follows from the presence, in RAPGAP, of additional diagrams involving gluons from the colourless exchange, leading to a large contribution from boson-gluon fusion. The partonic state in lowest-order BGF consists of a gluon (the 'pomeron remnant') and a $q\overline{q}$ pair in a colour-octet state. The fragmentation of this state allows for various string topologies, including two-string configurations, thus leading one to expect further similarities with large M_X hadron dissociation. The admixture of the BGF sub-process with the $q\overline{q}$ and QCD-Compton processes naturally explains the larger mean multiplicity and stronger fluctuations. The results for RAPGAP (fit 1), with a quark-dominated pomeron leading to $q\overline{q}$ parton states, are very similar to those of JETSET e^+e^-, as expected, and not shown.

The data can also be qualitatively understood in the photon dissociation picture of diffraction. The lowest-order ('aligned jet' [27, 135]) excitation ($\gamma^* \to q\overline{q}$) is dominant for $M_X^2 < Q^2$ and leads to a similar final state as in e^+e^- annihilation. In addition, higher order fluctuations (such as $\gamma^* \to q\overline{q}g$ where the gluon has low momentum), which resemble the BGF sub-process, contribute at larger M_X and effectively interact as an octet-octet colour dipole [136, 137]. Due to the octet colour charge at the dipole endpoints, such a system hadronises with a larger mean multiplicity than a $q\overline{q}$ state [138].

The LEPTO model with 'soft colour interactions' (dashed curve) is also seen to agree with the H1 data although it predicts somewhat larger multiplicity fluctuations above $M_X \sim 20$ GeV. This model too contains a sizable BGF contribution. However, here diffraction is viewed as a final state interaction and does not assume the existence of a pre-formed colour-neutral object in the proton.

The moments of the multiplicity distribution for particles with positive and negative rapidity ('forward' and 'backward', respectively) are displayed in Figs. 5.5d–f. The H1 data show no evidence for an asymmetry between the forward and backward hemispheres, in contrast to what is observed for the mean multiplicity measured in fixed-target μp DIS ($Q^2 > 4$ GeV^2) [91], where the influence of proton fragmentation on the backward hemisphere multiplicity distribution is known to be substantial. The μp data in the current fragmentation region, where the comparison with LRG is most relevant, agree well with the e^+e^- expectations, as expected for production dominated by quark jets. Here also, the LRG results are characterised by larger $\langle n \rangle$ and stronger fluctuations above $M_X \gtrsim 10$ GeV.

The RAPGAP and LEPTO models also predict forward-backward symmetry of the single hemisphere moments and describe the data adequately. The earlier noted differences with e^+e^- annihilation for the full phase space moments are also seen here.

The role of gluons in high-mass photon dissociation is prominent in all the DIS models considered, in contrast to models for hadronic soft diffraction where quark (diquark) fragmentation is dominant. The indication that the pomeron remnant has a large gluon content also opens interesting opportunities for comparison with gluon jet fragmentation in other processes.

5.2.2 Multiplicity distributions

The multiplicity distributions in full phase space have been measured, separately for negatively and positively charged tracks, in three intervals of M_X. The results are displayed in Fig. 5.6a–c in the form of a KNO distribution [49]. No significant difference is observed between the distributions for positively and negatively charged hadrons. The data are well reproduced by the RAPGAP model (solid curves) although there are indications that it underestimates the high-multiplicity tail of the distribution at large M_X. The comparison with JETSET e^+e^- predictions (at $\sqrt{s} = M_X$) shows that the multiplicity distribution is broader in the LRG data, indicative of stronger correlations among the hadrons. The predictions for LEPTO are similar to those of RAPGAP.

Figs 5.6d–f further illustrate the forward-backward symmetry of the system X, now for the all-charged multiplicity distribution. The RAPGAP and LEPTO predictions are also forward-backward symmetric but tend to fall below the data at large z. The single hemisphere distributions are closer to the e^+e^- expectations than in full phase space (cf. Figs. 5.6a–c). This difference can be understood as the effect of correlations between hadrons emitted in opposite hemispheres which, as will be shown below, are larger in DIS LRG data.

5.2.3 Rapidity spectra

The charged particle rapidity density in three intervals of M_X is shown in Fig. 5.7. The spectrum rises slowly with M_X in the central region and a rapidity plateau develops with increasing phase space. These features confirm earlier observations in hadron diffraction [130, 132, 139] that the diffractive system hadronises in a jet-like manner both in the forward and backward regions [124, 140, 141]. There is no evidence for a significant forward-backward asymmetry of the y spectra contrary to what is observed in μN interactions [65, 98].

The particle density in the central region is much larger in LRG events than in μN interactions for W values close to M_X. It is also larger than in e^+e^- annihilation (at $\sqrt{s} = M_X$) according to the JETSET expectation. The RAPGAP (fit 1) model curve is

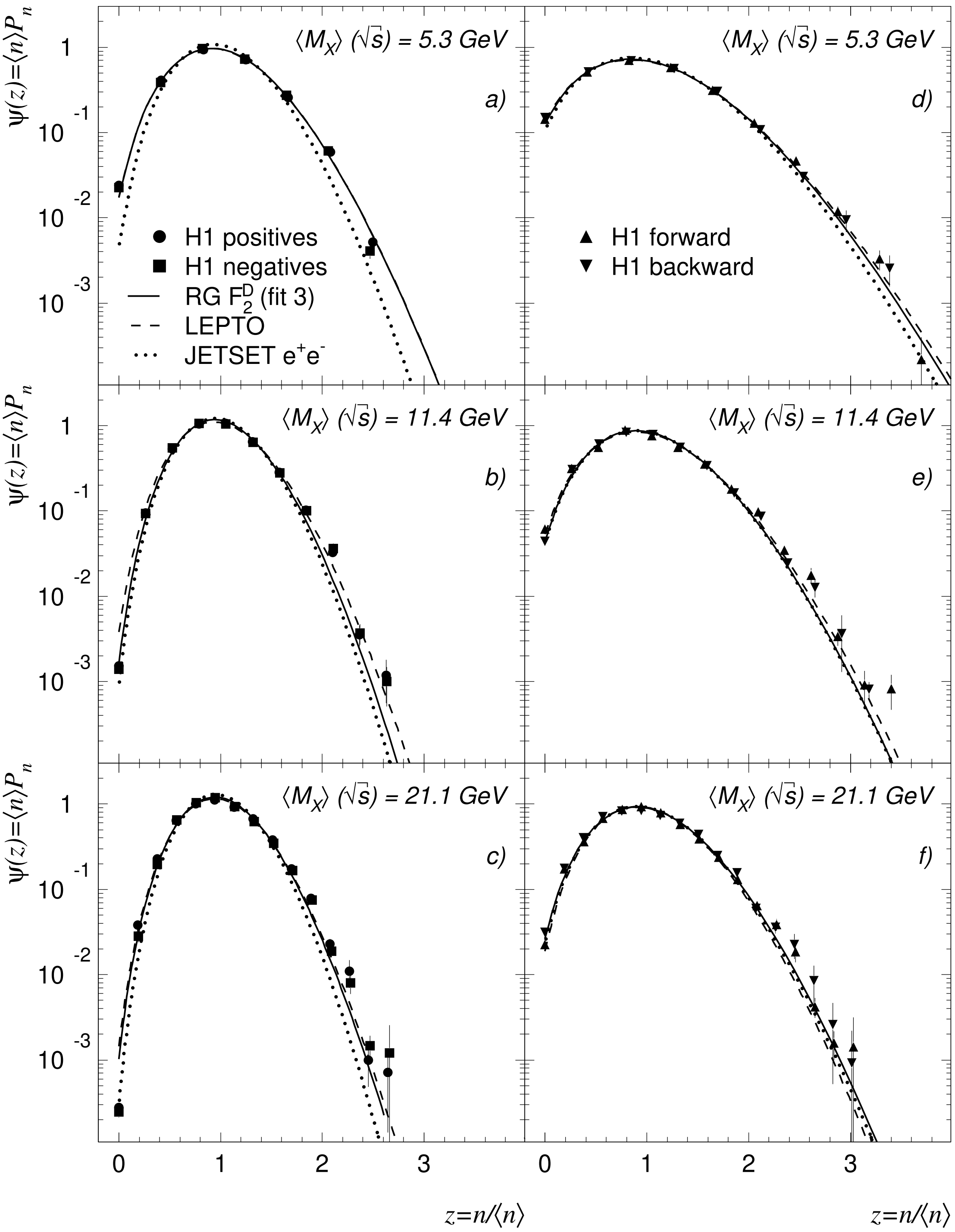


Figure 5.6 : Multiplicity distributions in diffractive DIS in KNO form
The multiplicity distributions in KNO form is plotted for three intervals in M_X (DIS data and Monte Carlo) and at $\sqrt{s} = \langle M_X \rangle$ (JETSET e^+e^-), in full phase space for positive ad negative particles separately (a–c), and for all charges in single hemispheres (d–f). The error bars show statistical errors only. Also shown are predictions of several Monte Carlo models (see text). The Monte Carlo curves are charge and forward-backward symmetric.

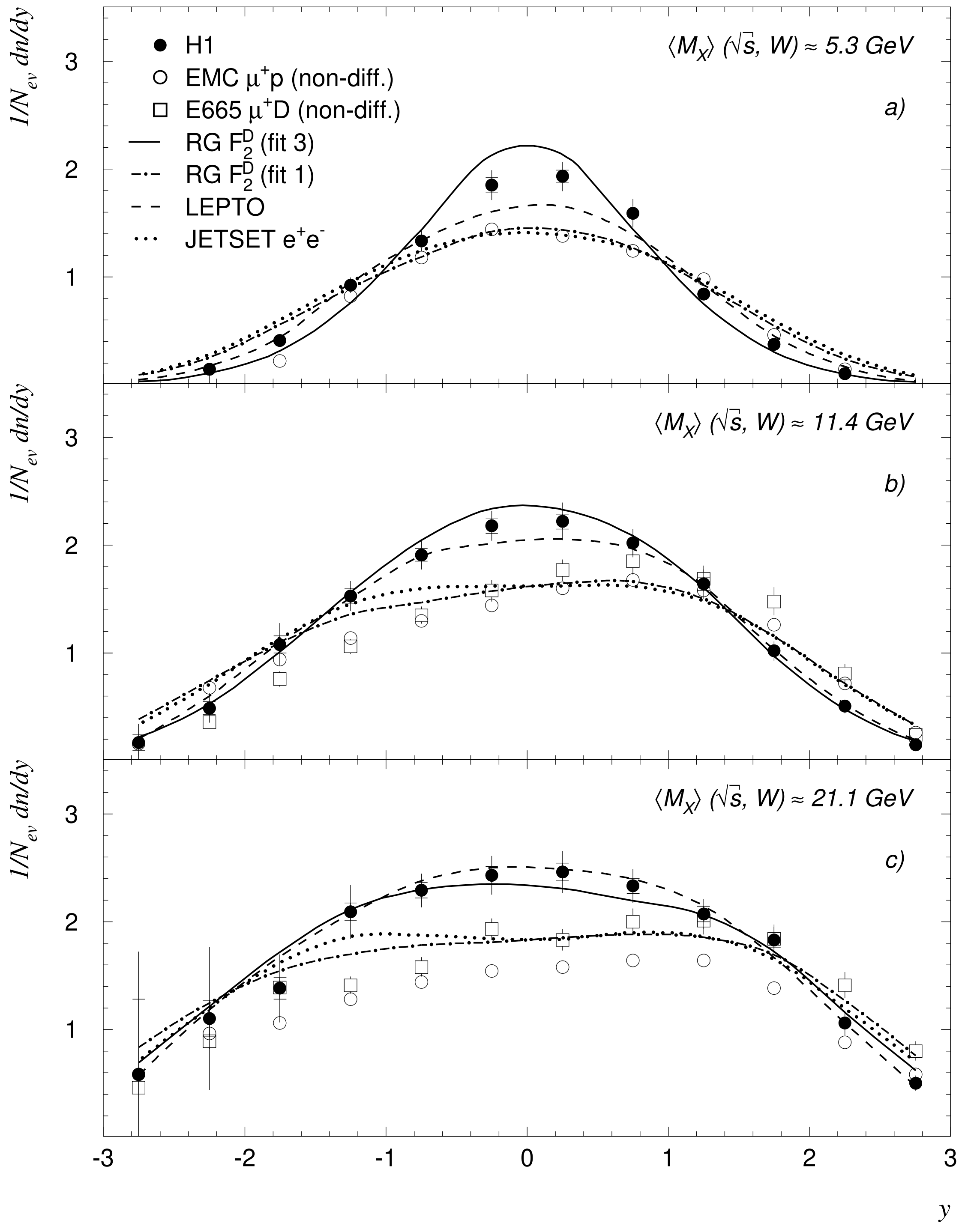


Figure 5.7 : Rapidity spectra in diffractive DIS
Charged particle rapidity spectra are plotted for three intervals in M_X (H1), at $W = \langle M_X \rangle$ (μN) and at $\sqrt{s} = \langle M_X \rangle$ (JETSET e^+e^-). The $\langle W \rangle$ values for EMC and E665 differ slightly from the ones indicated for H1 and are equal to 5.2, 11 and 19 GeV and 11.4 and 23.6 GeV, respectively. Also shown are predictions of several DIS Monte Carlo models (see text).

close to JETSET and the fixed-target data and predicts a too small particle density. Both RAPGAP with a leading-gluon distribution and LEPTO describe the rapidity spectra, except for small deviations at lowest M_X.

Figure 5.8a further compares the M_X dependence of the central particle density (defined as the mean multiplicity in the region $-0.5 < y < 0.5$) in LRG events to e^+e^- expectations, to that in μN collisions [65, 98], non-diffractive meson-proton [115] collisions and proton diffraction [142]. The particle density near $y = 0$ is seen to be larger in LRG events than in all other processes.

The excess particle production relative to that in e^+e^- and μN indicates that additional mechanisms besides hard and soft gluon Bremsstrahlung from quarks are needed (cf. Sec. 5.2.1). The comparison with non-diffractive meson-proton and high-mass proton diffraction further shows that the central particle density in processes which are believed [133] to involve two or more strings with colour triplet-antitriplet endpoints ($q\overline{q}$ and quark-diquark strings) is also significantly lower than in the LRG data. This, together with previous observations, argues in favour of models which attribute a higher gluonic content to the partonic system created in virtual photon dissociation than in the other processes.

An estimate of the importance of an additional gluonic component may be obtained by assuming that the particle density in the central region is a linear superposition of two contributions, one arising from $q\overline{q}$ fragmentation (including additional QCD radiation), the second from a colour octet-octet string configuration. This hypothesis is in line with expectations from the photon dissociation picture of diffractive DIS (see e.g. [137]). Using the EMC data for the former, and JETSET simulations of a colour-singlet gluon-gluon string for the latter, it is found that, at $\langle M_X \rangle = 11.4$ GeV, about equal contributions of the two components are needed to explain the particle density at mid-rapidity. This is consistent with the contribution of about 50% from boson-gluon fusion events to the total diffractive cross section estimated with the RAPGAP model; the latter value depends, however, on the chosen cut-off scheme [124].

