Measurement of the proton structure function at HERA using the ZEUS detector

Stanislaus Cornelius Maria Bentvelsen
geboren te Schipluiden

Measurement of the proton structure function at HERA using the ZEUS detector

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. Dr. P. W. M. de Meijer
ten overstaan van een door het college van dekanen ingestelde
commissie in het openbaar te verdedigen in de Aula der Universiteit
op donderdag 28 april 1994 te 10.30 uur

door

Stanislaus Cornelius Maria Bentvelsen
geboren te Schipluiden

Promotor: Prof. Dr. J. J. Engelen
Co-Promotor: Dr. P. M. Kooijman

Faculteit der Natuur- en Sterrenkunde

The work described in this thesis is part of the research programme of 'het Nationaal Instituut voor Kernfysica en Hoge–Energie Fysica (NIKHEF–H)' in Amsterdam, the Netherlands. The author was financially supported by the University of Amsterdam.

All nature wears
one universal grin
- H. Fielding

Contents

Preface

Measurements of deep-inelastic scattering have been seminal to our understanding of hadronic substructure. They have provided us with an insight not only in the constituents from which the proton is built, but also in the theory of the interactions between them. The advent of the HERA electron-proton collider, with its center of mass energy of 295 GeV, has allowed an extension of these measurements to a new kinematic regime. The structure of the proton is probed with a resolution that is up to 1000 times smaller than the diameter of the proton itself. Even constituents in the proton that carry only a fraction of 10^{-4} of the proton momentum, can be observed at HERA. Measurements of deep-inelastic scattering in this kinematic region are essential for a detailed understanding of the proton, as well as for the theory of strong interactions.

For the last four years, I have been part of the NIKHEF team working on the ZEUS experiment, which is located at one of the interaction regions at the HERA collider in Hamburg. In 1992 the first electron-proton collisions were observed and I stayed at Hamburg for one year to participate in the initial analysis of the recorded events. This analysis resulted in the determination of the proton structure function F_2 in the kinematic domain accessible at HERA, and it is the main subject of this PhD thesis.

Stan Bentvelsen, Amsterdam, March 1992.

Chapter 1

Review of Deep–Inelastic–Scattering

Several experimental methods are available to gain knowledge of Quantum-Chromo-Dynamics (QCD) in high energy physics. Utilizing electron-positron colliders, for example, the strong coupling constant and the Casimir factors of the colour gauge group can be determined from quark and gluon production. Although these experiments allow precision tests, it is not possible to study stable quark configurations like the nucleons.

With hadron colliders, the parton content of nucleons can be studied via for example the Drell-Yann process or through the production of jets. The cross sections of these processes are sensitive to the distributions of the parton constituents inside the nucleon and therefore reveal some of its structure. However, measurements at hadron colliders are difficult to perform due to the complex nature of both interacting particles.

An attractive way of studying QCD is by probing the parton content of a nucleon with a pointlike lepton. In deep-inelastic-scattering experiments a highly energetic lepton interacts with the nucleon, or, if the resolution is large enough, the lepton interacts directly with the constituents of the nucleon. Because the electro-weak interactions between the lepton and the constituents of the nucleon are well understood, this process reveals the internal structure of nucleons. In this chapter we summarize the ingredients of deep-inelastic-scattering processes.

1.1 Deep Inelastic Scattering

The study of deep-inelastic-scattering (DIS) is devoted to lepton–nucleon reactions at high energies. The HERA collider allows the study of high energy collisions between protons and electrons or positrons. Generally, the proton breaks up during the interaction, and a number of particles is produced in the final state. Two experimentally different classes of DIS are distinguished:

$$e^{\pm} + P \to e^{\pm} X \quad ; \quad e^{\pm} + P \to \overset{(-)}{\nu} X \tag{1.1}$$

where X represents the spray of particles that is produced by the break-up of the proton. In the first process the charge of the lepton is conserved, and the intermediate vector boson is neutral. This is referred to as neutral current (NC) DIS. In the second process the lepton converts to an (anti–) neutrino by the exchange of a charged intermediate vector boson. This process is referred to as charged current (CC) DIS. The two types of interactions are depicted in figure 1.1.

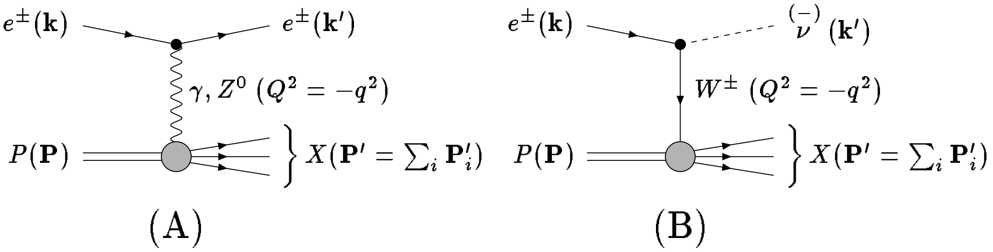


Figure 1.1: *Deep-inelastic-scattering. Figure (A) represents the neutral-current DIS, while figure (B) displays the charged-current variant. The definition of momenta is indicated between brackets.*

1.1.1 The DIS kinematic variables

The analysis and measurements treated in this thesis are restricted to inclusive NC DIS. This means that no attempt is made to identify structures or specific particles in the final state X. It also implies that in theoretical calculations one integrates over the complete hadronic final state. The analysis is also restricted to collisions with unpolarized beams. The kinematics of inclusive DIS is, at a given center-of-mass energy s, described by two independent variables. We introduce a number of suitable Lorentz–invariant variables and any combination of two is sufficient to describe the unpolarized inclusive DIS process completely. In the definitions of the variables we neglect the proton and electron masses.

The first Lorentz–invariant is the virtuality of the exchanged vector boson, q^2 or Mandelstam variable t. This variable sets the scale of the interaction, since the wavelength or resolution of the virtual boson is proportional to $1/|\mathbf{q}|$. For high momentum transfer the lepton-proton interaction becomes essentially an interaction between the lepton and a parton inside the proton. Because the interaction is purely spacelike, it is convenient to introduce an additional minus-sign in the definition of q^2. Following the notation given in figure 1.1 we define:

$$Q^2 = -q^2 = -(\mathbf{k}' - \mathbf{k})^2 = -(\mathbf{P} - \mathbf{P}')^2 = -t > 0. \tag{1.2}$$

The value for Q^2 of the exchanged boson ranges between 0 and s.

The second useful variable is known as elasticity or Björken-x. It is defined as

$$x = \frac{Q^2}{2\,\mathbf{P}\cdot\mathbf{q}} = \frac{-t}{u+s} \tag{1.3}$$

with the usual definition of the Mandelstam variable $u = (\mathbf{k}' - \mathbf{P})^2$. In elastic scattering, where the proton does not break up during the interaction, the value of x equals 1. In deep–inelastic scattering, i.e. $Q^2 \gg 1$ GeV^2, x is less than one. In lowest order perturbation theory with massless quarks, x can be identified as the momentum fraction carried by the struck quark. The parameter x plays an important role in the description of the structure of the proton.

The variables Q^2 and x can be related via y. In the proton rest frame y corresponds to the fraction of the lepton energy that is transferred from the lepton to the proton. In DIS y ranges from 0 to 1. The Lorentz-invariant expression for y is

$$y = \frac{\mathbf{P}\cdot\mathbf{q}}{\mathbf{P}\cdot\mathbf{k}} = \frac{Q^2}{s\,x} = \frac{u+s}{s}. \tag{1.4}$$

The invariant mass of the hadronic final state X, denoted by W, is deduced from momentum conservation at the hadronic vertex and can be related to x and Q^2 as follows:

$$W^2 = (\mathbf{P}')^2 = (\mathbf{P}+\mathbf{q})^2 = Q^2\frac{1-x}{x} = s+t+u. \tag{1.5}$$

1.1.2 DIS cross section and structure functions

We discuss the expression for the DIS cross section in the approximation that a single virtual gauge boson probes the nucleon. The unpolarized differential Born cross section can then be factorized in a leptonic part, denoted by $L^{\mu\nu}$ and a part that contains all information of the hadronic vertex, $W_{\mu\nu}$ [1]. The specific form of both tensors depends on the intermediate gauge boson V. Bearing in mind interference terms between the exchanged boson V, the double differential cross section as function of x and Q^2 is written as:

$$\frac{d^2\sigma^{VV'}(e^{\pm}p)}{dx\,dQ^2} = \frac{1}{4\pi}\left(\frac{\mathcal{P}(VV')}{Q^4\,x}\right)\frac{y}{s}\,L^{VV'\,\mu\nu}W^{VV'}_{\mu\nu}. \tag{1.6}$$

The factor $\mathcal{P}(VV')$ represents the propagator term of the boson V and V' in the cross section, relative to the photon propagator term $1/Q^4$.

In the standard model the leptonic tensor can be calculated in perturbation theory. Its explicit form depends on the vector and axial gauge coupling. Parametrizing the vertex of the boson V as $i\gamma^{\mu}(g_v + g_a\gamma^5)$ (similar for the V'), the leptonic tensor in lowest order becomes:

$$L^{VV'\mu\nu} = 2\,(g_v g_v' + g_a g_a')\left[k^{\mu}k'^{\nu} + k^{\nu}k'^{\mu} + (q^2/2)g^{\mu\nu}\right] - 2\,(g_v g_a' + g_v' g_a)\,\epsilon^{\mu\nu\kappa\lambda}k^{\kappa}k'^{\lambda}. \tag{1.7}$$

The hadronic tensor $W_{\mu\nu}(\mathbf{P},\mathbf{q})$ serves to parameterize the vertex of the proton and the vector boson. The tensor $W_{\mu\nu}(\mathbf{P},\mathbf{q})$ is symmetric in μ and ν and can only depend on the metric $g_{\mu\nu}$ and momenta p_μ, q_ν. In general, such a tensor has five terms, however, after imposing current conservation at the hadron vertex only three independent terms remain. Current conservation at the lepton vertex implies that terms proportional to $\mathbf{q}$ can be ignored. The most general expansion of $W_{\mu\nu}^{VV'}$ can be parametrized with the aid of three Lorentz-invariant structure functions $F_1^{VV'}$, $F_2^{VV'}$ and $F_3^{VV'}$[1] as:

$$W_{\mu\nu}^{VV'} = (-g_{\mu\nu})\, F_1^{VV'} + \left(\frac{p_\mu p_\nu}{\mathbf{P}\cdot\mathbf{q}}\right) F_2^{VV'} + \left(\epsilon_{\mu\nu\kappa\lambda}\frac{p^\kappa p^\lambda}{4\mathbf{P}\cdot\mathbf{q}}\right) F_3^{VV'}. \tag{1.8}$$

Note that the term proportional to $F_3^{VV'}$ violates parity. This immediately indicates that for photon exchange only the structure functions $F_1^{VV'}$ and $F_2^{VV'}$ survive.

Inserting the expansion of $W_{\mu\nu}^{VV'}$ in the cross section 1.6, we obtain the general expression for the NC DIS cross section

$$\frac{d\sigma^{NC}(e^\pm p)}{dx\,dQ^2} = \frac{4\pi\alpha^2}{x\,Q^2}\left[y^2 x F_1^{NC} + (1-y)F_2^{NC} \mp \left(y - \frac{y^2}{2}\right) x F_3^{NC}\right] \tag{1.9}$$

with the structure functions for γ-exchange, Z^0 exchange and the interference term

$$F_i^{NC} = F_i^{\gamma\gamma} + \mathcal{P}(\gamma Z^0)\, F_i^{\gamma Z^0} + \mathcal{P}(Z^0 Z^0)\, F_i^{Z^0 Z^0} \quad ; \; i = 1,2,3 \tag{1.10}$$

and the propagator terms

$$\begin{aligned} \mathcal{P}(\gamma Z^0) &= Q^2/\left(Q^2 + M_{Z^0}^2\right), \\ \mathcal{P}(Z^0 Z^0) &= Q^4/\left(Q^2 + M_{Z^0}^2\right)^2. \end{aligned} \tag{1.11}$$

Similarly we get for the general form of the CC DIS cross section, i.e. exchange of the virtual $W^\pm$:

$$\begin{aligned} \frac{d\sigma^{CC}(e^\pm p)}{dx\,dQ^2} &= \frac{\pi\alpha^2}{2x\,\sin^4\theta_W\,\left(Q^2 + M_W^2\right)^2} * \\ &\quad \left[y^2 x F_1^{WW} + (1-y)F_2^{WW} \mp \left(y - \frac{y^2}{2}\right) x F_3^{WW}\right]. \end{aligned} \tag{1.12}$$

All information of the hadron vertex for inclusive DIS is incorporated in the structure functions F_1, F_2 and F_3. Measurement of the structure functions is therefore crucial in the study of hadron substructure. The structure functions $F_i^{VV'}$ depend on two variables, commonly taken to be x and Q^2. The motivation of this choice becomes clear in the physical interpretation of the structure functions.

---

[1] Most often $W_{\mu\nu}$ is expanded into functions W_i [1]. Only after implementing the Bjorken limit and contraction with the leptonic tensor $L^{VV'\mu\nu}$ is identification with the structure functions F_i made.

The expression for the ep cross section is modified due to additional higher order electro-weak interactions. Photon emission of the lepton line realizes shifts of the kinematic variables, whereas virtual corrections enhance the Born cross-section. Both these effects can be parametrized in the radiative cross section $d\sigma_{RAD}/dxdQ^2$ as

$$\frac{d^2\sigma_{RAD}(x,Q^2)}{dx_s\,dQ_s^2} = \frac{d^2\sigma_{Born}(x_h = x_s, Q_h^2 = Q_s^2)}{dx_h\,dQ_h^2}\left(1+\delta_r(x_s,Q_s^2)\right) \tag{1.13}$$

with $\delta_r(x_s, Q_s^2)$ a complicated radiator function. The variables $(x,Q^2)_s$ represent the shifted variables with respect to the Born variables $(x,Q^2)_h$. We discuss the effect of the higher order electroweak interactions on the measurement of the Born cross section in more detail in section 1.6.

1.2 Partons in the proton

1.2.1 A short review of QCD

Quantum-Chromo-Dynamics (QCD) is a gauge field theory defined by its Lagrange density. The QCD Lagrangian is, apart from the gauge fixing terms, invariant under local $SU(3)$ gauge transformations. The Lagrangian provides a description of the dynamics of quarks, the spinor fields which, when they act as free fields, satisfy the Dirac equation. The quarks appear in nature in three colours and five different flavours, u d s c b. Experimental searches in progress to find the sixth flavour, t, which makes the electro-weak standard model consistent. The extension to mutual interactions is made in QCD by introducing 8 coloured gluon fields, the gauge fields, to restore local $SU(3)$ invariance. The non-Abelian nature of QCD is exhibited in the fact that the gluons carry colour charge and hence interact not only with the coloured quarks, but also with each other.

The phenomenology of QCD is fundamentally different from its Abelian counterpart QED (Quantum-Electro-Dynamics), due to the gauge-field self-interactions, which are absent in QED. Whereas in the renormalizable QED theory the electric charge increases as one penetrates the virtual charge cloud of an electron, in QCD the observed colour coupling of the quark decreases as one penetrates the colour cloud. This anti-screening effect in QCD leads to the notion of asymptotic freedom; the interaction becomes weaker at shorter distances.

This picture is quantified when renormalization is performed in perturbation theory. A renormalization scale μ_R^2 inevitably needs to be introduced to regularize the ultraviolet singularities. At the same time no physical observables (cross sections) may depend on this artificially introduced parameter. From this requirement it follows, via the renormalization group equation, that the effective strong coupling α_s becomes in perturbation theory a function of μ_R^2. In DIS one chooses this scale to be equal to Q^2, and this means that the couplings at two different values of Q^2 are related. The standard expression of the dimensionless $\alpha_s(Q^2)$ is given in reference to a scale Λ. In first order of QCD perturbation theory it reads:

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_1 \ln(Q^2/\Lambda^2)} + \mathcal{O}\left(\frac{1}{\ln^2(Q^2/\Lambda^2)}\right) \qquad \beta_1 = 11 - 2n_f/3, \tag{1.14}$$

where n_f is the number of open flavours of quarks.

The fundamental free scale parameter of QCD, Λ, sets the scale at which the interactions become strong. Indeed, the measured value of $\Lambda \sim 100 - 300$ MeV is roughly equivalent to the size of the hadrons. The definition of Λ can however be controversial. One most often determines its value from measurements of α_s from which one concludes that the value of Λ depends on the number of flavours n_f involved. Also, multiplying Λ with a constant factor gives a correction in α_s which is one order higher in perturbation theory and can therefore be neglected. One has to agree on the actual *formula* of α_s to define a unique value for Λ. Finally, in second and higher order of perturbation theory, Λ also depends on the factorization scheme (see sect 1.2.3).

Besides the above-mentioned ultraviolet divergences, which appear when the momenta in the Feynman loop-integrals go to infinity, calculations in perturbative QCD suffer from two more types of divergences [2]. These are the infra-red (IR) divergences and collinear singularities.

The IR divergences show up in the calculation of Feynman diagrams when in real and virtual gluon amplitudes the gluon momenta go to zero. However, the IR singularities cancel between the virtual and the real gluon graphs. In an inclusive measurement, which implies that one integrates over all present momenta in the final state, the IR divergency is no longer present.

The collinear divergences or mass singularities appear whenever the momenta of quarks or gluons become parallel to each other (which is only possible for coupling between massless particles). In DIS, the collinear divergences of the final state are also cancelled between the real and virtual graphs, for example in processes like

$$\gamma^\star q \rightarrow X \quad ; \quad \gamma^\star g \rightarrow X \tag{1.15}$$

there are no collinear divergences in the final state. The only collinear divergences left over are the ones associated with the initial state. These are absorbed via mass factorization in the parton distribution functions.

The quantities that do not suffer from the soft, long range behaviour of QCD and hence are not IR or collinear divergent are called *infrared-safe.* In QCD, only infrared-safe quantities can be calculated more accurately order by order in perturbation theory. The non-infrared-safe, collinear divergent quantities are absorbed such that all physical observables are finite. In DIS this is done in the theoretical framework of factorization, which we briefly discuss after the introduction of the parton model.

1.2.2 Quark-parton model and factorization

Experimentally, the discovery of the Ω^-, predicted by the $SU(3)$ quark-flavour model of Gell-Mann [3] provided confirmation that hadrons are built of smaller constituents. At that time however, those constituents (or quarks) were purely mathematical concepts and not yet considered as a new, underlying form of matter. This attitude changed after the first deep-inelastic-scattering (DIS) experiments. The famous experiment performed at SLAC [4], where evidence for structure inside a single nucleon was obtained, was from a physical point of view very similar to the one performed by Rutherford much earlier this century. Rutherford found evidence for the nucleus inside the atom whereas the SLAC experiment found evidence for partons inside the nucleus. These partons were later identified with quarks. In 1979 evidence for the existence of gluons was provided at DESY in Hamburg via the observation of three-jet events in e^+e^- annihilation [5].

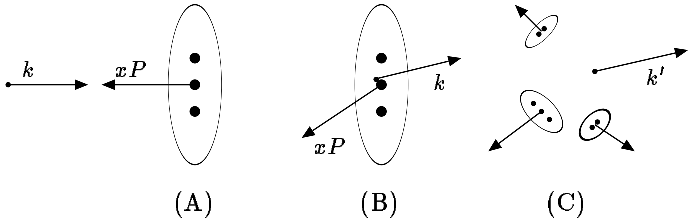


Figure 1.2: *Schematic parton model picture in deep inelastic scattering. Figure (A) is before the hard interaction, and the proton is Lorentz-contracted and undergoes time dilatation. The 'frozen' hadron consists of universal partons distributions. In figure (B) the hard scattering takes place between the electron and one parton, the others are spectators. After the hard collision in (C) hadronization takes place.*

The (QCD-improved) quark-parton model can be applied, with varying degrees of success, to any hadronic cross section involving large momentum transfer. The ingredients of this parton model for deep-inelastic scattering can be summarized as follows:

- All hadrons (including the proton) consist of partons. In QCD the partons are identified with the quarks and gluons.
- The scattering between the electron and the proton proceeds via the electron - parton scattering.
- The parton is pointlike, and hence has no form factor.
- The parton stays on-shell during the elastic interaction which means that if the transverse momentum of the quark can be neglected compared to the virtuality of the intermediate boson then:

  - Exchange interactions between the struck parton and the spectator partons can be neglected (these 'higher twist' effects vanish like $(1/Q^2)^i$, $i \geq 1$).
  - The scattering is incoherent between various partons in the proton.

These last two points are an assumption only. They can (heuristically) be justified with the physical concept of *factorization.*

Factorization can be understood intuitively by the following picture. In the infinite momentum frame the virtual partons are in a state of momentum fraction xP(see figure 1.2A). This state is characterized by a lifetime τ in the proton rest frame. In the center of mass system the lifetime is larger due to time dilatation. Combining this with the Lorentz-contraction of the proton (indicated in the figure by a disk shape) this means the time it takes the photon to interact with the parton vanishes in the center of mass frame as $s \longrightarrow \infty$. The photon effectively sees the partons 'frozen' during interaction (figure 1.2B), and possible long range 'initial state' interactions happen too early to affect the photon interaction.

After the collision, anything may happen as the electron recedes. Quark pairs will be created and hadronization takes place (figure 1.2C). Those 'final state' interactions or hadronization happen too late to influence the hard parton scattering.

This factorization of hadron cross sections between hard parton interaction and the soft, long range configuration of partons inside the hadrons is shown to be valid in all orders of QCD perturbation theory. However, beyond the leading order, the notion of 'long range' and 'short range' interactions becomes ambiguous, and this freedom is expressed by the factorization scale[2] μ^2. Factorization implies that if Q^2 is large the hadron cross sections can be written in terms of

- a hard, short distance, calculable cross section $\hat{\sigma}^{fV}(x, Q^2; \mu^2)$, which is infrared safe, that describes the interactions between the elementary leptons and quarks. This cross section $\hat{\sigma}$ depends on the intermediate electroweak vector boson V, on the parton f and on a factorization scale, which is denoted by μ^2. The hard cross section is however independent of all long range effects, in particular it is independent of the hadron to which the parton belongs.

- the soft, long range parton distribution functions $q_f^h(z, Q^2; \mu^2)$ in which the collinear divergences are absorbed. It is specific for the particular hadron type, however, the parton distribution functions are *universal*, i.e. independent of the hard scattering and intermediate vector boson V. The parton distribution functions $q_f^h(z, Q^2; \mu^2)$ describe the probability that at the scale Q^2 the parton f has a fraction z of the momentum of the hadron h.

---

[2] For a complete description of scale dependences, the remaining initial state mass singularities imply an additional dependence on a mass singularity scale μ_S^2 for the hard parton cross section and the parton distribution functions. However, this scale is set equal to the factorization scale, $\mu_S^2 = \mu^2$

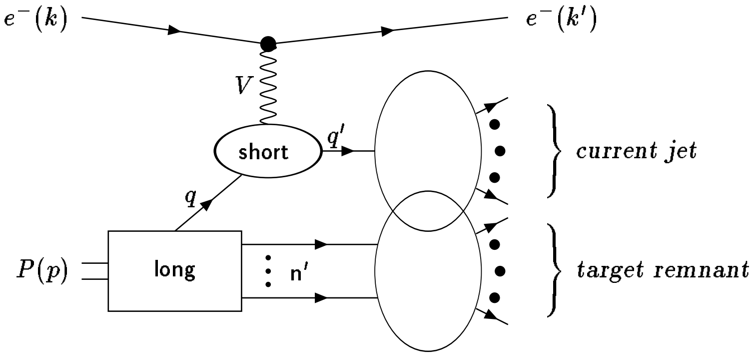


Figure 1.3: *This picture schematically indicates the short and long range processes of DIS. The short-distance interaction between the vector boson V and the quark q is specific for the process. It factorizes with the long distance parton distribution functions. After the hard interaction the hadronization process takes place.*

There is no quantum interference between the long and short range effects. The inclusive hadron cross section is then the convolution between the hard cross section and the parton distribution functions (PDF's), see also figure 1.3:

$$\sigma^{h\,V}(x,Q^2) = \sum_f \left[\hat{\sigma}^{fV}(Q^2;\mu^2) \otimes q_f^h(Q^2;\mu^2)\right](x) \tag{1.16}$$

where the convolution integral is defined as

$$\left[A \otimes B\right](x) = \int_x^1 \frac{dz}{z}\, A(z)\, B\left(\frac{x}{z}\right). \tag{1.17}$$

In lowest order the convolution between the hard interaction $\hat{\sigma}^{fV}(x,Q^2)$ and the parton distribution function $q_f^h(x,Q^2)$ becomes a simple multiplication.

The concept of factorization gives QCD its predictive power. On the one hand factorization relates various hadron cross sections. When the PDF's $q_f^h(x,Q^2;\mu^2)$ are known, all hadron cross sections are related via the corresponding hard scattering $\hat{\sigma}$. For example, by measuring the PDF's in DIS experiments, the cross section for the Drell-Yan process in hadron-hadron scattering can be predicted. Also the various DIS experiments are related by factorization. When the PDF's are obtained from electron-proton DIS experiments, the neutrino-proton or muon-neutron DIS cross sections can be calculated, or vice versa.

On the other hand, QCD predicts the scale evolution in Q^2 of the PDF's. This implies that measurements of hadron cross sections at a particular value for the scale Q_0^2 determine the cross section at a different value for the scale Q^2.

1.2.3 Factorization schemes in DIS

For Deep-Inelastic Scattering (DIS) factorization implies that the structure functions can be written as a convolution of the PDF's with the hard interactions. The latter are for DIS denoted by the coefficient functions $C_i^{fVV'}$. We restrict ourselves to the proton structure functions, so that we drop the index h:

$$F_i^{VV'}(x,Q^2) = \sum_{n=1,\ldots} \left(\frac{\alpha_s(Q^2)}{2\pi}\right)^{n-1} \sum_{f=q,\bar{q},G} \left[C_i^{fVV'(n)}(Q^2;\mu^2) \otimes q_f^{(n)}(Q^2;\mu^2)\right](x) \tag{1.18}$$

where the index n represents the order of perturbation theory in the strong coupling constant $\alpha_s(Q^2)$. The freedom of the choice of the scale μ^2 beyond the leading-order approximation is exhibited in the different factorization schemes. In the literature several factorization schemes circulate. From a theoretical point of view the $\overline{MS}$ scheme is attractive, since constant terms in the dimensional regularization of the hard scattering are absorbed in the long range PDF's, by a suitable choice of μ^2.

The most convenient scheme in deep-inelastic-scattering for F_2 measurements is called 'DIS-scheme'. In this scheme, the coefficient functions for the electromagnetic structure function $F_2^{\gamma\gamma}$, are by definition zero for second and higher orders in perturbation theory. This means that $C_2^{f\gamma\gamma(i)} = 0$ for $i > 1$, and the expression of $F_2^{\gamma\gamma}$ in terms of the parton distribution functions is not altered beyond the leading-order approximation.

1.2.4 Short distance NC DIS

For partons with spin $\frac{1}{2}$, explicit calculation of the hard scattering shows that the coefficient function for the electromagnetic structure function $F_2^{\gamma\gamma}$ becomes $C_2^{f\gamma\gamma(1)}(x) = e_f^2\delta(x-1)$ (the coefficient functions for the gluon are zero). Hence $F_2^{\gamma\gamma}$ is the charge-weighted sum of the quark-distribution functions

$$F_2^{\gamma\gamma}(x,Q^2) = \sum_{f=u,d,s,c,t,b} e_f^2 \left[x q_f(x,Q^2) + x\bar{q}_f(x,Q^2)\right] \tag{1.19}$$

Here e_f is the electric charge divided to the unit charge for quark flavour f. In the DIS scheme this expression is by definition valid to all orders of perturbation theory.

The expressions for the structure functions F_1^{NC}, F_2^{NC} and F_3^{NC}, including the Z^0-exchange, become in lowest order QCD in terms of the PDF's:

$$F_1^{NC}(x,Q^2) \quad = \quad F_2^{NC}(x,Q^2)/2x \tag{1.20}$$

$$F_2^{NC}(x,Q^2) = \sum_{f=u,d,s,c,t,b} A_f(Q^2)\left[xq_f(x,Q^2) + x\bar{q}_f(x,Q^2)\right] \quad (1.21)$$

$$xF_3^{NC}(x,Q^2) = \sum_{f=u,d,s,c,t,b} B_f(Q^2)\left[xq_f(x,Q^2) - x\bar{q}_f(x,Q^2)\right] \quad (1.22)$$

For unpolarized electrons the flavour dependent coefficients A_f and B_f are given by [6]:

$$\begin{aligned} A_f(Q^2) &= e_f^2 - 2e_f v_e v_f \mathcal{P}(\gamma Z^0) + (v_e^2 + a_e^2)(v_f^2 + a_f^2)\mathcal{P}(Z^0 Z^0) \\ B_f(Q^2) &= 2e_f a_e a_f \mathcal{P}(\gamma Z^0) + 4v_e v_f a_e a_f \mathcal{P}(Z^0 Z^0) \end{aligned} \quad (1.23)$$

where the propagator terms $\mathcal{P}$ are defined in (1.11). Note that the γ-exchange, Z^0 exchange and interference term are included in the definition of A_f and B_f (compare with equation 1.10). The weak coupling constants v_f and a_f defined in the electro-weak standard model as:

$$v_f = \frac{T_{3f} - 2e_f \sin^2\theta_W}{\sin(2\theta_W)}, \qquad a_f = \frac{T_{3f}}{\sin(2\theta_W)}. \quad (1.24)$$

The third component of the weak isospin operator yields $T_{3f} = \frac{1}{2}$ for neutrino's and the up-type quarks (u,c,t) and $T_{3f} = -\frac{1}{2}$ for electrons and the down-type quarks (d,s,b).

Relation 1.20 is known as the Callan-Gross [7] relation, and is the consequence of the spin $s = \frac{1}{2}$ of the quarks. The Callan-Gross relation is however violated in higher orders of perturbation theory. Effectively this violation is brought about by a contribution from longitudinally polarized photons. One quantifies this contribution by introducing the longitudinal structure function F_L as

$$F_L(x,Q^2) = F_2(x,Q^2) - 2xF_1(x,Q^2) \quad (1.25)$$

which vanishes in leading order. In second order the longitudinal structure function can be calculated using the relation

$$F_L(x,Q^2) = \frac{\alpha_s(Q^2)}{2\pi} \sum_{f=q,\bar{q},G} \left[\left(C_2^{f\gamma\gamma(2)} - 2xC_1^{f\gamma\gamma(2)}\right) \otimes q_f(Q^2)\right](x) \quad (1.26)$$

The coefficient functions are calculated from second order perturbative QCD [8]. We discuss the full expression and size of F_L and its effect on the measurement of F_2 in chapter 6. Rewriting equation 1.9 in terms of the longitudinal structure function yields

$$\frac{d\sigma^{NC}(e^\pm p)}{dx\,dQ^2} = \frac{2\pi\alpha^2}{x\,Q^2} Y_+ \left[F_2 - \frac{y^2}{Y_+}F_L \mp \frac{Y_-}{Y_+}xF_3\right] \quad (1.27)$$

with $Y_\pm = 1 \pm (1-y)^2$.

1.2.5 Parton evolution

Neglecting the possibility of heavy quarks, the QCD Lagrangian contains no parameters with a mass dimension. Only via higher order of perturbation in QCD a massive regularization parameter needs to be introduced in the theory. From dimensional arguments it can therefore be argued directly that in lowest order QCD the parton distributions do not depend on the massive scale Q^2. The experimentally observed 'scaling' of the structure functions in DIS experiments was the first indication that hadrons consist of pointlike objects, the partons [9].

Small violations of this scaling behaviour of the structure functions, which have been verified experimentally, are generated by radiative QCD corrections. The dynamics of QCD allows interactions between quarks and gluons via the elementary processes as gluon emission, $q \longrightarrow qg$, $\bar{q} \longrightarrow \bar{q}g$ and $g \longrightarrow gg$, and creation of sea quark-anti-quark pairs $g \longrightarrow q\bar{q}$. Calculation of these processes shows that they are logarithmically UV divergent. These logarithmic divergences, parametrized by a scale $\mu_R^2 = Q_0^2$, are absorbed in the definition of the parton distribution functions. Moving to a different scale Q^2, the parton distribution functions are redefined to absorb the logarithmic terms in Q^2. Hence the parton distributions at different values of Q^2 are related.

To demonstrate this mechanism, consider the coefficient function $C_2^{(2)}(x, Q_0^2)$, and let Q_0^2 be the scale at which the divergences are absorbed into the parton distributions. If we now consider another scale Q^2, a logarithmic term containing Q^2 appears in the calculation of the hard scattering QCD diagrams. The coefficient function, in which this term appears, can be factored out in a term containing the logarithmic term explicitly, $P(x)$, and a finite term $f(x)$:

$$C_2^{(2)} = \Delta t \; P(x) + f(x) \quad ; \quad \Delta t = \ln(Q^2/Q_0^2) \tag{1.28}$$

The structure function is then expressed in terms of the lowest and leading order coefficient functions as (summation over f implied)

$$\begin{aligned} F_2(x,t) &= e_f^2 x \left\{ q_f(x,t_0) + \frac{\alpha_s}{2\pi} \left[q_f(t_0) \otimes (\Delta t P + f) \right](x) \right\} \qquad (1.29) \\ &= e_f^2 x \left\{ q_f(x,t_0) + \Delta q_f(x,t) + \frac{\alpha_s}{2\pi} \left[q_f(t_0) \otimes f \right](x) \right\} \end{aligned}$$

where in the second line, the part of the coefficient function that is proportional to Δt is reabsorbed in the modification of the parton distribution at t_0, denoted by $\Delta q_f(t)$. The relationship between the $\Delta q_f(t)$ and $q_f(t_0)$ can now be deduced:

$$\Delta q_f(x,t) = \frac{\alpha_s}{2\pi} \Delta t \left[q_f(t_0) \otimes P \right](x) \tag{1.30}$$

This equation states that the scale transition from t_0 to t can be replaced by an effective t dependence of the parton distribution. From equation 1.30 one obtains

$$\frac{dq_f(x,t)}{dt} = \frac{\alpha_s}{2\pi} \left[q_f(t) \otimes P \right](x) \tag{1.31}$$

This is the generic form of the Gribov-Lipatov-Altarelli-Parisi [10] (GLAP) integro-differential equations. The scale Q_0^2 is usually set to Λ^2. In an extensive treatment of all QCD diagrams, the functions P are split into contributions stemming from each type of parton interaction and hence the gluon distribution functions come into play. In effect the GLAP equations are coupled for the quark and gluon distributions. In leading order the complete GLAP equations become:

$$\frac{dq_f(x,t)}{dt} = \frac{\alpha_s(Q^2)}{2\pi}\left[q_f(t)\otimes P_{qq} + G(t)\otimes P_{qG}\right](x) \tag{1.32}$$

$$\frac{dG(x,t)}{dt} = \frac{\alpha_s(Q^2)}{2\pi}\left[\sum_f q_f(t)\otimes P_{Gq} + G(t)\otimes P_{GG}\right](x) \tag{1.33}$$

The functions $P_{ij}(z)$ are the splitting functions. Explicit calculations of the leading order QCD diagrams show that[3]:

$$P^0_{qq}(z) = \frac{4}{3}\left[\frac{1+z^2}{1-z}\right]_+ \tag{1.35}$$

$$P^0_{qG}(z) = \frac{1}{2}\left[z^2+(1-z)^2\right] \tag{1.36}$$

$$P^0_{Gq}(z) = \frac{4}{3}\left[\frac{1+(1-z)^2}{z}\right] \tag{1.37}$$

$$P^0_{GG}(z) = 6\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \left(11 - \frac{n_f}{3}\right)\delta(1-z) \tag{1.38}$$

The physical interpretation of the splitting functions P_{ij} is given by the variation of the parton densities. For a variation in unit t, $(\alpha_s/2\pi)P_{ij}(y)$ is the probability of finding a parton i inside parton j with a fraction y of the parent momentum.

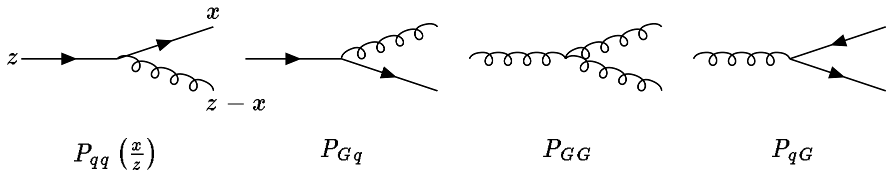


Figure 1.4: *The lowest order splitting functions. The splitting function* $(\alpha_s/2\pi)P_{qq}\left(\frac{x}{z}\right)$ *is the probability that by an increase of unit* t *a parton* q *with momentum fraction* z *emits a gluon and goes down to momentum* x*. The same for the splitting functions* P_{Gq}*,* P_{GG} *and* P_{qG}*. The last splitting function dynamically generates the sea quark distribution.*

The GLAP equations express the fact that a quark or gluon with momentum fraction x could have originated from a parent quark with momentum fraction z,

---

[3] The 'plus-prescription' regularizes the divergences of the splitting functions at $z = 1$. It is defined as

$$\int_x^1 dx\,\phi(x)\,[F(x)]_+ = \int_x^1 dx\,[\phi(x)-\phi(1)]\,F(x) - \phi(1)\int_0^x dx\,F(x) \tag{1.34}$$

where $x < z < 1$ via the splitting function $P_{qq}\left(\frac{x}{z}\right)$, see figure 1.4. Hence in the GLAP equation one integrates over the momentum fraction z from x to 1. One can interpret the equations as how the momentum distribution of the quark changes as Q^2 is increased, due to an increase of the resolution of the photon with which the partons are probed.

The GLAP equations predict a decrease of the PDF's at high values of x when the scale Q^2 increases, and an increase at low values of x. This can be pictured physically by a gain-loss competition of the parton distribution at a particular value of x. For example, increasing the scale Q^2, the distribution increases due to gluon radiation of partons with originally higher momentum $z > x$ via the splitting function $P_{qq}\left(\frac{x}{z}\right)$, or quark-anti-quark production via the splitting $P_{qG}\left(\frac{x}{z}\right)$. However, the parton distribution decreases because the partons with momentum fraction x can resolve into a parton with lower momentum fraction $x' < x$ and a gluon with momentum $x - x'$. At high values of x the second process dominates, and effectively the parton distributions decrease as a function of Q^2. At low values of x, more partons are gained by the first process and the parton distribution function increases as a function of Q^2, see figure 1.5.

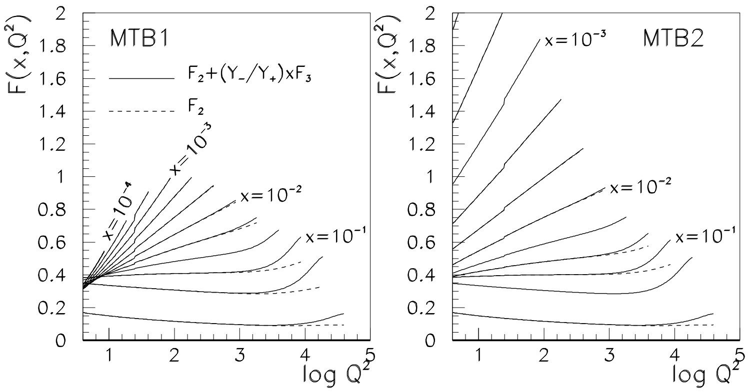


Figure 1.5: *The structure functions* $F_2 + \frac{Y_-}{Y_+}xF_3$*, as function of* Q_0^2 *for various values of* x*. The structure functions show the scaling violations. In the left and right plot we compare two different parametrizations of the parton distribution functions. See for a discussion of the parametrizations section 1.4.*

1.2.6 Low x phenomena

At small values of x both the gluon and quark distributions grow rapidly due to the splitting of partons in the GLAP equations. In this region most quarks are sea quarks and originate from small-x gluons. The asymptotic growth of the gluon distribution for very small x and large Q^2 is governed by [13]

$$xG(x,Q^2) \sim e^{\sqrt{\ln\left(\frac{\ln Q^2/\Lambda^2}{\ln Q_0^2/\Lambda^2}\right)\ln\left(\frac{1}{x}\right)}} \tag{1.39}$$

which indicates that the gluon distribution should exhibit an effective power increase for $x \longrightarrow 0$.

However, the gluon splitting function is singular for $x \longrightarrow 0$, see equation 1.38. This indicates that in higher order perturbation theory the terms which contain singular parts $\frac{1}{x}$ become important at low x. In fact, for moderate Q^2, they contribute more than the leading logarithmic terms in Q^2. As an alternative to the GLAP equations, which resum the leading logarithmic terms in Q^2, the Balitsky, Fadin, Kuraev, Lipatov (BFKL) [11] equation resums the singular $\frac{1}{x}$ terms. The BFKL equation is in effect an evolution equation in x, schematically expressed for the gluon distribution by

$$x\frac{dG(x,t)}{dx} = K \otimes G = -\lambda\ G(x,t) \tag{1.40}$$

with K a complicated gluon radiation function. The convolution is an integral over intermediate values[4] of Q^2, starting at a scale k_0^2. The range of validity of the BFKL equations is for small x and moderate Q^2. The maximum eigenvalue λ of the BFKL equation determines the gluon distribution as $x \longrightarrow 0$, for then $xG(x,t) \sim x^{-\lambda_{max}}$. Explicit calculations of the BFKL equation showed that $\lambda_{max} = 12\alpha_s \ln 2/\pi \sim 0.5$ for fixed α_s, or ranges between $0.4 < \lambda_{max} < 0.5$ for running α_s, depending on the cut-off scale k_0^2. The BFKL equation is the subject of many theoretical discussions [12]. Numerical simulation shows that for the low x behaviour of the gluon distribution the BFKL equation gives surprisingly similar results as the GLAP evolution.

Since at low x the gluon density is large, the gluons start to overlap in the proton at low x and recombine via the QCD process $gg \longrightarrow g$. This process limits the growth of the gluon density and ultimately ensures that unitarity of the DIS cross section is not violated. Gribov, Levin and Ryskin [14] have added a non-linear term in the evolution equations to describe this effect. The process of gluon recombination is not included in the (linear) GLAP or BFKL equations. The starting scale of the recombination process is not known. It is parametrized with a radius R in which the gluons recombine. It has been suggested that the proton need not be filled uniformly with gluons, but gluon recombination might start locally in the proton, for values of R smaller than the proton radius. These local high density partons in the proton are named 'hot-spots'. For even higher densities towards $x \longrightarrow 0$ the proton saturates and explicit QCD calculation become increasingly difficult, if not impossible.

---

[4] The dominant process in the BFKL evolution equation is radiation of a ladder of many soft gluons in successive steps. The BFKL convolution integrates over the intermediate values of Q_i^2, with $Q_{i+1}^2 > Q_i^2$, for each ladder step i.

We schematically sketch the validity of the various equations in in the (x, Q^2) phase space in figure 1.6.

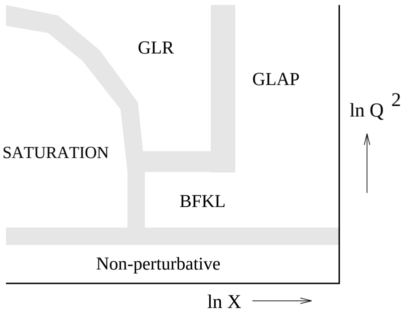


Figure 1.6: *A schematic sketch of the validity of various evolution equations in the (x, Q^2) phase space. For a discussion of the various equations, see text.*

1.3 Pre-HERA DIS data

Experimental data for lepton-hadron and hadron-hadron cross sections are used to measure the parton densities. The lepton-hadron DIS experiments provide the most accurate determination of the parton densities, although measurements of prompt photons (e.g. WA70 [15]) and lepton pair production (e.g. E605 [16]) in $p\bar{p}$ collisions are also used in the determination of the parton densities.

The pre-HERA DIS experiments have been performed at SLAC, CERN and FNAL. These are all fixed target experiments with either electron, muon or (anti-) neutrino beams.

The SLAC [17] experiments used electron beams in the energy range $3.65 < E_e < 20$ GeV in collision with hydrogen, deuterium and heavy nuclear targets. At CERN, a number of fixed target DIS experiments with muon beams have been performed. These are the EMC, BCDMS, NMC and the presently running SMC experiments [18, 19]. The beam energies range between $100 < E_\mu < 280$ GeV, and a variety of targets is used. The CERN experiments CDHSW, WA25 and WA59 [20] used wide band (anti-) neutrino beams on deuterium, iron and neon targets. At FNAL, the CCFR [21] experiment also measured DIS with an (anti-) neutrino beam on an iron target.

The experiments with an electron or muon beam on a hydrogen target measure the proton structure function F_2^p. If in addition data from a deuterium target is obtained, the neutron structure function F_2^n can be determined. Using heavy nuclear targets, one measures the average of the proton and neutron structure functions, the nucleon structure function. Using data of the neutrino experiments, the valence quark distributions can be determined via measurement of the structure function F_3, since this structure function is proportional to $q(x, Q^2) - \bar{q}(x, Q^2)$.

The maximum center-of-mass energy reached by the fixed target experiments is approximately $\sqrt{s} = \sqrt{2m_p E_{\mu,e}} \sim 25$ GeV. The kinematic range in x and Q^2 that is explored by the experiments is limited by the beam energies and by the requirement that the final state is measurable. The region in x that is explored by the fixed target experiments ranges approximately between $10^{-2} < x < 0.9$, and in Q^2 between $0.1 < Q^2 < 200$ GeV^2. For a complete overview of the status of the fixed target experiments we refer to [22].

1.4 Parametrizations of F_2

All parametrizations of the structure functions are based on the QCD evolution of the universal quark-distributions. The boundary conditions for the evolution in Q^2 are the parton distributions as a function of x at a particular value $Q^2 = Q_0^2$. These PDF's are evolved using the GLAP equations and thus yield a parametrization of the PDF's at all x and Q^2. The aim is to chose the parton distributions at Q_0^2 such that the predicted cross sections match all measured lepton-hadron and hadron-hadron data. In practice this means that at Q_0^2 a functional form for the quark-distributions is chosen, which contains a number of 'free', tunable parameters. The corresponding hadron cross sections and structure functions are computed. In a fit procedure the parameters are adjusted to obtain a parametrization of the parton densities that match the existing data [23].

There are a number of groups that perform these global fit procedures. They all use the next-to-leading (NLO) QCD predictions for the evolution and, for DIS, use the NLO convolutions with the hard parton scattering to obtain the structure functions. We briefly summarize the characteristics of three most commonly used phenomenological fits.

1.4.1 MRS parametrizations

The MRS parametrizations [24] are obtained from a global next-to-leading order QCD fit, starting from $Q_0^2 = 4$ GeV^2. Results from the latest NMC and CCFR data are included in three sets, the $S0'$, $D0'$ and $D-'$ parametrizations (the parametrizations without primes were obtained using preliminary NMC results). Special attention is paid to the Gottfried Sum Rule (GSR), which has been measured by the NMC experiment. This sumrule yields a prediction for the integrated difference of the proton and neutron structure functions F_2^p and F_2^n as a function of the integrated difference between the $\bar{u}$ and $\bar{d}$ parton densities:

$$I_{GSR} \equiv \int_0^1 \frac{dx}{x}\,(F_2^p - F_2^n) = \frac{1}{3} + \frac{2}{3}\int_0^1 dx\,(\bar{u} - \bar{d}) \tag{1.41}$$

Experimentally NMC finds:

$$I_{GSR}(NMC) = \int_{0.004}^{0.8} \frac{dx}{x}\,(F_2^p - F_2^n) = 0.227 \pm 0.007 \pm 0.014 \tag{1.42}$$

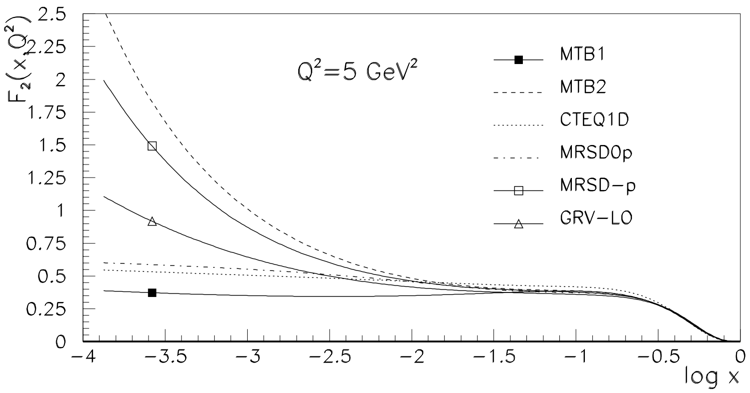


Figure 1.7: *Various parametrizations of* F_2 *as function of* x *at* $Q^2 = 5$ *GeV*^{2}.

Since the measured $I_{GSR}(NMC)$ is significantly smaller than $\frac{1}{3}$, the NMC measurement indicates that $\bar{d} > \bar{u}$. The CCFR measurement show that roughly $\bar{s} = \frac{1}{4}(\bar{u}+\bar{d})$, hence MRS parametrize the sea-quark distributions at Q_0^2 as:

$$\begin{aligned} 2\bar{u} &= 0.4\,S - \Delta \\ 2\bar{d} &= 0.4\,S + \Delta \\ 2\bar{s} &= 0.2\,S \end{aligned} \tag{1.43}$$

with $S(x)$ the total sea-quark distribution. With a non-zero $\Delta(x)$ distribution the NMC measurement of the GSR is incorporated. In their nomenclature this parametrization is labeled by 'D' (in $D0'$ and $D-'$). A zero $\Delta(x)$ distribution does not reproduce the NMC measurement of the GSR, but still provides an acceptable fit for the other measurements (it should be noted that the measurement of the GSR is dominated by the low x behaviour of the structure functions). This parametrization is labeled by 'S' (*same*, as in the $S0'$ parametrization).

At $Q^2 = Q_0^2$ the distributions for the sum of the up and down valence, the down valence, the total sea, Δ and the gluon are parametrized with the functional forms

$$xq(x, Q_0^2) = Ax^{\delta}(1-x)^{\eta}(1+\epsilon\sqrt{x}+\gamma x) \tag{1.44}$$

The charm and bottom quark distributions are generated dynamically by the evolution equations taking $c(x, Q_0^2) = b(x, Q_0^2) = 0$. In total there are 16–19 free parameters in the fit procedure.

The low x behaviour of the gluon and sea distributions is parametrized by the same exponent δ_G, such that $xS(x, Q_0^2) \sim x^{\delta_G}$ and $xG(x, Q_0^2) \sim x^{\delta_G}$ as $x \rightarrow 0$. This is justified by the fact that the low x behaviour is driven by the gluon distribution. With the pre-HERA measurements the parameter δ_G is not restricted.

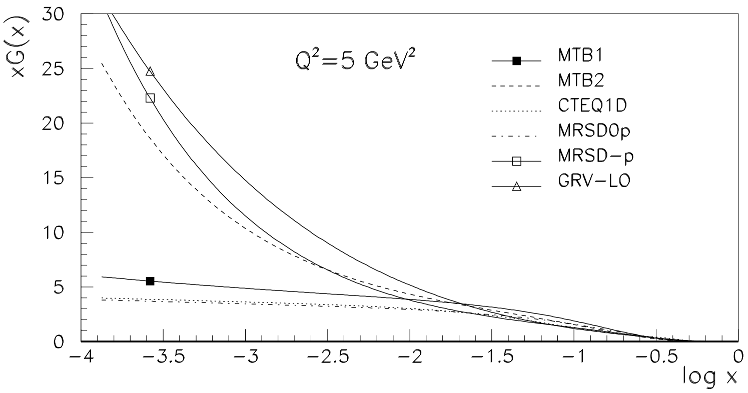


Figure 1.8: *Various parametrizations of the gluon distribution as function of* x *at* $Q^2 = 5$ GeV^2.

Hence the MRS group produced two parametrizations with extreme values of δ_G. One with a 'flat' sea and gluon distribution, with $\delta_G = 0$ (the $D0'$ parametrization), and one with the extreme singular sea and gluon distributions, $\delta_G = -\frac{1}{2}$ (the $D-'$ parametrization). The NLO MRS$D0'$ and MRS$D-'$ parametrizations of the F_2 structure function and the gluon distribution are shown in the DIS scheme in figures 1.7 and 1.8 respectively.

1.4.2 CTEQ parametrizations

The CTEQ, or former MT, parametrizations are based on NLO QCD evolution of the PDF parametrizations. In this thesis we use two extreme parametrizations of the parton distributions, MTB1 and MTB2 [25]. They are the result of a global analysis of the BCDMS, CDHSW, E288 and E605 data. At $Q_0^2 = 4$ GeV^2 each quark flavour is parametrized with the functional form

$$xq(x, Q_0^2) = e^{A_0} x^{A_1} (1-x)^{A_2} \ln^{A_3}\left(1 + \frac{1}{x}\right) \tag{1.45}$$

In the MTB1 and MTB2 no flavour differentiation is assumed, and a $SU(3)$ symmetric sea distribution of the $\bar{u}$, $\bar{d}$ and $\bar{s}$ is enforced. The MTB1 parametrization is characterized by a flat behaviour of the gluon and sea quarks, whereas MTB2 is steeply rising towards low values of x, see figure 1.8. These parametrizations are particularly useful because they set the extremes of the structure function F_2 expectations at HERA. We plot the F_2 structure function of the MTB1 and MTB2 parametrization in figure 1.7.

The more recent fits of the CTEQ group [26] include the CCFR and NMC results. The $SU(3)$ flavour symmetry is no longer imposed, and $\bar{u}$, $\bar{d}$ and $\bar{s}$ are evolved separately. In total approximately 30 free, tunable parameters are included. As a result they found that the best fit to the data is obtained with a $\bar{s}$ which is much softer than the $\bar{u}$ and $\bar{d}$ distributions, and that $\bar{d}-\bar{u}$ changes sign as a function of x. In figures 1.7 and 1.8 we show the NLO F_2 and gluon parametrization in the DIS scheme, the CTEQ1D parametrization.

1.4.3 GRV parametrizations

A somewhat different approach in the parametrization of the structure functions is taken by GRV [27]. They noticed that the PDF's become unphysical (i.e. negative) when the GLAP equations are used to evolve the PDF's backwards to lower values of $Q^2 < Q_0^2$. As an alternative, they postulate that at a low scale Q_0^2 the nucleon consists almost entirely of valence-like quarks and valence-like gluons. The region $x \ll 1$ is not subjected to a parametrization but fixed by demanding $xq(x,\mu^2) \longrightarrow 0$ as $x \longrightarrow 0$. This is achieved by using the functional form

$$xq(x,\mu^2) = Ax^{\alpha}(1-x)^{\beta} \tag{1.46}$$

with $\alpha > 0$, for the gluon, the valence and $\bar{u} = \bar{d}$ sea quarks. The remaining sea quarks are assumed to be zero at Q_0^2. The total of approximately 10 free parameters of the PDF's are fitted to the data. They found a convenient parametrization of the data when the scale Q_0^2 was chosen to be $Q_0^2 = 0.25$ GeV^2 (for the leading order GLAP evolution equations). The growth of the PDF's at low values of x and large Q^2 is generated dynamically and the behaviour of the PDF's is completely determined by the GLAP equations. The GRV parametrization has no *free* parameters, except for the value Q_0^2, to describe the low x behaviour.

Analysis repeated with the NLO QCD evolution indicated that perturbative stability is achieved, despite the fact that the GRV parametrization start the evolution already at $Q_0^2 = 0.25$ GeV^2. In figures 1.7 and 1.8 we plot the F_2 and gluon distribution in LO QCD evolution. The figure shows that evolution from $Q_0^2 = 0.25$ GeV^2 to $Q^2 = 5$ GeV^2 already predicts a rising of F_2 towards low values of x.

1.5 DIS at HERA

HERA extends the lever arm of the measurement of the parton distributions towards lower values of x and higher values of Q^2. During 1992, HERA operated with an electron beam of energy $E_e = 26.7$ GeV and a proton beam with an energy $E_p = 820$ GeV. The center-of-mass energy at HERA is

$$\sqrt{s} = 2\sqrt{E_e E_p} = 296 \text{ GeV}. \tag{1.47}$$

For neutral current DIS, the electron will enter the detectors if Q^2 is above a few GeV^2. The region in x and Q^2 is bounded by $y = 1$. The measurable region in x is

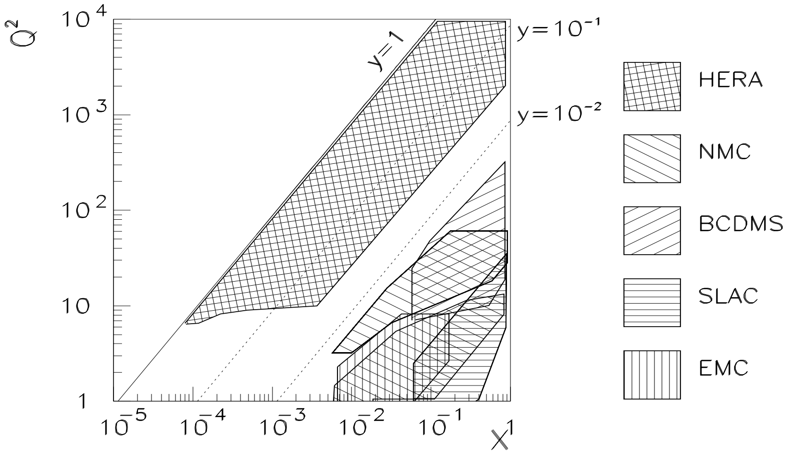


Figure 1.9: *Approximate regions in the* (x, Q^2) *that are covered by the various DIS experiments. The lines* $y = 1$, $y = 0.1$ *and* $y = 0.01$ *are drawn for the HERA cms. The fixed target experiments explored the low* Q^2 *and high* x *region. HERA, with its high cms energy, accesses the region of very small* x *and high* Q^2 *values.*

therefore $10^{-4} < x < 1$. The maximum Q^2 that can be reached is $Q^2_{max} = 8.75 \cdot 10^4$ GeV^2.

The region in the (x, Q^2) phase space that is accessible with the HERA collider is drawn in figure 1.9. We also show the accessible regions for some fixed target experiments. As will be explained, in the region of low values of y, the determination of the cross section becomes inaccurate. Hence there is at present no overlap with fixed target experiments. In order to obtain overlap, it is proposed to lower the beam energies.

The HERA collider allows exploration of a kinematic region two orders of magnitude lower in x and two orders of magnitude higher in Q^2 than previous experiments. In the low x region the predictions of the structure function, and hence of the cross section, become different. In figure 1.10 we present the cross sections for the NC and CC DIS for $Q^2 > 10$ GeV^2, integrated over the indicated bins. In each bin the upper number corresponds to the high parametrization of MTB2, and the lower number to MTB1. Whereas in the high x region the two parametrizations predict more or less the same cross section, in the low x region the predictions differ by more than a factor of three. The total cross section for $Q^2 > 10$ GeV^2 is ~80 *nb* for MTB1 and ~127 *nb* for MTB2 for neutral current DIS. The charged current cross section is much lower, ~69.3 *pb* and ~71.3 *pb* respectively. It is in the low x region that a measurement of the neutral current DIS cross section, even with

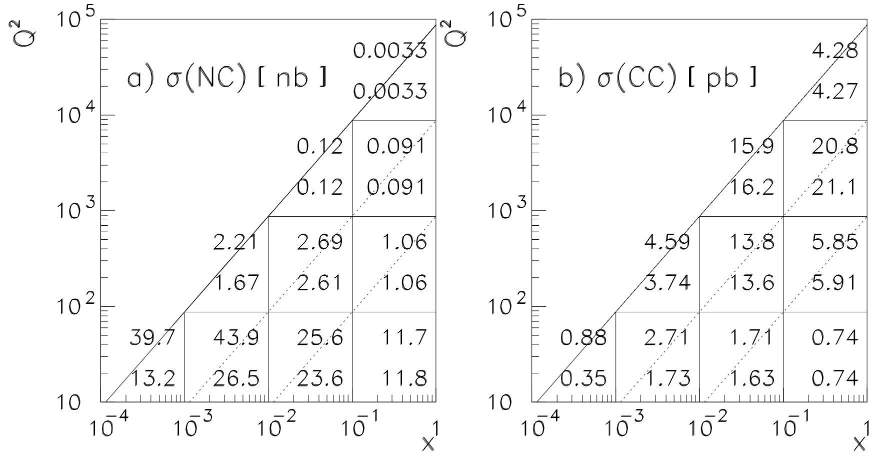


Figure 1.10: *Integrated Born cross sections over the indicated bins in* x *and* Q^2 *for the neutral current DIS (plot a) and charged current DIS (plot b). Note that in plot (a) the cross section is listed in nb, in plot (b) in pb. The first number in each bin is the prediction according to the* $MTB1$ *parametrization, the second number for the* $MTB2$ *parametrization.*

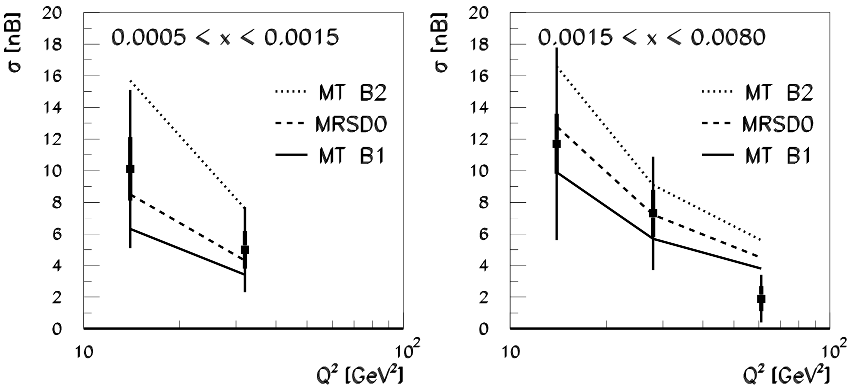


Figure 1.11: *The measured Born cross section, integrated over the bins, of the initial 1992 running period in 2* x*-bins. The four bins-edges for the three bins in* Q^2 *are at* $10, 20, 40, 100$ GeV^2 *respectively. The systematic error (outer error bar) is added linearly to the statistical error (inner error bar).*

limited statistics, provides new information about the structure of the proton.

In an initial series of runs at HERA in summer 1992, the ZEUS detector recorded its first neutral current DIS data. The results have been published in ref [49]. This data sample corresponded to an integrated luminosity of 2.4 nb^{-1}. In figure 1.11 the measured cross sections, integrated over the bins (indicated in the figure), are compared with the predictions of three different parton distribution parametrizations. The errors are too large to favor any of them. In this thesis we present the analysis of the F_2 structure function with a larger data sample in detail, and show that some parametrizations can be excluded.

1.6 Radiative corrections

The expression for the cross section of the ep-interaction is modified due to additional electroweak interactions. Therefore these additional contributions necessary needs to be separated from the lowest order Born cross section, which contains the most direct information of the internal structure of the proton, see section 1.1.2.

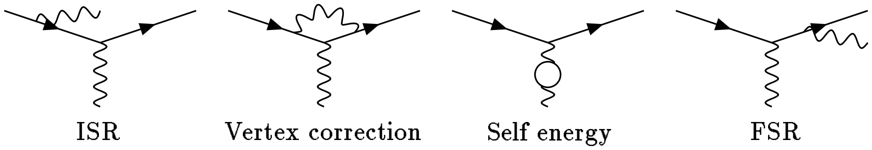


Figure 1.12: *The four types of QED radiation to the lepton line.*

The corrections originate from the emission of additional real and virtual bosons from either the lepton or quark line. The quarkonic corrections can be absorbed into the quark distribution functions, very much like the gluon radiation, whereas corrections from interference between the lepton and quark line proved to be small [28]. This leaves the leptonic corrections which can be classified further into virtual corrections and emission of real photons. The dominant contribution to the virtual corrections originates from the QED vertex correction and fermion contributions to the γ and Z_0 self energies. In the leading logarithmic approximation the region in phase space where the emitted, real photon is collinear with the initial (final) lepton is called initial (final) state radiation ISR (FSR), see figure 1.12. There is a third contribution, called 'Compton scattering (CS)', which can be viewed as a hard Compton scattering of the lepton on a quasi-real photon originating from the incoming proton[5].

In general, when a photon is emitted from the lepton line, the emission of momentum will shift the kinematics of the event. The kinematic variables with which the proton is probed, $(x, y, Q^2)_h$, are defined at the hadronic vertex of the scattering

---

[5] The Compton scattering is described by the same Feynman diagrams as the ones for ISR and FSR, and is a result of the approximation to handle the large cross section in the phase space region where Q_h^2 is small.

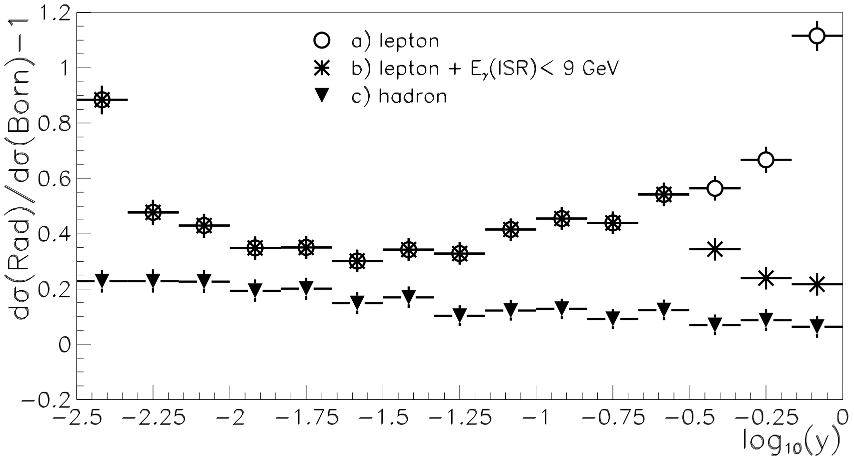


Figure 1.13: *Radiative corrections for non-detector-smeared kinematic variables as function of y, integrated over all values of x. In the figure we plot the ratio $d\sigma(RAD)/d\sigma(Born) - 1$. In* (a) *the variables $(x, Q^2)_l$ are used. The same in* (b) *but in addition a cut $E_\gamma < 9$ GeV is made. Figure* (c) *gives the radiative corrections for $(x, Q^2)_h$.*

amplitude, in contrast to the 'apparent' variables $(x, y, Q^2)_l$, determined from the final state lepton (without the emitted photon):

$$\begin{array}{llll} Q_l^2 & = -(\mathbf{k}-\mathbf{k}')^2 & Q_h^2 & = -(\mathbf{P}-\mathbf{P}')^2 = -(\mathbf{k}-\mathbf{k}'-\mathbf{l})^2 \\ x_l & = Q_l^2/(2\mathbf{P}\cdot(\mathbf{k}-\mathbf{k}')) & x_h & = Q_h^2/(2\mathbf{P}\cdot(\mathbf{P}'-\mathbf{P})) \end{array} \tag{1.48}$$

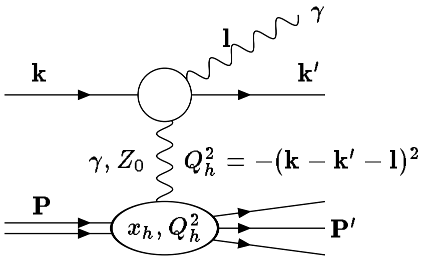


Figure 1.14: *Notation for radiative lepton nucleon scattering.*

The difference between the two sets of variables is due to the photon radiation only. Whereas $x_h \geq x_l$ (and $y_h \leq y_l$), the hadronic momentum transfer Q_h^2 can be smaller or larger than the leptonic Q_l^2. Since the magnitude of the cross section is determined by the photon propagator $1/Q_h^2$, photon emission that results in $Q_h^2 \ll$

Q_l^2 is enhanced compared to the Born cross section $1/Q_l^2$. This is the main reason that radiative corrections to the cross section as a function of leptonic variables can get large, whereas corrections as a function of the true hadronic variables remain of moderate size. Hard ISR shifts the measured value of Q_h to a lower value compared to Q_l. This implies that for the unfolding procedure of F_2 it is necessary to have knowledge of the structure function for values $Q_h^2 = Q_l^2$ down to $Q_h^2 = 0$, which is outside the scope of the experiment. Several parametrizations of the structure functions exist which extend to low values of Q_h^2 and have a smooth transition between the deep-inelastic structure functions and soft photo-production.