The RAPGAP (fit 3) predictions for the central particle density are shown in Fig. 5.8a. They are compatible with the LRG data only above $M_X \gtrsim 10$ GeV and are nearly M_X independent. The LEPTO model on the other hand predicts a rather stronger dependence on M_X, closer to the tendency observed in the H1 data. The enhanced particle density, both in RAPGAP (fit 3) and in LEPTO, are related to the large contribution from boson-gluon fusion. The RAPGAP (fit 1) model version follows closely the JETSET e^+e^- prediction.

In order to investigate the sensitivity of the results to possible contributions from non-diffractive processes, the analysis has been repeated changing the cut $x_{I\!P} < 0.05$ to $x_{I\!P} < 0.025$. Within errors, no significant effect on the results was observed.

5.2.4 Forward-backward correlations

In this section, differences between the LRG events and final states in other processes are further examined through a measurement of the correlation between hadrons emitted in opposite event hemispheres. These so-called 'forward-backward' correlations are known to be sensitive to finer details of the fragmentation process and, in particular, to the presence in an inclusive event sample of several distinct sub-classes of events [143].

In previous experiments, the forward-backward correlation was analysed by studying the regression between the forward multiplicity, n_F, and the backward multiplicity, n_B. The correlation is usually well parameterised by a simple linear dependence:

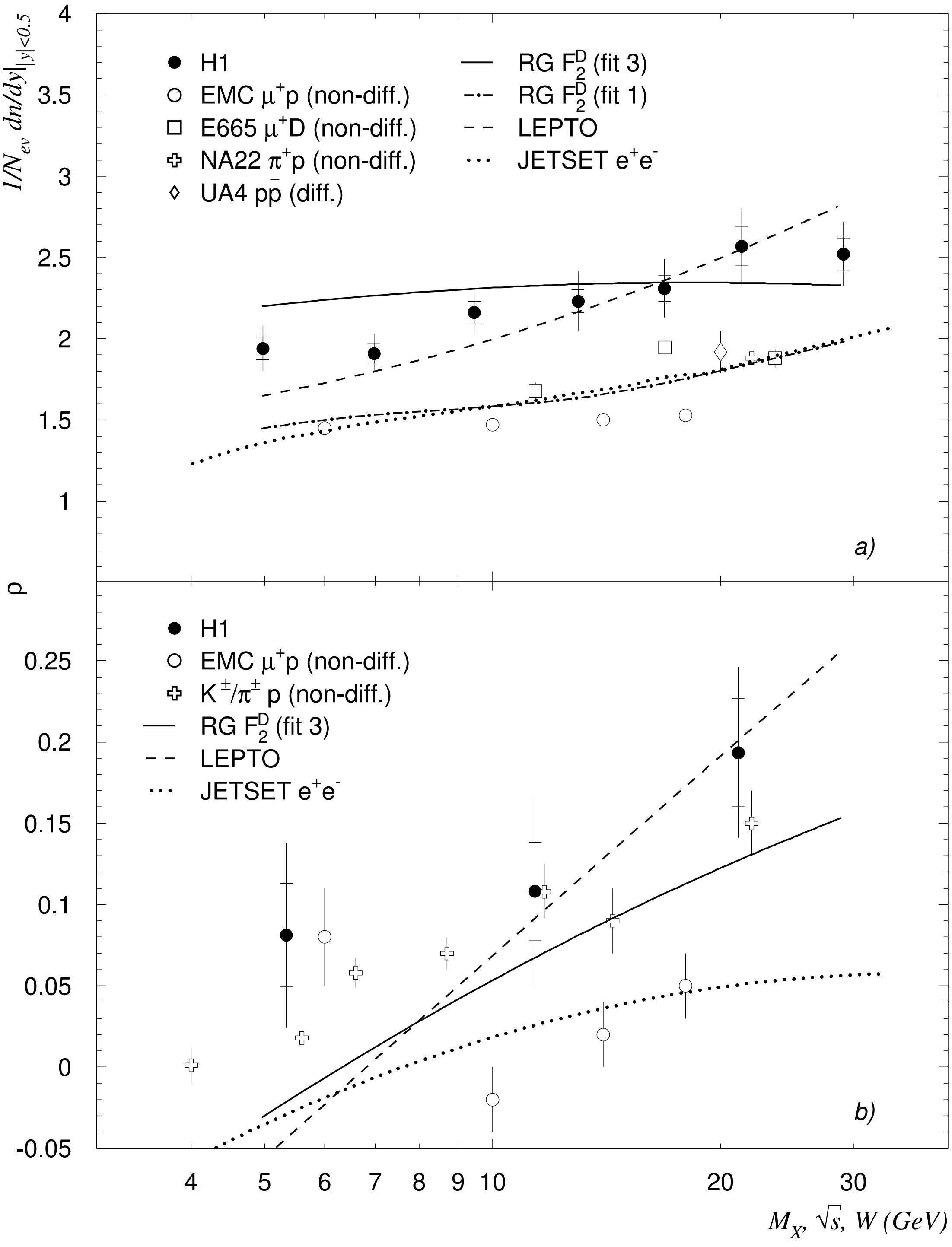


Figure 5.8 : Central rapidity density and forward-backward correlation in diffractive DIS

(a) The central region (-0.5 < y < 0.5) charged particle density is plotted as a function of M_X; (b) The parameter ρ (H1) and b (others), which reflects the correlation between the number of particles in the forward and backward hemispheres, is plotted as a function of M_X (H1 and hadron diffraction), $W = \langle M_X \rangle$ (μp) and $\sqrt{s}$ (non-diffractive hadron-hadron collisions and JETSET e^+e^-). Also shown are predictions of several DIS Monte Carlo models (see text).

$$\langle n_F \rangle = a + b \cdot n_B. \tag{5.14}$$

For reasons of statistics, matrix techniques as used in other works [60–62, 88, 144] to unfold the two-dimensional forward and backward multiplicity distributions have not been employed. Instead, the forward-backward correlation parameter is estimated from the separately unfolded and corrected multiplicity distributions in full phase space, in the forward and in the backward hemispheres. Exploiting the relation between the dispersion for the full phase space (D) and that for the forward and backward hemispheres (D_F and D_B), one can define the correlation parameter ρ as:

$$\rho = \frac{D^2 - D_F^2 - D_B^2}{2 D_F D_B}. \tag{5.15}$$

The parameter ρ is identical to the slope b in eqn. 5.14 in the case of forward-backward symmetric systems [47].

Figure 5.8b shows the parameter ρ in three intervals of M_X for the LRG data. Also shown are data on the parameter b for μp collisions [91], for non-diffractive $\pi^{\pm}/K^{\pm}p$ collisions compiled in [47] and JETSET predictions for e^+e^-. Forward-backward correlations have not previously been measured in diffractively produced final states. In spite of the large errors, there is clear evidence for stronger correlation in LRG events than observed in e^+e^- annihilation [88] and in the μp data for energies above $\gtrsim$ 10 GeV. At lower energy, phase space effects are important and mask possible differences in dynamics. The correlation strength in diffractive DIS is comparable to that in meson-proton interactions.

At LEP, where a value of $b \sim 0.1$ is measured, OPAL finds that the small correlation observed in an inclusive sample of e^+e^- events is primarily due to the superposition of events with distinct number of jets and, therefore, different average charged multiplicity. Sub-classes of n-jet events ($n \geq 2$) show no or even negative correlations [60–62, 144]. In νp and $\overline{\nu}p$ reactions [104] no clear evidence for correlations is observed. These data therefore show that forward-backward correlations are small at energies where the production mechanism is believed to be dominated by single-string $q\overline{q}$ or quark-diquark fragmentation.

In contrast, abundant data from hadron-hadron interactions, compiled in [47], which cover the range $10 \leq \sqrt{s} \leq 900$ GeV, show that the correlation increases logarithmically with energy, with b as large as 0.65 ± 0.01 at $\sqrt{s} = 900$ GeV. The strength and energy dependence of the effect is attributed to strong event-to-event fluctuations of the particle density as occur e.g. in the multi-string Dual Parton Model due to fluctuations in the number of strings overlapping in phase space [143].

The observation of forward-backward correlations in LRG events with a strength comparable to that in soft hadron interactions adds further support to the view that the inclusive sample of DIS LRG events is a mixture of states with distinct hadronisation properties. To disentangle their precise nature and relative contribution, more differential studies will be needed, however.

In present models for diffractive DIS, distinct production processes are readily identified and related, either to differences in parton composition and absorption probability of virtual photon Fock-states, or to quark- and gluon-initiated hard interactions of a colourless exchange. That a mixture of such contributions leads to significant forward-backward correlations is demonstrated by the model predictions for RAPGAP-F_2^D (fit 3) (solid line) and LEPTO (dashed) which are close to the H1 data for $M_X \gtrsim 10$ GeV. The large difference between these DIS models and JETSET for e^+e^- illustrates the sensitivity of this correlation measure to differences in the dynamics of these two processes.

5.3 Summary

The charged particle multiplicity structure of large rapidity gap events of the type $\gamma^* p \to XY$ in deep inelastic scattering at HERA has been measured. The major fraction of these events is generally interpreted as due to diffractive dissociation of the virtual photon on the proton, $\gamma^* p \to Xp$.

Multiplicity distributions, lower-order moments, rapidity spectra and correlations between hadrons emitted in opposite hemispheres in the rest frame of the system X have been presented as a function of the invariant mass M_X.

The data have been compared with e^+e^- annihilation (at $\sqrt{s} = M_X$), lepton-nucleon data in a W range comparable to the M_X range in the H1 data, with hadro-produced diffractive final states, and also with data from hadron-hadron non-diffractive collisions at $\sqrt{s} \sim M_X$. The main observations are the following:

- The mean total charged particle multiplicity $\langle n \rangle$ is a function of M_X and increases proportional to $\log^2 M_X$. The inclusive rapidity spectrum is forward-backward symmetric in the rest-frame of X. A plateau develops with increasing M_X. Both $\langle n \rangle$ (for $M_X \gtrsim 10$ GeV) and the particle density near $y = 0$ are larger than in DIS at comparable values of W, than in e^+e^- annihilation at $\sqrt{s} = M_X$ and than in hadro-produced diffractive final states. The central region density is also higher than in non-diffractive collisions at $\sqrt{s} = M_X$.

- For $M_X \gtrsim 10$ GeV, particle number fluctuations are larger than in e^+e^- annihilation and than in the current fragmentation region of lepton-nucleon interactions at comparable values of $\sqrt{s}$ and W, respectively. Also the forward-backward multiplicity correlations are larger and of comparable strength as those measured in hadron interactions at $\sqrt{s} = M_X$.

The distinctive characteristics of large rapidity gap events mentioned can be globally understood if it is assumed that the photon dissociation mechanism involves a mixture of different partonic states wherein gluons play an increasingly important role as M_X increases. A large contribution from gluon-rich states is also required to explain the size and the steep increase, at fixed Q^2, of the diffractive as well as the total virtual photon-proton cross section at very small Bjorken-x [136, 145].

Good agreement with the data is achieved by a model which assumes that the diffractive process is initiated by the interaction of a point-like virtual photon with a gluon-dominated colour-singlet object emitted from the proton, as is suggested by a perturbative QCD-Regge analysis based on DGLAP evolution of the diffractive structure function. Possibly significant deviations are seen in the large-n tail of the multiplicity distribution at high M_X.

A model with soft colour interactions which rearrange the colour topology after a normal DIS scattering also describes the data although the multiplicity fluctuations are somewhat overestimated for M_X larger than about 20 GeV.

The present analysis adds new support for the conclusion, derived from studies of event shapes [124, 141] and from a study of energy flow and single particle momentum spectra [140] in large rapidity gap events in H1, that gluons play a prominent role in deep inelastic diffraction.

6

Conclusion

The results presented in this work are based on a sample of deep inelastic collisions of the type $ep \rightarrow eX$. This sample has been collected with the H1 detector at the HERA collider in 1994 and comprises about 50 000 collision events, corresponding to an integrated luminosity of 1.3 pb^{-1}. About 10% of the events exhibit a large rapidity gap in the final state, adjacent to the proton beam direction. These are mainly attributed to diffractive photon dissociation and are analysed separately.

The purpose of this study was to learn more about how the hadronic final state in deep inelastic ep scattering is formed. Several theoretical approaches exist, which result in a variety of Monte Carlo models and analytical predictions. However, most conclusions were drawn from data-to-data comparisons, so that they are not biassed by current theoretical views. The hadronic final state was studied in terms of the multiplicity structure. The (charged) particle multiplicity is one of the basic observables characterising the final state and its fluctuations are the direct result of underlying dynamical processes. Therefore the charged particle multiplicity distribution and its moments have been reconstructed in various domains of phase space.

Non-diffractive deep inelastic scattering

The multiplicity distribution of the final state in non-diffractive deep inelastic scattering was analysed in the hadronic centre-of-mass system. In that frame, the total available energy is W.

The evolution with W of the lower moments of the multiplicity distribution in the current hemisphere shows some noteworthy similarities with single hemisphere data in e^+e^- annihilations, indicating that the universality of the quark fragmentation process is approximately valid. The average multiplicity exhibits a faster-than-logarithmic increase with centre-of-mass energy. This suggests that Feynman-scaling is violated and that the rapidity plateau slowly rises with energy. This can be understood in QCD where additional gluon radiation leads to a faster increase of the multiplicity, dampened however by soft-gluon interference.

Analytical predictions for the parton multiplicity from perturbative quantum chromodynamics, which have already proven their usefulness for e^+e^- annihilations, were confirmed and tested at higher centre-of-mass energies than presently available at LEP. Correlations between particles, e.g. measured by the second order factorial moment R_2, are however still sensitive to higher order QCD corrections and/or non-perturbative hadronisation effects.

In the central rapidity region, the average multiplicity grows more slowly with energy than in e^+e^- annihilation and is more compatible with the evolution of the mean multiplicity in hadron-hadron collisions. This provides evidence that the proton remnant, an extended coloured object, effectively suppresses gluon radiation. The correlation between particles in the central rapidity region are smaller in DIS than in non-diffractive hadron-hadron collisions where the presence of two quark-diquark strings (as modelled in the Dual Parton Model) results in larger multiplicity fluctuations.

Diffractive deep inelastic scattering

In the study of diffractive photon dissociation in deep inelastic scattering, the emphasis shifted from the parton cascade and hadronisation phase to the hard subprocess. The question now was, which Fock-states of the virtual photon play an important role, or, equivalently in the proton infinite momentum frame, what the parton content of the pomeron is. The multiplicity structure was studied in the rest frame of the dissociated photon, i.e. the $\gamma^* I\!P$ centre-of-mass system, where the total available energy is M_X.

The main observations are that, although the evolution with energy of the *total* charged multiplicity in diffractive DIS is very similar to e^+e^- annihilation and meson diffraction, the particle density near $y = 0$ is much larger in diffractive DIS than in all other processes, including non-diffractive hadron-hadron interactions. Moreover, at high masses of the photon dissociation system, the multiplicity fluctuations are also larger in diffractive DIS than in e^+e^- annihilation but comparable to non-diffractive hadron-hadron interactions. This can be seen from the dispersion and the second order factorial moment, but also from the forward-backward correlation.

These observations can be understood in a picture where multiple strings cause long-range correlations between final state particles and where these string configurations include octet-octet dipoles which result in a larger central particle density due to the octet colour charge at the dipole endpoints.

One may thus conclude that higher Fock-states and/or a gluon dominated pomeron play a significant role in deep inelastic diffraction.