The effect of radiation is shown in figure 1.13, where the radiative corrections to the neutral current DIS, $d\sigma(RAD)/d\sigma(Born) - 1$, is drawn. As can be seen, the correction to the cross section determined with the lepton variables $(x, y, Q^2)_l$ (line -a-) peaks at high and low values of y_l.

However, it is possible to define a cut on experimentally obtained variables, the δ-cut as will be described in section 4.4.4, which in effect removes events with a catastrophic initial state photon, $E_\gamma(ISR) > 9$ GeV [29]. This cut largely removes the uncertainty stemming from the structure functions at low values of Q_h^2, as it reduces the corrections at high values of y_l considerably (figure 1.13, line -b-).

When the variables from the hadronic vertex are used (line -c-) the enhancement of the cross section due to the purely virtual corrections is visible.

1.7 Event generators

To investigate efficiencies, biases and acceptances of DIS in the experimental environment, the need for event simulation is essential. The event generators simulate with Monte Carlo techniques the DIS events, and produce four-vectors of the complete final state. The four vectors are fed into the detector simulation in order to obtain the detector response functions.

The event generators simulate events according to the factorization scheme. The cross section for partons involved in the hard interactions are calculated with the matrix elements. A separate FORTRAN library [30] is needed for input of the parton distribution functions. For a satisfactory description of the final state, this simulation is performed in two steps. In a first, intermediate step the partons with large virtuality produce QCD inspired cascades of partons, the parton showers, until a lower limit on the virtuality is reached. In the second step the hadronization of the low virtuality partons is simulated. The hadronization includes decays of unstable hadrons.

HERACLES [31] is an *ep* event generator at the parton level, that includes the complete calculation of first order electroweak radiation. It simulates the Born-level together with the complete one-loop virtual corrections, the initial and final state photon radiation and inelastic Compton scattering. Because the program provides the radiative photon emission on an event by event basis, the corrections for radiative effects, i.e. for example the shifts from $(x, Q^2)_l$ to $(x, Q^2)_h$, are indistinguishable from other detector response effects. Hence using the HERACLES event generator, detector smearing and the effect of real photon emission is corrected simultaneously. HERACLES needs an interface with a program that simulates the QCD cascades and the hadronization.

LEPTO [32] is also an event generator at the parton level. The matrix elements of the first order QCD processes, the gluon radiation $\gamma^\star g \longrightarrow qg$ and boson-gluon fusion $\gamma^\star g \longrightarrow q\bar{q}$, can be included optionally. Higher order QCD processes have been implemented using parton showers. The program does however not include radiative corrections. LEPTO needs an interface with JETSET for hadronization.

ARIADNE [33] is a program that simulates the QCD cascades or parton showers using a colour dipole model. In this model, a colour dipole is formed between the scattered quark and the proton remnant, which acts as an antenna and emits partons. Since the emitted partons also carry colour, new dipoles are created from which softer partons are emitted.

HERWIG [34] simulates initial and final state QCD cascades, where angular ordering in the gluon emission is required. The emission stops if the virtuality of the gluons becomes low. The programs keeps track of the colour flow, and forms colourless objects from the generated quarks and gluons.

JETSET [35] is the implementation of the LUND string fragmentation model. The string represents the colour flow between the partons and fragments into hadrons. The program contains many phenomenological parameters that control this process. They are tuned to make the final state distribution agree with previous fixed target DIS data and data from e^+e^- machines.

From the above it is clear that many combinations of these programs are possible (even including the fragmentation of the colourless objects in HERWIG using JETSET). In this thesis we use the combination of HERACLES with ARIADNE as default. Other combinations are used to estimate systematic effects of radiation and fragmentation.

Chapter 2

HERA and ZEUS

In this chapter we briefly discuss the HERA electron-proton collider and its performance in 1992. We explain the principle of measuring the luminosity of the colliding beams and discuss the 1992 integrated luminosity for the ZEUS experiment. The ZEUS detector, its properties and general layout, is summarized thereafter. We review the high resolution calorimeter, the main component used in the subsequent analysis, in somewhat more detail.

2.1 The 'Hadron-Elektron-Ring-Anlage' (HERA)

date	Milestone
jul 1981	HERA proposal $E_e : 10 - 30$ GeV, $E_p : 300 - 820$ GeV.
apr 1984	Approval of HERA
jul 1986	Recommendation to approve H1 and ZEUS by the Phys. research committee
aug 1987	HERA tunnel drilled.
aug 1988	Electrons circulating in HERA.
apr 1990	First proton octant at superconducting temperatures.
nov 1990	Last superconducting dipole installed. HERA 'complete'.
apr 1991	Protons injected and stored.
jun 1991	Superconducting cavities for e-beam, 30 GeV reached.
oct 1991	First collision 12 and 26 GeV $E_e \times 480$ GeV E_p.
sep 1991	Cosmic runs for the detectors.
may 1992	First HERA collisions with 26.7×820 GeV. Beginning of physics runs.
sep 1992	Observation of 60% transverse electron polarization.

Table 2.1: *Milestones in the completion of HERA [40].*

HERA (Hadron-Elektron-Ring-Anlage) is an electron proton colliding beam machine with the design center-of-mass energy of $\sqrt{s} = 314$ GeV, which is more than

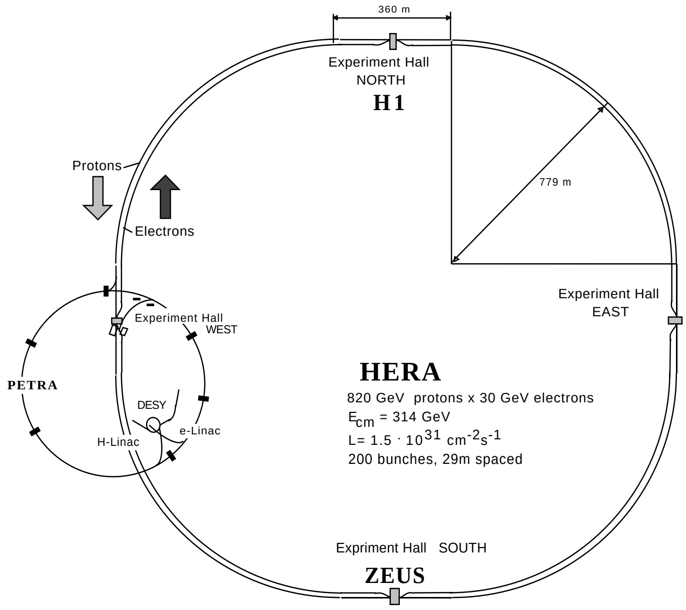


Figure 2.1: *The HERA accelerator complex at DESY, Hamburg.*

an order of magnitude higher than is reached by the fixed target experiments. Its construction was approved in April 1984, and eight years later, in 1992, the first collisions between colliding electrons and protons were observed in the two detectors, ZEUS and H1. The most important historical milestones in the construction of HERA are listed in table 2.1.

The HERA ring is located partially on, partially around the DESY (Deutsches-Elektron-Synchrotron) laboratory in Hamburg, Germany. HERA consists of a tunnel in the form of an approximate circle with circumference of 6335 m (see figure 2.1). Inside the tunnel the electron and proton beams are stored in two separate beampipes. The proton-ring has 104 cells of superconducting magnets for bending and focus. One of the straight sections contains a 52 MHz and a 208 MHz RF cavity system, that accelerates the protons up to 820 GeV. The electron ring consists of 416 normal magnet modules. Acceleration to about 27 GeV is achieved by 82 normal cavities and the additional 3 GeV to reach the design energy of 30 GeV for the electrons, is achieved by 16 additional superconducting cavities. The proton and electron rings intersect at four points to allow for ep interactions.

The electrons and protons are preaccelerated before injection into the HERA machine. Two LINAC's (LINear Accelerators) accelerate electrons or positrons up to 220 and 450 MeV respectively. The electron synchrotron DESY II preaccelerates the electrons to the injection energy for PETRA of 7 GeV. PETRA, a ring of 2300 m circumference, is an injector for electrons or positrons in HERA at $E_e = 12$ GeV.

parameter	typical value in 1992	design value
proton energy E_p	820 GeV	820 GeV
electron energy E_e	26.7 GeV	30 GeV
number of bunches	10	210
time between crossings	96 ns	96 ns
P injection energy	40 GeV	40 GeV
E injection energy	12 GeV	12 GeV
P current/bunch	0.2 mA	0.8 mA
E current/bunch	0.2 mA	0.3 mA
E bunch length	1-2 cm	0.8 cm
P bunch length	45 cm	11 cm
max luminosity	$1.5 \cdot 10^{29}\ cm^{-2}s^{-1}$	$1.5 \cdot 10^{31}\ cm^{-2}s^{-1}$
$\int_{year} dt\ \mathcal{L}$	30 nb^{-1}	100 pb^{-1}

Table 2.2: *For some HERA parameters the typical values of 1992 are compared with the design values.*

The protons are accelerated to an energy of 40 GeV in the same PETRA ring before they are injection into HERA. A new proton synchrotron of 317 m circumference, DESY III, was constructed to reach the 7.5 GeV proton energy which is needed for injection in PETRA.

The H1 and ZEUS experiments are located in two of the intersection regions, the north and south hall respectively. In figure 2.1 we show a layout of the DESY accelerator complex.

HERA operates with 220 electron and proton buckets, separated in time by 96 ns. According to the design, 200 buckets are filled with particles, and these bunches contain approximately $4 \cdot 10^{10}$ electrons or 10^{12} protons. With these numbers an instantaneous luminosity of $1.5 \cdot 10^{31}\ cm^{-2}s^{-1}$ can be reached, leading to the integrated design luminosity of $\sim$ 100 pb^{-1}/year. The high rate of the bunch crossing, 10.4 MHz, has profound implications on the technical requirements of the two detectors.

2.1.1 Performance of HERA in 1992

In May 1992, operation of HERA started for physics with electron energy $E_e = 26.7$ GeV and proton energy $E_p = 820$ GeV. Every 2-3 hours a new electron fill took place. The lifetime of the proton beam was generally much larger, typically 50 hours. The time consumed between the electron beam dump, refilling of electrons, stabilization and beam focus was typically several hours. The total electron current in the 10 bunches was 1–2 mA, the proton current $\sim$2 mA. A comparison between the design values and the typical values reached in 1992 for the most important HERA parameters are listed in table 2.2.

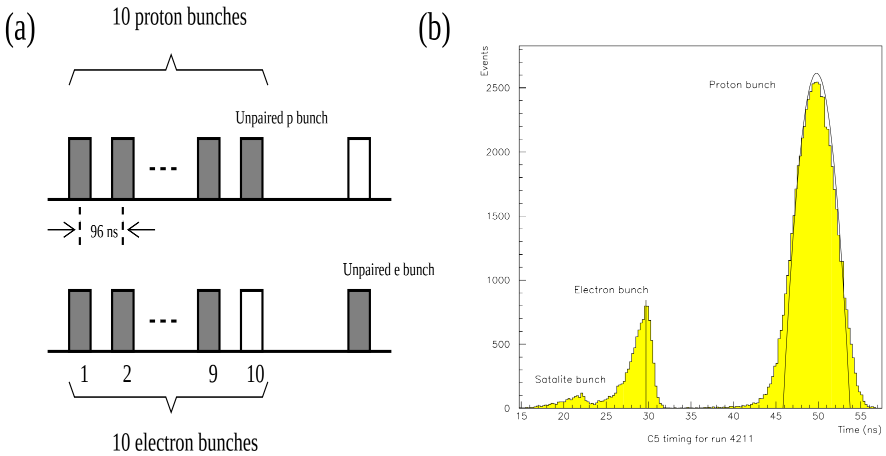


Figure 2.2: *In figure (a) we show the electron-proton bunch structure of HERA in 1992. The bunches are hatched and the buckets without particles are left blank. The unpaired e- and p-bunches are the pilot bunches, used for beam related background studies.*
In (b) the **C5** *counter histogram for a typical run in 1992 is plotted. The proton and electron arrival times are clearly visible. The electron satellite bunch is shifted by approximately 8 ns with respect to the electron arrival time. The time offset on the* x-axis *is arbitrary.*

During most of the 1992 running period, 10 electron and proton bunches were accelerated, separated in time by the design value. Nine pairs of electron and proton bunches were used for *ep* collisions. In addition one proton and one electron bunch was left unpaired. These so called ‘pilot’-bunches allow studies of beam related backgrounds. A schematic view of the bunch configuration is shown in figure 2.2a.

Figure 2.2b shows a timing distribution from a scintillator counter (**C5**), placed at approximately 3 *m* from the intersection point. At this position two clear separated signals can be observed, one originating from the halo particles accompanying the proton bunch, and one from the electron bunch. The distribution of the electron bunch shows a separate structure, the ‘satellite peak’. It is due to not fully efficient containment inside the nominal RF bucket. Approximately 12% of the electron beam current is therefore delayed by 8 *ns*. This delay is equivalent to a shift in the nominal interaction point of 1.2 *m*, and hence these events did not enter our data sample. Implications on the luminosity measurement from these satellite events is discussed in the next section.

2.1.2 Coordinate system

In the global HERA and ZEUS coordinate system the initial state protons move towards the positive **z**-axis. The origin of the ZEUS coordinate system is defined at the nominal interaction point. The coordinates **x** and **y** are defined in figure 2.3.

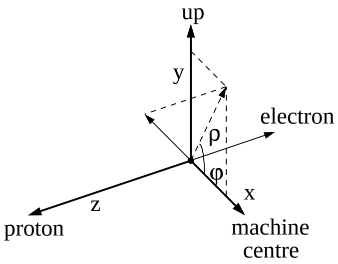


Figure 2.3: *Global HERA and ZEUS coordinate system.*

2.2 Luminosity measurement

For the measurement of the luminosity at ZEUS, two separate detectors, LUMI-γ and LUMI-e, are installed at $\mathbf{z} = -106\ m$ and $\mathbf{z} = -34.7\ m$ respectively. They are used to measure the rate of small angle, hard photon bremsstrahlung that is produced in the Beithe-Heitler (B-H) [38] process

$$ep \longrightarrow e'p\gamma. \tag{2.1}$$

The theoretically predicted value of the cross section for this process is large and known to a high precision. The radiative corrections are calculated to be small, -0.3% [39]. At full HERA luminosity the rate of this process is of the order of 50 kHz. This indicates that this Bethe-Heitler process provides an excellent way of determining the luminosity at HERA.

A schematic view of the two LUMI detectors is given in figure 2.4. Via a thin beam-pipe exit window the hard photons pass a carbon-lead filter shield ($1X_0$), that removes the softer synchrotron radiation, and deposit their energy in the γ-counter. The LUMI-γ counter is a lead-scintillator calorimeter. The energy-resolution of the γ-counter is measured to be $\sigma(E_\gamma) = 0.18\sqrt{E_\gamma}$ under test beam conditions. The position of the photon can be determined with a resolution better than 3 mm. The acceptance of this detector, for photons with $E_\gamma > 5$ GeV emitted collinearly to the electron at the event vertex, is 98%.

The LUMI-e calorimeter measures the $E_{e'}$ spectrum for electrons with a scattering angle $\theta_e \leq 6\ mrad$. The electrons with decreased energy are bent by the machine bending magnets, leave the electron beam pipe and hit the lead-scintillator calorimeter. The acceptance of the LUMI-e is around 70% for electrons in the energy range

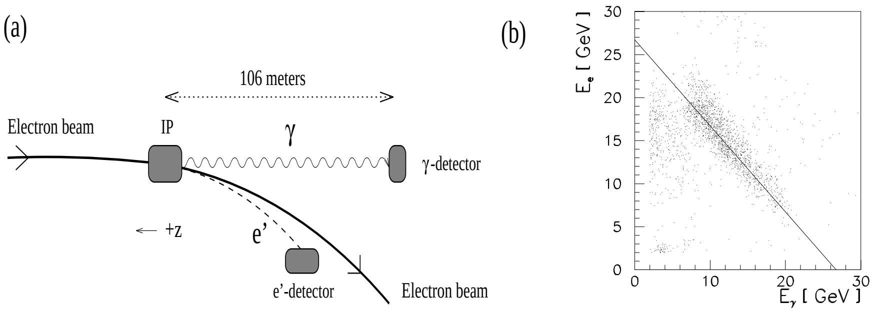


Figure 2.4: *Plot (a) shows a schematic view of the luminosity detectors. In the interaction point (IP) the Beithe-Heitler reaction takes place. The photon γ is detected at $z = -106$ m in the γ-counter. The electron e' is bend due to energy-loss in the reaction and detected in the e'-detector. The proton beam line is not drawn in this figure.*
Plot (b) shows the measured energy in the LUMI-e counter versus the energy deposit in the LUMI-γ counter. The line $E_{e'} + E_\gamma = 26.7$ GeV is drawn.

$0.35E_e < E_{e'} < 0.65E_e$ GeV. The sum of the energy-depositions in the LUMI-γ and the LUMI-e for the bremsstrahlung events is equal to the initial electron beam energy. A sample of this type of events is shown in figure 2.4b.

Because the acceptance is less well understood in the LUMI-e, in 1992 the luminosity was determined using the LUMI-γ only. The coincidence rate of the LUMI-e and LUMI-γ, $R(\gamma \cdot e)$, served only as a cross check in the determination of the luminosity.

The luminosity is determined from the measured rate $R_\gamma[\text{meas}]$ of hard photons, $E_e > 5$ GeV, in the LUMI-γ. The measured rate suffers from a number of backgrounds, which need to be subtracted. The first source of background is the bremsstrahlung of electrons with residual gas in the beam-pipe. This type of background is estimated using the pilot bunches. The number of bremsstrahlung events from the pilot bunch is scaled with the ratio of the total electron current I_e^{tot} and the pilot current I_e^{pilot} to obtain the estimated total electron-gas bremsstrahlung rate. The second background stems from the satellite bunches. The rate-fraction that belongs to these satellite bunches, $\sim 6\%$, is determined from the current fraction in the satellite and the known variation of specific luminosity as a function of the vertex position. This rate is also subtracted. Finally, an overall factor f is applied that corrects for multiple B-H processes in a single bunch crossing:

$$R_\gamma[\text{corr}] = \left(R_\gamma^{tot}[\text{meas}] - \frac{I_e^{tot}}{I_e^{pilot}} R_\gamma^{pilot} - R_\gamma^{satellite} \right) f \tag{2.2}$$

This corrected rate is then converted to a value for the luminosity using the observed B-H cross section that is corrected for detector response, σ_γ^{MC}. The detector

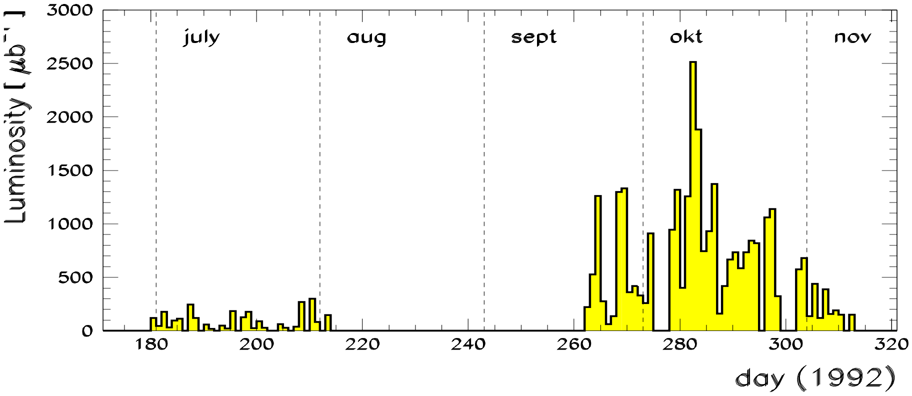


Figure 2.5: *The luminosity as determined offline for the year 1992.*

response function, $r(E_\gamma)$, is obtained with a Monte Carlo simulation:

$$\begin{aligned} \mathcal{L} &= R_\gamma[\text{corr}]/\sigma_\gamma^{MC} \\ \sigma_\gamma^{MC} &= \int \frac{d\sigma(\text{B-H})}{dE_\gamma} r(E_\gamma) dE_\gamma . \end{aligned} \tag{2.3}$$

2.2.1 Luminosity in 1992

HERA delivered luminosity for physics analysis to the experiments during two periods in 1992. The first period started on May 31 1992 and coincided with the first HERA collisions. This so called 'summer'-period ended on august, 2nd. The second, 'fall'-period started on September 19th and lasted until November 9th. Considerable improvements in the HERA operation were made between the summer- and fall-periods. Better beam tuning and larger beam currents increased the instantaneous luminosity with approximately a factor of 10. The proton bunch length decreased from 1 *m* to ~45 *cm*. Figure 2.5 show the integrated luminosity/day collected in the ZEUS detector versus the day in 1992. These data are based on on-line luminosity measurements and are accurate to approximately 10%.

During the summer period, ZEUS collected 2.1 nb^{-1} of useful data. The DIS events from this period have been analyzed and published in [49]. The luminosity that ZEUS collected during the fall period in 165 useful runs was:

$$\int_{fall\ period} dt\, \mathcal{L} = 24.7\, nb^{-1} \pm 5\% \tag{2.4}$$

$$\text{Main sources of error:} \quad \left\{ \begin{array}{ll} 1\% & \text{electron-gas subtraction} \\ 3\% & \gamma\text{-energy calibration} \\ 3\% & \gamma\text{-acceptance} \\ 1\% & \text{satellite bunch subtraction} \end{array} \right. \tag{2.5}$$

The error on the luminosity measurement is mainly due to the systematic uncertainties that we list. The statistical error is negligible.

The determination of the proton structure function F_2 from this running period is published in [50] and is the main subject of this thesis. Note that the total collected luminosity in 1992 is still more than three orders of magnitude below the design luminosity/year (the 1993 running period showed already an increase in luminosity of one order of magnitude).

2.3 The ZEUS detector

Approximately 450 physicists from 52 institutes and 11 countries participate in the ZEUS collaboration. The ZEUS detector, one of the two experiments at the HERA collider, is located in the south hall of the HERA ring, thirty meters underground. The ZEUS detector measures approximately 20 by 11 by 12 meters and has a total weight of about 3600 tons. A cross section of the detector along the beam line is shown in figure 2.6. Figure 2.7 shows a cross section of the ZEUS detector perpendicular to the beam line. In this section we briefly describe the detector components, proceeding outwards from the interaction point. For a full and detailed overview we refer to [57, 58].

The Inner Tracking System consists of the vertex detector (**VTX**), the central drift chamber (**CTD**), the forward and rear tracking detectors (**FDET** and **RTD**) and the magnet coil (**MAGNET**). This system enables the accurate reconstruction of the direction and momentum of charged particles with a coverage of almost 4π *sr*. The tracks are also used for particle identification and for the determination of the primary and secondary vertices.

- **VTX**: The purpose of the vertex detector is to measure the event vertex and possibly secondary vertices, to improve the momentum and angular resolution of charged particles and to contribute in pattern recognition. It is a cylindrical drift chamber with 1.6 *m* long wires parallel to the beam line. The spatial point resolution in the $r - \phi$ plane ranges between 35 and 55 μm.

- **CTD**: Information from the central drift chamber is used to reconstruct the direction and momentum of charged particles with high precision, and to measure the energy loss dE/dx that is used for particle identification. The **CTD** jet chamber consists of 72 cylindrical drift chamber layers organized into nine superlayers. Five superlayes have wires parallel to the beam line, four are stereo layers which have wires with a small angle with respect to the beam line. The z positions of the hits are measured with a resolution between 1.0 and 1.4 *mm* using the stereo layers, and 4 *cm* using the z-by-timing information of the remaining superlayers. The design hit resolution in the $r - \phi$ plane is 120 μm, giving a momentum resolution of the **CTD** $\sigma(p)/p = 0.0021p \oplus 0.0029$ for a charged particle that traverses all superlayers. Particles with polar direction

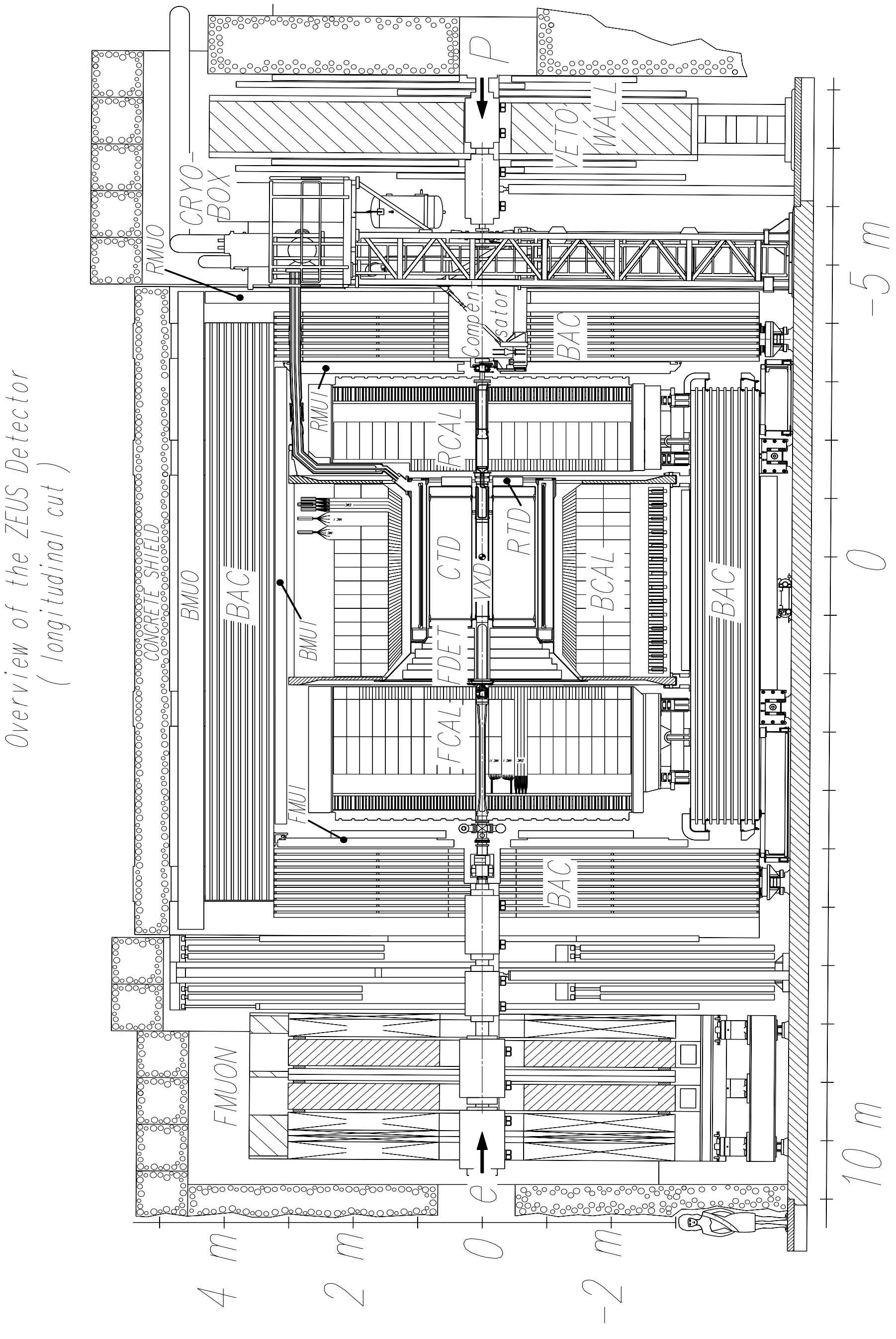


Figure 2.6: *Cross section of the ZEUS detector along the beam line.*

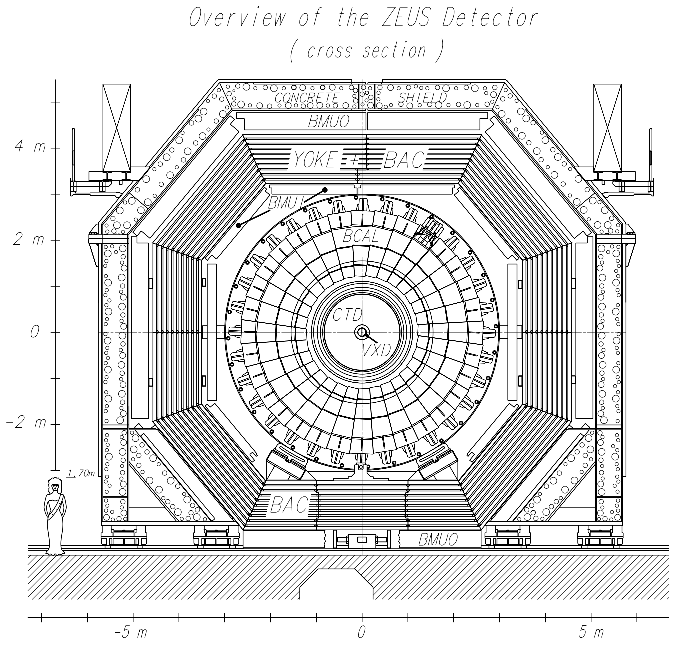


Figure 2.7: *Cross section of the ZEUS detector perpendicular to the beam line.*

in the range $15^o < \theta < 164^o$ cross at least one **CTD** superlayer. Information of both the **VTX** and the **CTD** leads to a vertex resolution in **x**,**y** between 30 and 80 μm.

During the 1992 runs, superlayers one, three and five (numbering superlayers from inside out) were instrumented with z-by-timing readout only. This resulted in the vertex resolution of approximately 4.5 cm in z, and ~ 1 mm in $r - \phi$.

- **FDET** and **RTD**: The measurement of tracks in the very forward or backward region is completed with the forward and rear tracking detectors. The **FDET** in the forward region consists of three planar chambers, each with three layers of drift cells, and modules of transition radiation chambers positioned between them. The **FDET** covers the range $7.5^o < \theta < 28^o$. The **RTD** in the rear direction, consists of a single chamber with three layers of drift cells, and covers the polar region $160^o < \theta < 170^o$.

  In 1992 the **FDET** and **RTD** were not instrumented with readout electronics, and are therefore not used in the present analysis.

- **MAGNET**: A superconducting solonoid, positioned around the **CTD**, supplies the magnetic field that bends charged particles in the $r - \phi$ plane and hence allows a momentum measurement. The influence of the magnetic field on the beams is compensated by a special magnet (Compensator) with an opposite magnetic field. In 1992, the **MAGNET** produced a magnetic field of 1.45 Tesla.

The Calorimeters surround the inner tracking devices and consist of the high resolution uranium calorimeter (**CAL**) and a backing calorimeter (**BAC**). Inside the **CAL** the hadron electron separator (**HES**) is mounted. The calorimeters measure the energy of charged or neutral particles by total absorption.

- **CAL**: The depleted-uranium scintillator calorimeter is divided into an electromagnetic and hadronic calorimeter, with a total depth varying between 5 and 7 interaction lengths λ_{int}. We describe this important subdetector in more detail in the next section.

- **BAC**: Energy leakage of the uranium calorimeter is absorbed and detected in the backing calorimeter, which consists of iron plates and aluminium proportional tubes. The iron plates are also used as a return path for the magnetic flux.

- **HES**: Separation between hadronic and electromagnetic energy deposits in the **CAL** is enhanced with the hadron electron separator. It provides information about the shower development in the **CAL**. The **HES** consists of a plane of $3 \times 3\ cm^2$ silicon diodes, which are placed in the **CAL** at a depth of $3.3\ X_0$. In 1992, the **HES** was partially installed in the **RCAL** only.

The MUON detectors consist of rear (**RMUON**), barrel (**BMUON**) and forward (**FMUON**) muon chambers. They are positioned at the inside and outside of the iron yoke and measure, with information of the inner tracking devices, tracks of muons that traverse the calorimeter.

The VETO wall and C5 counter signal particles that enter the detector from the rear direction. The **VETO** wall is a large iron wall ($8 \times 7.6 \times 0.9\ m$) covered on both sides with scintillation counters, positioned 7 m from the interaction point in the upstream proton direction. It shields the detector against particles from the proton beam halo and provides a veto of the beam-gas interactions that induce a trigger in the main detector.

The **C5** counter is positioned around the beam-pipe at 312 cm from the IP in the upstream proton direction, near the C5 collimator. It consists of two planes of scintillator separated by 5 cm of lead. The time information of the **C5** counter is crucial to measure the beam arrival times (see also section 2.1.1) and to reject proton beam-gas events.

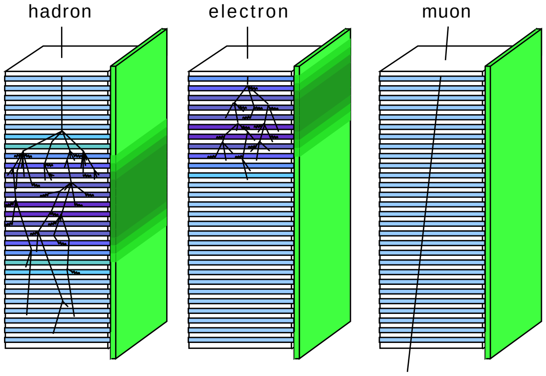


Figure 2.8: *The different types of shower shapes in the calorimeter. The scintillator-uranium sandwich structure is drawn for three towers with the wave-length shifters on the right side. The penetrating hadrons produce a large shower. The electron shower is small. The light signal of minimum ionising muons is equally spread over the depth of the tower.*

2.4 The ZEUS calorimeter

The ZEUS calorimeter is a sampling calorimeter which consists of layers of depleted uranium (DU), which act as the absorber, and scintillator as the active material. Highly energetic neutral or charged particles that enter the calorimeter interact with the uranium, and create in a cascade large numbers of secondary particles, the particle showers. The ionization energy loss of the shower particles in the active layers is measured in the form of light. The light propagates by internal reflection to wave-length shifters. The wave-length shifters transport the accumulated light to photo-multipliers (PMT's), which convert it to electric pulses. The pulse height in the PMT's is proportional to the number of shower particles, and hence is a measure for the incoming particle energy.

The shower shape inside the sandwich uranium-scintillator structure is determined by the cascade process. The origin of the cascade process for impacting hadrons is different from impacting leptons and photons. Whereas the latter undergo only QED type of interactions like bremsstrahlung, Compton scattering and pair production, the hadrons undergo predominantly nuclear interactions. The electromagnetic showers are small. 95% of the energy is contained laterally within two times the Molière radius, which is typically $R_M = 2\ cm$. The electromagnetic shower penetrates the ZEUS calorimeter to a maximum depth of about 25 radiation lengths X_0.

The hadron shower shape is spread more laterally and penetrates deeper in the calorimeter. A parametrization of the hadronic shower development [43] shows that for an 10 GeV hadron 95% of the transverse energy is contained in a cylinder with radius of approximately 20 cm, and penetrates up to an equivalent of 150

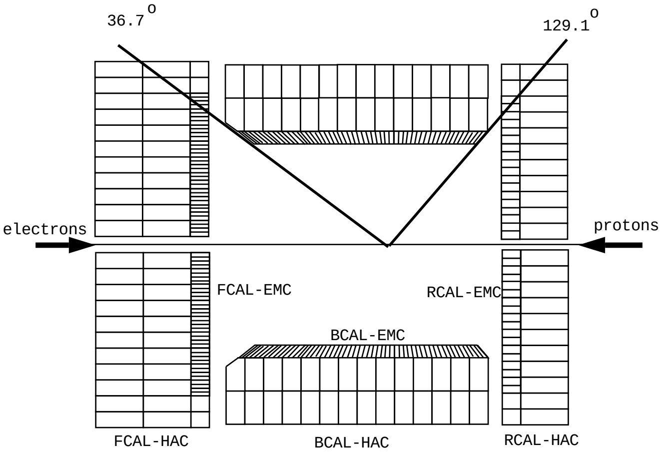


Figure 2.9: *Layout of the ZEUS uranium calorimeter.*

X_0[1]. Highly energetic muons behave as minimum ionizing particles. They loose only a small amount of energy, which is proportional to the number of uranium layers they traverse. The difference in the shower development for electromagnetic showers, hadron showers and muon penetration is used for particle identification. We picture the difference of the various types of showers schematically in figure 2.8.

An important feature of the ZEUS calorimeter is the fact that it is compensating. This means that the amount of light that is produced, and hence the signal that is measured, is equal for the electromagnetic and hadronic component of showers. Since each hadron shower has an electromagnetic component with large energy fluctuations, compensation considerably improves the energy resolution of impacting hadrons. Compensation in the ZEUS calorimeter is achieved by a suitable choice of relative thickness of the uranium and scintillator planes. After an extensive research program [43, 44, 58], compensation is realized with an accuracy better than 2%.

2.4.1 Mechanical layout

The sandwich structure for the ZEUS calorimeter consists of 2.6 *mm* thick scintillator planes, which are made of SCSN-38, a material that is very stable against aging and radiation damage, and planes of depleted uranium (DU). The DU planes are wrapped in stainless steel, and actually consist of an alloy of 98.4% ^{238}U, 1.4% Nb (to make the alloy harder), and $\leq 0.2\%$ ^{235}U.

The calorimeter consists of three main parts. These are the forward calorimeter (FCAL), the barrel (BCAL) and rear calorimeter (RCAL), shown in figure 2.9. All three parts of the CAL are designed similarly, which guarantees an uniform response

---

[1] For the ZEUS calorimeter the mean free path of the hadronic interactions (interaction length λ_{int}), corresponds roughly to $\lambda_{int} \sim 25X_0$

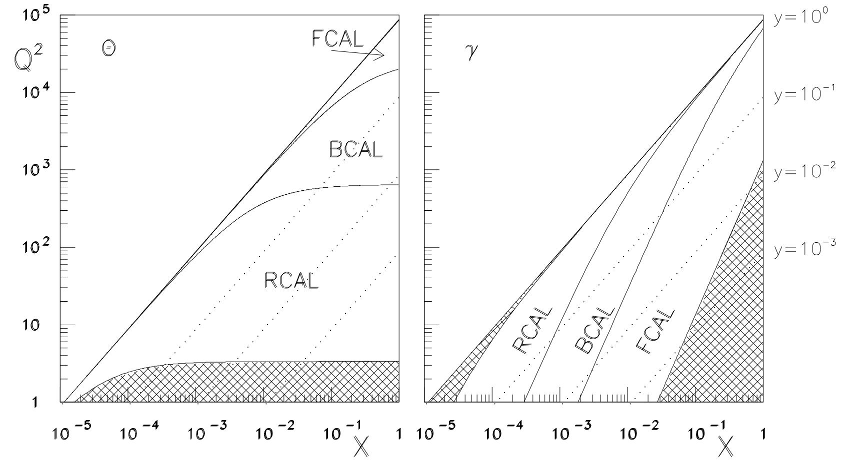


Figure 2.10: ***Regions in (x, Q^2) phase space that are covered by the FCAL, BCAL and RCAL for the electron scattering angle θ (left) and the quarks scattering angle γ (right). The hatched areas correspond to the beam holes.***

to impacting particles in 99.6% of the solid angle. Only the small (but important) beam-pipe hole in the FCAL and RCAL, which measures 20×20 cm^2, spoils the complete hermeticity of the CAL.

For lowest order DIS events, we plot the regions in electron scattering angle θ and quarks angle γ that are covered by the FCAL, BCAL and RCAL in the (x, Q^2) phase space. The beam hole is indicated by the hatched area. We discuss the consequences for the (x, Q^2) measurement in detail in the next chapter.

The mechanical layout for the FCAL and RCAL is very similar. Both the FCAL and RCAL consist of 23 separate modules, each with a width of 20 *cm* and a height varying between 220 and 460 *cm*, that are mounted vertically next to each other. The fronts of the modules form an approximate circle facing the IP. The first 25 layers of scintillator and DU, closest to the interaction point, form the electromagnetic calorimeter (FCAL-EMC and RCAL-EMC). This calorimeter has a depth of 25.9 X_0, and absorbs all electromagnetically showering particles. The hadronic calorimeter, following the EMC section, is segmented in one unit in the RCAL (RCAL-HAC), and two units in the FCAL (FCAL-HAC1 and FCAL-HAC2). Each HAC section has a depth varying between 40 and 80 layers. The cell size for the hadronic sections is 20×20 cm^2. The granularity of the EMC cells is larger in the FCAL (20×5 cm^2) compared to the RCAL (20×10 cm^2). Hence the HAC1 and HAC2 cells in the FCAL are preceded by 4 EMC cells (except for regions that are shadowed by the BCAL), whereas each HAC1 cell of the RCAL is preceded by only 2 EMC cells (also with the exception of regions that are shadowed by the BCAL).

The BCAL consists of 32 wedge shaped modules, placed symmetrically around the beam axis, as can be seen in figure 2.7. In order to prevent particles penetrating

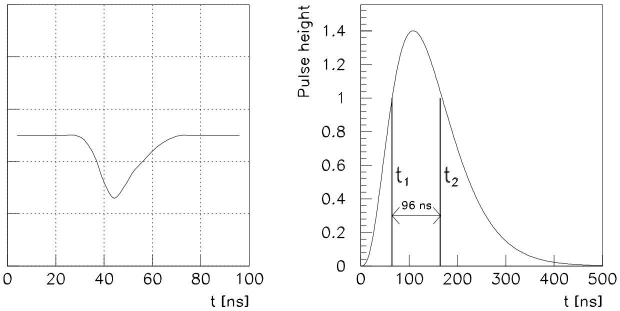


Figure 2.11: *On the left plot the profile of the original PMT signal pulse. On the right plot the same pulse after analogue shaping.*

in cracks, the modules are not pointing to the beam axis exactly, but are tilted by 2.5^o. The inner radius of the modules is 122 *cm*. The EMC section of the BCAL consists of 53 BEMC cells, $5.06 \times 23.4\ cm^2$, in one BCAL module, each of which is projective in polar angle. The BCAL-EMC section consists of 21 layers of uranium/scintillator. The HAC section of the BCAL (BCAL-HAC) is segmented in two units of 49 layers, and is non-projective.

Each calorimeter cell is equipped with wave-length shifter bars on two sides, and hence each cell is read-out by two PMT's. In total the ZEUS calorimeter has 11836 PMT's for 5919 calorimeter cells.

2.4.2 Readout and calibration

The electronic signal of the PMT's is sent to the analogue cards, mounted on the endbeam of each module. These electronics shape the pulse signal of the PMT, as shown in figure 2.11, and the shaped pulse is sampled with a 96 *ns* clock and stored in an analogue pipeline. When the event is accepted by the first level trigger, the samples are digitized. The digitizing electronics reconstruct the PMT pulseheight and calculate the energy and arrival time [44] from these pulseheights at times t_1 and t_2 (see figure 2.11). The sample at time t_0 gauges the baseline and allows correction for possible pile-up.

There are several methods to calibrate the reconstruction of energy and time of the PMT signals. For example, a charge injection system simulates precisely defined charge deposits of the PMT, which are fed into the analogue cards. In this way the shaping and sampling of the PMT signals are calibrated. In a different calibration system, the laser flash system, laser light is injected just in front of the PMT's. Since the laser is able to produce a large range of light intensity, the system tests the linearity of the PMT and readout electronics. This system is also very important in calibration of the PMT time, and individual time offsets for each channel are determined.

One additional important tool in the calibration of the uranium calorimeter is

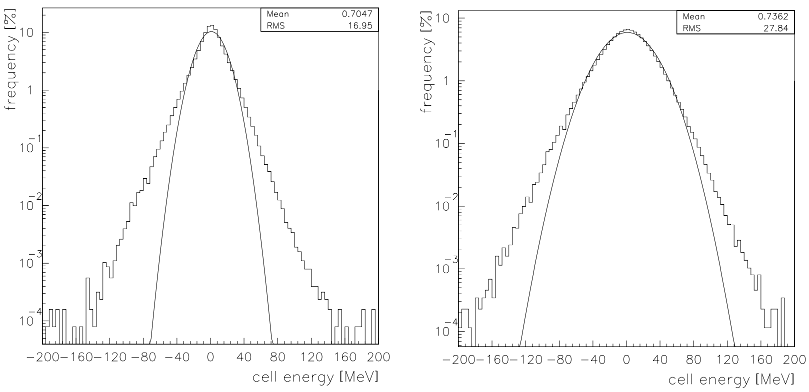


Figure 2.12: ***Spectra of the cell energy noise. In the left plot for the EMC cells, in the right plot for the HAC cells. A Gaussian fit through the histogram is plotted to guide the eye.***

the natural radioactivity of the uranium itself. The natural radioactivity induces a very stable, small readout signal, the uranium noise (UNO). A large number of systematic uncertainties, such as cell and wave-length shifter imperfections, are cancelled in the ratio of the shower signal to the UNO signal. The gains of all PMT's is set such that the UNO signal for all equally sized cells is equal. The short term variation of the UNO signal is below 0.5% [45], and hence the short term calibration fluctuations are small. Longer term variations (a few days) on the UNO signal are larger, ~3%, which are due to the PMT gain variations. With the UNO signal this effect can be corrected. Also very long term effects, such as radiation damage of the CAL, are monitored with the UNO signal.

The absolute energy scale, i.e. the conversion between reconstructed pulse heights and corresponding particle energies, has been obtained for a subset of the FCAL, BCAL and RCAL modules from test beam data at CERN [41] and FNAL [42]. The module to module variations were in general less than 1%, so there was no problem in transferring the energy scale to other, not calibrated, modules.

2.4.3 Performance of the calorimeter

In the test beam configuration, the resolution of the ZEUS calorimeter was measured to be $\sigma(E)/E = 0.18/\sqrt{E}$ for electromagnetic showers and $\sigma(E)/E = 0.35/\sqrt{E}$ for hadronic showers, with E in units of GeV. The contribution from noise to the resolution is small. The angular resolution of the impact position is in first order determined by the cell sizes. However, the resolution improves considerably when using information from both the PMT's for each cell. As will be shown in chapter 4, a resolution for impacting electrons of 0.4–1.2 *cm* is achieved using information from the calorimeter.

The signal of the UNO produces a small amount of noise in the calorimeter.

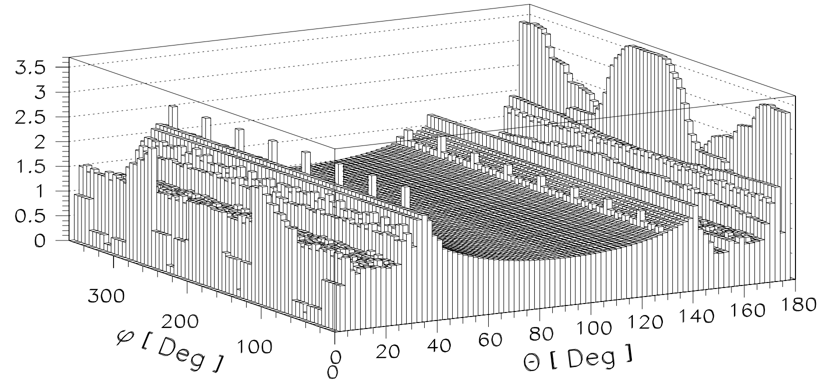


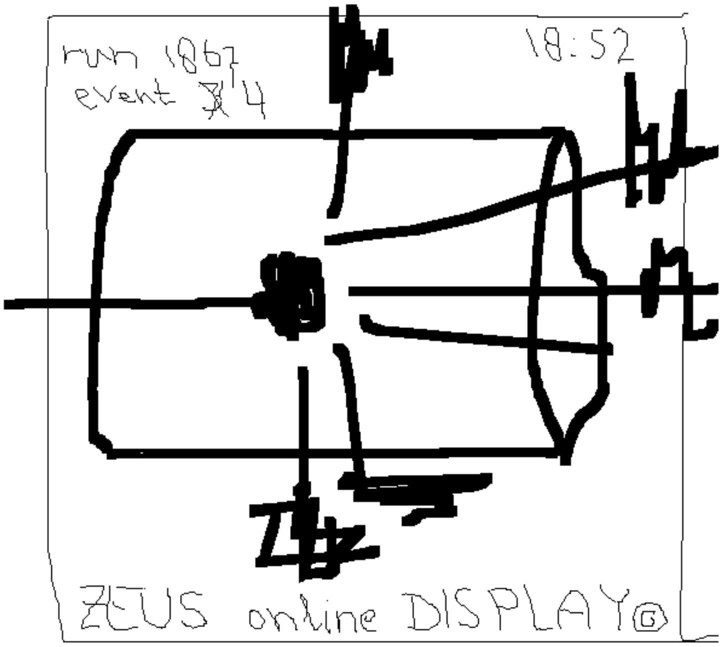
Figure 2.13: *Dead material in front of the calorimeter as a function of the polar angle θ and the azimuthal angle φ. The units are radiation lengths X_0.*

The UNO calibrated cell energies for empty events (i.e. no impacting particles) are shown in figure 2.12 for EMC and HAC cells separately. The figure shows the fluctuation of the calibrated cell energies due to the UNO signal. The fluctuation is larger in HAC cells because these are bigger in size and accumulate more UNO. A zero-suppression on the cells is applied to remove the UNO signals. Detailed study [69] showed that the requirement

$$E(EMC) > 60 \quad ; \quad E(HAC) > 110 \text{ MeV} \tag{2.6}$$

gives the best separation between UNO and the signal of impacting particles.

The excellent performance of the CAL under ideal test beam conditions is somewhat degraded in the presence of material between the interaction point and the surface of the CAL in the experiment. In the ZEUS experiment, the amount of material in front of the CAL fluctuates strongly. In figure 2.13 we plot the present best estimate of the amount of material, in units of X_0, as a function of the polar angle θ and azimuthal angle ϕ as seen from the interaction point. Especially for large values of θ, the amount of material is large and fluctuates strongly.


run 1867
event 34
18:52
ZEUS online DISPLAY ©

Chapter 3

Reconstruction of (x, Q^2) in neutral current scattering

An accurate reconstruction of both x and Q^2 is of vital importance for structure functions measurements at HERA. In this chapter we present various methods to reconstruct x and Q^2 for inclusive neutral current deep inelastic scattering. The conventional reconstruction methods either use detector information of the final state electron or the hadron flow only. We utilize the redundancy of the electron and hadron system to generate a number of additional methods to obtain x and Q^2. The properties of the most promising reconstruction methods are discussed. The Monte Carlo simulation of the ZEUS detector is used to investigate the accuracy of these methods.

3.1 Suitable variables for (x, Q^2) reconstruction

For measurement of the Lorentz-invariant variables x, y and Q^2 for inclusive NC DIS we have the final state electron and hadron flow at our disposal. The four-momentum vectors of the initial and final state electron are denoted by $\mathbf{k}_e$ and $\mathbf{k}'_e$ respectively. The initial state proton is denoted by $\mathbf{P}_p$, and the final state hadron flow by $\mathbf{P}'_p$. In the ZEUS coordinate system (see section 2.1.2), these four-momenta can be written as:

$$\mathbf{k}_e = \begin{pmatrix} A \\ 0 \\ 0 \\ -A \end{pmatrix}, \quad \mathbf{k}'_e = \begin{pmatrix} E \\ E\sin\theta\cos\varphi \\ E\sin\theta\sin\varphi \\ E\cos\theta \end{pmatrix}, \quad \mathbf{P}_p = \begin{pmatrix} P \\ 0 \\ 0 \\ P \end{pmatrix}, \quad \mathbf{P}'_p = \begin{pmatrix} E_h \\ P_{x\,h} \\ P_{y\,h} \\ P_{z\,h} \end{pmatrix}, \tag{3.1}$$

where for $\mathbf{P}'_p$ a summation is assumed over all particles h that are produced in the final state X.

For a general, model independent description of *ep* scattering, we require that the measured kinematic variables x, y and Q^2 have the following two important properties:

- The extracted variables should be independent of final state fragmentation. For an inclusive measurement it is irrelevant whether the hadron flow produces jet-like objects or not. The jet mass is a consequence of effects of fragmentation and gluon radiation, and hence an inclusive determination of the kinematic variables should be independent of fragmentation mass-effects.
- The extracted variables should be insensitive to particle losses in the dead area along the forward beam direction, the beam-pipe. This means that the contribution to the extracted variables of particles that originate from the diquark should effectively be suppressed.

It is clear that the four-momentum of the final state electron satisfies these requirements and that therefore the energy E and polar angle θ are suited to describe the kinematics of the event.

For the hadron flow $\mathbf{P}'_p$, we show that the transverse momentum $P_{t\,h}$ and the difference of the hadron-flow energy and z-momentum, $(E - P_z)_h$, where we assume summation over h, also satisfy the above requirements. Consider therefore the case where a massless parton with momentum fraction x interacts with the virtual photon, and the case where a parton with modified momentum fraction ξ generates a jet-mass M^2, see figure 3.1.

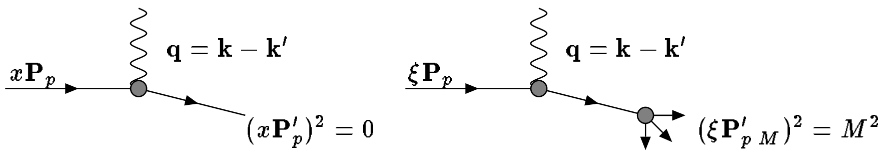


Figure 3.1: *The hadron vertex for massless particles (left) with final state* $\mathbf{P}'_p$ *and for creation of a jet-mass* M^2 *(right) with final state* $\mathbf{P}'_{p\,M}$.

From energy conservation in both cases, x is related to ξ as follows:

$$\left.\begin{array}{l}(\mathbf{q}+x\mathbf{P}_p)^2 = (x\mathbf{P}'_p)^2 = 0\\ (\mathbf{q}+\xi\mathbf{P}_p)^2 = (\xi\mathbf{P}'_{p\,M})^2 = M^2\end{array}\right\} \rightarrow \ \xi = x\left(1+\frac{M^2}{Q^2}\right). \tag{3.2}$$

Momentum conservation at the vertex yields a relation between the energy, transverse momentum and longitudinal z-momentum of the massless case and the massive 'jet-like' case.

$$\left.\begin{array}{l}\mathbf{q}+x\mathbf{P} = \mathbf{P}'_p\\ \mathbf{q}+\xi\mathbf{P} = \mathbf{P}'_{p\,M}\end{array}\right\} \rightarrow \left\{\begin{array}{l}E_M = E + xP\frac{M^2}{Q^2}\\ P_{t\,M} = P_t\\ P_{z\,M} = P_z + xP\frac{M^2}{Q^2}.\end{array}\right. \tag{3.3}$$

From the right hand side of equation 3.3 we conclude that besides the transverse momentum $P_{t\,h}$, also the combination $(E-P_z)_h$ is independent of the jet-mass M^2. These two variables are in addition insensitive to energy losses in the forward beampipe, since then the transverse momentum is approximately zero and the energy almost equal to the z-momentum. They are therefore well suited to reconstruct x, y and Q^2.

3.2 The conventional reconstruction methods

The most straightforward method to reconstruct (x, Q^2) uses information from the scattered electron. The energy E and angle θ of the final state lepton are converted to (x, Q^2) using the relations 1.3 and 1.2. This method of extracting (x, Q^2), which we call the 'electron–method', is used in all fixed target experiments and is summarized briefly in section 3.4.2.

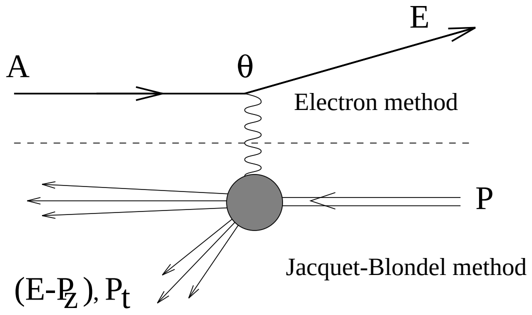


Figure 3.2: *The kinematic variables from the electron side or the hadron flow.*

The other possibility for almost completely hermetic detectors in neutral-current ep scattering is, as we saw, the determination of x and Q^2 from the hadron flow measurement, see figure 3.2. In charged-current DIS, where the electron is converted to a neutrino in the final state and escapes detection, the hadron flow is the only handle to reconstruct the kinematic variables. The method that effectively uses the variables $(E - P_z)_h$ and P_t of the hadron flow, is the 'Jacquet–Blondel (JB) reconstruction' [51, 52]. For HERA kinematics it is more suitable to express y and Q^2 in terms of laboratory variables and then to deduce x. From relations 1.3 and 1.2 for the hadron side follows:

$$\begin{aligned} y_{JB} &= \frac{(E_i - P_z)_h}{2A}, \\ Q^2_{JB} &= \frac{P^2_{t\,h}}{1 - y_{JB}}, \\ x_{JB} &= \frac{Q^2_{JB}}{s\, y_{JB}}. \end{aligned} \tag{3.4}$$

Reconstruction of x_{JB} and Q^2_{JB} is performed with summation of the complete hadron flow, without any jet identification or hypothesis on the proton structure. Particles of the hadron flow that are emitted in the forward direction contribute only little.

Before discussing the two reconstruction methods in more detail, we first rewrite the hadron flow quantities $(E - P_z)_h$ and $P_{t\,h}$ in terms of an angle γ and an energy F.

3.3 Parametrization of the hadron flow

In the naive quark model, the hadron flow originates from the fragmented struck quark and remaining diquark. It is the fragmented struck quark, the current jet, that determines the kinematics of the event on the hadron side, whereas the contribution to the hadron flow stemming from the diquark (the target remnant or spectator jet) can obscure this measurement. For this reason we introduce a notation of the hadron flow that describes the hard electron-quark interaction in a more direct way. To this end the hadron flow is split and the current jet momentum is described by an energy F and polar angle γ. In the naive quark model these are the energy and polar angle of the scattered quark. In this notation the kinematics of inclusive DIS electron-proton scattering is replaced by the kinematics of the final state electron and struck quark, denoted by the four parameters $\{E, \theta, F, \gamma\}$ as follows:

E	Energy of the final electron
θ	Polar angle of the final electron
F	Energy of the scattered (struck) quark
γ	Polar angle of the scattered (struck) quark
A	Energy of the initial electron (26.7 GeV)
P	Energy of the initial proton (820.0 GeV)

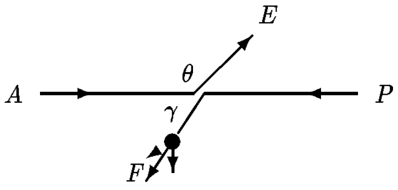


Figure 3.3: *Definition of the variables* $\{E, \theta, F, \gamma\}$

From the requirement that the extracted variables x, y and Q^2 should be insensitive to fragmentation effects and particle loss in the forward beam-pipe we argued that the quantities $(E - P_z)_h$ and $P_{t\,h}$ have to be used in the determination of the kinematic variables. It is very useful to construct from the quantities $(E - P_z)_h$ and $P_{t\,h}$ an energy F and direction $\cos\gamma$ of a *massless* object, that describes the hadron flow completely and satisfies the above defined requirements [53]. We define:

$$F \equiv \frac{P_{t\,h}^2 + (E - P_z)_h^2}{2(E - P_z)_h}, \qquad (3.5)$$

$$\cos\gamma \equiv \frac{P_{t\,h}^2 - (E - P_z)_h^2}{P_{t\,h}^2 + (E - P_z)_h^2}, \qquad (3.6)$$

were summation of the hadrons h is assumed. In this construction of F and γ we have made no assumption on the structure of the scattering process. Therefore, *the inclusive ep DIS can be described by, without any assumptions or approximations, the scattering between the electron and a massless object, parametrized by the above defined quantities F and γ.*

The quantities F and γ are related to the hadron flow energy and z-momentum as follows:

$$\begin{aligned} F \equiv & \ \tfrac{1}{2}\left\{\frac{P_{t\,h}^2}{(E-P_z)_h} + (E-P_z)_h\right\} & = E_h - \frac{W^2}{2(E-P_z)_h} = E_h - (1-x)P, \\ F\cos\gamma \equiv & \ \tfrac{1}{2}\left\{\frac{P_{t\,h}^2}{(E-P_z)_h} - (E-P_z)_h\right\} & = P_{z\,h} - \frac{W^2}{2(E-P_z)_h} = P_{z\,h} - (1-x)P. \end{aligned}$$

We observe that effectively for F and $F\cos\gamma$ the energy and z-momentum of the diquark, $(1-x)P$, is subtracted from the hadron flow energy and z-momentum. In the naive quark model F and γ correspond to the energy and direction of the struck quark.

Beyond the naive quark model the quantities F and γ cannot be related to the energy and polar angle of the struck quark, but still provide an exact description of the kinematics of ep inclusive scattering.

With the above definitions of F and γ we can rewrite the Jacquet-Blondel reconstruction method as:

$$\begin{aligned} y_{JB} &= \frac{F}{2A}(1-\cos\gamma), \\ Q^2_{JB} &= \frac{F^2\sin^2\gamma}{1-y_{JB}}. \end{aligned} \tag{3.7}$$

A useful property of the construction of $\cos\gamma$ is the fact that the extracted value for $\cos\gamma$ is in first order stable against energy fluctuations of the hadron flow. If a fraction f_{mis} of the energy is lost, the measured values of $P_{t\,h}$ and $(E-P_z)_h$ are shifted by a factor $(1-f_{mis})$. Because $\cos\gamma$ is constructed from a ratio of these two quantities, the factor $(1-f_{mis})$ cancels in the expression for $\cos\gamma$. So to first order the angle is independent of energy loss and, what is more interesting, this means that it is independent of fluctuations on the energy measurement of the hadron flow.

3.4 Multiple reconstruction methods

The determination of kinematics with the electron–method or the Jacquet-Blondel method will lead to identical values for both x and Q^2 for non-radiative ep-scattering in a perfect detector. In other words, the final state lepton and hadron flow are not independent. This can be demonstrated with the polar diagram in figure 3.4, in which lines of constant (x, Q^2, y) (isolines) are drawn in the longitudinal and transverse momentum plane of the final state lepton and current jet. The longitudinal

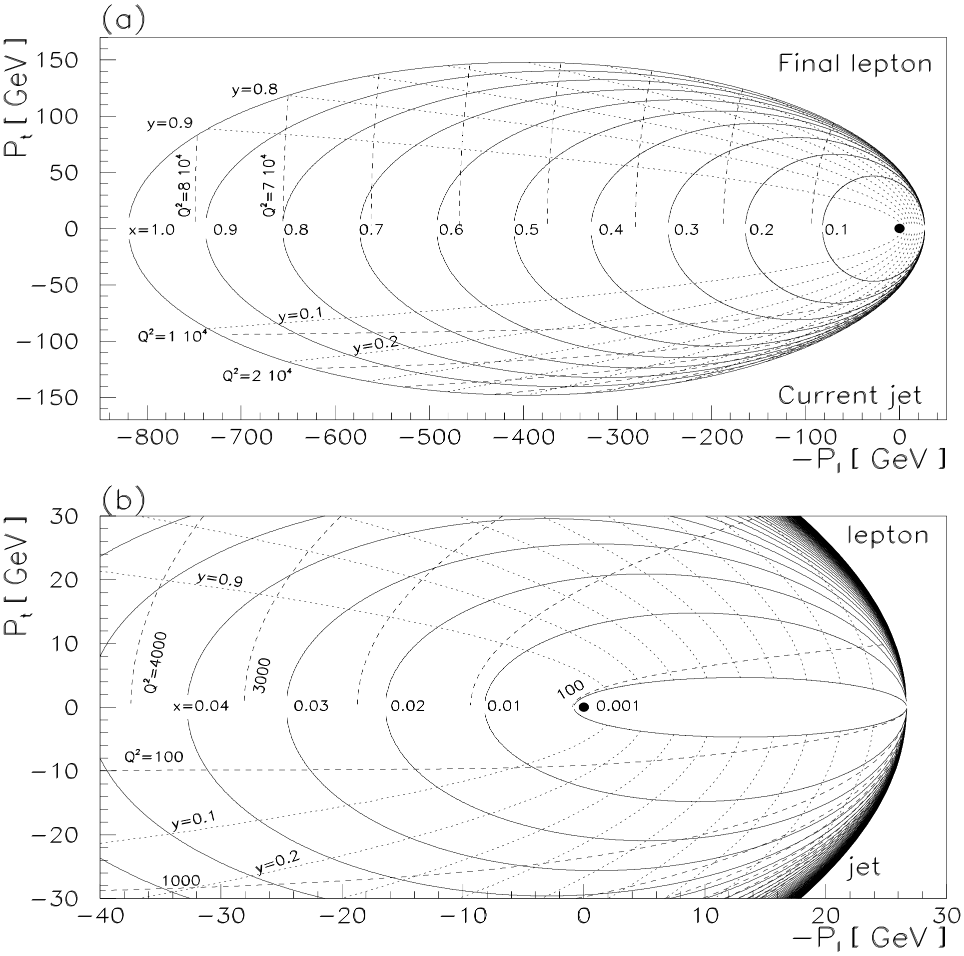


Figure 3.4: *Polar diagram of longitudinal momentum P_l and transverse momentum P_t for the final state lepton (upper half of plots) and current jet (lower half of plots) for ep kinematics with $A = 26.7$ GeV and $P = 820$ GeV. Plot (b) is a blow up of the low momentum region.*
The isolines of x are drawn with a straight line (-a-: $x = 0.1 - 1$ step 0.1, -b-: $x = 0.001$, $x = 0.01 - 1$ step 0.01). The dashed lines are the isolines in Q^2 (-a-: $Q^2 = 1 \cdot 10^4 - 8 \cdot 10^4$ step $1 \cdot 10^4$ GeV^2, -b-: $Q^2 = 100$, $Q^2 = 1000 - 10000$ step 1000 GeV^2). The dotted lines are the isolines in y ($y = 0.1 - 0.9$ step 0.1 for -a- and -b-).

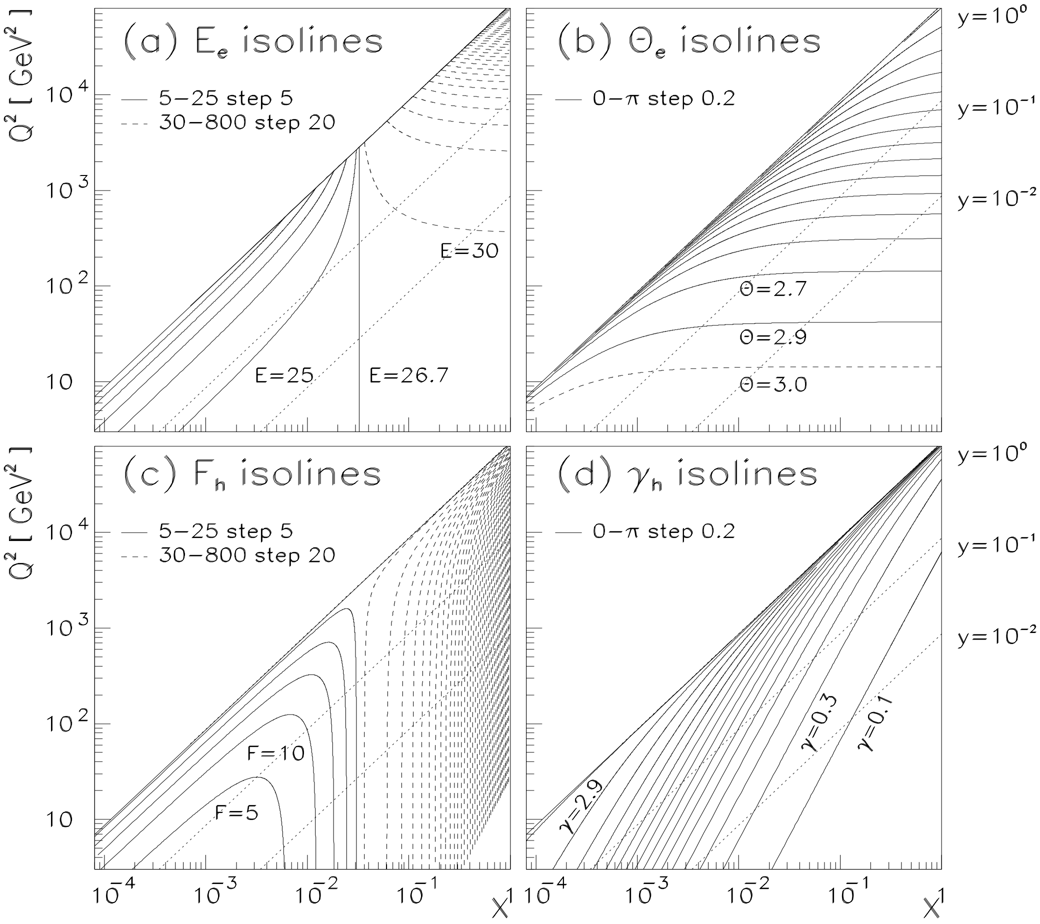


Figure 3.5: *Isolines of* $\{E, F, \theta, \gamma\}$ *in* (x, Q^2) *phase space for colliding ep beam kinematics with initial electron energy of 26.7 GeV and initial proton energy of 820 GeV.*

In plot (a) the isolines for the energy of the final state electron. The range of values for the isolines is given. In addition the isoline $E_e = A = 26.7$ *GeV is drawn which contains no information on* Q^2*.*

In plot (b) the isolines for the angle of the final state electron. In (d) for the energy F of the hadron system, or, in the naive quark model, of the quark energy. In (d) the isolines for the hadron angle γ_h *which corresponds to the quark angle in the naive quark model.*

momentum of the current jet is given by $P_l = F\cos\gamma$ and the transverse momentum by $P_t = F\sin\gamma$. Connecting a specific choice of (x, Q^2) with the origin gives the momentum vector P_t and P_l in the laboratory frame of the final lepton (upper half of the two plots), as well as the momentum vector of the current jet (lower half of the two plots). It is clear that this over-determined system provides redundancy for the reconstruction of the kinematic variables.