Appendix A

Unfolding experimental distributions

An algorithm for the unfolding of measured distributions using information from a detector simulation program is proposed*. The technique is based on an iterative method where successive approximations of the underlying distribution are smeared for detector effects and compared to the measured data. Propagation of statistical errors is ensured by a sampling technique which randomly varies the input distributions to the algorithm to reflect their error.

A.1 Introduction

The aim of many physics analyses is to estimate a distribution $f(x)$ of some property x of a physical object in an interval $[a, b]$. Unfortunately, this task is often complicated by the limited acceptance, efficiency and resolution of the detector apparatus and by the limited number of available observations. Instead of directly observing the variable x, one actually measures a variable y which is statistically related to x. The relation between the measured distribution $g(y)$ and the underlying distribution $f(x)$ after an infinite number of observations is given by the convolution of $f(x)$ with a detector response function $\rho(x, y)$:

$$g(y) = \int_a^b \rho(x, y) f(x) \, \mathrm{d}x. \tag{A.1}$$

$\rho(x, y)$ is the probability that an attempt to observe a physical object with underlying property x† results in the measurement of y.

Taking a measured distribution $g^{data}(y)$ as an estimate (in the real world $g^{data}(y)$ is not equal to $g(y)$ because of the limited number of observations available) for $g(y)$, one can get an estimate $f^{data}(x)$ for $f(x)$ by solving the integral Eq. (A.1). However , in typical applications the function $\rho(x, y)$ is ill-conditioned in such a way that, due to statistical fluctuations, straightforward procedures are highly unstable. To overcome this problem, unfolding techniques exist which control such instabilities (see e.g. [147]).

In this appendix an unfolding method is proposed that consists of an iterative algorithm where successive approximations $f^k(x)$ are fed into Eq. (A.1). The resulting $g^k(y)$

*This iterative matrix unfolding method is also detailed in [146].

†In the following this phrase is shortened to "an attempt to measure x".

are compared with the actually observed $g^{data}(y)$ and the iteration is repeated until a satisfactory description of the observed data is obtained. The method relies on generated events that underwent full detector simulation but is independent of the actual generator model used, if acceptance problems can be neglected. Although the method is applied here to the case of charged particle distributions [79, 123], it is suitable for a much wider range of problems. A comparison of this and other unfolding techniques can be found in [148].

A.2 Definitions and notations

In the case of binned distributions, the functions $f(x)$ and $g(y)$ are replaced by the finite arrays f_i and g_j with $i = 1, 2, \ldots n$ and $j = 1, 2, \ldots m$. Also the response function $\rho(x, y)$ has its discrete analogon ρ_{ij}. Equation (A.1) now reads:

$$g_j = \sum_{i=1}^{n} \rho_{ij} f_i. \tag{A.2}$$

The sums $\sum_j \rho_{ij}$ are not necessarily equal to 1, since it can happen that an attempt to observe x will not result in any measurement of y at all. Therefore the matrix element ρ_{ij} is factorised in two parts:

$$\rho_{ij} = \mu_{ij} \epsilon_i. \tag{A.3}$$

The efficiency ϵ_i is the probability that an attempt to observe x in bin i will result in any measurement of y. The migration matrix μ_{ij} is the probability that a successful attempt to observe x in bin i will result in the measurement of y in bin j. Now, the equation $\sum_j \mu_{ij} = 1$ holds. One can also introduce the distribution h_i, which is the number of attempts to observe x in bin i that lead to any measurement of y:

$$h_i = \epsilon_i f_i. \tag{A.4}$$

The probabilities μ_{ij} and ϵ_i have to be estimated from Monte Carlo simulated events. These estimates are calculated from the number of Monte Carlo 'observations' m_{ij}^{mc} of x in bin i where y has been 'measured' in bin j, from the number of Monte Carlo observations h_i^{mc} of x in bin i that resulted in any measurement of y and from the total number of Monte Carlo observations f_i^{mc} of x in bin i, using the formulae:

$$\mu_{ij}^{mc} = \frac{m_{ij}^{mc}}{h_i^{mc}} \qquad \text{and} \qquad \epsilon_i^{mc} = \frac{h_i^{mc}}{f_i^{mc}}. \tag{A.5}$$

Finally, the probability $\tilde{\mu}_{ij}^{mc}$ for a measurement of y in bin j, to have originated from an observation of x in bin i can be estimated for Monte Carlo events in the following way:

$$\tilde{\mu}_{ij}^{mc} = \frac{m_{ij}^{mc}}{g_j^{mc}}. \tag{A.6}$$

Note however, that this probability $\tilde{\mu}_{ij}^{mc}$ is highly dependent on the generator model, since obviously the probability for a measurement y to have originated from an observation of x depends on how many observations of x there are.

The set of Eqs. (A.2) can be both over- and under-constrained, depending on the actual values of n and m. In principle it can happen that no satisfactory (in view of the statistical errors) solution f_i^{data} can be found for a given ϵ_i^{mc}, μ_{ij}^{mc} and g_j^{data}. This means that the Monte Carlo detector simulation must be inadequate, since obviously the observed data g_j^{data} are correct and an underlying distribution f_i^{data} must exist. It can also happen that more than one solution exists. Then unfolding, i.e. inferring an estimate for f_i from a measurement g_j^{data} using the knowledge of ρ_{ij}^{mc}, typically corresponds to selecting one solution out of many for an under-constrained problem. In such a case one should use techniques like the 'Method of Maximum Entropy' [149] or the 'Method of Reduced Entropy' [150] where an algorithm is derived by exploiting consistency conditions that must be satisfied for reproducible experiments.

The algorithm proposed in this note assumes that $m \geq n$, which means that any solution for the underlying distribution f_i is unique. This can always be achieved by decreasing the number of bins n for the underlying distribution.

A.3 The unfolding algorithm

Although the probabilities μ_{ij}^{mc} and ϵ_i^{mc} are model-independent as well as can be expected, $\tilde{\mu}_{ij}^{mc}$ is not. Straightforward calculation of f_i^{data} is therefore not advisable. One can however calculate a first approximation of h_i^{data} using $\tilde{\mu}_{ij}^{mc}$ starting from the measured g_j^{data}:

$$h_i^{data(1)} = \sum_{j=1}^{m} \tilde{\mu}_{ij}^{mc} g_j^{data}. \tag{A.7}$$

This first approximation can then be used to weight the number of Monte Carlo observations m_{ij}^{mc}:

$$m_{ij}^{mc(1)} = m_{ij}^{mc} \frac{h_i^{data(1)}}{h_i^{mc}}. \tag{A.8}$$

The weighted number of Monte Carlo observations leading to the measurement of y in bin j is now given by:

$$g_j^{mc(1)} = \sum_{i=1}^{n} m_{ij}^{mc(1)}, \tag{A.9}$$

and the weighted probability for a measurement of y in bin j to have originated from an observation of x in bin i is:

$$\tilde{\mu}_{ij}^{mc(1)} = \frac{m_{ij}^{mc(1)}}{g_j^{mc(1)}}. \tag{A.10}$$

The probability $\tilde{\mu}_{ij}^{mc(1)}$ can now be used to calculate a second approximation $h_i^{data(2)}$. The k^{th} iteration is given by:

$$h_i^{data(k)} = \sum_{j=1}^{m} \tilde{\mu}_{ij}^{mc(k-1)} g_j^{data}, \tag{A.11}$$

$$m_{ij}^{mc(k)} = m_{ij}^{mc} \frac{h_i^{data(k)}}{h_i^{mc}}, \tag{A.12}$$

$$g_j^{mc(k)} = \sum_{i=1}^{n} m_{ij}^{mc(k)}, \tag{A.13}$$

$$\tilde{\mu}_{ij}^{mc(k)} = \frac{m_{ij}^{mc(k)}}{g_j^{mc(k)}}. \tag{A.14}$$

This iteration should be repeated until $g_j^{mc(k)}$ gives a satisfactory description of the measured data distribution g_j^{data}. This can be evaluated by calculating e.g.:

$$\chi^2 = \sum_{j=1}^{m} \frac{(g_j^{mc(k)} - g_j^{data})^2}{\sigma^2(g_j^{mc(k)}) + \sigma^2(g_j^{data})}. \tag{A.15}$$

$\sigma(g_j^{mc(k)})$ and $\sigma(g_j^{data})$ are the statistical errors on $g_j^{mc(k)}$ and g_j^{data}, respectively. In practice one could approximate $\sigma(g_j^{mc(k)})$ by using $\sigma(g_j^{mc})$ to save computing time.

Finally a correction for observations of x which did not lead to any measurement of y has to be applied:

$$f_i^{data} = \frac{h_i^{data(k+1)}}{\epsilon_i^{mc}}. \tag{A.16}$$

A.3.1 Model dependence

In spite of the fact that the Monte Carlo generator distributions are weighted, one should be very careful regarding the model dependence of the algorithm.

Especially when the correction includes an extrapolation outside the acceptance region of the detector, no unfolding method can be expected to be completely model independent. One could think of the situation where x is an integral over phase space and y is simply the amount of x that falls inside the acceptance region, with an otherwise perfect detector efficiency. The matrix $\mu_{ij}^{mc(k)}$ will correct for this loss but the weighting procedure will not change the fraction of x that falls outside the acceptance region. The resulting unfolded distribution f_i^{data} will therefore still depend on the amount of x that is produced outside the acceptance region by the Monte Carlo model.

Of course, the unfolding method crucially depends on the detector simulation. This simulation has to describe the detector efficiency in every detail since the weighting procedure will only correct wrong generator models, not a faulty simulation.

A.4 Propagation of errors

The straightforward propagation of statistical errors is a non-trivial task since both the measured data and the generated and simulated Monte Carlo distributions are estimated from a finite number of events. Moreover, because several measured data points are used to calculate a single unfolded data point, the final results will be correlated. Here, a sampling technique is proposed which randomly varies the input distributions to the algorithm to reflect their error.

When quoting an error for a measured data point one implicitly assumes that when repeating the experiment many times the distribution of results will have a mean equal to the initial value and a dispersion equal to the error. Instead of propagating the errors in the conventional way, one can simulate the effect of the statistical fluctuations by sampling all the input data according to an appropriate distribution, repeating the unfolding procedure and calculating the spread of the final results.

In the case of non-weighted counting distributions one should use the Poisson distribution, which is given by:

$$P_r = \frac{\lambda^r}{r!} e^{-\lambda}, \tag{A.17}$$

where λ is the actual measurement and P_r is the probability that when repeating the experiment one would get r as a result. After repeating the sampling N times, one ends up with a set $(f_i^{data})_k$ of unfolded results. The error should be calculated using:

$$\sigma^2(f_i^{data}) = \frac{1}{N-1} \sum_{k=1}^{N} \left((f_i^{data})_k - f_i^{data} \right)^2 . \tag{A.18}$$

The co-variance between unfolded results is given by:

$$\mathrm{cov}(f_i^{data}, f_j^{data}) = \frac{1}{N} \sum_{k=1}^{N} \left((f_i^{data})_k - f_i^{data} \right) \left((f_j^{data})_k - f_j^{data} \right) . \tag{A.19}$$

A.4.1 Weighted events

A complication will arise when using weighted events. The weighted number of events in each bin is now given by:

$$n = \sum_i w_i, \tag{A.20}$$

where the w_i are the event weights.

If one would repeat the experiment many times, the distribution of results would no longer follow a Poissonian. In fact, when weights larger than 1 are used, the distribution will be broader. The error is now given by:

$$\sigma^2 = \sum_i w_i^2 . \tag{A.21}$$

A practical solution is to use a Negative Binomial Distribution (NBD) instead of a Poisson distribution to do the sampling. The NBD is defined as:

$$P_n(k, \overline{n}) = \frac{k(k+1)\ldots(k+n-1)}{n!} \left(\frac{\overline{n}}{\overline{n}+k}\right)^n \left(\frac{k}{\overline{n}+k}\right)^k, \tag{A.22}$$

with parameters k (or $1/k$) and $\overline{n}$. The average $\langle n \rangle$ and the dispersion D of the NBD are related to the two parameters by:

$$\langle n \rangle = \overline{n} \qquad \text{and} \qquad \frac{D^2}{\langle n \rangle^2} = \frac{1}{\langle n \rangle} + \frac{1}{k}. \tag{A.23}$$

These parameters have to be calculated from each entry (and its error) of the weighted distribution.

If however the number of observations is sufficiently large, both the Poisson and the Negative Binomial distributions converge towards a Gaussian.

A.5 Example: Unfolding multiplicity distributions

In the following the case of charged particle multiplicity distributions is considered. The raw multiplicity distribution in a given kinematical region needs to be corrected for several effects. These include loss of events and particles due to limited geometrical acceptance and resolution of the tracking system, limited track finding efficiency, contamination by tracks from particle decays and interactions in the material of the detector which are assigned to the primary vertex, and also QED initial state radiation which affects the event kinematics.

Correction factors are obtained from Monte Carlo simulation by comparing the 'true' generated distributions before the detector simulation with the 'observed' distributions after this simulation followed by the same reconstruction, selection and analysis as the real data. The Monte Carlo generator used in this analysis is based on ARIADNE [40].

Figure A.1a shows a typical efficiency array ϵ_i. One can see that the detection efficiency drops drastically, as can be expected, for low multiplicity events. The bin $i = 0$ in this plot contains events with no tracks inside the pseudorapidity range $1 < \eta^* < 5$ in hadronic centre-of-mass frame, which does *not* mean that there are no tracks detected elsewhere in the detector. Figure A.1b shows the migration matrix μ_{ij} which in turn is used to calculate the particle reconstruction efficiency as shown in

Fig. A.1c and given by:

$$\sum_j \frac{j}{i} \mu_{ij}. \tag{A.24}$$

The large value for events with just one particle in the final state is artificial since events with no tracks detected in the Central Tracker are lost during the event selection. The overall efficiency for finding a genuine primary track is better than 95%. Depending on the phase space region considered, the correction that has to be applied will however become somewhat larger due to the geometrical acceptance of the Central Tracker.

One possibility to evaluate the performance of the algorithm described above is to look whether the successive Monte Carlo approximations $g_j^{mc(k)}$ converge to the observed

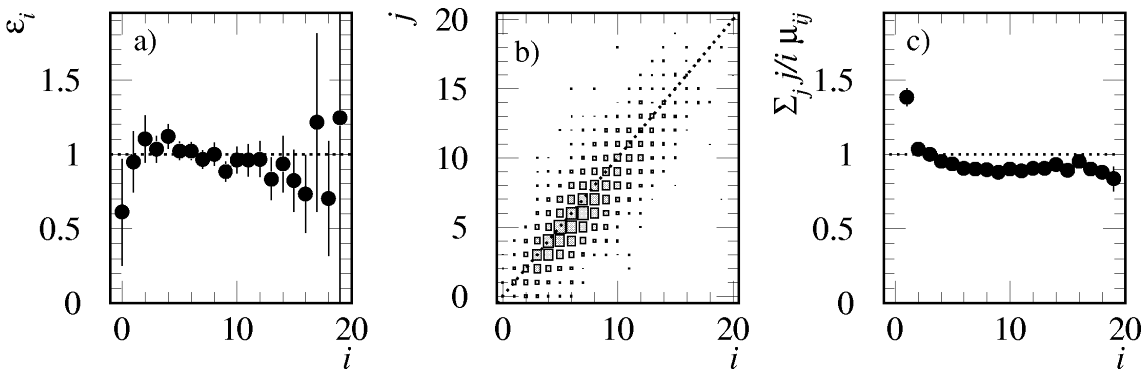


Figure A.1 : Efficiency array and migration matrix used in the iterative unfolding algorithm
a) and b) Efficiency array ϵ_i and migration matrix μ_{ij} used for the unfolding of the multiplicity distribution for particles with $1 < \eta^* < 5$ and events with $115 < W < 150$ GeV.
c) Particle reconstruction efficiency (including acceptance losses) versus the generated number of particles for the same kinematical region.

data g_j^{data}. This is shown in Fig. A.2. One should bear in mind that, if the Monte Carlo simulation of the detector is not perfect, no exact agreement with the data can be obtained.