3.4.1 Isolines

The measurement errors on $\{E, \theta, F, \gamma\}$ affect the kinematic variables as determined from the lepton or hadron side for non-radiative DIS events. The detector effects dictate the uncertainties on the determination of the variables $\{E, \theta, F, \gamma\}$.

The specific reconstruction method of (x, Q^2) determines the propagation of the measurement uncertainties on $\{E, \theta, F, \gamma\}$ to the reconstructed values of (x, Q^2). Properties of the transformations from $\{E, \theta, F, \gamma\}$ to $\left(x, Q^2\right)$ can be derived from the polar diagram of figure 3.4. In regions where the isolines in x (or Q^2) are close together, a very precise measurement of the momenta is needed to limit the uncertainty on x (Q^2). In regions where the isolines are distant, a larger error on the measurement of the momenta brings about only a small uncertainty in x or Q^2.

It is more instructive to invert the polar diagram. In figure 3.5 we present the isolines of the basic set $\{E, \theta, F, \gamma\}$ in the $\left(x, Q^2\right)$ phase space. We give the values for the equidistant isolines in the figure. The isolines provide an understanding of the effect of measurement errors on $\{E, \theta, F, \gamma\}$ in the $\left(x, Q^2\right)$ phase space. In contrast to the polar diagram of figure 3.4, the isolines of $\{E, \theta, F, \gamma\}$ in the (x, Q^2) phase space of figure 3.5 imply a good intrinsic resolution if they are close together. For these dense isolines, measurement errors on $\{E, \theta, F, \gamma\}$ lead to small uncertainties on x and Q^2. The intrinsic resolution is worse for isolines that are far apart because then a small measurement error on $\{E, \theta, F, \gamma\}$ corresponds to a large volume in the (x, Q^2) phase space.

In figure 3.5a we have drawn the isolines of the energy of the final state electron. The energy $E = A$ divides the phase space in x at $x = x_0 = A/P$. At this value for x the electron-quark CM frame coincides with the laboratory frame. The quark energy also equals A and the directions of the electron and quark are opposite (the value for Q^2 is determined from these directions). For lower values of x the electron always loses energy during the hard interaction. For the energy of the quark one has also $F < A$ in this case (plot 3.5c), due to the fact that the center of mass energy of the quark-lepton interaction $(\sqrt{s})_{lq}$ is less than $2A$. In the region $x > x_0$ both the energy of the final state electron and the quark energy are larger than the incident electron energy. The electron energy isolines E are very distant to each other in the $\left(x, Q^2\right)$ phase space for $E \sim A$, at low values of y. In this region of phase space the electron energy needs to be measured with high precision to reconstruct (x, Q^2).

In plot 3.5b we show the isolines of the polar angle of the final state electron, θ. From this plot we conclude that the electron angle measurement is powerful for the determination of Q^2, since these isolines are almost perpendicular to the Q^2 axis.

The intrinsic resolution on Q^2 from θ deteriorates at low values of θ. Only at high values of y the angle θ bears information on the value of x.

For the hadron flow or quark kinematics, we present in plot 3.5c and 3.5d the isolines in F and γ respectively. The quark energy discriminates between values of x, especially at high x, but contains hardly any information on Q^2. Due to the curved isolines for $F < A$, a measured quark energy in the vicinity of 26.7 GeV corresponds to a very large value of y or the specific value of $x = x_0$. The isolines of the quark polar angle γ are evenly spread in the entire phase space, except at low values of γ. For γ it should be noted that at very low values of x the quark points towards the rear part of the detector, at very high angles.

3.4.2 Novel reconstruction methods

To obtain the ultimate accuracy for both x and Q^2 it seems logical to use both the final state electron and hadron flow. Using the electron method one combines the measurement of E and θ to obtain a unique point in the (x, Q^2) phase space. The JB-method uses F and γ. However, it is also possible to combine two other quantities from the set $\{E, \theta, F, \gamma\}$ because *any combination of two of these defines a unique point in the* (x, Q^2) *plane* (except for ambiguities due to two possible solutions when one combines (E, γ) or (F, θ) for $x < x_0$). In this way it is possible to obtain six different reconstruction methods, each with its own dependence on kinematical variables and detector effects:

A1: $Q^2 = Q^2[E, \theta]$ $x = x[E, \theta]$ $y = y[E, \theta]$
A2: $Q^2 = Q^2[E, \gamma]$ $x = x[E, \gamma]$ $y = y[E, \gamma]$
A3: $Q^2 = Q^2[\theta, \gamma]$ $x = x[\theta, \gamma)$ $y = y[\theta, \gamma]$
A4: $Q^2 = Q^2[E, F]$ $x = x[E, F]$ $y = y[E, F]$
A5: $Q^2 = Q^2[F, \theta]$ $x = x[F, \theta]$ $y = y[F, \theta]$
A6: $Q^2 = Q^2[F, \gamma]$ $x = x[F, \gamma]$ $y = y[F, \gamma]$

In addition to the electron method (denoted here by A1) and the JB–method (A6) four new reconstruction methods for x and Q^2 are possible [53, 54]. We obtained these relations from momentum conservation in electron-quark scattering. The corresponding formulae are given in appendix A.

We investigated these reconstruction methods. They all respond differently to detector effects and it turns out that not all methods can compete for the most accurate reconstruction of (x, Q^2) . The electron method A1, Jacquet-Blondel method A6 and 'double-angle' method A3 are the most promising. In this 'double-angle' method the angles θ and γ are used to reconstruct x and Q^2.

In figure 3.6 we present the error ellipses for these three methods. It is clearly visible that the reconstruction methods behave differently in phase space. At $Q^2 = 10$ GeV^2 for example, the electron method measures x with good resolution for very low values of x. This corresponds to the low electron energy region, as viewed in the previous isoline plot 3.5a. Moving to higher values of x, we observe in the same

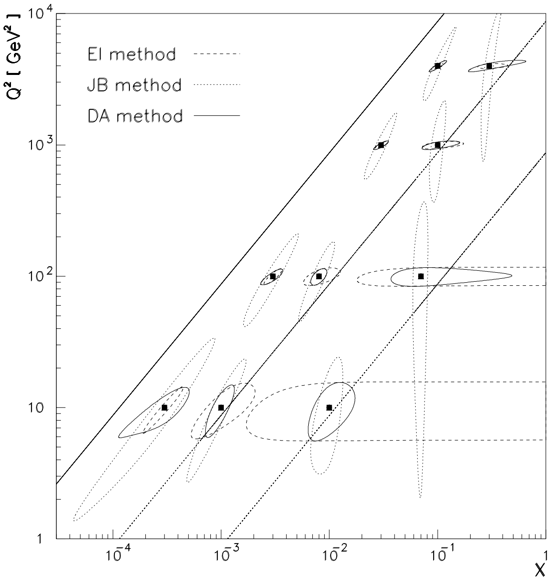


Figure 3.6: *Error ellipses for three reconstruction methods. The dashed line for the electron method, the dotted line for the Jacquet-Blondel method and the full line represents the double-angle method. The central value is denoted by the black square. We assumed* $\sigma_E = 0.2 \cdot \sqrt{E} + 0.01 \cdot E$, $\sigma_\theta = 30$ *mrad,* $\sigma_F = 0.8 \cdot \sqrt{F} + 0.05 \cdot F$ *and* $\sigma_\gamma = 150$ *mrad as error estimation. Besides the error ellipses, the lines* $y = 1$, $y = 0.1$ *and* $y = 0.01$ *are drawn.*

figure that the energy isolines become very distant to each other. A small error in the electron energy E has a large effect on the determination of x when the angle θ is fixed. This translates to large error ellipses in the x direction in figure 3.6. The resolution in Q^2, determined from the electron angle θ, is stable as a function of x and becomes somewhat better as Q^2 increases.

The Jacquet-Blondel method shows the opposite effect. The resolution in x is good at high values of x, and becomes worse at low values of x. But especially the resolution in Q^2 is poor compared to the electron method. First of all this is due to the fact that the measurement errors on F and γ of the hadron flow are larger than the errors on E and θ. But on top of that, for high values of y, F and γ isolines are parallel which implies a large intrinsic uncertainty in both x and Q^2.

The 'double-angle' method has the smallest error ellipses for large parts of phase space and therefore has a good intrinsic resolution. The isolines of θ and γ are almost perpendicular and equidistant in a large fraction of the phase space. Only at high values of y these two isolines become parallel and the determination of x and Q^2 deteriorates.

The reconstruction methods can also be combined. Particularly useful is the method that uses the $Q^2[E,\theta]$ from the electron variables and $y[F,\gamma]$ from the hadron flow, while x is computed with the relation $Q^2 = sxy$. This is called the 'mixed' reconstruction. Another useful mixture of combinations is the '$y\theta$' method, which uses the $Q^2[F,\theta]$ and $y[F,\gamma]$ variables.

For each method we determine the expression for the relative error in x and Q^2. These expressions serve to point at parts of phase space where the corresponding reconstruction method is less accurate. The ultimate test to determine the accuracy of the method will be done with the Monte Carlo simulation including electroweak radiative corrections, in the next section.

Electron method

In our notation the formulae for the kinematic variables for the electron method are:

$$\begin{aligned}
1 - y_{elec} &= \frac{E}{2A}(1-\cos\theta) \\
Q^2_{elec} &= \frac{E^2 \sin^2\theta}{1 - y_{elec}} \\
x_{elec} &= \frac{Q^2_{elec}}{s\, y_{elec}} &= x_0 \frac{E(1+\cos\theta)}{2A - E(1-\cos\theta)},
\end{aligned} \tag{3.8}$$

with

$$x_0 = A/P. \tag{3.9}$$

These relations follow directly from equations 1.2, 1.3, 1.4 and 3.1.

The dependence of the reconstructed x_{elec} and Q^2_{elec} on measurement uncertainties of the laboratory variables E and θ can be expressed analytically by the following partial derivatives that represent the relative errors on x and Q^2:

$$\begin{aligned}
\left.\frac{\partial Q^2}{Q^2}\right|_E &= \frac{dE}{E} & \left.\frac{\partial Q^2}{Q^2}\right|_\theta &= -\tan(\theta/2)\; d\theta \\
\left.\frac{\partial x}{x}\right|_E &= \frac{1}{y}\,\frac{dE}{E} & \left.\frac{\partial x}{x}\right|_\theta &= \left[\frac{x}{x_0} - 1\right] \tan(\theta/2)\; d\theta
\end{aligned} \tag{3.10}$$

From these relative errors we conclude that the resolution in Q^2_{elec} is very good, except for small scattering angles (i.e. $\theta \sim 180^o$), where the angular precision is the dominant term. The resolution in x however, is very poor at low values of y, due to the factor $1/y$.

In the electron–method, we can relate the angle θ to y_{elec} as follows:

$$\frac{1 - y_{elec}}{y_{elec}} = -\frac{x_{elec}}{x_0} \tan^2(\theta/2). \tag{3.11}$$

Jacquet–Blondel method

We summarize the Jacquet–Blondel method reconstruction method for the variables $(E-P_z)_h$ and $P_{t\,h}$ [55], and for the parameters F and γ:

$$\begin{aligned}
y_{JB} &= \frac{\sum_h (E-P_z)_h}{2A} &&= \frac{F}{2A}(1-\cos\gamma) \\
Q^2_{JB} &= \frac{\left(\sum_h P_{xh}\right)^2 + \left(\sum_h P_{yh}\right)^2}{1-y_{JB}} &&= \frac{F^2 \sin^2\gamma}{1-y_{JB}} \\
x_{JB} &= \frac{Q^2_{JB}}{s\, y_{JB}} &&= x_0 \frac{F(1+\cos\gamma)}{2A - F(1-\cos\gamma)},
\end{aligned} \tag{3.12}$$

where the sum runs over the all observed final state hadrons.

The relation between the angle γ and y_{JB} can be written for the Jacquet-Blondel method as:

$$\frac{1-y_{JB}}{y_{JB}} = \frac{x_{JB}}{x_0} \tan^2(\gamma/2). \tag{3.13}$$

The resolution of (x, Q^2) can also be expressed as a function of the resolution on F and γ by partial derivatives. The error propagation from (F,γ) to $(x, Q^2)_{JB}$ reads:

$$\begin{aligned}
\left.\frac{\partial Q^2}{Q^2}\right|_F &= \frac{2-y}{1-y}\,\frac{dF}{F} & \left.\frac{\partial Q^2}{Q^2}\right|_\gamma &= -\frac{x}{x_0}\left(\frac{y}{1-y}\right)\, d\gamma \\
\left.\frac{\partial x}{x}\right|_F &= \frac{1}{1-y}\,\frac{dF}{F} & \left.\frac{\partial x}{x}\right|_\gamma &= \left[\frac{x}{x_0}\left(\frac{1-y}{y}\right) - 1\right] \tan(\gamma/2)\, d\gamma.
\end{aligned} \tag{3.14}$$

The resolution in Q^2_{JB} is worse compared to Q^2_{elec} near $y = 1$, as a result of the term $1/(1-y)$. The same conclusion holds for the resolution in x. At low values of y the resolution depends on dF/F solely. The angular resolution only affects the precision on Q^2_{JB} at very low values of γ. Both very low and very high values of γ result in a poor determination of x_{JB}.

Double–Angle method

The double angle method only uses the angles θ and γ to reconstruct (x, Q^2) . This implies immediately that this novel reconstruction method does not depend on the energy scale intrinsically. This feature will turn out to be of fair importance. The double–angle method reads:

$$\begin{aligned}
y_{DA} &= \frac{\sin\theta(1-\cos\gamma)}{\sin\gamma + \sin\theta - \sin(\theta+\gamma)} \\
Q^2_{DA} &= 4A^2 \frac{\sin\gamma(1+\cos\theta)}{\sin\gamma + \sin\theta - \sin(\theta+\gamma)} \\
x_{DA} &= x_0 \frac{\sin\gamma + \sin\theta + \sin(\theta+\gamma)}{\sin\gamma + \sin\theta - \sin(\theta+\gamma)}.
\end{aligned} \tag{3.15}$$

This reconstruction method is mathematically exact since γ is obtained in a model independent way from the hadron flow and is insensitive to hadronization.

We obtain for the relative errors on x and Q^2 for this reconstruction method the following relations:

$$\begin{aligned} \left.\frac{\partial Q^2}{Q^2}\right|_\theta &= \frac{y-2}{\sin\theta}\, d\theta & \left.\frac{\partial Q^2}{Q^2}\right|_\gamma &= \frac{-y}{\sin\gamma}\, d\gamma \\ \left.\frac{\partial x}{x}\right|_\theta &= \frac{-1}{\sin\theta}\, d\theta & \left.\frac{\partial x}{x}\right|_\gamma &= \frac{-1}{\sin\gamma}\, d\gamma. \end{aligned} \tag{3.16}$$

Both at very small or very large scattering angles for the electron or hadron flow the resolution in x_{DA} and Q^2_{DA} degrades. Low values of θ correspond to low values of Q^2_{DA}. The resolution on Q^2_{DA} at low values of γ is increased due to the suppression by the y term in the numerator. The relation between the two angles and y_{DA} for the double angle method can be written as:

$$\frac{1-y_{DA}}{y_{DA}} = \frac{\tan(\gamma/2)}{\tan(\theta/2)}. \tag{3.17}$$

Mixed method

The mixed method utilizes the redundancy between the electron and hadron flow in a more direct way, and uses the well measured kinematic variables of the electron and Jacquet-Blondel method. The value of y is extracted with the Jacquet-Blondel method and Q^2 from the electron information, hence this method does not depend on the transverse momentum of the hadron flow. The value for x is then deduced from these. One then has:

$$\begin{aligned} y_{mix} &= y_{JB} \\ Q^2_{mix} &= Q^2_{elec} \\ x_{mix} &= \frac{Q^2_{elec}}{s\; y_{JB}} = x_0 \frac{E}{F}\frac{(1+\cos\theta)}{(1-\cos\gamma)}. \end{aligned} \tag{3.18}$$

We write the relative errors on x and Q^2 as follows:

$$\begin{aligned} \left.\frac{\partial x}{x}\right|_E &= \frac{dE}{E} & \left.\frac{\partial x}{x}\right|_\theta &= -\tan(\theta/2)\, d\theta \\ \left.\frac{\partial x}{x}\right|_F &= -\frac{dF}{F} & \left.\frac{\partial x}{x}\right|_\gamma &= \tan(\gamma/2)\, d\gamma \end{aligned} \tag{3.19}$$

These relations show a remarkable behaviour when they are compared with the electron or JB method. The factor $1/y$, present in the electron method, is absent in the relative error $\partial x/x$ from the mixed method. Also the factor $1/(1-y)$, which is responsible for large uncertainties in the reconstruction of x in the JB method at high values of y, is absent. So indeed the intrinsic resolution in x is better compared to both the electron and the Jacquet-Blondel method.

The $y\theta$–method

The $y\theta$-method is constructed with yet another mixture of the electron and hadron flow. The method uses the angle θ of the electron, which determines Q^2, and y_{JB} as the two independent variables. Also this method is independent of the transverse momentum of the hadron flow. Compared to the mixed method, the electron energy E is eliminated here. The $y\theta$ reconstruction method reads:

$$\begin{aligned} y_{y\theta} &= y_{JB} \\ Q^2_{y\theta} &= 4A^2(1-y_{JB})\left(\frac{1+\cos\theta}{1-\cos\theta}\right) \\ x_{y\theta} &= \frac{Q^2_{y\theta}}{s\ y_{y\theta}} \qquad = x_0\frac{1-y_{JB}}{y_{JB}}\left(\frac{1+\cos\theta}{1-\cos\theta}\right). \end{aligned} \tag{3.20}$$

We write the relative errors on x and Q^2 as functions of θ, γ and F as:

$$\begin{aligned} \frac{\partial x}{x}\Big|_\theta = \frac{\partial Q^2}{Q^2}\Big|_\theta &= \frac{2}{\sin\theta}\,d\theta & \frac{\partial Q^2}{Q^2}\Big|_\gamma &= \frac{-y}{(1-y)}\tan(\gamma/2)\,d\gamma \\ \frac{\partial x}{x}\Big|_F = \frac{\partial Q^2}{Q^2}\Big|_F - \frac{dF}{F} &= \frac{-1}{(1-y)}\frac{dF}{F} & \frac{\partial x}{x}\Big|_\gamma &= \frac{-1}{(1-y)}\tan(\gamma/2)\,d\gamma. \end{aligned} \tag{3.21}$$

The intrinsic resolution of the $y\theta$ method becomes poor for high values of y in both $x_{y\theta}$ and $Q^2_{y\theta}$. At low scattering angles θ the reconstruction of $x_{y\theta}$ and $Q^2_{y\theta}$ also has a larger uncertainty.

3.4.3 The current jet mass

The current jet mass is not a measure of the quark mass, but a consequence of the effects of fragmentation, gluon radiation and jet finding algorithms. Current jet masses are generally larger than 10 GeV and can well exceed 30 GeV. The jet mass should not be confused with the mass of the $\gamma^\star p$ system, which is denoted by W. In practice, the jet mass M_c is obtained by a jet finding algorithm. We will not do this, but we present the visible invariant hadronic jet mass, M_X. For a perfect, completely hermetic detector without beam pipe hole, this mass M_X equals the mass of the $\gamma^\star p$ system, W. In the other case, when the target remnant remains completely undetected (in the beam pipe of the detector) and the event produced one current jet, the visible jet mass equals the mass of this current jet.

The expression for $\cos\gamma$ in terms of the visible mass M_X becomes:

$$\cos\gamma = \frac{P^2_{t\,h} - E^2_h - P^2_{z\,h} + 2E_hP_{z\,h}}{P^2_{t\,h} + E^2_h + P^2_{z\,h} - 2E_hP_{z\,h}} = \frac{2E_hP_{z\,h} - 2P^2_{z\,h} - M^2_X}{2E^2_h - 2E_hP_{z\,h} - M^2_X} \tag{3.22}$$

If we neglect the current jet mass M_X we end up with:

$$\cos\gamma = \frac{P_{z\,h}}{E_h} = \frac{\sum_h E_h\cos\theta_h}{\sum_h E_h} \tag{3.23}$$

which is just the energy weighted average of the cosines of the particles in the detector. If we do not neglect the invariant mass M_X, it can be constructed with equation 3.22 as follows:

$$M_X^2 = 4A\ y_{JB}\ E_h\ \frac{(\cos\theta_h - \cos\gamma)}{(1-\cos\gamma)}\ ;\quad \cos\theta_h = \frac{P_{z\ h}}{E_h}. \tag{3.24}$$

with the sum running over the visible part of the complete final state hadron flow h, that is, the sum running over all particles in the final state that enter the detector.

3.4.4 Radiative events

In the previous sections we have ignored radiative corrections. A large contribution of initial state radiation from the electron can be expected. It is clear that the extraction of the kinematic variables using any of the measurements of the outgoing electron will be more sensitive to the effects of initial state radiation than the extraction using hadron variables only.

So far we have only used two of the four measured quantities at a time. If we ignore the jet energy for the time being as it is less accurate and more difficult to define we are left with the electron energy E and the two angles θ and γ.
Using energy and momentum conservation and the three measured quantities θ, E and γ it is possible to eliminate the initial electron energy or, equivalently, it is possible to compute the radiative photon energy E_γ. We get:

$$E_\gamma \quad = \quad A - E\frac{\sin\theta + \sin\gamma - \sin(\theta+\gamma)}{2\sin\gamma}, \tag{3.25}$$

such that the reconstruction method, insensitive to initial state radiation (RC) becomes:

$$y_{RC} \quad = \quad \frac{\sin\theta(1-\cos\gamma)}{\sin\gamma + \sin\theta - \sin(\theta+\gamma)} \tag{3.26}$$

$$x_{RC} \quad = \quad \frac{E}{P}\frac{\sin\gamma + \sin\theta + \sin(\theta+\gamma)}{2\sin\gamma} \tag{3.27}$$

$$Q^2_{RC} \quad = \quad E^2\frac{(\sin\theta + \sin\gamma - \sin(\theta+\gamma))(1+\cos\theta)}{\sin\gamma} \tag{3.28}$$

We have thus developed a method of extracting x and Q^2 which is *independent of initial state radiation.* It turns out that this method shows a marked increase in accuracy over the traditional methods of extracting the kinematic variables.

In figure 3.7 we present a selection of initial state radiation events that deposited the photon energy in the LUMI-γ detector. The histogram represents the full ZEUS Monte Carlo simulation and the points the 'fall 1992 data'. In the left plot, we show the distribution of the calculated energy $E_\gamma(E,\theta,\gamma)$. It can reach values up to 22 GeV and although the statistics is limited, the distribution of data and Monte Carlo tend to agree. In the right plot we show the difference between the energy

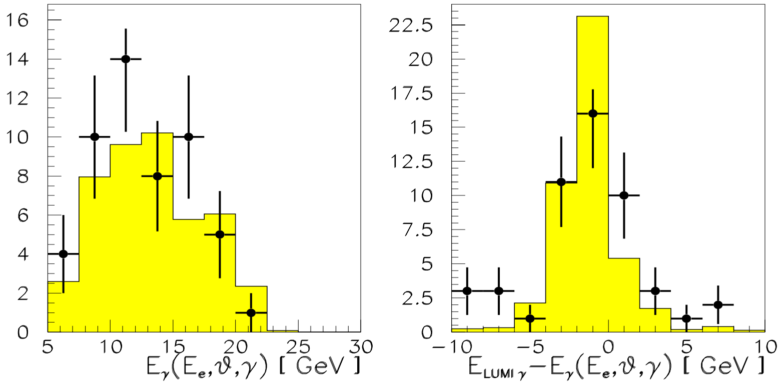


Figure 3.7: ***In the left plot we present the calculated photon energy E_γ from the MC simulation (histogram) and the data. In the right plot we show the difference of the calculated photon energy and the energy deposit in the LUMI-γ counter.***

measured in the LUMI-γ detector and the calculated photon energy $E_\gamma(E, \theta, \gamma)$. The distribution peaks at zero for both data and Monte Carlo. This result gives confidence that the kinematics, especially the angle γ, is determined correctly.

3.5 Resolution of the reconstruction methods

To investigate the accuracy with which the above formulae extract the x and Q^2 variables we used both a simple smearing model and the full ZEUS detector simulation. The smearing model covers the region of high Q^2 and high x values, the region in phase space where the cross section is small. The DIS event signatures are very clean and the need for high statistics is more urgent than detailed detector simulation. We reported results of this smearing model earlier in [53].

The full detector simulation is performed for the region of low x and low Q^2 values. At large values of Q^2 the resolution functions are approximated well by simple Gaussians, and the correlations between $\{E, \theta, F, \gamma\}$ are negligible. In contrast the resolution functions at low Q^2 are dominated by detector imperfections such as poor granularity and energy losses due to dead material in front of the detector, and leakage when the electron impacts the detector near the beam-pipe. At low values of x, reconstruction is further hampered by the fact that the hadron flow can overlap with the electron and so becomes critically dependent on the details of the electron identification. To investigate these effects a full detector simulation is necessary.

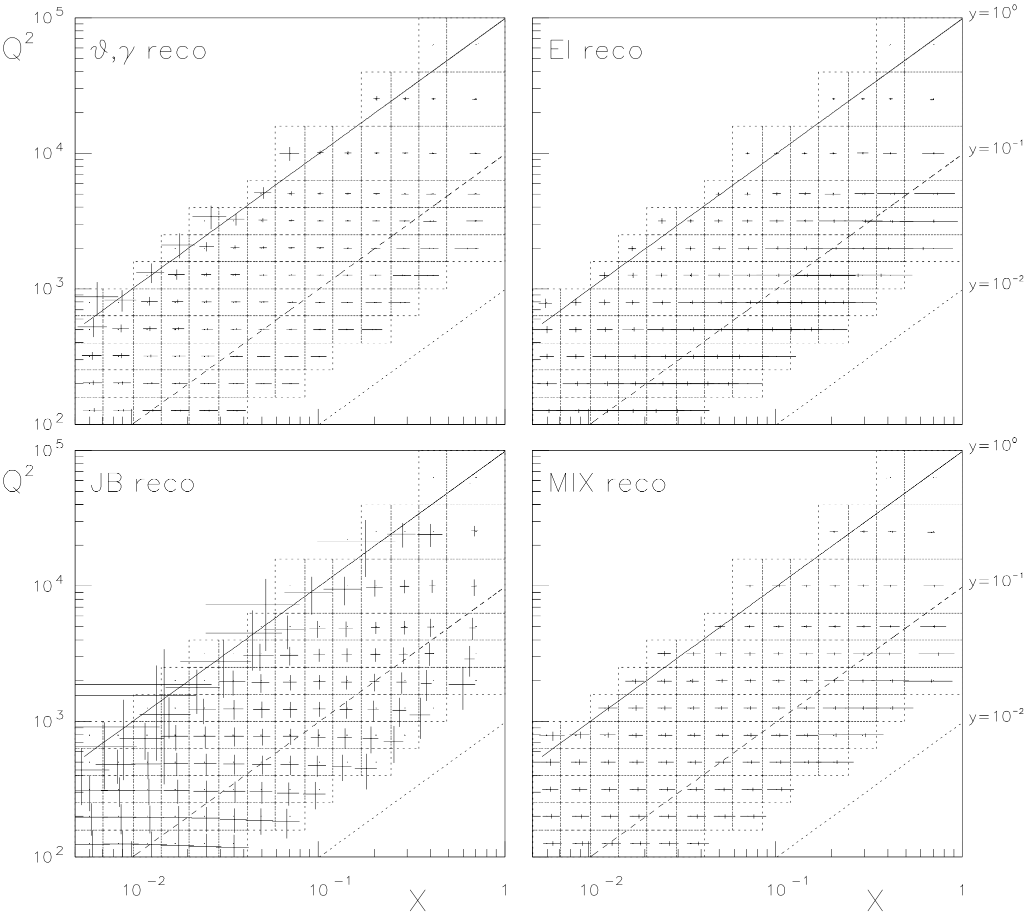


Figure 3.8: *Reconstruction of x and Q^2 according to the four different reconstruction methods. Indicated are the rms (bars in x and Q^2 direction) and bias (deviation from center of bin) of the Δx and ΔQ^2 distributions.*

3.5.1 Smearing model

We have used the DIS event generator LEPTO [32] to simulate the equivalent of 100 pb^{-1} non-radiative NC ep scattering events. The Eichten-Hinchliffe-Lane-Quigg (EHLQ) [56] parton distributions[1] served as input with $x > 0.005$ and $Q^2 > 100$ GeV^2. These events were then 'smeared' as follows. All particles entering the beam hole were removed. Subsequently the energy of electrons and photons were

---

[1] The somewhat outdated EHLQ parametrization is still in reasonable agreement with the present data in this part of the phase space. It is used since the resolution does not depend strongly on the structure function parametrization and it is implemented in LEPTO.

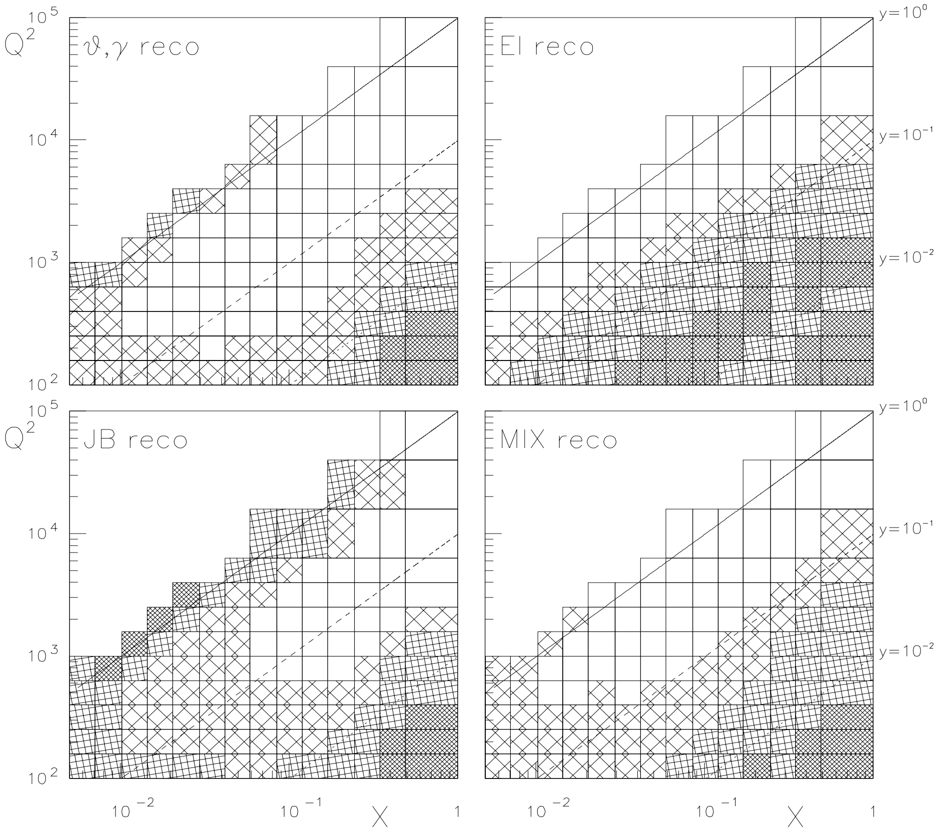


Figure 3.9: *Rms of distribution* $\Delta x/x$ *for each bin in phase space. In bins without hatching the rms is less than* 0.1, *cuts for different styles of hatching, from light to intense, were made for rms values from* 0.1 *to* 0.2, 0.2 *to* 0.4 *and* > 0.4.

smeared with a resolution $\sigma(E) = 18\%\sqrt{E}$ and the hadrons were smeared according to $\sigma(E) = 35\%\sqrt{E}$. The polar angles of the hadrons were smeared with $\sigma(\theta) = 35$ mrad (approximately the resolution as determined by the cell size of the ZEUS calorimeter). The scattered electron was treated separately in that the expected resolution of the tracking detectors was taken into account.

In figure (3.8) we show the (x, Q^2) phase space with a logarithmic binning for both x and Q^2. For high x as well as high Q^2, the bins are large because of the low statistics in this part of phase space. In each bin the distributions Δx (defined as $x_{rec} - x_{MC}$) and ΔQ^2 (defined similarly) were determined. In the figure deviations from the center of the bin reflect the bias of those Δx and ΔQ^2 distributions. The 'error bars' in both directions represent the 'root-mean-square' (rms) of the distributions. Hence the ideal reconstruction would give a point in the center of the bin.

We left out the low-y part ($y < 0.05$) for all four reconstruction methods because there the rms becomes too large and would make the figure unreadable.

As one observes the Q^2 reconstruction is almost always quite accurate except for the Jacquet-Blondel method and for very high y values in the double angle reconstruction. However, it should be noted that for large parts of the phase space the double angle method is superior to the other methods.

Perhaps more interesting is the reconstruction of x. Figure (3.8) shows that the electron reconstruction of x is poor for low y values, in accordance with equation 3.10. The double angle and mixed reconstruction are the most accurate, although the x reconstruction with Jacquet-Blondel at low y values is also good. To illustrate the differences of the x reconstruction for the four methods more clearly, we show figure (3.9) where now the rms of the relative error in x, $\Delta x/x$, is plotted. One observes that compared to the conventional electron reconstruction one *gains more than one order of magnitude in* y using a mixture of hadron and electron measurements. For very low y-values the double angle reconstruction is somewhat better as compared to the mixed reconstruction and *vice versa* for high y-values.

3.5.2 Full ZEUS detector simulation

To investigate the resolution at low Q^2 values, we generated 50k NC DIS events with $Q_l^2 > 4$ GeV^2, with the event generator HERACLES that simulates photon radiation of the lepton realistically (see also section 1.6), together with the ARIADNE colour dipole model for the hadronic fragmentation. The structure function parametrization used was MTB1. This sample of $\sim$ 230 nb^{-1} was then passed through the detailed ZEUS detector simulation. The reconstruction of the kinematic variables was done with the central tracking and calorimeter information, for which the actual configuration during the 1992 data taking runs was incorporated, and deviates in no respect from reconstruction of kinematics for real ZEUS data. In the next chapter we discuss the use of this detector information to obtain the event vertex, the final state electron energy and angle, and the hadron flow at length. As it turns out, a number of selection criteria are needed to suppress the non-DIS background events in the data sample, and the Monte Carlo simulated events are subjected to the same criteria to obtain realistic acceptance corrections. Also a number of criteria are used to increase the accuracy to reconstruct (x, Q^2). We explain all selection criteria in the next chapter, but the effects on the reconstruction of (x, Q^2) are investigated already here. A summary of the selection cuts can be found in section 4.9.

In figure 3.10 we present the resolutions in γ and F, as obtained from the calorimeter cell information. In both graphs the difference of the reconstructed and the generated values are shown. The dashed lines correspond to the complete sample, that is, those events that leave an electron signal in the detector which

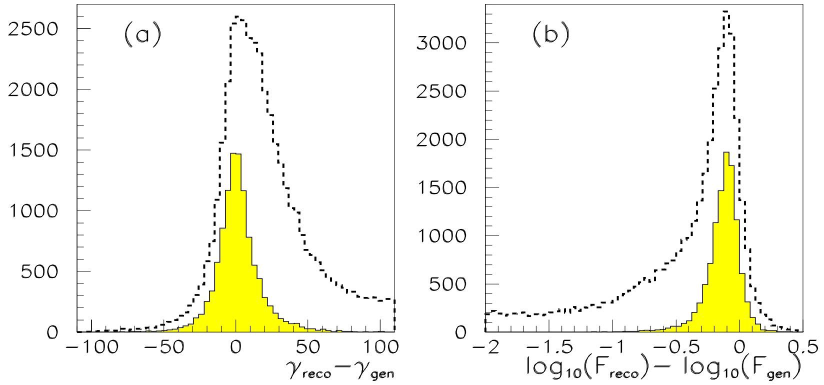


Figure 3.10: ***Resolution of γ and F. In plot (a) we show the resolution of $\Delta\gamma = \gamma_{reco} - \gamma_{gen}$ for $Q_h^2 > 10$ GeV^2 (dashed line), and in the same plot the resolution $\Delta\gamma$ after applying the selection filters (shaded histogram). In (b) the same for $\Delta F = \log(F_{reco}) - \log(F_{gen})$.***

is recognized as such. The distributions are biased towards higher reconstructed angles γ and lower energies F. In the same figure we plot the distributions after the selection criteria have been applied. The bias in γ disappeared and the width of the distribution also decreased. The bias in F becomes less but has not disappeared completely, which indicates that a fraction of the energy is lost due to the dead material in front of the calorimeter.

Together with the reconstruction of the electron variables, the kinematic variables x and Q^2 are determined according to the various methods. In figure 3.11 we show the mean migration of the events in the (x, Q^2) phase space for four reconstruction methods. In these plots, where we suppressed the events with large photon radiation, the phase space is binned and for each bin an arrow is drawn from the mean value of the generated to mean value of the reconstructed x and Q^2. Hence the direction and length of the arrow indicate the mean migration of the events in the phase space.

For the electron method, the migration is small for large values of y, whereas for low y-values the events migrate to lower values of x. The resolution in Q^2 is good for the complete phase space. In the Jacquet-Blondel method the value of Q^2 is reconstructed too low, due to the shift in F. For the events which have the current jet in the beam-pipe, at low values of y, the events migrate to the isoline of γ that corresponds to the edge of the beam pipe. The double-angle method shows a large increase in accuracy in x for lower values of y, if one compares it to the electron method. The resolution in Q^2 is similar to the electron method, since both use information of the electron polar angle θ. In the $y\theta$ method the resolution in x

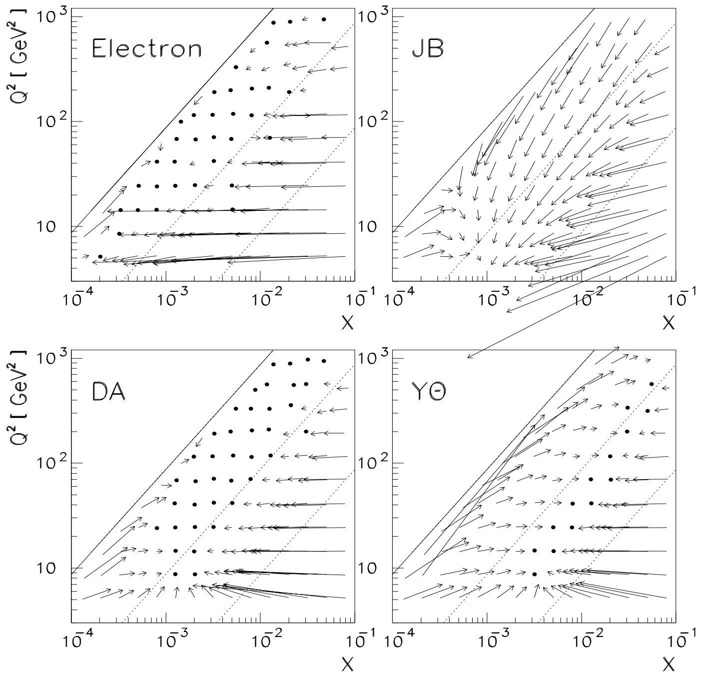


Figure 3.11: *Migration of the reconstructed kinematic variables. The base of the arrow is positioned at the mean generated values of x and Q^2 and it points at the mean reconstructed value for x and Q^2. Events that radiated a photon from the lepton line with $E_\gamma > 1$ GeV are removed in this plot.*

is the best at low values of y, but is worse at high values of y, in the region where y and θ isolines become parallel.

A different way of presenting the resolution of the kinematic variables is given in the plots of figure 3.12. For the five reconstruction methods the generated versus reconstructed values of x, y and Q^2 are plotted.

The behaviour of the resolution in Q^2 is already discussed; it is accurate for the methods that use the electron angle information and is somewhat poorer for the Jacquet-Blondel reconstruction. The reconstruction of y is interesting, since in all methods the value of x is constructed via $x = Q^2/sy$. Note that this reconstruction of y is identical for the JB, MIX and $y\theta$ method.

The reconstructed y_{elec} is most often larger than the generated y, due to electron energy-loss in the dead material and photon energy radiation. Reconstruction of y_{elec} is hardly possible below $y_{elec} \sim 0.01$, and this is the origin of the enormous spread and migration in x_{elec}. The spread of $y_{_{JB}}$ is smaller, but also the value of

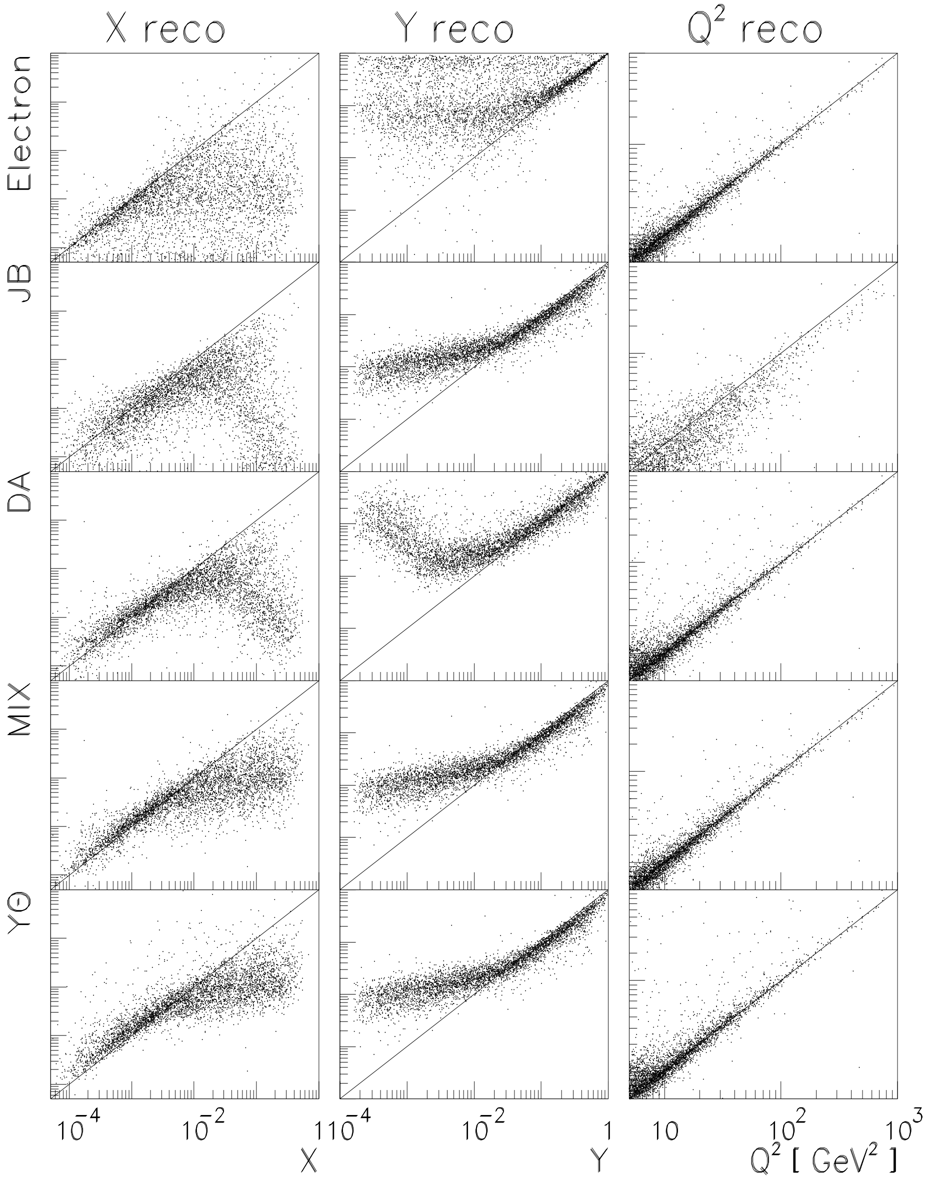


Figure 3.12: *Resolution of x, y and Q^2 for the five reconstruction methods. On the horizontal axis we plot the generated values, on the vertical axis the reconstructed values for x, y and Q^2. A line that corresponds to perfect reconstruction is drawn to guide the eye. No selection criteria are applied.*

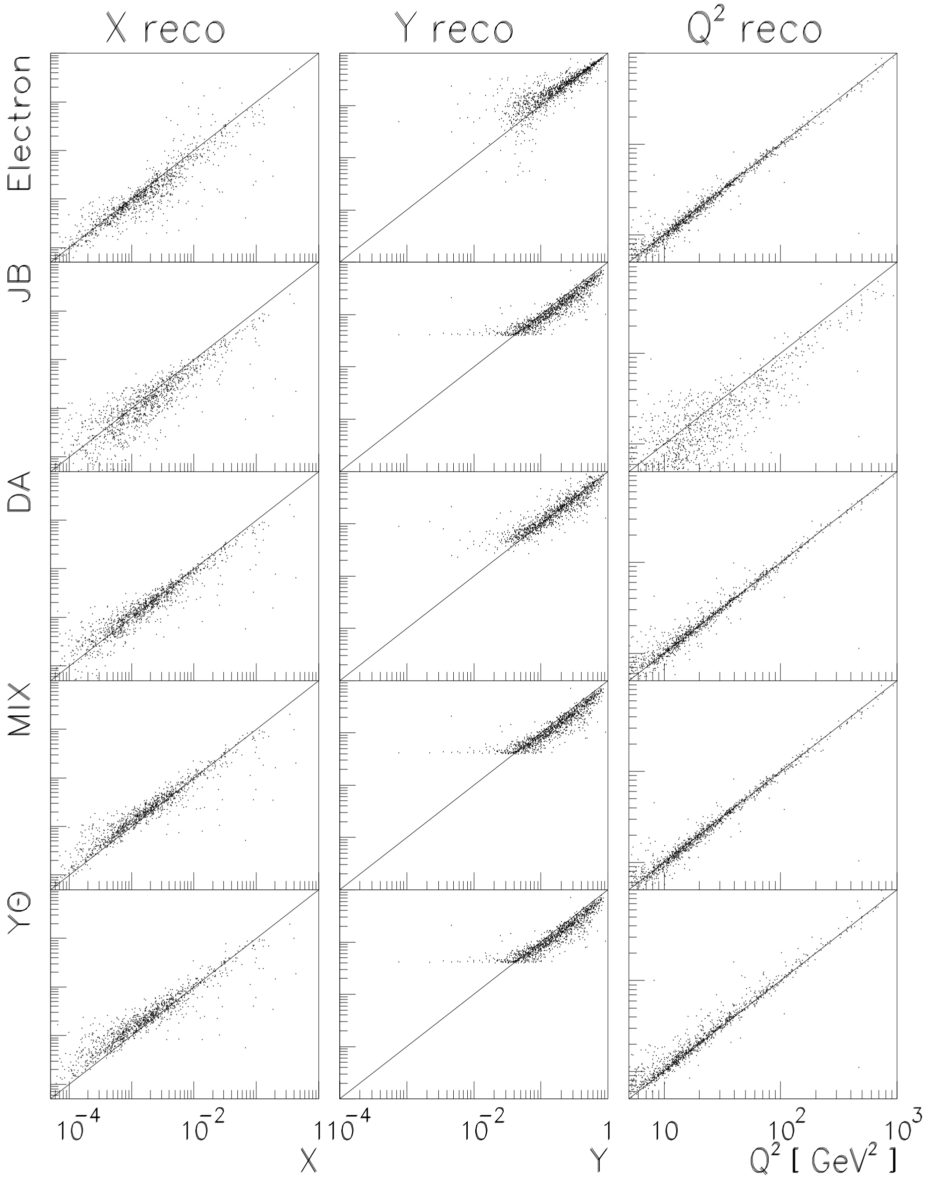


Figure 3.13: *Resolution of x, y and Q^2 for the five reconstruction methods. The same plots as in the previous figure are drawn, but now with all selection criteria imposed. The selection criteria are listed in section 4.9.*

y_{JB} does not end up below $y_{JB} \sim 0.007$. For these very low values of y the current jet disappears in the forward beam pipe, and the Jacquet-Blondel method picks up calorimeter cell noise which is interpreted as real energy deposits. The cell noise of an 'empty' calorimeter corresponds to a mean $y_{JB} \sim 0.007$ (see also section 2.4.3).

Due to the specific algorithm to obtain γ, the cell noise of an 'empty' detector is effectively weighted with $(E - P_z)$ and pulls the angle γ towards the rear direction. This is equivalent to large values of y_{DA}. This is indeed visible in the resolution plot of the y_{DA} where the low generated values of y are reconstructed at high reconstructed values y_{DA}.

The reconstruction of x is most accurate for the double-angle, MIX and $y\theta$ methods. All these methods are decreasingly accurate for higher values of x in this region of phase space.

In figure 3.13 the same resolution plots are shown, but now with all selection filters imposed. Although many events are rejected, the resolution of the remaining sample is much better compared to the complete sample. The typical resolution of Q^2_{DA} is 25%, independent of Q^2_{DA}. The relative resolution in x_{DA} varies smoothly with x_{DA}, and is 25% for $x_{DA} = 10^{-2}$, 60% for $x_{DA} = 10^{-3}$, and 80% for $x_{DA} = 10^{-4}$. Where the smearing model and the full simulation overlap, the results are in agreement.

3.6 Summary

In this chapter we presented the reconstruction of the kinematic variables for inclusive neutral current ep scattering. Via introduction of F and γ, the quantities that describe the hadron flow, together with the final state electron energy and angle, six reconstruction methods for x and Q^2 are possible. Two of them, the electron and Jacquet-Blondel method, have been proposed and used earlier. Of the novel methods, the one that uses the electron and hadron angles only (double-angle method) is very promising. We showed that the success of the double angle reconstruction is also due to the fact that the hadron angle γ is stable against energy losses in the direction of the jet. Since x is reconstructed from the relation $Q^2 = sxy$, the methods can also be combined, and all in all five methods were discussed in detail.

The methods were tested and compared with Monte Carlo simulation in the high Q^2 region with a simple smearing model and with the full ZEUS detector simulation in the low Q^2 region. The new methods, especially the 'double-angle' method, show a marked increase in accuracy when compared to the conventional methods, in particular in the low y region whereas in the high y region comparable results are obtained.

Chapter 4

DIS data selection

Only very few of the huge number of triggers that were recorded during the initial running periods in 1992 actually correspond to genuine neutral current DIS events. For these events the electron is deflected through the interaction with the proton such that it strikes the calorimeter and deposits its energy there. The bulk of triggers, however, are due to of background events of various types which have to be recognized and removed from the sample.

In this chapter the general selection procedure is described. A short summary of the ZEUS on-line trigger system is given. The definition of the Data Summary Tape (DST), used to store the results of the first step in off-line data reduction, is discussed.

Further background rejection on the DST sample, as well as the requirements to improve the accuracy of the kinematic reconstruction of the DIS events, are discussed at length. The reduction factors of the various selections are listed. Finally an estimate of the remaining background in the sample is given.

4.1 Selection procedure

The selection of 'gold plated neutral current events' is formulated in terms of a number of criteria for recognizing and rejecting background events. To this end we catalogue the various backgrounds and discuss signatures for recognizing each category.

The first requirement is simply that all components of the ZEUS detector which are used in the analysis are active and functioning. Events which were recorded when this requirement was not fulfilled were rejected and the luminosity was corrected accordingly.

The second requirement, most obvious but difficult to achieve, is to actually select colliding-beam interactions, that is, interactions between a 26.7 GeV electron and a 820 GeV proton. The *non-colliding beam* interactions give rise to the highest rates in the ZEUS detector. In particular, the cross section for collisions between protons of 820 GeV and restgas molecules in the beampipe is huge. The rate of this type of interactions is estimated to be approximately 0.5 kHz/m. The spray of particles that is produced during these proton-gas interactions, or products from secondary interactions with accelerator elements such as magnets and collimators, can enter the calorimeter and deposit energy, thereby producing a trigger signal. It is important to recognize the beam-gas interactions in order to remove them from the sample. In general, the proton-gas interactions occur all along the proton beamline, but the signatures in the detector depend on their vertices. First, the particles emanating from proton-gas interactions that have the vertex downstream (positive $\mathbf{z}$ position) of the main detector do not enter the calorimeter and apart from the fact that they reduce the proton beam life time, are completely harmless. Second, the interactions can occur inside the detector. Those background events can deposit a large amount of energy in the forward calorimeter FCAL, in the direction of the incident proton, whereas the rear part, RCAL, remains empty. This type of events is recognized with the help of energy-momentum conservation. The event is rejected if the visible momentum is inconsistent with DIS (see section 4.4.4). The third class of proton-gas interactions occurs upstream of the detector and the spray of particles can hit both the RCAL and the FCAL. Monte Carlo studies show [60] that beam-gas interactions with vertices up to $\mathbf{z} = \mathcal{O}(-100\ m)$ are visible in the calorimeter. The total rate seen in the detector is of the order of $\mathcal{O}(50\ kHz)$, approximately five orders of magnitude larger than the rate for beam-beam collisions. This large class of proton-gas interactions is rejected with the beam monitor time counter **C5** and the calorimeter timing information. The vetowall is also used to reject this type of background.

A source of the electron beam induced background is the collinear synchrotron radiation from the electrons. However, most of this type of radiation occurs at large curvature of the HERA ring, far away from the experiments. Shielding with a set of masks and collimators in the electron beam reduces the remaining fraction of this type of background considerably. The electrons can also collide with nuclei of remaining rest-gas in the beampipe. The collisions of such background events are analogous to the fixed target events with $s = 26.7$ GeV, and need to be rejected. This type of background events is especially dangerous when they occur inside the main detector close to the nominal event vertex. When the electron is deflected such that it enters the main detector, the signatures of these events are difficult to distinguish from nominal DIS events. However, the rate of these electron-gas events can be estimated with the electron pilot bunches, and proved to be small (see section 4.8).

Separate classes of non-colliding beam background are the cosmic muon and beam halo muon events. They are discussed in section 4.5.5.

Requirement three, applied after rejecting beam-gas background, has the purpose of recognizing other than neutral current DIS events. Most important is the background from *photo-production.* Here the electron scatters only little during the interaction and disappears down the beampipe. The background originates from the fact that occasionally the electron finding algorithm wrongly identifies a scattered electron in the hadronic activity of the event. Although this happens rarely, the total cross section for photoproduction as measured by ZEUS [48] is $153 \pm 16 \pm 32$ μb, approximately $1.5 \cdot 10^3$ times the anticipated cross section for DIS events with $Q^2 > 10$ GeV^2, and hence photoproduction background is a serious problem. Most of the photoproduction events are rejected by selecting only the events in which the final state electron is identified. Also momentum conservation is used to reject this type of events. But, as it turns out, this type of events even contaminates the final sample of DIS events and the amount has to be estimated (section 4.8).

After the selection of DIS neutral current events, the kinematics has to be reconstructed reliably. For this the electron has to be well contained in the calorimeter, for which a fiducial cut around the beampipe is defined and applied. Also at very low values of y_h, the noise of the calorimeter cells dominates hadronic signal, and a cut on the value of the estimator of y, $y_{_{JB}}$, is applied. These last types of filter reduce phase space and statistics of the sample but the quality of the measurable quantities increases so that the structure function can be determined reliably.

4.2 Trigger, event reconstruction and DST

4.2.1 On-line trigger setup

The data were accumulated with the ZEUS on-line trigger system. The trigger system aims at a rejection factor of 10^5 while maintaining full efficiency for DIS events and other interesting physics events. The design ZEUS trigger system operates on three levels, that reduce the beam crossing rates (10 MHz) to accepted trigger rates of 1 kHz, 100 Hz and 5 Hz respectively [46, 59, 60, 61]. Besides triggers induced by tuned selection algorithms, each trigger level accepts scaled down random triggers for test purposes.

- **First Level Trigger (FLT)** For each beam crossing a FLT decision is made. The analogue signals of the components are pipelined to allow a trigger decision time of 4.42 μs. For the 1992 runs the calorimeter and **C5**-counter were used for the FLT. For the FLT the calorimeter was divided into 448 non-overlapping trigger towers for both the EMC and HAC sections. Each trigger tower was typically formed by a pair of adjacent cells. Whenever the energy deposit in one of the towers exceeded a programmable threshold, listed in table 4.1, a FLT-accept was issued. The thresholds generally decrease with increasing distance to the beampipe. The **C5** counter was used to veto triggers that were inconsistent with the beam arrival times. The experimental dead-time due to trigger and read-out was less than 0.1%.

- **Second Level Trigger (SLT)** The SLT has acces to the digitized data, which implies that the calculations on the SLT are more precise compared to

CAL	section	Number of towers	EMC thres	HAC thres
FCAL	beampipe	4	50	70
	inner	20	20	25
	outer	88	10	10
RCAL	beampipe	4	10	2.5
	other	108	1	1
BCAL		224	2.5	1

Table 4.1: *The calorimeter FLT thresholds in units of GeV. The thresholds and the number of towers with that threshold are listed. The CAL is divided into towers next to the beampipe (beampipe), a ring surrounding them (inner), and the remaining towers (outer).*

the FLT. For the calorimeter SLT, timing information is available. In 1992, the SLT was commissioned and tested for the most part, although a rejection of background events based on the timing information was implemented towards the end of the running period. A detailed discussion of background event rejection using the time information of the CAL is given in section 4.4.2.

-1

- **Third Level Trigger (TLT)** The TLT has access to the full data of all components. The TLT rejected clear 'spark' events. A discussion of these 'spark' events follows. The rejection using the timing information of the CAL was tightened. Several algorithms were implemented to reject cosmic muon events and beam halo muons. The events that were accepted by the TLT decision, were written to tape.

In the fall period of 1992, a total of $4.2 \cdot 10^6$ triggers were recorded on tape.

4.2.2 Event reconstruction

The ZEus PHYsics Reconstruction program (ZEPHYR) organizes the raw data into ADAMO [62] data structures with ZEBRA memory management for further analysis. In a first phase, ZEPHYR reconstructs the calibrated energies and track segments for each detector component separately.

In the calorimeter first phase event reconstruction, the variations of the PMT gains are repaired using the UNO (uranium noise) signals. During the running period, the UNO signal is measured regularly. The variation of the the latest measurement of the UNO I_{UNO}^{meas} from the nominal UNO current I_{UNO}^{nom} is used to correct the PMT raw energy. The calibrated energy of each calorimeter cell is obtained by adding the corrected energies of the two corresponding PMT's:

$$E_{cell} = \frac{I_{UNO}^{nom}}{I_{UNO}^{meas}} E_{PMT}(left) + \frac{I_{UNO}^{nom}}{I_{UNO}^{meas}} E_{PMT}(right) \tag{4.1}$$

For each cell the energy *imbalance* is determined as the difference of the two calibrated PMT energies. The requirement on the minimum energy of the cell in order to remove noise, described previously in equation 2.6, is applied at this stage.

Since each cell in the calorimeter is read-out by two PMT's, the system provides some redundancy. If one of the PMT signals is lacking, due to failure of the front end electronics, a correction is made. The cell energy is then obtained by doubling the PMT energy of the remaining PMT, and the imbalance is set to zero. The number of these 'bad' channels is required to be below 1-2% of the total number of PMT's. For a larger fraction of 'bad' channels, access to the detector was made to repair the front end electronics. The probability that both PMT's of one cell fail, which means a 'hole' in the detector, was very small ($\leq 0.1\%$).

Also offsets in the measured time of the cells are included off-line. We discuss the calorimeter timing in detail in section 4.4.2.

For the CTD and vertex detector, ZEPHYR identified in the first phase track segments from the calibrated hits. The track reconstruction code is able to find tracks using standard techniques of seed finding, pattern recognition and track fitting. At least four hits were required to create a track. The efficiency of track finding was shown to be high in the region of good acceptance.

In subsequent phases the information of detector components are connected to identify global quantities. The global track matching extrapolates the track segments of the CTD, vertex detector and muon chamber over the sub-detector boundaries, and matches the track segments to each other. The global tracks are used to identify the position of the event vertex. The information of the tracking devices and the calorimeter are connected. In the calorimeter, calorimeter cluster algorithms merge cells to identify 'single particle' showers. It is also aimed to merge fragmentation results of single partons in global clustering algorithms. These jets play an important role in the global event reconstruction.

The ultimate aim of the event reconstruction is to reconstruct the 4-vectors of the produced particles in the interactions. Due to limited information of the tracking devices in 1992, and the complicated structure of the *ep* collisions, this goal was not fully reached. However, this did not prevent an accurate measurement of inclusive *ep* scattering, as will be shown in this and the following chapters.

4.2.3 Data Summary Tape

After full reconstruction of the recorded events, a Data Summary Tape (DST) is created for each physics analysis at ZEUS. The physics groups supply appropriate routines that flag the interesting events (additionally, events are flagged for testing purposes i.e. scaled down triggers etc). All events that are flagged by at least one of the groups define the DST. The aim of this filter stage is to reduce the recorded sample by a large fraction via rejection of the definitely non-interesting triggers.

The DST for the NC structure function group [63] selected approximately 5.5% of the recorded triggers. It resulted from the following tests on the events:

- **Trigger bit.** The characteristic of NC is that the final state electron enters the CAL and deposit its energy in the EMC section. Therefore the event had to be triggered by either the RCAL(EMC) or the BCAL(EMC) trigger on the first level. Events that were triggered by the FCAL only, were rejected.
- **Minimum energy.** It was required that the calorimeter parts contain a certain minimal amount of (calibrated) energy. For the BCAL(EMC) triggers the energy should be $E_{BCAL(EMC)} \geq 5$ GeV, and for the RCAL(EMC) triggers $E_{RCAL(EMC)} \geq 2$ GeV. If both sections fired only one of the minimum energy conditions had to be fulfilled.
- **Timing.** The requirement on the calorimeter timing, as mentioned in the discussion of the on-line trigger, is tightened again. Also the beam monitor counter **C5** is checked again.
- **Muons.** Clearly identified muons in a calorimeter section were rejected.

After this filter stage, the DIS DST sample contained 228491 events (approximately 9 μb) for the fall 1992 period. They were subjected to detailed off-line analysis.

4.3 Monte Carlo simulation

The complete chain between the *ep* collisions and reconstruction of the events is duplicated in Monte Carlo simulation programs. These are indispensable for a correct understanding of the detector response and background events. The simulated events are used to understand the background events and to correct the data for acceptance and detector smearing. Event generators produce four-vectors according to the desired type of *ep* scattering. The four-vectors are fed into the GEANT based, ZEUS Monte Carlo simulation program. It provides an accurate description of the complete ZEUS experiment, which include a detailed configuration of all

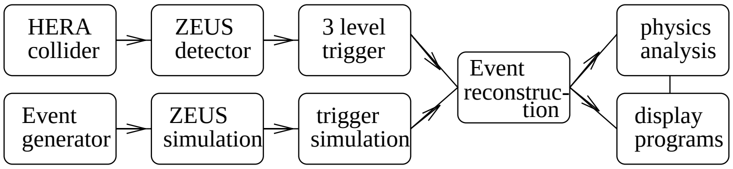


Figure 4.1: *Overview of the data and simulation chain.*

sub-detectors with both active and dead material, and it simulates the response and read-out on the deepest level possible. The simulation of the sub-detectors took into account their status in 1992. For the calorimeter, the shower routines were adapted and modified such that the results of the ZEUS test data were reproduced by the simulation. The calorimeter simulation also includes the uranium and electronic noise. An overview of the data and simulation chain is presented in figure 4.1.

In principle, the output of this simulation program is indistinguishable from the data and is processed by the next module, the one that simulates the ZEUS trigger system. Here the Monte Carlo events that would have been rejected by the trigger system are identified and marked.

The ZEUS simulation software provides a complete and realistic sample of DIS events with a desired structure function parametrization, including radiative effects. It also provides samples of background events. Both the data and the simulated events can be examined and compared with help of display programs. For ZEUS, the two-dimensional LAZE [64] (see for example figure 4.13) and three-dimensional GAZE [65] (see figure 4.19) display program were developed.

4.4 Background rejection

A number of algorithms have been developed to recognize background events. The most important ingredients used were recognition of apparatus failure, calorimeter timing, conservation of four-momentum and presence of an electron. They will be discussed here.

4.4.1 Apparatus failure

In the first year of data taking, not all of the ZEUS subdetectors were always operational. For inclusive DIS measurement functioning of the central drift chamber, the calorimeter, the magnet and the luminosity system is essential. For each run the status of the subdetectors was examined. The CTD should be at full high voltage, the magnet should be switched on, the calorimeter had to be operational and luminosity information had to be available. Triggers that were recorded while these requirements were not fulfilled were excluded from the sample. The luminosity recorded during these periods of detector failure was then also removed.

During the fall period the DST event selection rate went up drastically for some runs due to some 'overefficient (hot)' readout trigger channels in the BCAL. The reason for this was that occasionally the calorimeter photo-multipliers suffered sudden discharges between the (increasing) static charge on the PMT and PMT shielding. These 'sparks' occur mostly in the BCAL-EMC section. The discharge induces a large signal in the affected PMT that is interpreted as an energy deposit. This "energy" is then large enough to fulfill the trigger requirements. In the RCAL and FCAL sections this effect was anticipated and the shielding was put on high voltage to avoid such discharges. Spark events are characterized by a large energy imbalance of a calorimeter cell, because only one of the two cell-PMT's discharge, while all other calorimeter cells remained empty. The spark events are not beam correlated and occur random in time. Spark events are easily recognized and spark-rejection algorithms were implemented in the trigger (the TLT as well as the SLT) in the latter part of the running period.

4.4.2 Timing

Besides the energy deposit itself, the time the particle enters the calorimeter cell is extracted from the pulse of the PMT. The resolution of this timing measurement improves with increasing energy deposit as $\sigma(t) \simeq 500 + 1500/\sqrt{E}\ ps$, which translates to approximately one nanosecond for a few GeV. Time offsets on the individual PMT's are adjusted such that for interactions originating from the nominal interaction point and at the nominal bunch crossing time, the calorimeter cell time equals zero. These offsets are determined through a combination of laser runs, charge injection runs and geometrical considerations. During the running period they are constantly monitored and adjusted if necessary.

The PMT time is used to construct a time of the global sections of the calorimeter (the R/B/FCAL) by calculating an energy weighted average of the individual PMT's in the corresponding sections. This is done at the trigger level already, but in the off-line selection filter the most stringent PMT time weight calculation is made:

$$t_{R/B/FCAL} = \frac{\sum_{i_{PMT(R/B/FCAL)}} w_i\, t_i}{\sum_{i_{PMT(R/B/FCAL)}} w_i}; \qquad \begin{array}{l} w_i = \min\{E_{PMT}, 2\ \mathrm{GeV}\}, \\ w_i > 0.2\ \mathrm{GeV}, \end{array} \tag{4.2}$$

with the additional requirement that at least two PMT's contribute to the time measurement. Events that do not satisfy this requirement do not have a time measurement in that particular calorimeter section.