Another cross-check of the method was done by taking the 'data' produced by the LEPTO 6.3 Monte Carlo generator [39] as input to the algorithm. Again the information from the generator based on ARIADNE was used to do the unfolding. Figure A.3 shows the unfolded distributions compared to the true underlying ones. Despite the substantial differences between both generators, the result is in good agreement with the underlying LEPTO distribution.

A.6 Conclusion

The unfolding method proposed here performs rather well in the case of particle multiplicity distributions. Convergence is reached rapidly and a cross-check using different Monte Carlo generators shows that the unfolded result is in good agreement with the true underlying distribution. The method of sampling is an easy, general and practical way to do the error propagation.

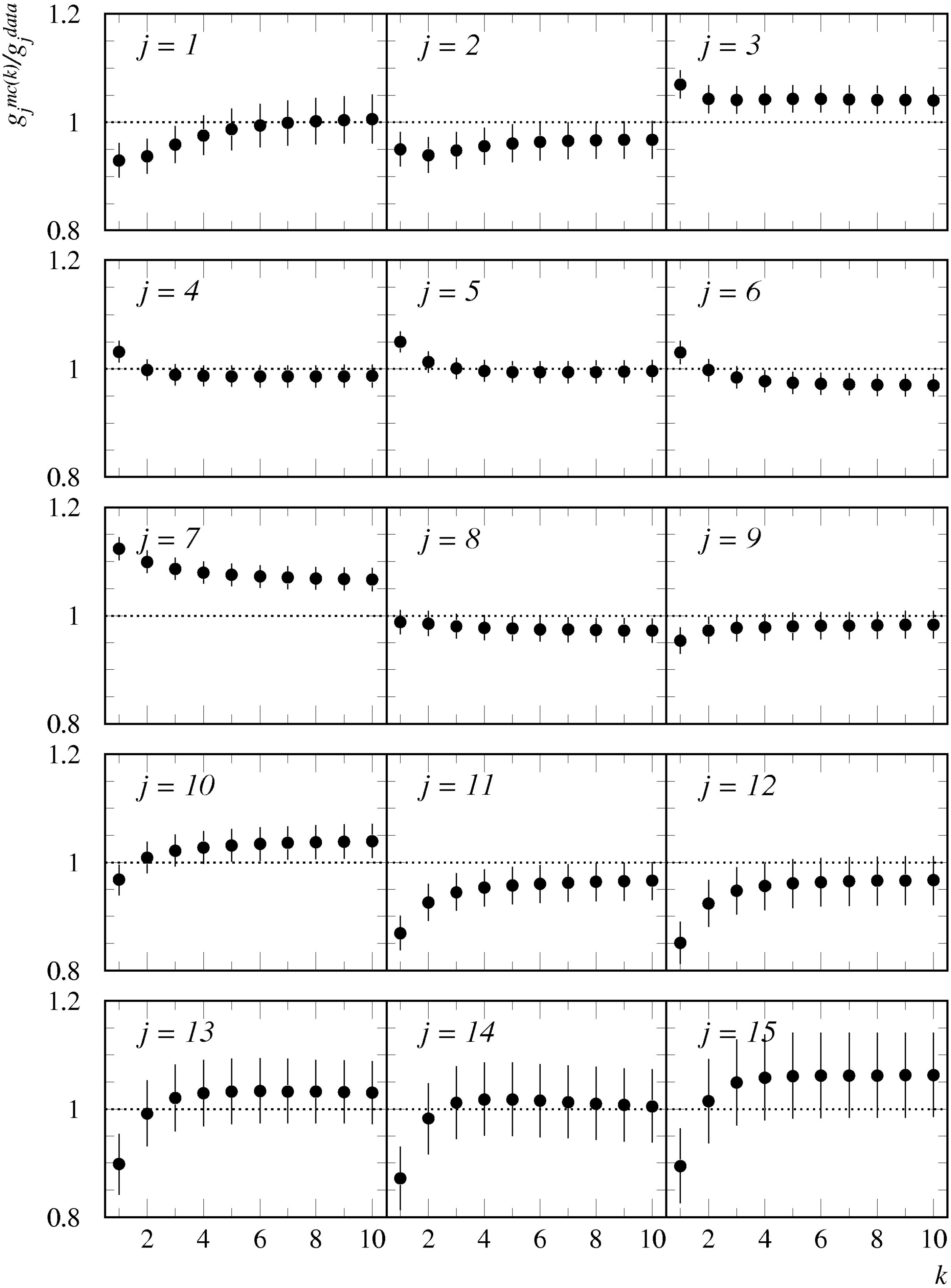


Figure A.2 : Convergence of the iterative unfolding method
Successive Monte Carlo approximations $g_j^{mc(k)}$, normalised to the observed data g_j^{data}, are shown for events with different number of observed particles (j). Results from the same kinematical region as in Fig. A.1 are used.

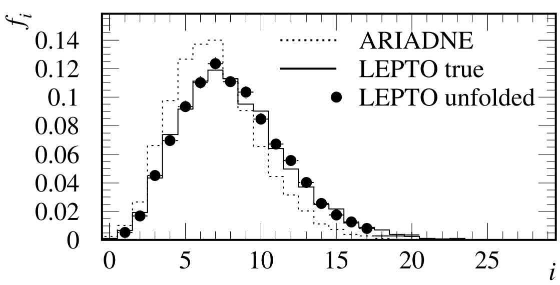


Figure A.3 : Test of the iterative unfolding method using two different Monte Carlo generators

The comparison is shown between the underlying distributions of ARIADNE and LEPTO (histograms) and the result when LEPTO 'data' are unfolded using ARIADNE for the correction (dots). The distributions were generated for the same kinematical region as in Fig. A.1.

Appendix B

Tables

All results discussed in Chs. 4 and 5 are listed here for easy reference. They can also be retrieved from the HEPDATA database (http://durpdg.dur.ac.uk/HEPDATA).

When two errors are listed, the first is the statistical error and the second is the systematical uncertainty. If only one error is given, it is the total error (quadratic sum of statistical and systematical errors).

All results are fully corrected for detector acceptance and efficiency, migration in the kinematical plane and in (pseudo-) rapidity. Decay products of K^0_S, Λ and $\overline{\Lambda}$ and from weakly decaying particles with a lifetime longer than $8 \cdot 10^{-9}$ s are excluded.

In non-diffractive DIS the multiplicity structure has been measured for the hadronic final state (excluding the scattered electron), while in diffractive DIS only the photon dissociation system X is considered, which is separated from the (excited) proton system Y by the largest rapidity gap in the event.

The kinematical bins used in the analyses of the multiplicity structure are listed in Tabs. 4.3 and 5.1.

80 GeV < W < 115 GeV

n	P_n (%) $1 < \eta^* < 2$	$1 < \eta^* < 3$	$1 < \eta^* < 4$	$1 < \eta^* < 5$
0	15.79 ± 0.52 ± 2.18	3.26 ± 0.22 ± 1.58	0.78 ± 0.09 ± 1.66	0.40* ± 0.08
1	22.55 ± 0.50 ± 1.92	8.01 ± 0.27 ± 1.29	2.73 ± 0.18 ± 0.94	1.64 ± 0.16 ± 0.39
2	20.62 ± 0.46 ± 1.24	12.84 ± 0.35 ± 1.30	5.73 ± 0.28 ± 1.45	3.82 ± 0.22 ± 1.27
3	15.96 ± 0.43 ± 2.29	14.03 ± 0.35 ± 1.21	9.55 ± 0.29 ± 1.46	7.59 ± 0.28 ± 1.06
4	10.21 ± 0.37 ± 1.06	13.38 ± 0.32 ± 2.22	11.11 ± 0.32 ± 1.21	10.08 ± 0.34 ± 1.18
5	6.07 ± 0.30 ± 1.54	11.73 ± 0.35 ± 0.46	12.76 ± 0.35 ± 2.22	12.30 ± 0.34 ± 0.89
6	3.85 ± 0.26 ± 0.48	9.84 ± 0.34 ± 0.71	12.19 ± 0.32 ± 1.28	12.91 ± 0.35 ± 1.29
7	2.24 ± 0.16 ± 0.62	7.94 ± 0.33 ± 0.52	11.17 ± 0.37 ± 0.96	12.27 ± 0.35 ± 1.86
8	1.15 ± 0.12 ± 0.42	5.91 ± 0.32 ± 1.20	9.22 ± 0.31 ± 0.48	10.65 ± 0.33 ± 0.86
9	0.68 ± 0.10 ± 0.31	4.00 ± 0.21 ± 0.82	7.72 ± 0.33 ± 1.37	8.69 ± 0.30 ± 1.08
10	0.49 ± 0.12 ± 0.24	3.21 ± 0.19 ± 0.56	5.69 ± 0.24 ± 1.31	6.38 ± 0.26 ± 0.99
11	0.19 ± 0.06 ± 0.26	2.22 ± 0.18 ± 1.06	3.95 ± 0.22 ± 1.00	4.79 ± 0.25 ± 1.23
12	0.07 ± 0.04 ± 0.13	1.49 ± 0.14 ± 0.66	2.76 ± 0.23 ± 0.96	3.20 ± 0.24 ± 0.97
13	0.05 ± 0.03 ± 0.06	0.90 ± 0.15 ± 0.53	1.79 ± 0.18 ± 0.80	2.05 ± 0.19 ± 0.69
14		0.54 ± 0.11 ± 0.53	1.22 ± 0.15 ± 0.62	1.38 ± 0.16 ± 0.74
15		0.36 ± 0.08 ± 0.57	0.67 ± 0.12 ± 0.68	0.86 ± 0.14 ± 0.66
16		0.14 ± 0.07 ± 0.19	0.40 ± 0.11 ± 0.43	0.40 ± 0.08 ± 0.24
17		0.08 ± 0.05 ± 0.10	0.24 ± 0.11 ± 0.33	0.27 ± 0.13 ± 0.37
18			0.13 ± 0.08 ± 0.19	0.10 ± 0.05 ± 0.11

115 GeV < W < 150 GeV

n	P_n (%) $1 < \eta^* < 2$	$1 < \eta^* < 3$	$1 < \eta^* < 4$	$1 < \eta^* < 5$
0	16.80 ± 0.62 ± 0.55	4.52 ± 0.35 ± 1.19	1.09 ± 0.18 ± 0.77	0.35* ± 0.08
1	21.55 ± 0.54 ± 1.05	8.00 ± 0.35 ± 1.21	2.57 ± 0.19 ± 1.12	1.17 ± 0.13 ± 0.62
2	19.69 ± 0.50 ± 0.53	12.15 ± 0.36 ± 0.70	4.72 ± 0.24 ± 0.70	2.58 ± 0.14 ± 1.26
3	15.19 ± 0.48 ± 0.77	12.96 ± 0.36 ± 0.91	8.09 ± 0.29 ± 0.74	5.63 ± 0.23 ± 0.69
4	10.54 ± 0.36 ± 0.31	12.74 ± 0.38 ± 0.74	9.96 ± 0.33 ± 0.95	7.97 ± 0.24 ± 1.03
5	6.76 ± 0.34 ± 0.75	11.09 ± 0.32 ± 1.63	11.23 ± 0.36 ± 0.95	10.36 ± 0.31 ± 0.82
6	4.17 ± 0.27 ± 0.50	9.35 ± 0.31 ± 0.70	11.27 ± 0.33 ± 1.37	11.44 ± 0.37 ± 0.56
7	2.55 ± 0.20 ± 0.76	7.31 ± 0.27 ± 2.06	10.32 ± 0.31 ± 0.25	11.96 ± 0.35 ± 1.18
8	1.29 ± 0.14 ± 0.48	6.01 ± 0.29 ± 0.93	9.21 ± 0.34 ± 0.52	10.41 ± 0.32 ± 0.66
9	0.64 ± 0.10 ± 0.62	4.53 ± 0.24 ± 1.38	8.03 ± 0.30 ± 0.29	9.65 ± 0.36 ± 0.74
10	0.31 ± 0.07 ± 0.14	3.76 ± 0.25 ± 0.53	6.51 ± 0.29 ± 0.87	7.88 ± 0.28 ± 0.49
11	0.20 ± 0.04 ± 0.07	2.57 ± 0.23 ± 0.86	5.32 ± 0.26 ± 0.88	6.15 ± 0.23 ± 0.80
12	0.12 ± 0.03 ± 0.15	1.90 ± 0.18 ± 0.77	3.84 ± 0.23 ± 0.88	4.96 ± 0.27 ± 1.35
13	0.07 ± 0.03 ± 0.07	1.05 ± 0.16 ± 0.29	2.96 ± 0.23 ± 0.53	3.51 ± 0.25 ± 0.50
14	0.03 ± 0.02 ± 0.04	0.77 ± 0.12 ± 0.54	1.81 ± 0.18 ± 0.65	2.15 ± 0.21 ± 0.45
15		0.56 ± 0.10 ± 0.48	1.25 ± 0.16 ± 0.33	1.47 ± 0.17 ± 0.43
16		0.31 ± 0.08 ± 0.24	0.76 ± 0.12 ± 0.62	0.93 ± 0.14 ± 0.92
17		0.18 ± 0.06 ± 0.28	0.46 ± 0.10 ± 0.29	0.72 ± 0.17 ± 0.62
18		0.12 ± 0.06 ± 0.12	0.25 ± 0.09 ± 0.20	0.25 ± 0.08 ± 0.34
19			0.14 ± 0.06 ± 0.23	0.17 ± 0.07 ± 0.24
20			0.08 ± 0.05 ± 0.11	0.12 ± 0.08 ± 0.13
21				0.03 ± 0.02 ± 0.03

Table B.1 : Charged particle multiplicity distributions of the hadronic final state in non-diffractive DIS

The fully corrected multiplicity distribution P_n (%) is listed for different regions of the hadronic centre-of-mass energy, W, and centre-of-mass pseudorapidity, η^*.