In contrast to the FCAL, the energy deposit in the RCAL from an upstream proton-gas interaction occur early compared to the interactions originating from the $z = 0$ vertex. The time difference equals two times the distance traveled from $z = 0$ to the RCAL, approximately 11 *ns*, and is well measured. However, this picture is somewhat complicated by the size of the proton bunch length (the size of the electron bunch length can be neglected). The proton bunch length broadens the time distribution of the particles that enter in the FCAL. For *ep*-interactions the time width of the RCAL is insensitive to the proton bunch length since it depends on the electron bunch length only. But for upstream proton-gas interactions both the FCAL and RCAL time distributions are affected by the proton bunch length. This can be seen in figure 4.2, where we plot t_{RCAL} versus $t_{FCAL} - t_{RCAL}$ for the DST accepted events. The two peaks separate clearly between events that occurred upstream and inside the detector. The peak located at the origin ($t_{RCAL} = 0, t_{FCAL} - t_{RCAL} = 0$) of the graph represents the *ep* interactions. The signal of the *ep*-interactions is tilted compared to the upstream proton-gas interactions due to the different effects of the proton bunch length in *ep*-interactions for t_{RCAL}

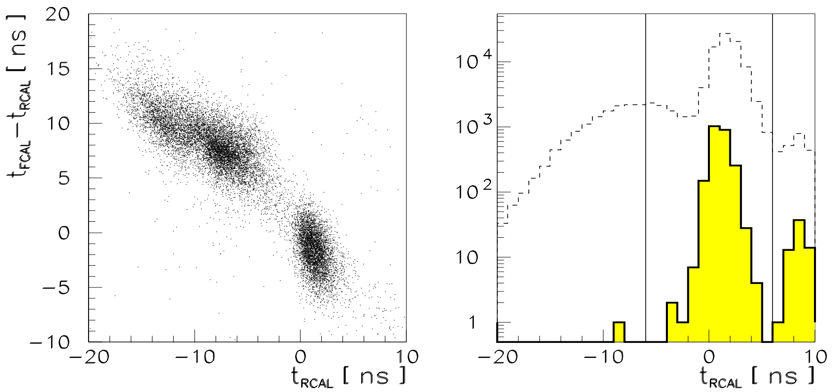


Figure 4.2: *The calorimeter time distribution. In the left plot, the interactions with upstream vertices are clearly separated from the interactions with vertices inside the main detector. In the right plot, the RCAL time distribution for the DST sample is drawn, denoted with a dotted line. The histogram represents events that pass all, except timing, selection filters. The cut values on t_{RCAL} are denoted with lines.*

and $t_{FCAL} - t_{RCAL}$. It should be noted that the distribution of the proton-gas interactions is already reduced considerably by the (somewhat looser) timing cuts performed at the trigger level.

A second complication to the timing distributions are the run to run differences of the HERA proton and electron bunch crossing times. This time difference is due to a shift of the bunch position with respect to the HERA machine RF. The electron and proton bunch offsets are measured with the **C5** beam counter monitor for each run. The offsets can therefore not be corrected on-line. In the off-line analysis a run-dependent time-offset for the three calorimeter sections is made. Because the FCAL-time is sensitive to the proton arrival time, the offset for t_{FCAL} equals the proton arrival time. For the t_{RCAL} the offset of the electron arrival time is incorporated. The BCAL is affected by both and t_{BCAL} is shifted by the average of the two.

Events were rejected if they have a time measurement in a calorimeter section and correspondingly:

$$\begin{aligned} \mid t_{FCAL} \mid &> 6\, ns, \\ \mid t_{RCAL} \mid &> 6\, ns, \\ \mid t_{FCAL} - t_{RCAL} \mid &> 6\, ns. \end{aligned} \tag{4.3}$$

Timing was not simulated in the ZEUS Monte Carlo. Hence acceptance corrections are not made for lost *ep*-events that occur inside the main detector. The fraction

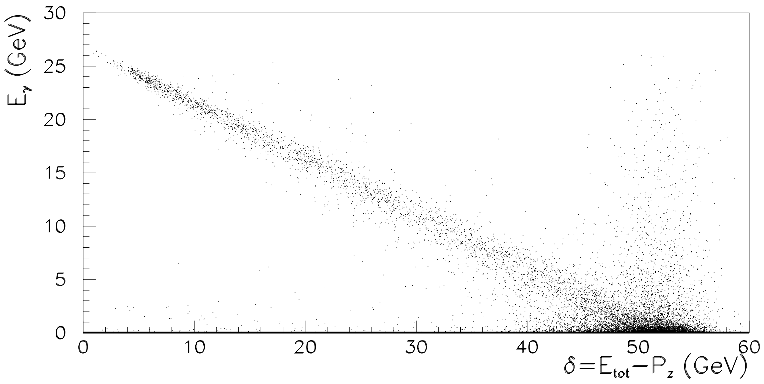


Figure 4.3: *The δ-distribution for DIS events. In the figure the correlation between δ and the radiated photon energy E_γ is plotted.*

of wrongly rejected events is estimated from a Gaussian fit of the events inside the main detector. The tail of the fitted Gaussian, cut by the time filter, contains only a fraction of $\sim 10^{-4}$ of the accepted events.

The BCAL time is used to reduce cosmic muon events. For this purpose we split the BCAL into an upper and lower half and we compute the times separately. A time-difference between the two halves signals cosmic muon events that penetrate the detector from above or below. We then rejected the events for which:

$$-10 < t_{BCAL} < 20\, ns \quad \wedge \quad |\, t_{BCAL}(\text{up}) - t_{BCAL}(\text{down})\,| < 10\, ns. \tag{4.4}$$

4.4.3 The vetowall

The vetowall signals particles that enter the detector from the upstream direction. If a coincidence of the inner and outer vetowall counters is measured, a particle has penetrated through the vetowall which signals beam gas events. These events were removed from the sample. There was a large overlap between the events that were rejected with the vetowall signal and the events failing the requirement on the timing of the calorimeter.

4.4.4 Momentum conservation

An important property of the ZEUS detector, is its almost complete hermetic enclosure with respect to both electromagnetic and hadronic energy deposit. Nevertheless, for DIS a large fraction of the proton remnant energy is lost in the forward beampipe. Hence the visible energy or longitudinal momentum are not conserved separately. But the effect of energy loss in the forward direction can be minimized by taking the difference of the energy and longitudinal momentum. We define δ as

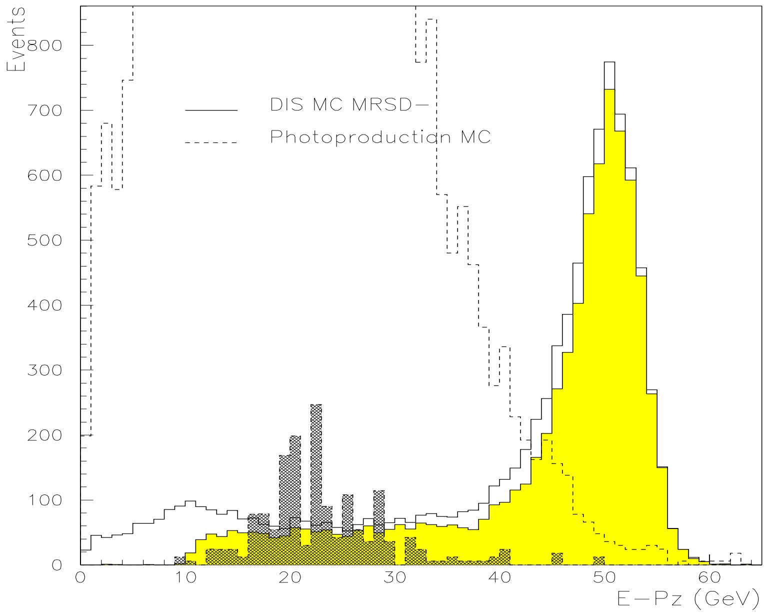


Figure 4.4: *The δ-distribution for photoproduction events compared with DIS. The dashed histogram shows the distribution for photoproduction events before any cut is applied. In the solid histogram the events are shown which survive the electron identification algorithm (*EF1*, see the next section) with $E_e > 5$ GeV. The solid line is the DIS MC sample and the events in the shaded histogram survive the electron finding algorithm.*

this difference:

$$\delta \equiv E - P_z = \sum_i E_i(1 - \cos\theta_i), \tag{4.5}$$

where the sum runs over all calorimeter cells and θ_i is the angle with respect to the incoming proton. For an absolutely hermetic detector, δ should equal two times the energy of the electron beam, $2A$, as can be deduced from the conservation laws of E and P_z. Particles lost in the beampipe in the proton direction contribute only little to δ since θ_i is small for them, in particular, δ is very insensitive to the proton remnant which clusters around the proton beam direction.

Since δ is a global quantity it is independent of any electron finding routine. For DIS however, where the scattered electron is recognized, the sum over the calorimeter-cells can be split into cells which belong to the electron and cells which belong to the hadronic final state. We can rewrite δ, using the formulae of section 3.4.2, as a sum of contributions from the electron and hadron flow:

$$\delta = \delta_{elec} + \delta_{hadron} = 2A(1 - \Delta y); \qquad \Delta y = y_{elec} - y_{JB}. \tag{4.6}$$

Hence, δ measures the difference in y of the electron and hadron system. To reject ill-contained DIS events or background events we required

$$35 < \delta < 60 \text{ GeV}. \tag{4.7}$$

In general, QED radiation from the electron spoils the equality of the kinematic variables for the electron and hadron system. This cut on δ is therefore a very powerful tool to remove hard initial state radiation [29], because δ is reduced by any energetic loss down the beampipe in the electron direction. Thus, for DIS events with emission of hard photons collinear with the initial electron (ISR) that disappear in the beampipe, δ is lowered:

$$\delta^{ISR}_{meas} = 2(A - E_\gamma). \tag{4.8}$$

In figure 4.3 we plot the correlation between δ and the photon energy, deduced from the MC simulation. The sharp correlation between the low values of δ and hard photon radiation is entirely due to ISR where the photon disappears down the beampipe. The hard photon radiation that does not alter the value of δ appreciably, originates from final state radiation. The selection requirement $\delta > 35$ GeV is equivalent to a cut on the unobserved photon of $E^{ISR}_\gamma \sim 9$ GeV. We have explained the importance of this cut in section 1.6.

Secondly, a cut on δ is utilized to reduce the photoproduction background. For these events the value of δ is lower, because the scattered electron remains undetected in the rear beampipe. In figure 4.4 the δ distribution for photoproduction MC events is compared with DIS MC. Most of the photoproduction is removed by the requirement that an electron be found (see the next section). The majority of remaining events are rejected with $\delta > 35$ GeV.

The third type of unwanted events that are rejected with the cut on δ are the proton beam-gas events that have the vertex inside the main detector. They deposit energy in the forward direction only, and are also characterized by low values of δ.

In figure 4.5 we show the δ-distribution for the data. Besides the full DST and the DST after applying timing cuts, the δ distributions for two electron finding algorithms, discussed in the next section, are shown. The peak at $2A$, containing DIS, is clearly visible and is more pronounced after a final state electron is required. Note that we used a logarithmic scale here. Beam-gas events populate the histogram at very low values of δ.

4.4.5 Electron finding

A neutral current DIS event is primarily characterized by a scattered electron in the detector. For high Q^2-events the high energetic electron-signal in both the tracking devices and the calorimeter is very clear and separated in space from the hadronic activity. The electron finding algorithms are therefore very efficient for this part of phase space. However, for low-x events (with high values of y) the scattered electron is less energetic and a possible overlap between the generally broad current

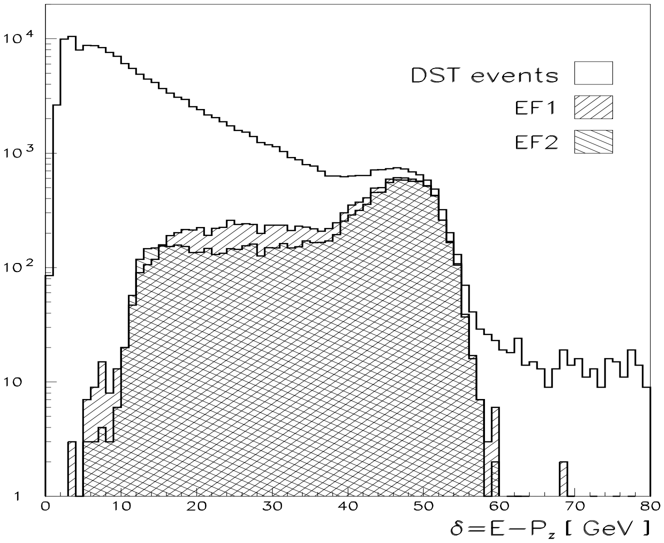


Figure 4.5: *The δ-distribution for data. In the open histogram the distribution on the DST level is drawn. The light hatched histogram is the δ distribution for the events that pass the timing criteria. We also plot the δ distribution for events in which a final state electron is identified. Two electron finding algorithms are used,* EF1 *and* EF2, *that show a different distribution at low values of δ, which is the region that is dominated by photoproduction and beam-gas events. The electron finding algorithms will be discussed in the next section.*

jet and the electron signal can occur, which makes the identification more difficult. Because electron identification is further complicated by photoproduction events, we use in this analysis two algorithms, which we call EF1 and EF2.

The two electron identification algorithms are based on calorimeter information only. The electron produces a shallow and narrow, pure electromagnetic shower in the EMC section, generally distant from other energy deposits. The calorimeter is unable to distinguish between electrons, positrons and photons, but these can be disentangled with information from the tracking devices. Hadrons, such as pions, generally deposit energy in the HAC sections of the CAL and are spread lateraly, whereas (minimum ionizing) muons deposit a small amount of energy and the pulse height is proportional to the traversed CAL depth. Therefore the calorimeter is searched for EMC cells that have $E_{EMC} > 1$ GeV. These cells are defined to be the 'seed' cells for further electron identification. Various cones around the seed cell are constructed and the purpose of the cones is to measure the probability that

Cell type	Cone	EF1 cone [mrad]	EF2 cone [mrad]
EMC	inner	250	100
EMC	outer	400	300
HAC1	inner	300	200
HAC1	outer	500	400
HAC2	inner	300	200
HAC2	outer	500	400

Table 4.2: *Table of the cone sizes used for the* EF1 *and* EF2 *electron finding algorithms. For an explanation, see text.*

the energy deposit originated from the impact of an electron. If so, in addition the cones are used to obtain all calorimeter cells that acquired energy from the electron.

For this purpose, the calorimeter was mapped onto a sphere. Only seed cells with the highest energy are kept when the seed cells are adjacent or less than 12° apart. The energy imbalance of the seed cell is used to reject spark-cells in the event. The unnormalized probability or 'quality-factor', that indicates whether the seed cell is hit by an electron or hadron c.q. muon, is the product of probabilities that are defined by the energy distribution in the cones. The following combinations of cones are used to define these probabilities:

- Energy weighted radius of the EMC(inner) cone,
- Ratio EMC(outer)/EMC(inner) cones,
- Ratio HAC1(inner)/EMC(inner) cones,
- Energy in HAC1(outer) cone,
- Ratio HAC2(inner)/EMC(inner) cones,
- Energy in HAC2(outer) cone,

with the values of the cone-angles given in table 4.2 for the two algorithms. The reason to use two concentric cones with different radii in the determination of the quality-factor is because the electron-signal is narrow compared to the hadron signal. Therefore the difference of energy contained in the inner and outer cones cannot be too large for electrons. In fact this implies that the electron signal is isolated, the degree of isolation being determined by the opening angle of the outer cone.

The value of the quality-factor was used to discriminate between incident electrons and hadrons or muons.

The energy deposit is said to have originated from an electron if it passes a certain cut on the quality-factor[1]. The cut on the quality-factor is determined

---

[1] In fact, the EF1 algorithm in addition required that the number of cells that belong to the electron is not too big and that the ratio of HAC to total energy is small.

from both calorimeter testbeam data, recorded at CERN [41, 42], and Monte Carlo samples.

The cells that belong to the electron are subsequently defined as all cells that lie inside the corresponding cones (EMC(inner)+EMC(outer)+HAC1(inner)+ HAC2 (inner)), and the energy of the electron is defined to be the accumulated energies of these cones. If more electrons were found in a single event, the one with the highest quality-factor for EF1 and the one with highest p_T for EF2 was defined to be the scattered beam electron.

Due to the different sizes of the cones the two algorithms are complementary with respect to the efficiency and purity of the identified electron sample. We define these quantities for MC events as:

$$\begin{aligned} \text{Efficiency} &= \frac{dN(\text{correctly identified})}{dN(\text{generated, with FLT trigger})} \\ \text{Purity} &= \frac{dN(\text{correctly identified})}{dN(\text{identified})} \end{aligned} \tag{4.9}$$

and the argument 'with FLT trigger' indicates that the event is triggered by the First Level Trigger simulation. Figure 4.6 shows the efficiencies for the two electron finding algorithms for DIS events, generated with HERACLES and ARIADNE as functions of the generated electron energy and angle, as well as Q_h^2 and x_h. To remove the hard initial state radiation, we required that $\delta > 35$ GeV. For both finders, the low energy electrons are problematic to select, as is seen in figure 4.6a. The EF2-finder is more efficient in this low E_e region. This is at the same time the region of high y values, where the hadronic activity is close to the scattered beam electron. The cone sizes of the EF2-finder are smaller and hence the isolation criterion, that is mentioned before, is easier to fulfill in this region. Both finders show the same behaviour as a function of polar angle θ (figure 4.6b). The efficiencies as a function of the kinematic variables x_h and Q_h^2 are shown in figure 4.6c and 4.6d. At very low values of x_h the electron finding algorithms efficiency decreases, especially for EF1, whereas the efficiency is flat as function of Q_h^2.

The purities for both electron finding algorithms are comparable and close to one when applied to DIS Monte Carlo events. To study purity it is therefore more meaningful to investigate the number of (wrongly) recognized electrons in the hadronic activity of photoproduction Monte Carlo events. As it turns out, the EF1 algorithm finds fewer electrons in this sample. This is explained by the fact that the cones of EF1 are bigger than those of EF2 and the electromagnetic cluster, if existent at all in the hadronic activity, has to be more isolated to be defined as originating from an electron. Therefore EF1 has a somewhat higher purity. This can also be observed in the δ distribution of the data, figure 4.5 of the previous section, where EF1 finds less electrons than EF2 in the photoproduction region $\delta < 35$ GeV. We postpone a detailed discussion of the remaining photoproduction in the sample after all selection cuts are applied to the end of this chapter.

As mentioned above, the efficiency of the electron identification decreases for

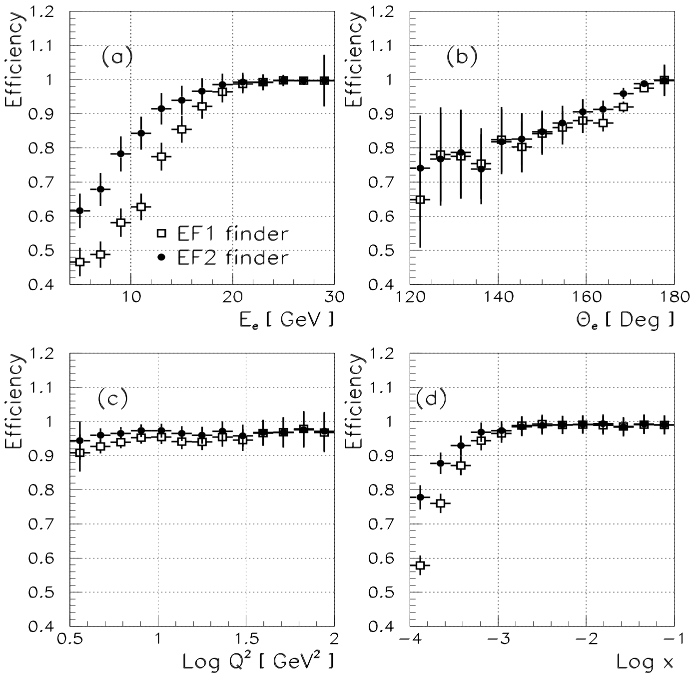


Figure 4.6: *Efficiencies for the two electron finding algorithms* EF1 *(open boxes) and* EF2 *(closed circles). In these plots we imposed a the requirement* $\delta > 35$ *GeV. In* (a) *the efficiency is shown as a function of the generated electron energy* E_e*, in* (b) *as a function of the generated* θ_e*. The efficiency as function of* $\log(Q^2_h)$ *and* $\log(x_h)$ *is plotted in* (c) *and* (d) *respectively.*

low electron energy. We therefore require that

$$E_e > 5 \text{ GeV}. \tag{4.10}$$

for both algorithms. This cut on E_e also increases the purity of the electron sample.

4.5 Reconstruction accuracy improvements

In the main part of the analysis, we reconstruct (x, Q^2) with the double angle method. This brought about a large effort to reconstruct the vertex of the event and impact point of the scattered electron as accurately as possible.

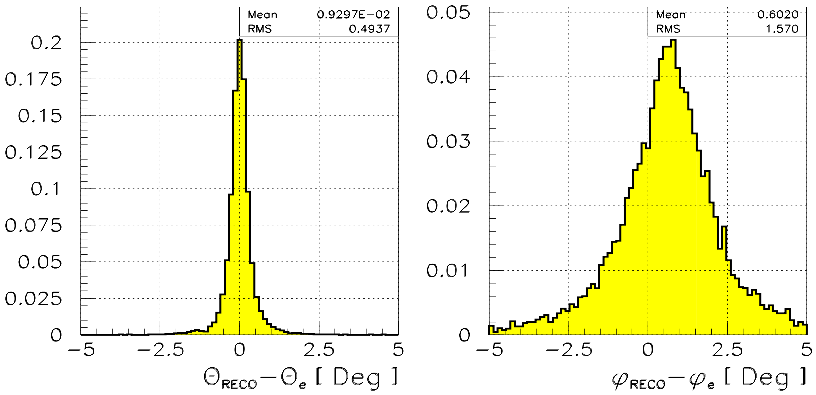


Figure 4.7: *Resolution of the scattered electron in θ and ϕ. The vertex reconstruction is assumed to be perfect.*

4.5.1 Electron position reconstruction and box cut

Only information from the calorimeter is used to reconstruct the impact position of the scattered electron. Of all the electrons that enter the calorimeter, approximately 1% impact either the BCAL or FCAL, the remaining 99% the RCAL. Impact position reconstruction in the RCAL is therefore of special importance, where the cell-sizes are 20×10 cm^2 in the $\mathbf{x}$ and $\mathbf{y}$ direction respectively. In the determination of $\mathbf{x}$ we use the energy deposit in each of the two PMT's, E_{left} and E_{right}. Cell energy-sharing of the cell with highest deposit energy $E_{highest}$ and neighbouring cell $E_{neighbour}$ is used for the position reconstruction in $\mathbf{y}$. The functions

$$\begin{aligned} \mathbf{x}_{impact} &= \mathbf{X}\left(\frac{E_{left} - E_{right}}{E_{left} + E_{right}}\right) \\ \mathbf{y}_{impact} &= \mathbf{Y}\left(\frac{E_{neighbour}}{E_{highest} + E_{neighbour}}\right) \end{aligned} \tag{4.11}$$

depend on the transverse and longitudinal electromagnetic shower shapes. The functions $\mathbf{x}_{impact}$ and $\mathbf{y}_{impact}$ are tuned with the CERN test beam data. Since the configuration of dead material in front of the calorimeter is not precisely known, determination of these functions in the ZEUS environment from Monte Carlo data is difficult. The functions for the impact position were therefore tuned with the HES, 3×3 cm^2 silicon diodes at a depth of 3.3 X_0, that is installed in part of the RCAL. In 1992 the HES was installed in 2 RCAL modules only. The comparison of the HES position and the reconstructed CAL position gives a resolution of about 1.2 *cm* in $\mathbf{x}$ and 0.4 to 1.2 *cm* in $\mathbf{y}$.

We show the resolution of the generated and reconstructed polar and azimuthal angles for DIS Monte Carlo events in figure 4.7. We included no vertex position

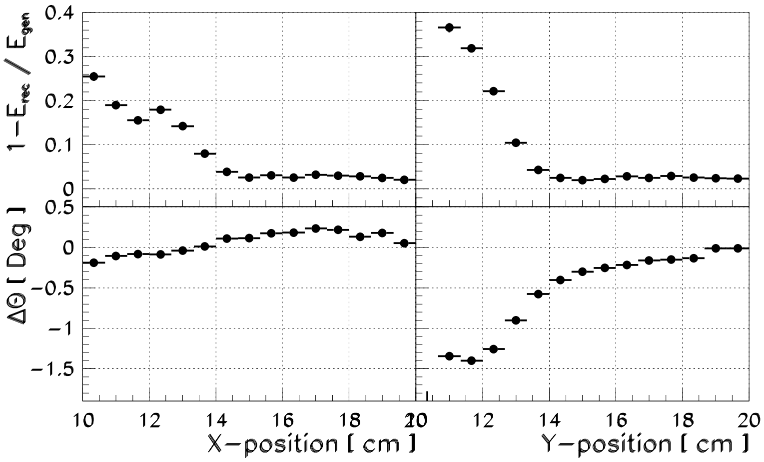


Figure 4.8: *Motivation for the box-cut. The top two figures show the fractional energy leakage* $1 - E_{rec}/E_{gen}$ *as function of the generated impact positions* **x** *and* **y** *respectively. Below the same for the difference between the reconstructed and generated azimuthal* $\Delta\theta = \theta_{rec} - \theta_{gen}$.

uncertainty in this plot. The resolution in θ is $\sigma(\theta) \sim 0.5^o$ without bias whereas $\sigma(\phi) \sim 1.5^o$ with a bias of 1.5^o. The larger resolution and bias in ϕ is consistent with the magnetic field of 1.5 Tesla.

If the impact point is close to the beampipe, a fraction of the electron energy is lost through leakage. The shower is not fully contained in the uranium calorimeter. In the top half of figure 4.8 the leakage factor $1 - E_{rec}/E_{gen}$ is drawn as a function of **x** and **y** respectively for the cells immediately surrounding the beampipe. Note that in **x** the scan from 10 to 20 *cm* covers only a half EMC cell, whereas the scan in **y** covers the complete EMC cell. Below 15 *cm* the energy leakage increases rapidly for both **x** and **y**.

The reconstruction of the polar angle θ in these regions is plotted in the same figure, lower half. The reconstruction of **x**, as given in equation 4.11, is not very sensitive to the energy leakage because it does not use information from cells other than the one with highest energy. The resolution of θ is fairly constant as function of **x**. As stated in formula 4.11, neighbouring cells supply the information to reconstruct **y**. Since the neighbouring cells are lacking for the cells adjacent to the beampipe, the reconstruction in **y** becomes worse in this region. In fact, for electrons that enter the cells adjacent to the beampipe in the **y** direction, the impact **y** position is put at the center of the cell for $\mathbf{y}_{gen} < 15$ *cm*. Hence the angle resolution decreases linearly in **y** from the cell center to the edge of the beampipe for these cells.

To circumvent the large energy leakages and poor determination of the angle of the scattered electron, we made a symmetric fiducial cut around the RCAL

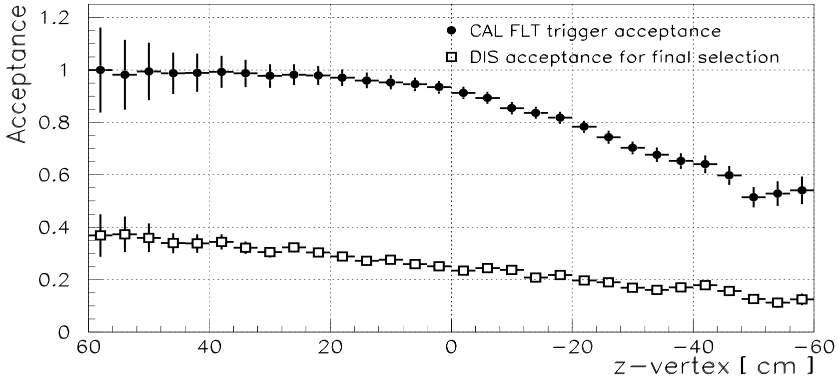


Figure 4.9: *The acceptance as function of z-vertex. In the figure both the acceptance for the calorimeter FLT and the acceptance for the final selection is drawn. The DIS Monte Carlo is generated with $Q_l^2 > 4$ GeV^2. The FCAL is towards the left side, the RCAL towards the right side.*

beampipe. The motivation for this fiducial beampipe cut is further justified by the fact that the amount of dead material in front of the RCAL is very high in this region (see also figure 2.13), and a slight mismatch between the amount of dead material simulated in the Monte Carlo and in the ZEUS detector causes large differences in energy and position determination for data and Monte Carlo. We therefore chose the fiducial cut:

$$|\mathbf{x}| > 16\ cm \quad \vee \quad |\mathbf{y}| > 16\ cm \tag{4.12}$$

4.5.2 Vertex determination

It is important to measure the position of the event vertex, not only because it influences the angles of the emanating particles and hence the reconstruction of (x, Q^2), but also because it influences the acceptance of DIS events. The acceptance dependence of Monte Carlo DIS events as function of the generated vertex is shown in figure 4.9. Both the acceptance of the calorimeter first Level Trigger (CAL FLT) and the acceptance of the complete set of selection filters are shown. Due to the cross-section propagator term, the events with low values for Q^2 or small electron scattering angles occur most frequently. This implies that the CAL FLT acceptance for DIS events (generated with $Q_l^2 > 4$ GeV) drops for vertices towards the RCAL, because there the electrons with small scattering angles are more likely to disappear in the beampipe. After all selections are imposed, the acceptance for DIS events changes approximately with 1% for a change in the z-vertex position of 4 *cm*. Therefore the Monte Carlo z-vertex distribution, with which corrections for acceptance are made, should correspond closely to the z-vertex distribution of the data.

For events that leave tracks in the tracking devices, the event vertex position can be reconstructed using this information. During the 1992 running period, the central tracking device (CTD) was instrumented only partially; only superlayers one, three and five were instrumented with z by timing readout electronics (see section 2.3). The vertex detector was fully equipped. The individual hits from both the CTD and vertex detector are fitted to obtain tracks. The well reconstructed tracks, obtained with a χ^2-cut on the tracks, were then used to fit the z-position of the vertex. Also on this vertex fit a requirement on the quality is made. The χ^2_{vertex} is required to be $\chi^2_{vertex} < 10$. In addition the number of tracks that participate in the vertex-fit should be two or more. For events that have a reconstructed vertex, the resulting resolution of the vertex z-position is approximately 4.5 *cm*. The efficiency of finding a vertex from tacking information is given in the first row of table 4.3 for the data and the DIS Monte Carlo.

	Data ($Q^2_{DA} > 10$ GeV)	DIS MC ($Q^2_{DA} > 10$ GeV)
Tracking vertex	87.4% (88.3 %)	96.3% (96.2 %)
CAL time vertex	9.1% (9.3 %)	3.6% (3.7 %)
No vertex	3.5% (2.4 %)	0.1% (0.1 %)

Table 4.3: *The efficiencies for finding an event vertex from tracking information and timing. In the left column the data and in the right column the DIS Monte Carlo efficiencies are listed, both after having imposed the final selection cuts. In parentheses the efficiencies with the additional restriction that* $Q^2_{DA} > 10$ *GeV.*

The z-position of the vertex can not be reconstructed with the tracking devices for all accepted DIS events. The probability that charged particles cross the active volume of the tracking device changes with the kinematics of the event. For example, the final state electron hits the first superlayer of the CTD only when $\theta < 168^o$, which corresponds to $Q^2 \sim 20$ GeV^2 (at low values of y, see the 'isolines'–plot 3.5b). The particles of the hadron flow have an decreased probability of passing the CTD at low values of the hadron angle γ, as the jet points in the direction close to the FCAL beampipe. Hence especially at low values of y the efficiency of finding a tracking vertex is low.

For the events that do not have a tracking vertex, the calorimeter timing information is used to obtain the event vertex. The time information of the calorimeter PMT's is already used to reject proton beam-gas events but can be utilized even further to determine the vertex position. As mentioned, the electron interacts with a proton at the center of the proton bunch at $z = 0$ in the nominal case. In this case the emanating particles from the collision arrive at the calorimeter at time $t = 0$

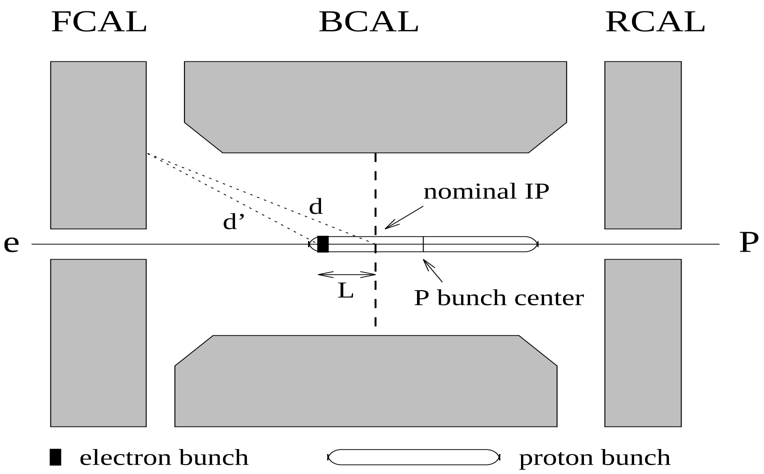


Figure 4.10: *Schematic view of the detector with the electron and proton bunches. The ep interaction can take place in different depths in the proton bunch. In this case the interaction occurs at* $t = -L/c < 0$*, before both bunch–centers arrive at the nominal IP. For particles that hit the FCAL, the shift* L *of the vertex results in addition to a difference of travel distance. The nominal travel distance* d *(for which the PMT times are adjusted to* $t = 0$*) becomes* d'*. The total time difference between the average interaction and the interaction in this case is then* $t = -L/c - (d - d')/c$*. This time difference is used to extract the shift* L *of the nominal vertex.*

for all cells.

Deviations from the nominal vertex are reflected in the FCAL PMT time[2] measurement. This is explained in figure 4.10 for a shift L of the vertex in the direction of the FCAL. In this case the arrival time of the particles in the calorimeter cells is early because of the shorter path length. Moreover the time of the cells close to the FCAL beampipe is earlier compared to the time of the calorimeter cells further away from the beampipe because the difference in the distance traveled $(d - d')$ is smaller in the latter case. Taking this effect into account, each cell with an energy deposit estimates the vertex shift L from the nominal interaction point. The event vertex was then calculated by taking the energy weighted average of the vertices calculated from each cell.

Only cells with more than 400 MeV energy deposit were included in the calculation of the average event vertex. The resolution of the vertex from the calorimeter timing depends on the total energy in the FCAL. For $E_{FCAL} > 5$ GeV, the resolution of the difference of the tracking–vertices and timing–vertex equals $\sigma(\mathrm{z}_{track} -$

---

[2] The electron bunch length is assumed to be pointlike in contrast with the proton bunch length. As explained in section 4.4.2, only the FCAL time measurement is affected by the proton bunch length.