* *The value of P_0 in the domain $1 < \eta^* < 5$ is not measured but taken from the DJANGO 6.0 Monte Carlo generator.*

150 GeV $< W <$ 185 GeV

n	P_n (%) $1 < \eta^* < 2$	$1 < \eta^* < 3$	$1 < \eta^* < 4$	$1 < \eta^* < 5$
0	16.01 ± 0.58 ± 5.43	4.24 ± 0.36 ± 3.19	1.13 ± 0.15 ± 1.51	0.28* ± 0.08
1	19.56 ± 0.63 ± 1.10	8.32 ± 0.41 ± 1.39	2.56 ± 0.20 ± 1.55	1.09 ± 0.13 ± 1.03
2	19.73 ± 0.54 ± 1.58	10.99 ± 0.40 ± 0.67	4.69 ± 0.24 ± 2.38	2.21 ± 0.14 ± 0.71
3	16.01 ± 0.60 ± 1.37	11.75 ± 0.37 ± 1.30	6.81 ± 0.30 ± 1.00	4.18 ± 0.20 ± 1.28
4	11.19 ± 0.50 ± 1.20	11.71 ± 0.41 ± 1.27	8.77 ± 0.35 ± 1.02	6.61 ± 0.23 ± 2.08
5	7.21 ± 0.34 ± 0.70	11.37 ± 0.41 ± 1.49	9.87 ± 0.38 ± 1.56	8.24 ± 0.31 ± 1.36
6	4.11 ± 0.30 ± 0.81	9.39 ± 0.39 ± 1.08	10.47 ± 0.34 ± 0.32	9.77 ± 0.34 ± 1.20
7	2.80 ± 0.25 ± 0.93	7.98 ± 0.35 ± 0.57	9.68 ± 0.33 ± 0.94	11.04 ± 0.33 ± 0.89
8	1.40 ± 0.17 ± 0.81	6.27 ± 0.38 ± 0.52	9.10 ± 0.34 ± 0.57	10.87 ± 0.34 ± 0.66
9	0.91 ± 0.15 ± 0.48	5.36 ± 0.35 ± 1.07	8.10 ± 0.34 ± 1.05	9.89 ± 0.35 ± 1.13
10	0.33 ± 0.08 ± 0.23	3.64 ± 0.25 ± 1.03	6.90 ± 0.31 ± 0.81	8.18 ± 0.35 ± 0.44
11	0.31 ± 0.07 ± 0.18	2.87 ± 0.22 ± 0.28	5.76 ± 0.30 ± 0.72	7.07 ± 0.31 ± 0.56
12	0.10 ± 0.04 ± 0.10	1.82 ± 0.20 ± 0.73	4.92 ± 0.28 ± 1.17	5.76 ± 0.34 ± 1.15
13	0.13 ± 0.07 ± 0.18	1.24 ± 0.15 ± 0.54	3.59 ± 0.25 ± 1.32	4.88 ± 0.31 ± 1.50
14	0.08 ± 0.06 ± 0.14	1.08 ± 0.16 ± 0.89	2.71 ± 0.24 ± 0.97	3.42 ± 0.26 ± 0.89
15		0.78 ± 0.15 ± 0.61	1.59 ± 0.16 ± 0.72	2.43 ± 0.25 ± 1.00
16		0.43 ± 0.09 ± 0.40	1.11 ± 0.12 ± 0.93	1.31 ± 0.17 ± 0.78
17		0.20 ± 0.06 ± 0.13	0.86 ± 0.13 ± 0.46	1.07 ± 0.19 ± 0.64
18		0.18 ± 0.07 ± 0.13	0.49 ± 0.13 ± 0.32	0.55 ± 0.14 ± 0.51
19		0.13 ± 0.07 ± 0.24	0.31 ± 0.11 ± 0.63	0.43 ± 0.13 ± 0.50
20		0.06 ± 0.05 ± 0.10	0.09 ± 0.05 ± 0.14	0.12 ± 0.05 ± 0.51
21			0.17 ± 0.09 ± 0.30	0.18 ± 0.08 ± 0.32
22			0.13 ± 0.09 ± 0.26	0.16 ± 0.11 ± 0.31

185 GeV $< W <$ 220 GeV

n	P_n (%) $1 < \eta^* < 2$	$1 < \eta^* < 3$	$1 < \eta^* < 4$	$1 < \eta^* < 5$
0	15.76 ± 0.68 ± 3.11	4.24 ± 0.42 ± 3.18	1.07 ± 0.24 ± 2.96	0.27* ± 0.09
1	19.66 ± 0.76 ± 0.78	8.00 ± 0.45 ± 0.48	2.81 ± 0.30 ± 1.27	0.96 ± 0.13 ± 1.96
2	19.56 ± 0.62 ± 1.75	10.74 ± 0.57 ± 1.52	5.22 ± 0.32 ± 0.73	2.37 ± 0.23 ± 0.68
3	15.76 ± 0.61 ± 2.03	11.51 ± 0.47 ± 0.69	6.70 ± 0.33 ± 0.50	3.73 ± 0.24 ± 1.12
4	11.08 ± 0.58 ± 1.85	11.76 ± 0.48 ± 2.08	7.98 ± 0.37 ± 0.87	5.47 ± 0.27 ± 0.99
5	6.95 ± 0.45 ± 0.86	11.65 ± 0.53 ± 1.27	9.18 ± 0.34 ± 0.58	7.50 ± 0.30 ± 0.83
6	4.63 ± 0.38 ± 0.97	9.74 ± 0.47 ± 1.62	9.07 ± 0.40 ± 0.87	8.49 ± 0.35 ± 1.24
7	3.20 ± 0.32 ± 1.14	7.96 ± 0.43 ± 0.91	10.18 ± 0.50 ± 0.63	10.08 ± 0.39 ± 1.61
8	1.36 ± 0.16 ± 0.74	6.83 ± 0.44 ± 1.06	9.10 ± 0.41 ± 0.68	10.54 ± 0.46 ± 1.18
9	0.95 ± 0.15 ± 0.79	5.12 ± 0.34 ± 1.16	7.95 ± 0.34 ± 1.57	10.40 ± 0.43 ± 1.71
10	0.47 ± 0.12 ± 0.31	3.70 ± 0.30 ± 1.24	7.50 ± 0.34 ± 1.83	9.20 ± 0.44 ± 1.00
11	0.33 ± 0.10 ± 0.50	2.52 ± 0.24 ± 0.88	6.10 ± 0.38 ± 0.78	8.03 ± 0.40 ± 1.12
12	0.11 ± 0.04 ± 0.09	1.79 ± 0.19 ± 0.76	4.98 ± 0.33 ± 1.30	6.48 ± 0.30 ± 0.99
13	0.05 ± 0.04 ± 0.10	1.39 ± 0.17 ± 0.92	3.46 ± 0.26 ± 0.77	4.78 ± 0.29 ± 0.95
14	0.03 ± 0.02 ± 0.05	1.11 ± 0.16 ± 1.23	2.60 ± 0.22 ± 0.75	3.81 ± 0.27 ± 1.22
15		0.66 ± 0.13 ± 0.33	2.03 ± 0.21 ± 0.58	2.32 ± 0.21 ± 0.78
16		0.44 ± 0.10 ± 0.27	1.29 ± 0.16 ± 0.80	2.13 ± 0.22 ± 1.08
17		0.29 ± 0.10 ± 0.19	1.00 ± 0.16 ± 0.65	1.14 ± 0.14 ± 0.67
18		0.14 ± 0.06 ± 0.19	0.71 ± 0.15 ± 0.22	0.83 ± 0.12 ± 0.44
19		0.10 ± 0.05 ± 0.14	0.39 ± 0.09 ± 0.35	0.64 ± 0.14 ± 0.29
20		0.06 ± 0.04 ± 0.29	0.21 ± 0.07 ± 0.18	0.30 ± 0.07 ± 0.39
21		0.05 ± 0.04 ± 0.10	0.17 ± 0.10 ± 0.26	0.13 ± 0.07 ± 0.18
22			0.11 ± 0.04 ± 0.13	0.19 ± 0.08 ± 0.26
23				0.03 ± 0.02 ± 0.05

Table B.1: (continued)

$\eta^* > 0$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	8.98 ± 0.07 ± 0.61	10.00 ± 0.07 ± 0.66	10.88 ± 0.09 ± 0.72	11.35 ± 0.09 ± 0.67
D_2	3.89 ± 0.07 ± 0.23	4.16 ± 0.06 ± 0.26	4.49 ± 0.08 ± 0.36	4.56 ± 0.08 ± 0.30
D_3	3.43 ± 0.16 ± 0.22	3.41 ± 0.15 ± 0.21	3.61 ± 0.26 ± 0.55	3.67 ± 0.25 ± 0.45
D_4	5.39 ± 0.15 ± 0.38	5.58 ± 0.13 ± 0.23	6.05 ± 0.21 ± 0.71	6.26 ± 0.19 ± 0.49
C_2	1.18 ± 0.01 ± 0.04	1.17 ± 0.01 ± 0.01	1.17 ± 0.01 ± 0.02	1.16 ± 0.01 ± 0.02
C_3	1.62 ± 0.02 ± 0.14	1.56 ± 0.01 ± 0.06	1.54 ± 0.02 ± 0.08	1.51 ± 0.02 ± 0.07
C_4	2.48 ± 0.07 ± 0.40	2.29 ± 0.05 ± 0.17	2.26 ± 0.07 ± 0.21	2.19 ± 0.06 ± 0.19
R_2	1.07 ± 0.01 ± 0.03	1.07 ± 0.01 ± 0.01	1.07 ± 0.01 ± 0.02	1.07 ± 0.01 ± 0.02
R_3	1.24 ± 0.02 ± 0.10	1.22 ± 0.01 ± 0.04	1.24 ± 0.02 ± 0.07	1.22 ± 0.02 ± 0.06

$1 < \eta^* < 2$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	2.46 ± 0.03 ± 0.10	2.50 ± 0.03 ± 0.12	2.63 ± 0.04 ± 0.18	2.66 ± 0.04 ± 0.18
D_2	2.07 ± 0.05 ± 0.08	2.13 ± 0.03 ± 0.10	2.22 ± 0.05 ± 0.12	2.22 ± 0.03 ± 0.14
D_3	2.21 ± 0.11 ± 0.10	2.24 ± 0.05 ± 0.14	2.39 ± 0.11 ± 0.18	2.31 ± 0.07 ± 0.12
D_4	3.08 ± 0.16 ± 0.15	3.15 ± 0.07 ± 0.17	3.39 ± 0.16 ± 0.26	3.24 ± 0.09 ± 0.12
C_2	1.71 ± 0.02 ± 0.02	1.72 ± 0.02 ± 0.04	1.71 ± 0.02 ± 0.05	1.69 ± 0.02 ± 0.03
C_3	3.87 ± 0.16 ± 0.18	3.90 ± 0.11 ± 0.24	3.89 ± 0.17 ± 0.18	3.75 ± 0.12 ± 0.20
C_4	10.67 ± 0.93 ± 1.23	10.77 ± 0.60 ± 1.17	11.03 ± 1.09 ± 0.63	10.02 ± 0.65 ± 1.22
R_2	1.30 ± 0.02 ± 0.03	1.32 ± 0.01 ± 0.04	1.33 ± 0.02 ± 0.04	1.32 ± 0.02 ± 0.03
R_3	2.11 ± 0.14 ± 0.15	2.15 ± 0.08 ± 0.19	2.22 ± 0.14 ± 0.14	2.11 ± 0.09 ± 0.13
K_3	0.187 ± 0.070 ± 0.097	0.167 ± 0.040 ± 0.063	0.226 ± 0.079 ± 0.085	0.147 ± 0.043 ± 0.101

$1 < \eta^* < 3$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	4.90 ± 0.04 ± 0.18	5.06 ± 0.05 ± 0.27	5.32 ± 0.07 ± 0.34	5.35 ± 0.06 ± 0.35
D_2	3.10 ± 0.04 ± 0.13	3.37 ± 0.04 ± 0.21	3.57 ± 0.06 ± 0.26	3.58 ± 0.07 ± 0.30
D_3	2.91 ± 0.10 ± 0.25	3.18 ± 0.09 ± 0.21	3.45 ± 0.12 ± 0.43	3.54 ± 0.20 ± 0.34
D_4	4.27 ± 0.11 ± 0.28	4.62 ± 0.10 ± 0.27	5.01 ± 0.13 ± 0.45	5.18 ± 0.28 ± 0.45
C_2	1.40 ± 0.01 ± 0.03	1.44 ± 0.01 ± 0.02	1.45 ± 0.01 ± 0.02	1.44 ± 0.01 ± 0.02
C_3	2.41 ± 0.04 ± 0.16	2.57 ± 0.05 ± 0.10	2.62 ± 0.07 ± 0.14	2.62 ± 0.09 ± 0.09
C_4	4.83 ± 0.19 ± 0.64	5.34 ± 0.20 ± 0.38	5.57 ± 0.28 ± 0.68	5.71 ± 0.47 ± 0.34
R_2	1.19 ± 0.01 ± 0.03	1.24 ± 0.01 ± 0.02	1.26 ± 0.01 ± 0.02	1.26 ± 0.01 ± 0.03
R_3	1.64 ± 0.04 ± 0.13	1.80 ± 0.04 ± 0.09	1.87 ± 0.06 ± 0.13	1.88 ± 0.08 ± 0.11
K_3	0.048 ± 0.015 ± 0.044	0.063 ± 0.016 ± 0.033	0.088 ± 0.022 ± 0.077	0.108 ± 0.041 ± 0.036

$1 < \eta^* < 4$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	6.45 ± 0.06 ± 0.33	7.00 ± 0.05 ± 0.35	7.51 ± 0.06 ± 0.51	7.66 ± 0.07 ± 0.47
D_2	3.23 ± 0.09 ± 0.14	3.57 ± 0.04 ± 0.23	3.93 ± 0.06 ± 0.25	4.06 ± 0.06 ± 0.20
D_3	2.71 ± 0.28 ± 0.21	2.89 ± 0.11 ± 0.23	3.28 ± 0.14 ± 0.50	3.32 ± 0.16 ± 0.28
D_4	4.34 ± 0.27 ± 0.24	4.72 ± 0.10 ± 0.20	5.29 ± 0.14 ± 0.44	5.41 ± 0.14 ± 0.22
C_2	1.25 ± 0.01 ± 0.03	1.26 ± 0.01 ± 0.02	1.27 ± 0.01 ± 0.02	1.28 ± 0.01 ± 0.02
C_3	1.82 ± 0.05 ± 0.13	1.85 ± 0.02 ± 0.07	1.90 ± 0.03 ± 0.08	1.92 ± 0.03 ± 0.08
C_4	3.01 ± 0.20 ± 0.38	3.04 ± 0.08 ± 0.22	3.22 ± 0.11 ± 0.27	3.26 ± 0.12 ± 0.26
R_2	1.09 ± 0.01 ± 0.03	1.11 ± 0.01 ± 0.02	1.14 ± 0.01 ± 0.02	1.15 ± 0.01 ± 0.02
R_3	1.29 ± 0.05 ± 0.09	1.35 ± 0.02 ± 0.06	1.43 ± 0.03 ± 0.07	1.45 ± 0.03 ± 0.06

Table B.2 : Moments of the charged particle multiplicity distribution of the hadronic final state in non-diffractive DIS

Moments and cumulants of the unfolded multiplicity distribution are listed for different ranges in hadronic centre-of-mass energy, W, and centre-of-mass pseudorapidity, η^*.

$1 < \eta^* < 5$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	6.90 ± 0.05 ± 0.34	7.72 ± 0.05 ± 0.42	8.45 ± 0.06 ± 0.58	8.81 ± 0.08 ± 0.54
D_2	3.15 ± 0.05 ± 0.21	3.45 ± 0.04 ± 0.25	3.77 ± 0.06 ± 0.25	3.86 ± 0.05 ± 0.22
D_3	2.56 ± 0.12 ± 0.33	2.74 ± 0.13 ± 0.19	3.02 ± 0.18 ± 0.54	2.95 ± 0.19 ± 0.41
D_4	4.23 ± 0.11 ± 0.37	4.59 ± 0.10 ± 0.29	5.08 ± 0.15 ± 0.45	5.17 ± 0.14 ± 0.29
C_2	1.20 ± 0.01 ± 0.03	1.19 ± 0.01 ± 0.02	1.19 ± 0.01 ± 0.02	1.19 ± 0.01 ± 0.02
C_3	1.67 ± 0.02 ± 0.12	1.64 ± 0.02 ± 0.06	1.64 ± 0.02 ± 0.08	1.61 ± 0.02 ± 0.08
C_4	2.60 ± 0.07 ± 0.35	2.50 ± 0.06 ± 0.16	2.50 ± 0.08 ± 0.24	2.42 ± 0.07 ± 0.21
R_2	1.06 ± 0.01 ± 0.02	1.07 ± 0.01 ± 0.02	1.08 ± 0.01 ± 0.02	1.07 ± 0.01 ± 0.02
R_3	1.19 ± 0.02 ± 0.09	1.21 ± 0.01 ± 0.05	1.24 ± 0.02 ± 0.07	1.23 ± 0.02 ± 0.06

$2 < \eta^* < 3$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	2.46 ± 0.03 ± 0.17	2.61 ± 0.03 ± 0.20	2.73 ± 0.04 ± 0.27	2.71 ± 0.04 ± 0.20
D_2	1.92 ± 0.03 ± 0.19	2.11 ± 0.03 ± 0.16	2.24 ± 0.04 ± 0.20	2.22 ± 0.03 ± 0.18
D_3	1.86 ± 0.05 ± 0.10	2.07 ± 0.06 ± 0.14	2.23 ± 0.08 ± 0.13	2.28 ± 0.07 ± 0.22
D_4	2.67 ± 0.06 ± 0.12	2.94 ± 0.08 ± 0.21	3.14 ± 0.09 ± 0.18	3.20 ± 0.09 ± 0.25
C_2	1.61 ± 0.01 ± 0.03	1.65 ± 0.01 ± 0.03	1.67 ± 0.01 ± 0.05	1.67 ± 0.02 ± 0.06
C_3	3.26 ± 0.08 ± 0.16	3.47 ± 0.10 ± 0.15	3.57 ± 0.10 ± 0.25	3.60 ± 0.11 ± 0.30
C_4	7.77 ± 0.35 ± 0.79	8.57 ± 0.48 ± 0.60	8.99 ± 0.52 ± 1.20	9.30 ± 0.56 ± 1.19
R_2	1.20 ± 0.01 ± 0.02	1.27 ± 0.01 ± 0.04	1.30 ± 0.01 ± 0.05	1.30 ± 0.02 ± 0.06
R_3	1.63 ± 0.05 ± 0.10	1.86 ± 0.08 ± 0.13	1.99 ± 0.08 ± 0.14	2.02 ± 0.09 ± 0.25
K_3	0.018 ± 0.021 ± 0.041	0.043 ± 0.030 ± 0.035	0.072 ± 0.040 ± 0.062	0.121 ± 0.038 ± 0.076

$3 < \eta^* < 4$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	1.54 ± 0.02 ± 0.18	1.93 ± 0.02 ± 0.17	2.17 ± 0.02 ± 0.18	2.29 ± 0.03 ± 0.17
D_2	1.32 ± 0.01 ± 0.14	1.48 ± 0.02 ± 0.14	1.59 ± 0.02 ± 0.09	1.67 ± 0.03 ± 0.10
D_3	1.23 ± 0.03 ± 0.07	1.33 ± 0.03 ± 0.09	1.39 ± 0.05 ± 0.04	1.54 ± 0.05 ± 0.16
D_4	1.79 ± 0.03 ± 0.11	2.01 ± 0.04 ± 0.10	2.14 ± 0.06 ± 0.06	2.31 ± 0.06 ± 0.17
C_2	1.73 ± 0.01 ± 0.11	1.58 ± 0.01 ± 0.10	1.53 ± 0.01 ± 0.06	1.53 ± 0.01 ± 0.10
C_3	3.70 ± 0.08 ± 0.55	3.08 ± 0.07 ± 0.47	2.87 ± 0.07 ± 0.26	2.91 ± 0.08 ± 0.46
C_4	9.24 ± 0.38 ± 2.29	6.99 ± 0.28 ± 1.72	6.21 ± 0.31 ± 1.00	6.47 ± 0.34 ± 1.70
R_2	1.08 ± 0.01 ± 0.04	1.07 ± 0.01 ± 0.07	1.07 ± 0.01 ± 0.03	1.09 ± 0.01 ± 0.08
R_3	1.17 ± 0.05 ± 0.15	1.16 ± 0.04 ± 0.19	1.18 ± 0.05 ± 0.08	1.28 ± 0.05 ± 0.24
K_3	0.079 ± 0.019 ± 0.062	0.045 ± 0.014 ± 0.040	0.054 ± 0.018 ± 0.033	0.011 ± 0.017 ± 0.068

$4 < \eta^* < 5$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\langle n \rangle$	0.46 ± 0.01 ± 0.04	0.71 ± 0.01 ± 0.04	0.93 ± 0.01 ± 0.09	1.17 ± 0.01 ± 0.09
D_2	0.71 ± 0.01 ± 0.02	0.88 ± 0.01 ± 0.08	0.99 ± 0.01 ± 0.06	1.09 ± 0.01 ± 0.19
D_3	0.83 ± 0.02 ± 0.02	0.93 ± 0.01 ± 0.03	0.98 ± 0.02 ± 0.05	1.01 ± 0.02 ± 0.01
D_4	1.09 ± 0.03 ± 0.04	1.26 ± 0.02 ± 0.03	1.37 ± 0.03 ± 0.07	1.47 ± 0.02 ± 0.03
C_2	3.34 ± 0.06 ± 0.11	2.52 ± 0.04 ± 0.18	2.13 ± 0.02 ± 0.09	1.86 ± 0.01 ± 0.11
C_3	13.78 ± 0.61 ± 1.06	7.85 ± 0.30 ± 1.18	5.58 ± 0.17 ± 0.59	4.23 ± 0.10 ± 0.55
C_4	68.80 ± 5.70 ± 9.55	28.93 ± 1.85 ± 6.83	17.21 ± 0.87 ± 3.47	11.24 ± 0.46 ± 2.25
R_2	1.19 ± 0.04 ± 0.06	1.12 ± 0.03 ± 0.10	1.05 ± 0.02 ± 0.05	1.01 ± 0.01 ± 0.05
R_3	1.41 ± 0.24 ± 0.29	1.16 ± 0.09 ± 0.23	1.02 ± 0.07 ± 0.19	0.92 ± 0.04 ± 0.09
K_3	0.159 ± 0.144 ± 0.138	0.221 ± 0.041 ± 0.136	0.150 ± 0.028 ± 0.104	0.109 ± 0.017 ± 0.061

Table B.2: (continued)

$1 < \eta^* < 2$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\overline{n}$	2.52 ± 0.10	2.58 ± 0.13	2.72 ± 0.15	2.75 ± 0.19
$1/k$	0.285 ± 0.080	0.288 ± 0.056	0.269 ± 0.034	0.270 ± 0.064
χ^2/NDF	29.4/13	23.0/14	33.0/15	27.1/14
m	2.44 ± 0.16	2.49 ± 0.16	2.66 ± 0.20	2.67 ± 0.13
d	0.791 ± 0.092	0.812 ± 0.077	0.743 ± 0.094	0.768 ± 0.122
c	2.6 ± 3.4	3.1 ± 1.2	1.9 ± 1.7	2.5 ± 7.5
χ^2/NDF	21.8/12	8.9/13	24.9/14	19.8/13

$1 < \eta^* < 3$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\overline{n}$	4.85 ± 0.16	5.12 ± 0.26	5.36 ± 0.25	5.41 ± 0.20
$1/k$	0.161 ± 0.039	0.235 ± 0.017	0.256 ± 0.028	0.245 ± 0.034
χ^2/NDF	121.6/19	54.5/19	61.5/21	44.7/24
m	4.79 ± 0.24	5.05 ± 0.30	5.25 ± 0.20	5.32 ± 0.17
d	0.634 ± 0.017	0.644 ± 0.080	0.662 ± 0.074	0.641 ± 0.060
c	4.5 ± 2.2	2.0 ± 6.5	2.1 ± 5.5	1.8 ± 6.3
χ^2/NDF	71.1/18	42.7/18	43.6/20	31.1/23

$1 < \eta^* < 4$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\overline{n}$	6.41 ± 0.28	7.00 ± 0.39	7.52 ± 0.46	7.71 ± 0.30
$1/k$	0.085 ± 0.018	0.102 ± 0.014	0.132 ± 0.026	0.142 ± 0.020
χ^2/NDF	96.8/21	112.4/21	88.6/25	44.7/24
m	6.40 ± 0.33	6.98 ± 0.37	7.45 ± 0.41	7.59 ± 0.28
d	0.475 ± 0.009	0.484 ± 0.011	0.505 ± 0.023	0.527 ± 0.022
c	2.7 ± 1.3	2.7 ± 2.5	2.8 ± 7.1	4.5 ± 5.9
χ^2/NDF	62.6/20	69.6/20	43.8/24	32.8/23

$1 < \eta^* < 5$

W	80 – 115 GeV	115 – 150 GeV	150 – 185 GeV	185 – 220 GeV
$\overline{n}$	6.90 ± 0.33	7.73 ± 0.41	8.44 ± 0.49	8.88 ± 0.34
$1/k$	0.067 ± 0.011	0.068 ± 0.012	0.074 ± 0.013	0.069 ± 0.020
χ^2/NDF	44.7/19	57.0/21	80.8/24	63.6/25
m	6.89 ± 0.29	7.72 ± 0.41	8.40 ± 0.48	8.78 ± 0.33
d	0.433 ± 0.010	0.422 ± 0.007	0.422 ± 0.015	0.422 ± 0.020
c	1.7 ± 1.5	1.8 ± 2.0	2.4 ± 5.1	4.1 ± 6.1
χ^2/NDF	41.6/18	52.8/20	58.4/23	27.7/24

Table B.3 : Results of Negative Binomial and Lognormal fits to the charged particle multiplicity distribution of the hadronic final state in non-diffractive DIS
The best-fit parameters of the Negative Binomial Distribution ($\overline{n}$ and $1/k$) and the Lognormal Distribution (m, d, and c) are listed together with their resulting χ^2/NDF, for different ranges in hadronic centre-of-mass energy, W, and centre-of-mass pseudorapidity, η^*.

3 GeV $< M_X <$ 8 GeV

n	P_n(%) forward	backward	positives	negatives
0	5.87 ± 0.48 ± 1.42	6.38 ± 0.53 ± 2.42	1.00 ± 0.10 ± 0.35	0.93 ± 0.09 ± 0.40
1	20.96 ± 1.08 ± 4.24	21.63 ± 1.07 ± 4.63	17.14 ± 0.94 ± 4.34	15.92 ± 1.09 ± 4.01
2	28.85 ± 1.02 ± 5.44	29.13 ± 1.07 ± 5.74	38.88 ± 1.45 ± 7.45	39.40 ± 1.72 ± 7.37
3	23.59 ± 0.97 ± 4.11	23.63 ± 0.88 ± 4.01	29.53 ± 1.25 ± 4.85	29.76 ± 1.11 ± 4.97
4	12.86 ± 0.72 ± 2.33	12.87 ± 0.62 ± 1.96	10.74 ± 0.67 ± 2.56	11.29 ± 0.66 ± 2.38
5	5.32 ± 0.36 ± 1.47	4.52 ± 0.30 ± 1.32	2.46 ± 0.26 ± 1.13	2.51 ± 0.29 ± 1.21
6	1.89 ± 0.18 ± 0.83	1.29 ± 0.15 ± 0.62	0.21 ± 0.04 ± 0.43	0.16 ± 0.03 ± 0.46
7	0.48 ± 0.07 ± 0.36	0.40 ± 0.11 ± 0.22		
8	0.13 ± 0.03 ± 0.17	0.10 ± 0.04 ± 0.05		

8 GeV $< M_X <$ 15 GeV

n	P_n(%) forward	backward	positives	negatives
0	1.60 ± 0.13 ± 0.70	1.17 ± 0.08 ± 0.36	0.03 ± 0.01 ± 0.05	0.03 ± 0.01 ± 0.05
1	8.17 ± 0.58 ± 2.41	8.12 ± 0.48 ± 2.05	2.49 ± 0.22 ± 0.91	2.43 ± 0.25 ± 1.01
2	14.55 ± 0.76 ± 3.19	16.17 ± 0.65 ± 4.02	14.04 ± 0.64 ± 3.52	14.34 ± 0.79 ± 3.69
3	23.24 ± 1.03 ± 4.20	22.27 ± 0.73 ± 5.04	27.54 ± 0.99 ± 5.79	27.82 ± 1.06 ± 5.72
4	19.93 ± 0.82 ± 3.42	20.85 ± 0.76 ± 4.20	27.91 ± 0.89 ± 5.35	27.63 ± 0.89 ± 5.26
5	14.45 ± 0.66 ± 2.48	14.79 ± 0.63 ± 2.87	16.90 ± 0.71 ± 3.09	16.65 ± 0.65 ± 2.99
6	9.28 ± 0.48 ± 2.00	8.93 ± 0.47 ± 1.68	7.40 ± 0.38 ± 2.12	7.33 ± 0.39 ± 1.95
7	4.69 ± 0.26 ± 1.38	4.29 ± 0.28 ± 1.31	2.66 ± 0.19 ± 1.35	2.64 ± 0.21 ± 1.31
8	2.55 ± 0.21 ± 0.80	2.28 ± 0.24 ± 0.91	0.85 ± 0.10 ± 0.57	0.95 ± 0.13 ± 0.52
9	0.90 ± 0.10 ± 0.72	0.64 ± 0.08 ± 0.66	0.09 ± 0.02 ± 0.28	0.09 ± 0.02 ± 0.34
10	0.46 ± 0.09 ± 0.31	0.33 ± 0.08 ± 0.32		
11	0.08 ± 0.01 ± 0.31	0.09 ± 0.06 ± 0.13		

15 GeV $< M_X <$ 30 GeV

n	P_n(%) forward	backward	positives	negatives
0	0.42 ± 0.04 ± 0.24	0.59 ± 0.04 ± 0.19	0.00 ± 0.01 ± 0.01	0.00 ± 0.01 ± 0.01
1	3.30 ± 0.39 ± 1.16	3.30 ± 0.23 ± 1.01	0.71 ± 0.08 ± 0.33	0.53 ± 0.08 ± 0.19
2	6.85 ± 0.54 ± 2.09	7.55 ± 0.35 ± 2.37	4.29 ± 0.37 ± 1.63	3.69 ± 0.35 ± 1.20
3	12.82 ± 0.76 ± 3.22	13.51 ± 0.50 ± 4.36	12.16 ± 0.63 ± 3.88	12.40 ± 0.72 ± 3.83
4	16.29 ± 0.81 ± 3.94	15.64 ± 0.54 ± 3.67	19.09 ± 0.75 ± 5.08	19.77 ± 0.78 ± 4.92
5	17.61 ± 0.74 ± 4.20	15.80 ± 0.53 ± 3.56	21.24 ± 0.83 ± 4.76	22.44 ± 0.82 ± 4.86
6	14.97 ± 0.67 ± 2.84	13.91 ± 0.51 ± 2.70	17.25 ± 0.69 ± 3.73	17.64 ± 0.69 ± 3.45
7	10.85 ± 0.52 ± 2.10	11.25 ± 0.45 ± 2.28	12.61 ± 0.59 ± 2.21	11.78 ± 0.46 ± 2.30
8	7.41 ± 0.38 ± 1.30	8.26 ± 0.38 ± 1.52	7.19 ± 0.38 ± 2.18	6.56 ± 0.36 ± 1.89
9	4.50 ± 0.25 ± 1.05	4.71 ± 0.25 ± 1.78	3.26 ± 0.23 ± 1.33	3.16 ± 0.23 ± 1.41
10	2.48 ± 0.16 ± 0.64	2.96 ± 0.23 ± 1.07	1.47 ± 0.16 ± 0.79	1.42 ± 0.16 ± 0.64
11	1.21 ± 0.13 ± 0.53	1.13 ± 0.08 ± 1.05	0.42 ± 0.06 ± 0.54	0.35 ± 0.05 ± 0.61
12	0.73 ± 0.11 ± 0.33	0.67 ± 0.10 ± 0.58	0.20 ± 0.07 ± 0.15	0.15 ± 0.03 ± 0.10
13	0.34 ± 0.08 ± 0.19	0.42 ± 0.14 ± 0.26		
14	0.07 ± 0.02 ± 0.14	0.16 ± 0.07 ± 0.12		
15	0.03 ± 0.01 ± 0.08	0.04 ± 0.03 ± 0.10		

Table B.4 : Charged particle multiplicity distributions of the photon dissociation system X in diffractive DIS

The fully corrected multiplicity distribution P_n (%) is listed for the forward and backward hemispheres, as well as for positive and negative particles.

Full phase space

M_X (GeV)	$\langle n \rangle$	D	R_2
5.0	4.66 ± 0.08 ± 0.23	1.86 ± 0.05 ± 0.14	0.944 ± 0.009 ± 0.022
7.0	5.72 ± 0.09 ± 0.21	2.03 ± 0.05 ± 0.15	0.951 ± 0.006 ± 0.017
9.5	6.87 ± 0.08 ± 0.28	2.43 ± 0.05 ± 0.16	0.979 ± 0.005 ± 0.013
13.0	8.18 ± 0.11 ± 0.39	2.88 ± 0.05 ± 0.19	1.001 ± 0.004 ± 0.011
16.9	9.49 ± 0.16 ± 0.44	3.21 ± 0.08 ± 0.25	1.009 ± 0.006 ± 0.014
21.3	10.85 ± 0.16 ± 0.54	3.47 ± 0.11 ± 0.25	1.010 ± 0.006 ± 0.012
29.1	12.32 ± 0.18 ± 0.72	4.28 ± 0.12 ± 0.27	1.039 ± 0.006 ± 0.016

Forward hemisphere

M_X (GeV)	$\langle n \rangle$	D	R_2
5.0	2.32 ± 0.05 ± 0.12	1.29 ± 0.03 ± 0.14	0.877 ± 0.014 ± 0.058
7.0	3.00 ± 0.07 ± 0.13	1.51 ± 0.03 ± 0.20	0.919 ± 0.011 ± 0.031
9.5	3.52 ± 0.06 ± 0.15	1.76 ± 0.04 ± 0.20	0.966 ± 0.011 ± 0.028
13.0	4.08 ± 0.06 ± 0.19	1.95 ± 0.04 ± 0.17	0.984 ± 0.009 ± 0.036
16.9	4.71 ± 0.10 ± 0.20	2.12 ± 0.05 ± 0.12	0.991 ± 0.010 ± 0.019
21.3	5.36 ± 0.11 ± 0.23	2.22 ± 0.06 ± 0.13	0.985 ± 0.010 ± 0.018
29.1	6.28 ± 0.11 ± 0.32	2.76 ± 0.06 ± 0.17	1.034 ± 0.009 ± 0.022

Backward hemisphere

M_X (GeV)	$\langle n \rangle$	D	R_2
5.0	2.29 ± 0.05 ± 0.13	1.31 ± 0.03 ± 0.18	0.894 ± 0.016 ± 0.038
7.0	2.74 ± 0.05 ± 0.11	1.40 ± 0.03 ± 0.08	0.897 ± 0.013 ± 0.030
9.5	3.33 ± 0.05 ± 0.18	1.60 ± 0.02 ± 0.16	0.930 ± 0.007 ± 0.041
13.0	4.13 ± 0.07 ± 0.25	1.91 ± 0.03 ± 0.23	0.972 ± 0.007 ± 0.027
16.9	4.70 ± 0.09 ± 0.40	2.16 ± 0.04 ± 0.25	0.998 ± 0.008 ± 0.031
21.3	5.53 ± 0.11 ± 0.42	2.39 ± 0.04 ± 0.16	1.006 ± 0.008 ± 0.015
29.1	6.06 ± 0.12 ± 0.47	2.81 ± 0.08 ± 0.32	1.049 ± 0.010 ± 0.020

Table B.5 : Lower moments of the charged particle multiplicity distribution of the photon dissociation system X in diffractive DIS
The average multiplicity $\langle n \rangle$, dispersion D and normalised second factorial moment R_2 are listed for the full phase space, as well as for the forward and backward hemispheres separately.

y	$\mathrm{d}n/\mathrm{d}y$		
	$3 < M_X < 8$ GeV	$8 < M_X < 15$ GeV	$15 < M_X < 30$ GeV
-3.0 – -2.5	0.03 ± 0.01 ± 0.02	0.17 ± 0.07 ± 0.16	0.58 ± 0.69 ± 0.90
-2.5 – -2.0	0.14 ± 0.02 ± 0.02	0.48 ± 0.05 ± 0.12	1.09 ± 0.16 ± 0.63
-2.0 – -1.5	0.41 ± 0.02 ± 0.04	1.08 ± 0.07 ± 0.18	1.38 ± 0.10 ± 0.29
-1.5 – -1.0	0.91 ± 0.04 ± 0.06	1.53 ± 0.06 ± 0.12	2.09 ± 0.07 ± 0.24
-1.0 – -0.5	1.32 ± 0.07 ± 0.07	1.91 ± 0.06 ± 0.12	2.28 ± 0.07 ± 0.14
-0.5 – 0.0	1.84 ± 0.11 ± 0.11	2.17 ± 0.06 ± 0.11	2.43 ± 0.08 ± 0.15
0.0 – 0.5	1.92 ± 0.11 ± 0.11	2.21 ± 0.07 ± 0.15	2.46 ± 0.07 ± 0.18
0.5 – 1.0	1.58 ± 0.11 ± 0.11	2.01 ± 0.05 ± 0.11	2.33 ± 0.07 ± 0.14
1.0 – 1.5	0.84 ± 0.04 ± 0.04	1.64 ± 0.05 ± 0.11	2.06 ± 0.06 ± 0.12
1.5 – 2.0	0.37 ± 0.03 ± 0.03	1.02 ± 0.04 ± 0.07	1.82 ± 0.06 ± 0.10
2.0 – 2.5	0.10 ± 0.03 ± 0.03	0.51 ± 0.03 ± 0.05	1.06 ± 0.04 ± 0.09
2.5 – 3.0	0.04 ± 0.01 ± 0.01	0.14 ± 0.01 ± 0.04	0.50 ± 0.03 ± 0.06

Table B.6 : Charged particle rapidity spectra in diffractive DIS
Charged particle rapidity spectra, measured in the $\gamma I\!P$ centre-of-mass frame are listed for three different intervals of M_X.

M_X (GeV)	$\mathrm{d}n/\mathrm{d}y\|_{y=0}$
4.9	1.94 ± 0.07 ± 0.12
6.9	1.90 ± 0.05 ± 0.10
9.4	2.15 ± 0.07 ± 0.10
12.9	2.23 ± 0.06 ± 0.17
16.8	2.30 ± 0.08 ± 0.16
21.3	2.56 ± 0.11 ± 0.19
29.0	2.50 ± 0.10 ± 0.17

Table B.7 : Particle density in the central rapidity region in diffractive DIS
The particle density in the centre-of-mass rapidity region $-0.5 < y^* < 0.5$ is plotted as a function of the invariant mass of the photon dissociation cluster, M_X.

M_X	3 – 8 GeV	8 – 15 GeV	15 – 30 GeV
ρ	0.081 ± 0.032 ± 0.035	0.108 ± 0.030 ± 0.041	0.194 ± 0.033 ± 0.023

Table B.8 : Forward-backward correlation in diffractive DIS
The correlation between the number of particles in the forward and backward hemispheres, characterised by the parameter ρ, is listed for three different intervals of M_X.

Bibliography

[1] F. Halzen and A. D. Martin, *Quarks & Leptons: An introductory course in modern particle physics*, John Wiley & Sons, New York, 1984;
D. H. Perkins, *Introduction to high energy physics*, Addison-Wesley publishing co., London, 1987.

[2] H1 Collaboration, C. Adloff et al., Phys. Lett. B393 (1997) 492.

[3] H1 Collaboration, S. Aid et al., Nucl. Phys. B470 (1996) 3.

[4] ZEUS Collaboration, J. Breitweg et al., Phys. Lett. B407 (1997) 432.

[5] A. Martin, *Structure functions and small x physics*, Proceedings of the XXIst International Meeting on Fundamental Physics "Physics at HERA", Miraflores de la Sierra, Spain, 1993, Eds. F. Barreiro, L. Hervás and L. Labarga, p. 250.

[6] V. Del Duca, S. J. Brodsky and P. Hoyer, Phys. Rev. D46 (1992) 931.

[7] B. L. Ioffe, Phys. Lett. B30 (1969) 123.

[8] See e.g. J. Bartels et al., *An analysis of diffraction in deep-inelastic scattering*, hep-ph/9803497 and references therein.

[9] H. I. Miettinen and J. Pumplin, Phys. Rev. D18 (1978) 1696.

[10] J. Amaldi, M. Jacob and G. Matthiae, Ann. Rev. Nucl. Sci. 26 (1976) 385.

[11] ZEUS Collaboration, M. Derrick et al., Phys. Lett. B345 (1995) 576.

[12] Particle Data Group, L. Montanet et al., Phys. Rev. D50 (1994) 1173.

[13] Yu. L. Dokshitzer, Sov. Phys. JETP 46 (1997) 641;
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.

[14] Yu. L. Dokshitzer et al., *Basics of Perturbative QCD*, Editions Frontières, Gif-sur-Yvette, 1991.

[15] G. Gustafson and U. Petterson, Nucl. Phys. B306 (1988) 746;
G. Gustafson, Phys. Lett. B175 (1986) 453;
B. Andersson et al., Z. Phys. C43 (1989) 625.

[16] D. Amati and G. Veneziano, Phys. Lett. 83 (1979) 87.

[17] Ya. I. Azimov et al., Z. Phys. C27 (1985) 65; C31 (1986) 213.

[18] P. Dahlqvist, G. Gustafson and B. Andersson, Nucl. Phys. B328 (1989) 76;
G. Gustafson and A. Nilsson, Nucl. Phys. B355 (1991) 106;
G. Gustafson and C. Sjögren, Phys. Lett. B248 (1990) 430.

[19] B. Andersson et al., Z. Phys. C49 (1991) 79.

[20] OPAL Collaboration, M. Z. Akrawy et al., Phys. Lett. B247 (1990) 617.

[21] H1 Collaboration, S. Aid et al., Nucl. Phys. B445 (1995) 3.

[22] ZEUS Collaboration, M. Derrick et al., Z. Phys. C67 (1995) 93.

[23] G. Marchesini and B. R. Webber, Nucl. Phys. B238 (1984) 1; B310 (1988) 461.

[24] G. Marchesini et al., Comp. Phys. Comm. 67 (1992) 465.

[25] B. Andersson et al., Phys. Rep. 97 (1983) 33.

[26] R. P. Feynman and R. D. Field, Phys. Rev. D15 (1977) 2590; Nucl. Phys. B136 (1978) 1.

[27] J. D. Bjorken and J. Kogut, Phys. Rev. D8 (1973) 1341.

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

[29] H1 Collaboration, T. Ahmed et al., Phys. Lett. B429 (1994) 477.

[30] H1 Collaboration, C. Adloff et al., Z. Phys. C74 (1997) 221.

[31] H1 Collaboration, C. Adloff et al., Z. Phys. C76 (1997) 613.

[32] G. Ingelman and P. Schlein, Phys. Lett. B152 (1985) 256:
A. Fritzsch and K. H. Streng, Phys. Lett. B164 (1985) 395;
A. Donnachie and P. V. Landshoff, Phys. Lett. B191 (1987) 309; Nucl. Phys. B303 (1988) 634;
K. Golec-Biernat and J. Kwiecinski, Phys. Lett. B353 (1995) 329;
E. L. Berger et al., Nucl. Phys. 286 (1987) 704.

[33] P. Collins, *An introduction to Regge theory and high energy physics*, Cambridge University Press, Cambridge (1977).

[34] ZEUS Collaboration, J. Breitweg et al., Eur. Phys. J. C1 (1998) 81.

[35] W. Buchmüller and A. Hebecker, Phys. Lett. B355 (1995) 573;
A. Edin, G. Ingelman and J. Rathsman, Phys. Lett. B366 (1996) 371.

[36] T. Sjöstrand, Comp. Phys. Comm. 82 (1994) 74.

[37] DELPHI Collaboration, P. Abreu et al., Z. Phys. C73 (1996) 11.

[38] M. Glück, E. Reya and A. Vogt, Z. Phys. C67 (1995) 433.

[39] G. Ingelman, *LEPTO 6.1 – The Lund Monte Carlo for deep inelastic lepton-nucleon scattering*, Proc. of the Workshop “Physics at HERA”, Hamburg, Germany, 1991, Eds. W. Buchmüller and G. Ingelman, Vol. 3, p. 1366.

[40] L. Lönnblad, Comp. Phys. Comm. 71 (1992) 15.

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

[42] J. F. Owens, Phys. Rev. D30 (1984) 943.

[43] A. Białas and F. Hayot, Phys. Rev. D33 (1986) 39.

[44] E. A. De Wolf, I. M. Dremin and W. Kittel, Phys. Rep. 270 (1996) 1.

[45] A. H. Mueller, Phys. Rev. D4 (1971) 150.

[46] M. G. Kendal and A. Stuart, *The advanced theory of statistics*, Vol. 1, C. Griffen and Co., London 1969.

[47] NA22 Collaboration, V. V. Aivazyan et al., Z. Phys. C42 (1989) 533.

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

[49] Z. Koba, H. B. Nielsen and P. Olesen, Nucl. Phys. B40 (1972) 317.

[50] A. K. Wróblewski, Acta Phys. Pol. B4 (1973) 857;
A. J. Buras et al., Phys. Lett. B47 (1973) 251.

[51] A. I. Golokhvastov, Sov. J. Nucl. Phys. 27 (1978) 430; 30 (1979) 128.

[52] R. Szwed and G. Wrochna, Z. Phys. C29 (1985) 255.

[53] G. J. Alner et al., Phys. Rep. 154 (1987) 247.

[54] UA5 Collaboration, R. E. Ansorge et al., Z. Phys. C43 (1989) 357.

[55] A. M. Polyakov, Sov. Phys. JETP 32 (1971) 296; 33 (1971) 850;
S. J. Orfanidis and V. Rittenberg, Phys. Rev. D10 (1974) 2892;
G. Cohen-Tannoudji and W. Ochs, Z. Phys. C39 (1988) 513;
W. Ochs, Z. Phys. C23 (1984) 131.

[56] P. Carruthers and C. C. Shih, Int. J. Mod. Phys. A2 (1987) 1447.

[57] L. Van Hove and A. Giovannini, Acta Phys. Pol. B19 (1988) 917;
A. Giovannini and L. Van Hove, Z. Phys. C30 (1986) 391.

[58] A. Giovannini, Nucl. Phys. B161 (1979) 429.

[59] G. Marchesini, L. Trentadue and G. Veneziano, Nucl. Phys. B181 (1981) 335.

[60] DELPHI Collaboration, P. Abreu et al., Z. Phys. C50 (1991) 185.

[61] DELPHI Collaboration, P. Abreu et al., Z. Phys. C52 (1991) 271.

[62] DELPHI Collaboration, P. Abreu et al., Z. Phys. C56 (1992) 63.

[63] ALEPH Collaboration, D. Buskulic et al., Z. Phys. C69 (1995) 15.

[64] N. Schmitz, Int. J. Mod. Phys. A3 (1988) 1997; ibid. Int. J. Mod. Phys. A8 (1993) 1993.

[65] E665 Collaboration, M. R. Adams et al., Z. Phys. C61 (1994) 179.

[66] S. Carius and G. Ingelman, Phys. Lett. B252 (1990) 647;
R. Szwed, G. Wrochna and A. K. Wróblewski, Mod. Phys. Lett. A5 (1990) 1851.

[67] R. Szwed, G. Wrochna and A. K. Wróblewski, Mod. Phys. Lett. A6 (1991) 245.

[68] M. Gazdzicki et al., Mod. Phys. Lett. A6 (1991) 981.

[69] OPAL Collaboration, P. D. Acton et al., Z. Phys. C53 (1992) 539.

[70] ALEPH Collaboration, D. Decamp et al., Phys. Lett. B273 (1991) 181.

[71] WA21 Collaboration, G. T. Jones et al., Z. Phys. C54 (1992) 45.

[72] JADE Collaboration, W. Bartel et al., Phys. Lett. B91 (1980) 142;
D. P. Barber et al., Phys. Rev. Lett. 43 (1979) 830;
PLUTO Collaboration, Ch. Berger et al., Phys. Lett. B86 (1979) 418;
TASSO Collaboration, R. Brandelik et al., Phys. Lett. B86 (1979) 243.

[73] W. Bartel et al., Proceedings of the Workshop "Future Physics at HERA", edited by G. Ingelman, A. De Roeck, R. Klanner (Hamburg, Germany, 1996), p. 1095.

[74] H1 Collaboration, I. Abt et al., Nucl. Instr. and Meth. A386 (1997) 310.

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

[76] H1 BEMC Group, J. Ban et al., Nucl. Instr. and Meth. A372 (1996) 399.

[77] H. Bethe and W. Heitler, Proc. Roy. Soc. A146 (1934) 83.

[78] R. Brun et al., *GEANT 3*, CERN DD/EE/84-1.

[79] H1 Collaboration, S. Aid et al., Z. Phys. C72 (1996) 573.

[80] K. C. Hoeger, *Measurement of x, y, Q^2 in Neutral Current Events*, Proc. of the Workshop "Physics at HERA", Hamburg, Germany, 1991, Eds. W. Buchmüller and G. Ingelman, Vol. 1, p. 43.

[81] A. Blondel and F. Jacquet, Proc. of *Study of an ep facility for Europe*, Ed. U. Amaldi, DESY 79/48 (1979), p. 391.

[82] P. Van Esch, *Studie van de p_T-afhankelijkheid bij de factorisatie van diffractieve diep-inelastische werkzame doorsnede in een flux-factor en een diffractieve structuurfunctie, a.h.v. gebeurtenissen waargenomen in het H1 experiment*, Ph. D. thesis, Vrije Universiteit Brussel, unpublished.

[83] H1 Collaboration, S. Aid et al., Nucl. Phys. B480 (1996) 3.

[84] ZEUS Collaboration, M. Derrick et al., Z. Phys. C68 (1995) 29.

[85] AMY Collaboration, H. W. Zheng et al., Phys. Rev. D42 (1990) 737.

[86] M. Derrick et al., Phys. Rev. D34 (1986) 3304.

[87] G. S. Abrams et al., Phys. Rev. Lett. 64 (1990) 1334.

[88] TASSO Collaboration, W. Braunschweig et al., Z. Phys. C45 (1989) 193.

[89] DELPHI Collaboration, P. Abreu et al., Phys. Lett. B372 (1996) 172.

[90] A. De Angelis, Proc. Int. Europhysics Conf. on High Energy Physics, Brussels 1995, Eds. J. Lemonne, C .Vander Velde and F. Verbeure (World Scientific Singapore, 1996) p.63.

[91] EMC Collaboration, M. Arneodo et al., Nucl. Phys. B258 (1985) 249.

[92] D. Allasia et al., Z. Phys. C24 (1984) 119.

[93] S. Lupia and W. Ochs, Phys. Lett. B365 (1996) 339.

[94] B. R. Webber, Phys. Lett. B143 (1984) 501.

[95] M. Virchaux and A. Milsztajn, Phys. Lett. B274 (1992) 221.

[96] W. Ochs, Proc. XXIVth Int. Symp. on Multiparticle Dynamics, Vietri-sul-Mare, Italy, 1994, Eds. A. Giovannini, S. Lupia and R. Ugoccioni (World Scientific, Singapore 1995) p. 243.

[97] E665 Collaboration, M. R. Adams et al., Phys. Lett. B272 (1991) 163.

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

[99] H1 Collaboration, I. Abt et al., Z. Phys. C63 (1994) 377.

[100] ZEUS Collaboration, M. Derrick et al., Z. Phys. C70 (1996) 1.

[101] E. D. Malaza and B. R. Webber, Nucl. Phys. B267 (1986) 702; ibid. Phys. Lett. B149 (1984) 510.

[102] OPAL Collaboration, P. D. Acton et al., Phys. Lett. B287 (1992) 401.

[103] P. Allen et al., Nucl. Phys. B181 (1981) 385;
V. V. Amossov et al., Nucl. Phys. B203 (1982) 1;
J. Ballam et al., Phys. Lett. B56 (1975) 193;
B. Barlag et al., Z. Phys. C11 (1982) 283, C14 (1982) 281, erratum;
C. K. Chen et al., Nucl. Phys. B133 (1978) 13;
C. del Papa et al., Phys. Rev. D13 (1976) 2934;
M. Derrick et al., Phys. Lett. B91 (1980) 470;
EMC Collaboration, J. J. Aubert et al., Phys. Lett. B114 (1982) 373;
P. H. Garbincius et al., Phys. Rev. Lett. 32 (1974) 328;
B. Gibbard et al., Phys. Rev. D11 (1975) 2367;
A. J. Sadoff et al., Phys. Rev. Lett. 32 (1974) 955.

[104] H. Grässler et al., Nucl. Phys. B233 (1983) 269.

[105] EMC Collaboration, M. Arneodo et al., Phys. Lett. B165 (1985) 222; ibid. Z. Phys. C31 (1986) 1.

[106] R. N. Cahn, J. W. Cleymans and E. W. Colglazier, Phys. Lett. B43 (1973) 323.

[107] WA21 Collaboration, G. T. Jones et al., Z. Phys. C51 (1991) 11.

[108] S. Hegyi, Phys. Lett. B388 (1996) 837; ibid. Phys. Lett. B414 (1997) 210.

[109] O. G. Tchikilev, Phys. Lett. B388 (1996) 848; ibid. Phys. Lett. B393 (1997) 198.

[110] A. Ringwald, Nucl. Phys. B330 (1990) 1;
I. I. Balitsky and V. M. Braun, Phys. Lett. B314 (1993) 237;
A. Ringwald and F. Schrempp, Proc. Int. Sem. Quarks '94, Vladimir, Russia, 1994, Eds. D. Y. Grigoriev et al., p. 170 (hep-ph/9411217); ibid. in Proc. IXth Int. Sem. Quarks '96, Yaroslavl, Russia, 1996 (hep-ph/9610213).

[111] M. J. Gibbs, et al., Z. Phys. C66 (1995) 285.

[112] A. Capella et al., Phys. Lett. B337 (358) 1994.

[113] A. Capella et al., hep-ph/9712327.

[114] M. Plümmer and R. M. Weiner, Phys. Rev. D37 (1988) 3136.

[115] NA22 Collaboration, M. Adamus et al., Z. Phys. C37 (1988) 215.

[116] A. De Roeck, *Inclusive particle production in hadron-proton interactions at 250 GeV/c*, Ph. D. thesis, Universitaire Instelling Antwerpen, unpublished.

[117] A. Breakstone et al., Nuov. Cim. A102 (1989) 1199.

[118] UA1 Collaboration, C. Albajar et al., Nucl. Phys. B335 (1990) 261.

[119] H1 Collaboration, S. Aid et al., Phys. Lett. B356 (1995) 118.

[120] H. U. Bengtsson and T. Sjöstrand, Comp. Phys. Comm. 46 (1987) 43.

[121] A. Capella and J. Tran Thanh Van, Z. Phys. C23 (1984) 165.

[122] B. Andersson, G. Gufstafson and B. Nilsson-Almqvist, Nucl. Phys. B281 (1987) 289.

[123] H1 Collaboration, C. Adloff et al., *Multiplicity Structure of Hadronic Final States in Diffractive Deep Inelastic Scattering at HERA*, DESY 98-044, submitted to Eur. Phys. J. C.

[124] H1 Collaboration, C. Adloff et al., Eur. Phys. J. C1 (1998) 495.

[125] DELPHI Collaboration, P. Abreu et al., Phys. Lett. B275 (1992) 231.

[126] E665 Collaboration, M. R. Adams et al., Z. Phys. C61 (1994) 539.

[127] NA22 Collaboration, M. Begalli et al., Z. Phys. C55 (1992) 531.

[128] R. L. Cool et al., Phys. Rev. Lett. 48 (1982) 1451;
F. C. Winkelmann et al., Phys. Rev. Lett. 32 (1974) 121.

[129] S. P. Misra, A. R. Panda and B. K. Parida, Phys. Rev. Lett. 45 (1980) 322.

[130] NA22 Collaboration, M. Adamus et al., Z. Phys. C39 (1988) 301.

[131] (pp): M. G. Albrow et al., Nucl. Phys. B102 (1976) 275; S. J. Barish et al., Phys. Rev. Lett. 31 (1973) 1080; J. W. Chapman et al., Phys. Rev. Lett. 32 (1974) 257; F. T. Dao et al., Phys. Lett. B45 (1973) 399;
($p\overline{p}$): F. Grard et al., Phys. Lett. B59 (1975) 409; C. P. Ward et al., Nucl. Phys. B153 (1979) 299.

[132] UA4 Collaboration, D. Bernard et al., Phys. Lett. B166 (1986) 459.

[133] A. Capella et al., Phys. Lett. B81 (1979) 68; ibid. Phys. Rev. 236 (1994) 225.

[134] V. V. Innocente et al., Phys. Lett. B169 (1986) 285;
J. Ranft, Z. Phys. C33 (1987) 517;
S. Roesler, R. Engel and J. Ranft, Z. Phys. C59 (1993) 481.

[135] J. D. Bjorken, AIP Conference Proceedings No. 6, Particles and Fields Subseries No. 2, Eds. M. Bander, G. Shaw and D. Wong (AIP, New York, 1972).

[136] N. N. Nikolaev and B. G. Zakharov, Z. Phys. C49 (1991) 607; ibid. Z. Phys. C64 (1994) 631.

[137] M. Wüsthoff, Phys. Rev. D56 (1997) 4311.

[138] OPAL Collaboration, K. Ackerstaff et al., Eur. Phys. J. C1 (1998) 479.

[139] UA5 Collaboration, R. E. Ansorge et al., Z. Phys. C33 (1986) 175.

[140] H1 Collaboration, C. Adloff et al., *Hadron Production in Diffractive Deep Inelastic Scattering*, DESY 98-029 submitted to Phys. Lett. B.

[141] ZEUS Collaboration, J. Breitweg et al., *Event Shape Analysis of Deep Inelastic Scattering Events with a Large Rapidity Gap at HERA*, DESY 97-202, accepted by Phys. Lett. B.

[142] UA4 Collaboration, D. Bernard et al., Phys. Lett. B186 (1987) 227.

[143] A. Capella and J. Tran Thanh Van, Z. Phys. C18 (1983) 85;
A. Capella and A. Krzywicki, Phys. Rev. D18 (1978) 4120.

[144] OPAL Collaboration, R. Akers et al., Phys. Lett. B320 (1994) 417.

[145] W. Buchmüller and A. Hebecker, Phys. Lett. B355 (1995) 573;
A. Edin, G. Ingelman and J. Rathsman, Phys. Lett. B366 (1996) 371;
W. Buchmüller, M. F. McDermott and A. Hebecker, Nucl. Phys. B487 (1997) 283; Erratum Nucl. Phys. B500 (1997) 621.

[146] P. Van Mechelen, *Unfolding experimental distributions*, H1 note 10/96-496.

[147] V. Blobel, *Regularised unfolding for high-energy physics experiments*, OPAL TN361, DESY 84-118.

[148] D. Kant and G. Thompson, *Charged track multiplicity distributions in the Breit frame*, H1 note 04/96-475.

[149] E. T. Jaynes, Phys. Rev. 106 (1957) 620.

[150] Y. Tikochinsky, N. Z. Tishby and R. D. Levine, Phys. Rev. Lett. 52 (1984) 1357.

List of figures

List of tables

List of abbreviations

BEMC	Backward Electromagnetic Calorimeter
BGF	Boson-Gluon Fusion
BPC	Backward Proportional Chamber
BSET	BEMC Single Electron Trigger
CDM	Colour Dipole Model
CIP	Central-Inner Proportional chamber
CIZ	Central-Inner z-chamber
CJC	Central Jet Chamber
CMS	Centre-of-Mass System
COP	Central-Outer Proportional chamber
COZ	Central-Outer z-chamber
CT	Central Tracker
CTL1	Central Trigger Level 1
DESY	Deutshes Elektronen-Synchrotron
DGLAP	Dokshitzer-Gribov-Lipatov-Altarelli-Parisi
DIS	Deep Inelastic Scattering
DLLA	Double Leading-Log Approximation
DPM	Dual Parton Model
FT	Forward Tracker
HERA	Hadron-Elektron-Ring-Anlage
HNBD	H-function extension of the Negative Binomial Distribution
IRON	Instrumented Iron
KNO	Koba-Nielsen-Oleson (scaling)
KNO-G	Koba-Nielsen-Oleson-Golokhvastov (scaling)
L1/L2/L3/L4/L5	(trigger) Level 1/2/3/4/5
LAr	Liquid Argon (calorimeter)
LLA	Leading-Log Approximation
LND	Lognormal Distribution
LPHD	Local Parton-Hadron Duality
LRG	Large Rapidity Gap (event)
LUMI	Luminosity
MEAR	Matrix Elements + ARIADNE
MEPS	Matrix Elements + Parton Showers
MLLA	Modified Leading-Log Approximation
MNBD	Modified Negative Binomial Distribution
$\overline{\mathrm{MS}}$	modified Minimal Subtraction scheme
MWPC	Multi-Wire Proportional Chamber

NBD	Negative Binomial Distribution
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QPM	Quark-Parton Model
SCI	Soft Colour Interactions
ToF	Time-of-Flight