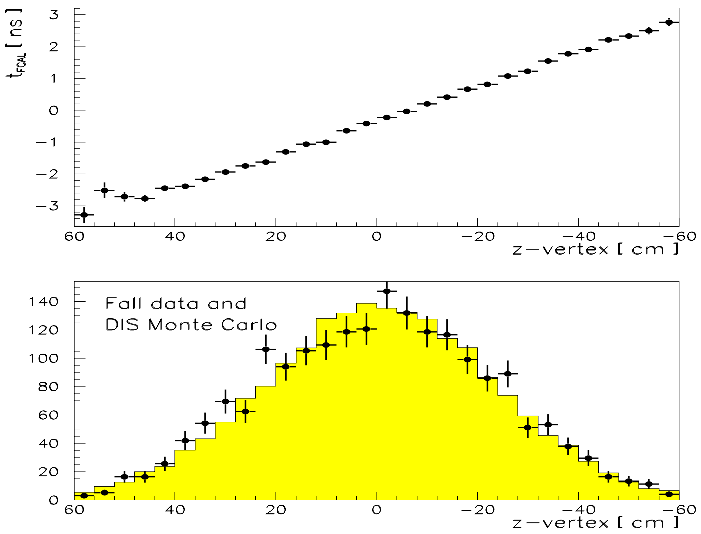


Figure 4.11: *In the figure above the z-vertex is compared with the FCAL timing t_{FCAL}. Below for the full sample the z-vertex is compared for the data with the DIS Monte Carlo sample.*

$z_{time}) \sim 12$ *cm*. The average FCAL time correlation with the z-position of the tracking vertices is shown in the top plot of figure 4.11. For those events that do not have a tracking vertex but have $E_{FCAL} > 5$ GeV, the calorimeter timing is used to calculate the event vertex position. The fraction of events that have the vertex from CAL timing information is listed in table 4.3.

In order to obtain realistic acceptance corrections, the input (generated) vertex distribution of the Monte Carlo events should correspond with the true vertex distribution of the data. However, as stated above, the acceptance for DIS events depends on the vertex position and hence the vertex distribution as obtained from the data DIS events will be biased towards large z and so cannot be used as input for the Monte Carlo simulation. But the data sample that contains the photo-production events is less dependent on the vertex position. For these events the electron disappears in the beampipe by definition and therefore the acceptance does not decrease for events with their vertex close to the RCAL. The less-biased vertex distribution of the data photo-production sample is therefore used to supply the input vertex distribution of the DIS Monte Carlo. The mean z position of this input distribution is at -6 *cm*. After the final selections are imposed, the reconstructed vertex for the data and the Monte Carlo are compared in the lower half of figure 4.11. The input mean z position of -6 *cm*, that correspond to the mean vertex position of the

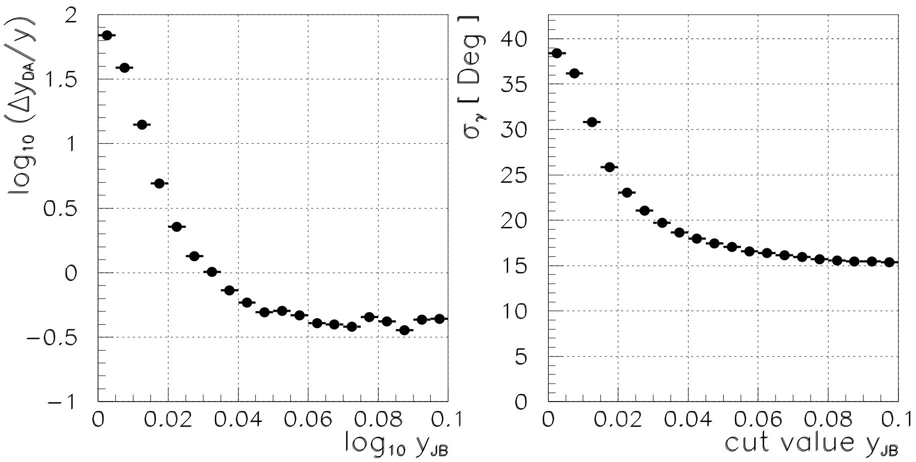


Figure 4.12: *Motivation for the cut on y_{JB}. Left the mean value of $(y_{DA} - y_{gen})/y_{gen}$ on a log-scale, as function of y_{JB}. At the right side the resulting RMS of $(\gamma_{rec} - \gamma_h)$ as a function of the value of the cut on y_{JB}.*

photo-production data sample, is shifted for the selected DIS MC events to 0 *cm*. The vertex distribution of the selected DIS MC events agrees with the observed vertex distribution of the DIS data.

Of the events that pass all filter selections, 3.5% do not have a tracking vertex or timing vertex, as seen in table 4.3, whereas this fraction is only 0.1% for the Monte Carlo. This difference is due to the 'diffractive'-type events that are observed in the data and are not anticipated in the Monte Carlo. They will be discussed in section 5.6. No events were rejected based on the lack of an event vertex, since for those events the vertex is set to zero, the nominal interaction point.

4.5.3 Hadronic activity

At very low values of y, where the current jet is close to the FCAL beampipe or even disappears in it completely, the remaining uranium noise in the calorimeter significantly contributes to the measurement of y_{JB} (or $\delta_{hadron} \equiv 2Ay_{JB}$). Especially noise energy in the rear direction contributes to y_{JB} due to the large corresponding angles. Besides the calorimeter cell noise, energy of the scattered electron that was deposited outside the cones which are used in the electron finding algorithms, contribute to y_{JB}. Therefore the measurement of y_{JB} is different for the two electron finding algorithm. Also the hadron angle γ, defined as

$$\cos\gamma = \frac{P_{t\,had}^2 - (2Ay_{JB})^2}{P_{t\,had}^2 + (2Ay_{JB})^2} \tag{4.13}$$

is affected by noise energy. At low values of y_{JB} the second term of the nominator and denominator is small. Hence, as $P_{t\,had}$ (the hadron transverse momentum) is also small, γ is determined inaccurately from a division of two small numbers with poor resolution. Since the value of y_{JB} increases due to uranium noise, the hadron angle shifts towards the rear direction.

From a data sample of random triggers, resulting in 'empty' events without impacting particles, the visible calorimeter cell noise was parametrized, see section 2.4.3. This cell noise translates to $y_{JB} \sim 0.007$. The cell noise was then simulated accordingly in the ZEUS Monte Carlo to study the effect of noise on the reconstruction of kinematic variables. Figure 4.12a shows the resolution $\log(y_{DA} - y_h)/y_h$ as function of y_{JB}. For low values of y_{JB} the resolution of y_{DA} is very poor. This is also seen in figure 4.12b, where we plot the RMS of the $(\gamma_{reco} - \gamma_{gen})$-distribution in the complete phase space as function of the value of the cut on y_{JB}. If only the events with

$$y_{JB} > 0.04 \tag{4.14}$$

are kept in the sample, the mean error on reconstructed γ is less than 20°. We therefore imposed this requirement on the events.

4.5.4 Fake FCAL electrons

A cut on the variable y_{elec}, the value of y that is reconstructed with information of the scattered electron, is added to remove electrons in the FCAL. Occasionally the electron finding algorithms identify electrons in the debris near the FCAL beampipe, at $\theta_e \sim 0^\circ$. The phase space for DIS events with the electron in the forward direction is very small and therefore the identified electrons are most probably fake. Primarily photoproduction events generate those fake events. To remove these events we required

$$y_{elec} < 0.95. \tag{4.15}$$

From the electron energy isoline plot in the (x, Q^2) phase space (figure 3.5) one observes that this cut does not affect electrons in the RCAL. Indeed for those electrons the requirement that $E_e > 5$ GeV already implies that $y_{elec} < 0.84$.

4.5.5 Selections with the eye

To judge events, one very important tool is the human eye. With help of visualization of events it is possible to detect background that is not recognized as such by computer programs. With the display programs GAZE or LAZE the events can be scanned and examined. All events that passed the previously defined selection cuts and have $Q^2_{DA} > 50$ GeV^2 were scanned. A few cosmic muon events were found that were not recognized by the programs. A beautiful example is shown in figure 4.13, where four cosmic muons penetrate the detector. For this figure the LAZE event display was used. Another type of unwanted events are the penetrating muons of the beam halo. They are characterized by the fact that the energy deposits in the

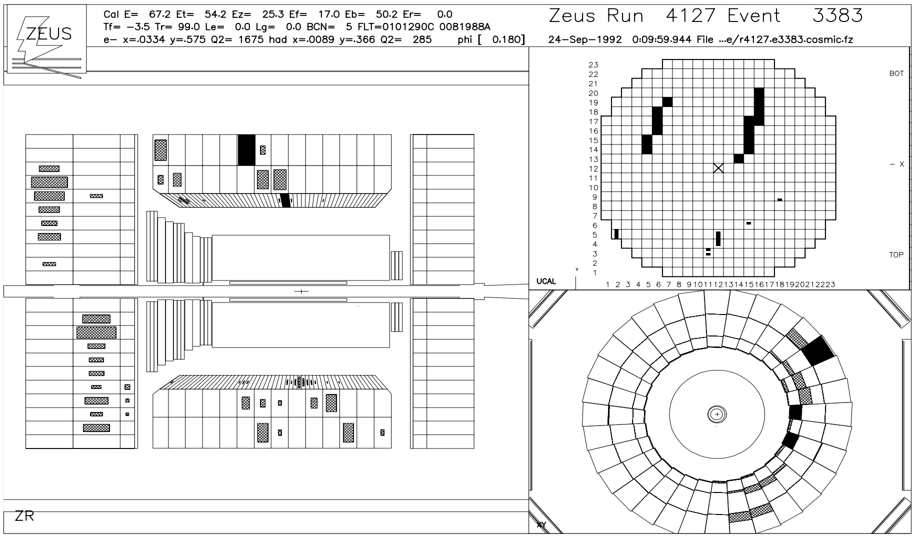


Figure 4.13: *A cosmic muon triggered event. In this case four muons penetrate the detector. On the left side the RZ view of the ZEUS detector is shown. On the lower plot of the other side the transverse view is drawn. The FCAL energy depositions are shown in the top right figure.*

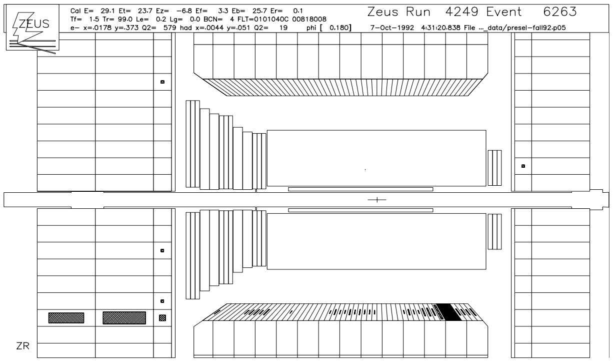


Figure 4.14: *The ZR view of a beam halo event.*

calorimeter are lined, parallel to the beam. In figure 4.14 we show an example of a beam halo muon.

A third type of unwanted events are the 'elastic Compton' events. They are not simulated in the Monte Carlo and therefore have to be removed from the sample. These 'elastic Compton' events are recognized via scanning of events that have $E_{FCAL} < 1$ GeV. In total, 31 such events were found in the preselected data files. They are characterized by only two electromagnetic clusters in the calorimeter.

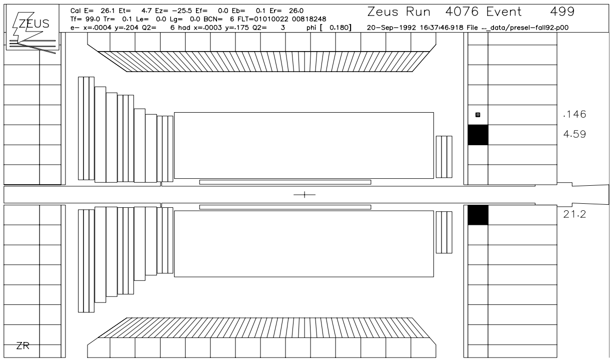


Figure 4.15: *The ZR view of an elastic Compton event with two electromagnetic deposits only. The energy of the two hit cells are 21.2 and 4.6 GeV respectively.*

The energy of the two clusters is approximately the electron beam energy. They are explained as radiative events that are described by the Compton-peak. The momentum of the exchanged virtual photon is too low to break the proton which escapes undetected in the forward beampipe. Figure 4.15 is an example of such an event.

4.6 Reduction factors

The filter selections discussed above, were applied to the 228491 DST events. The filters are not implemented at once on this large number of triggers, but, in order make the sample more manageable, were performed in two stages. The first filtering stage is the so called 'preselection', and its effect is listed in table 4.4.

In the preselection a loose cut of $\delta > 25$ GeV was applied. The value of this cut on δ was chosen such that it reduced the sample considerably but still allow for photo-production background studies after the preselection.

On the sample that passed the preselection (18913 events), we describe the effect of the 'final selection' in figure 4.16. Only in this final selection was it required that an electron be found. Hence the number of events that pass this final selection depend on the electron finding algorithm. The sample of events that pass the preselection and are subjected to the final selection is small enough to allow flexibility. The final selection represent the reconstruction quality improvement cuts. From figure 4.16 one reads that the number of events for which (x, Q^2) can be reconstructed reliably equals 2025 for the EF1 and 2442 for the EF2 electron finder.

4.7 Acceptances in (x, Q^2)

	Number of events	Rejected	Filter stage
Recorded	$4.2 \cdot 10^6$		Trigger
DST	228491	94.5 %	DST selection
Apparatus failure	218997	4.1 %	Preselection
Spark cut	116043	47.0 %	
Final timing	105623	9.0 %	
Vetowall	102649	2.8 %	
$\delta > 25$ GeV	19422	81.1 %	
Remaining muons	18913	2.6 %	
$35 < \delta < 60$ GeV	9328	50.7 %	Final selection

Table 4.4: *This table shows the background rejection cuts. The first column states the name of the filter, the second the number of events that survive this and all preceding filters. The third column gives the rejection percentage compared to the previous filter and the filter stage is given in the last column. For the final selection only the (electron identification independent) cut on δ is listed.*

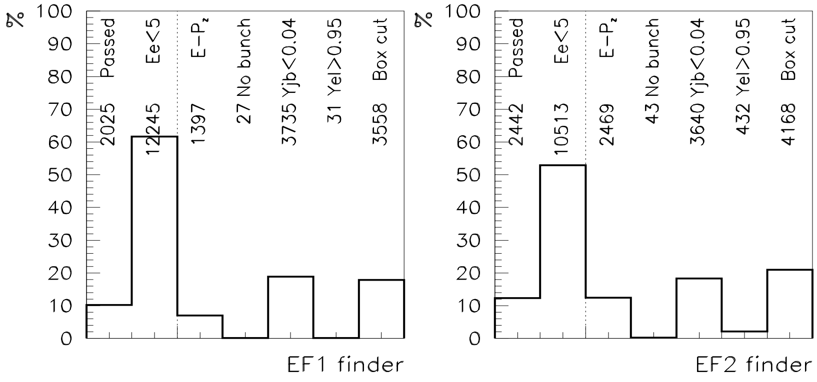


Figure 4.16: *Table for the final selection. The left histogram for electron finding algorithm* EF1, *the right one for* EF2. *The first entry shows the number of events that passed all cuts. The second entry gives the number events that are rejected by requiring an identified electron. Thereafter each entry represents a cut for which the number of rejected events is plotted, with the requirement that an electron is identified.*

The events that are removed from the data sample after the preselection level are plotted in figure 4.17. The effect of the main selection cuts on the preselected data

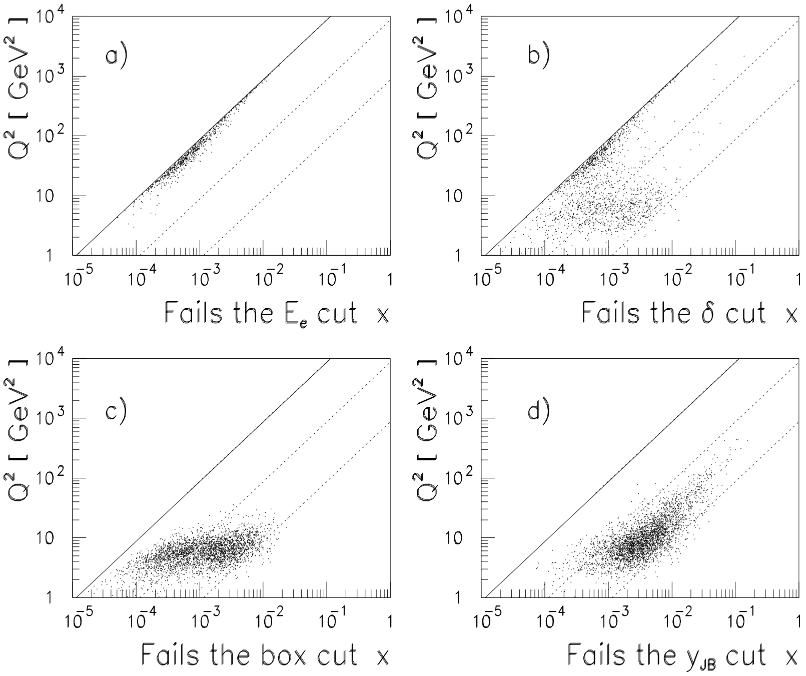


Figure 4.17: ***The effect of the important selection cuts in $(x, Q^2)_{DA}$ phase space, for data after the preselection. In (a) the events that fail the electron energy cut $E_e > 5$ GeV, in (b) the events with $\delta < 35$ GeV. The events that do not pass the box cut are plotted in (c) and the ones that have $y_{JB} < 0.04$ in (d).***

is displayed in the (x, Q^2) phase space. We used the double angle reconstruction method. In the figure, the events that are rejected by a particular cut, described below the figure, are plotted. The $E_e > 5$ GeV requirement removes events at high values of y, the box cut events at low values of Q^2_{DA} and the cut on δ in both regions. The cut on y_{JB} rejects events with low value of y_{DA}, as well as medium or even high values of y_{DA} with very little δ_{had}. For those events the direction of the angle γ is shifted towards the RCAL due to the calorimeter noise.

From the DIS Monte Carlo the acceptances of the selection filters can be estimated. In figure 4.18 the acceptance is plotted as a function of Q^2_h and x_h respectively, for both the CAL FLT trigger and the complete, final selection for Monte Carlo. Note that these plots depend strongly on the minimum Q^2 and x with which the events are generated. As a function of Q^2, the acceptance is integrated over

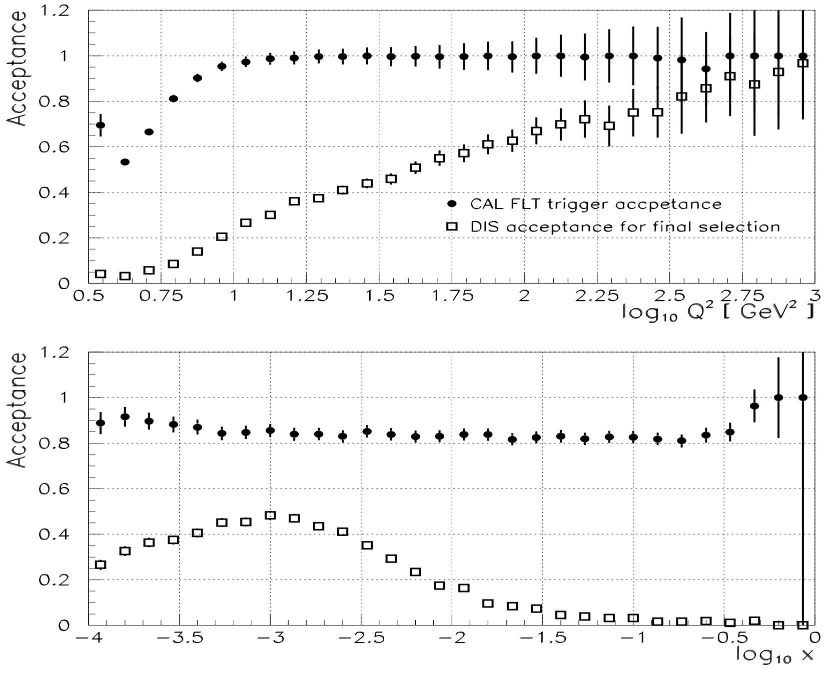


Figure 4.18: *In the figure above the acceptance as a function of Q_h^2 is given for both the calorimeter FLT and after the final selection. Below the same for the x_h-distribution.*

x and *vice versa* as function of x. As is seen from the figure, the FLT acceptance drops drastically for low Q_h^2 ($Q_h^2 < 10$ GeV^2). Therefore this region is potentially dangerous for the determination of the structure function. It is in this region where the experimental uncertainties on the values of the trigger–thresholds become non-negligible.

The acceptance for the final selection shows a strong correlation with Q_h^2, mainly due to the cut on $y_{_{JB}} > 0.04$. This acceptance curve is therefore well understood and in the determination of F_2 corrections will be made accordingly. The stability of this acceptance curve is studied extensively. The low values of the acceptances in x_h are caused by the integration of Q_h^2. As a function of x_h the FLT acceptance shows a flat behaviour over almost the complete range in x_h. However, after the final selections are applied, the acceptance in x is small for high values of x. The structure function can therefore not be measured at these high values of x. Also for very low values of x the acceptance curve drops. This is due to the requirement that the electron possesses a certain minimum amount of energy.

4.8 Remaining background in the sample

After applying all selection cuts, an estimate of the remaining background events in the sample is needed. Almost by definition this remaining background cannot be recognized on an event by event basis for then these events would have been rejected. In each bin where the structure function will be determined, the background will be subtracted statistically.

The non-colliding beam background is estimated using the information of the bunch crossing number. There were no events originating from the proton pilot bunch after the selection is applied, but four stemming from the electron pilot bunch. These remaining electron gas event are weighted by 9.73 when doing the background subtraction. This factor was obtained from the ratio of the electron current in the *ep* bunches to the current in the electron pilot bunch. Only one of them has $Q^2_{DA} > 10$ GeV^2 and enters in the (x, Q^2) structure function bins.

The major source of background one has to take into consideration is photoproduction. The amount of photoproduction background can be estimated with Monte Carlo or extracted from the data itself, using the δ distribution in the (x, Q^2) bins. We will use this last method for the photoproduction subtraction in the bins, but we postpone this discussion to section 6.3. For a Monte Carlo estimate of the photoproduction the PYTHIA event generator is used.

We used a sample of minimum bias PYTHIA resolved photoproduction events, with generated Q^2 up to 2 GeV^2. We determined the visible photoproduction background cross section, σ_{vis}, for the EF1 and EF2 electron finding algorithms. In PYTHIA, the Weizsäcker–Williams approximation, with p_t of the electron zero, is extended for resolved processes to generate non-zero electron p_t. A parametrization of the total $\gamma^* p$ cross section down to $Q^2 = 0$ GeV^2 was used to generate the scattered electron. Hadrons were generated using the minimum bias γp interaction scheme of PYTHIA. We reweighted the events to the fall 1992 luminosity and subjected them to the nominal selection cuts. We obtained for the visible photoproduction background cross section, integrated over the bins where F_2 will be determined:

$$\sigma_{vis}(\mathsf{EF1}) = 1.44 \pm 0.21 \text{ nb} \quad , \quad \sigma_{vis}(\mathsf{EF2}) = 4.00 \pm 0.36 \text{ nb}. \tag{4.16}$$

The contribution due to direct photon processes was estimated with the HERWIG Monte Carlo and was about 0.25 nb. for both finders.

As mentioned before, the photoproduction background depends on the electron finder, and is somewhat higher for EF2. This is the reason that we choose the EF1 finder as the nominal one in the remaining analysis.

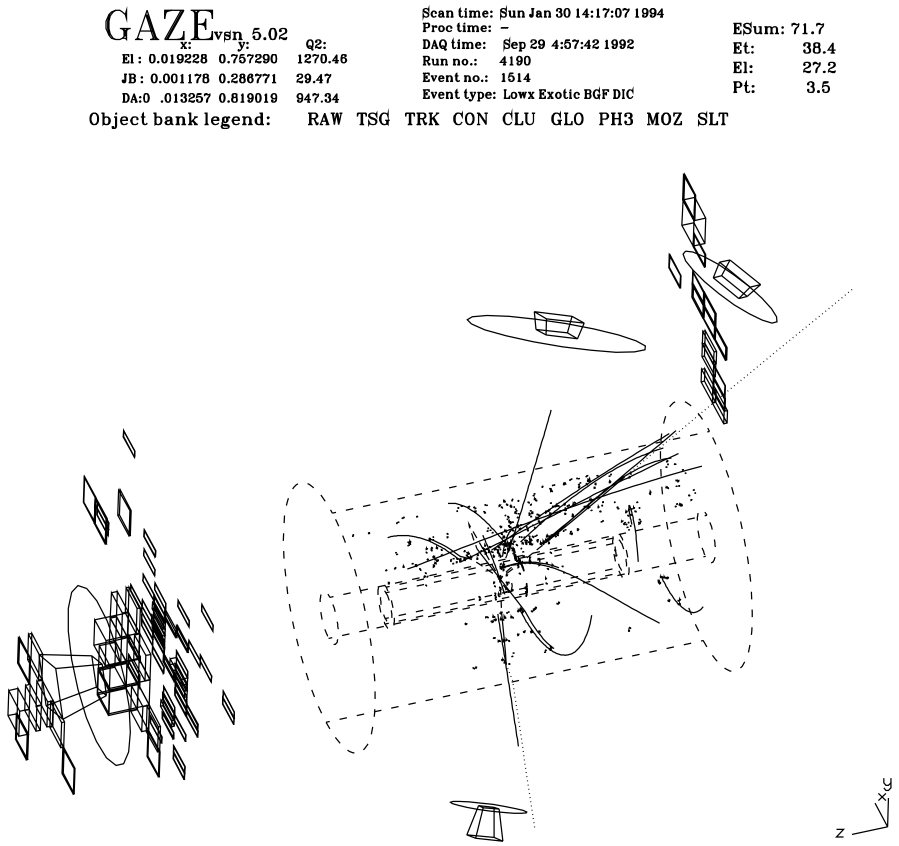
4.9 Summary

The main challenge at the outset of the measurement of the proton structure function at the HERA *ep* collider is to select the tiny fraction of well measurable DIS events from the huge number of beam crossings. At ZEUS, a sophisticated, three level layered trigger system is used for a first step on-line reduction. The accepted triggers are passed to the off-line analysis, where the DST sample is created.

We analyzed carefully the 228491 events that belong to the DST sample. This resulted in the invention of a number of selection criteria. The nominal selection cuts are summarized as:

- **Apparatus failure**: All events are rejected for which parts of the subdetectors were malfunctioning. Also the 'spark' events are rejected.
- **Timing**: Events are rejected if they possess a calorimeter time measurement with $\mid t_{FCAL} \mid \; > 6 \;\; ; \;\; \mid t_{RCAL} \mid \; > 6 \;\; ; \;\; \mid t_{FCAL} - t_{RCAL} \mid \; > 6 \, ns$.
- **Vetowall**: Events with a vetowall hit are rejected.
- **Momentum conservation**: Only events in the range $35 < \delta < 60$ GeV are accepted.
- **Identified electron**: Only events are accepted in which the EF1 finder identifies an electron with $E_e > 5$ GeV.
- **Hadronic activity**: The hadronic system should have $y_{_{JB}} > 0.04$.
- **Box cut**: Events are only accepted if the impact position of the electron satisfies $|\mathbf{x}| > 16 \;\; \vee \;\; |\mathbf{y}| > 16 \, cm$.
- **Fake FCAL electrons** are removed with $y_{elec} < 0.95$.
- **Scanning**: Identified cosmic muons, beam halo muons and elastic Compton scattering are removed.

When applied to the DST triggers, a very pure sample of DIS events is selected. In the 1992 fall period this DIS sample contained 2025 events. Only four events of this selected sample originate from the electron pilot bunch and none from the proton pilot bunch, which indicate that the beam-gas events are removed effectively. The photoproduction background that remained in the sample is estimated with a photo-production Monte Carlo simulation.



```
GAZE vsn 5.02                            Scan time:  Sun Jan 30 14:17:07 1994
       x:        y:          Q2:         Proc time:  -                          ESum: 71.7
El : 0.019228  0.757290   1270.46        DAQ time:    Sep 29  4:57:42 1992      Et:     38.4
JB : 0.001178  0.286771   29.47          Run no.:    4190                       El:     27.2
DA:0  .013257  0.819019   947.34         Event no.:  1514                       Pt:      3.5
                                         Event type:  Lowx Exotic BGF DIC
Object bank legend:     RAW  TSG  TRK  CON  CLU  GLO  PH3  MOZ  SLT
```

Figure 4.19: *A picture of a DIS event with the 3D display program* GAZE. *The outline of the CTD is drawn, including the hits and track segments. The calorimeter cells with an energy deposit above threshold are drawn. The result of the* ZEPHYR *cluster finding in the calorimeter is draw with circles.*

Chapter 5

Characteristics of the DIS events

In this chapter we examine the DIS candidates. In the first section (5.1) we take two events from the sample and examine their kinematic variables. In the next section we show elementary distributions, that should indicate that the sample consists of DIS events. Using the redundancy of the electron and hadron system, we observe a problem concerning the energy-scale. Since the double angle method is insensitive to the energy-scale, the kinematics of the events can still be reconstructed reliably. This is done in section 5.5. In the last section we describe a new class of DIS events that is found in the sample [66].

In many figures the data is compared with the full Zeus Monte Carlo. Unless stated differently, we use the HERACLES event generator with the ARIADNE fragmentation model (see also 4.3) to generate DIS events. The two extreme input structure function parametrizations MT$B1$ and MT$B2$ are used for comparison with the data. We normalize the cross sections of the Monte Carlo events to the luminosity of the data. The overall normalization uncertainty stemming from the luminosity measurement is not included.

5.1 Two 'typical' events

In this section we offer a closer look at two examples of candidate DIS events, i.e. events that pass all selection cuts. The first one, displayed in figure 5.1, is a typical low Q^2, low x event. The electron hits the calorimeter close to the RCAL beam pipe (**r**=23.4 *cm*) with energy $E_e = 22.0$ GeV. The hadron energy as a current jet is clearly visible we reconstruct the hadron energy $F = 3.5$ GeV at angle $\gamma_h = 78.9^o$. The vertex position is located at $\mathbf{z} = -6.8$ *cm*.

We show the (x, Q^2) reconstruction with the electron, Jacquet–Blondel and double angle method in the left plot of figure 5.3. Besides the (x, Q^2) points, the error-

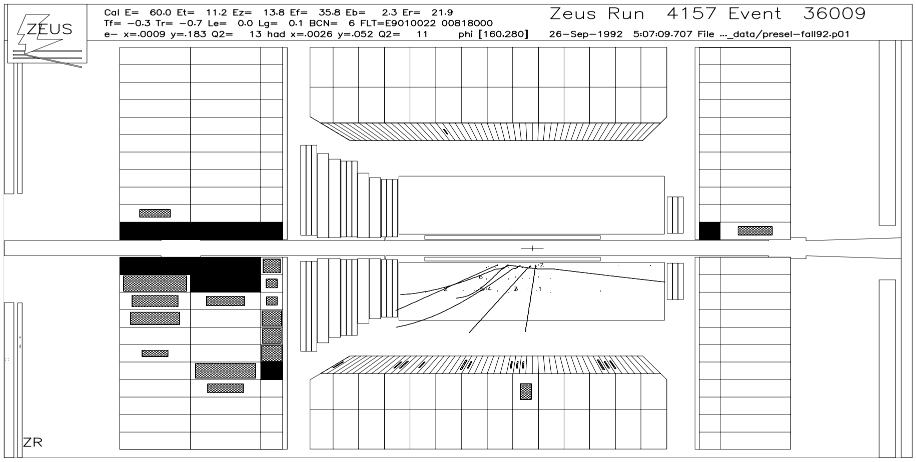


Figure 5.1: *An example of a typical DIS event. The double angle method measures* $Q^2_{DA} = 15.3$ *GeV*^{2}, $x_{DA} = 3.0 \cdot 10^{-3}$, $y_{DA} = 5.9 \cdot 10^{-2}$. *We draw the error-ellipses of this event in the left plot of figure 5.3.*

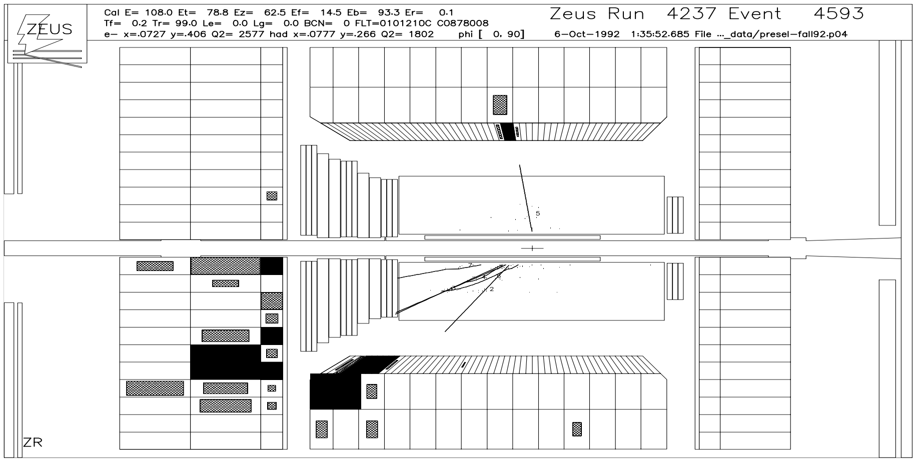


Figure 5.2: *An example of a beautiful high* Q^2 *event. The double angle method measures* $Q^2_{DA} = 2930$ *GeV*^{2}, $x_{DA} = 0.10$, $y_{DA} = 0.33$. *We draw the error-ellipses of this event in the right plot of figure 5.3.*

ellipses of the reconstruction methods are drawn. We use $\sigma_E = 0.2\sqrt{E} + 0.01E$, $\sigma_\theta = 0.03$, $\sigma_F = 0.8\sqrt{F} + 0.05F$ and $\sigma_\gamma = 0.15$ as a crude estimate of the errors[1].

---

[1] In the error estimate of $\{E, \theta, F, \gamma\}$ all detector effects have to be taken into account. The estimate of the energy resolution is especially difficult in the presence of dead material in front of the CAL and we did not correct for energy losses. The errors are therefore estimated conservatively.

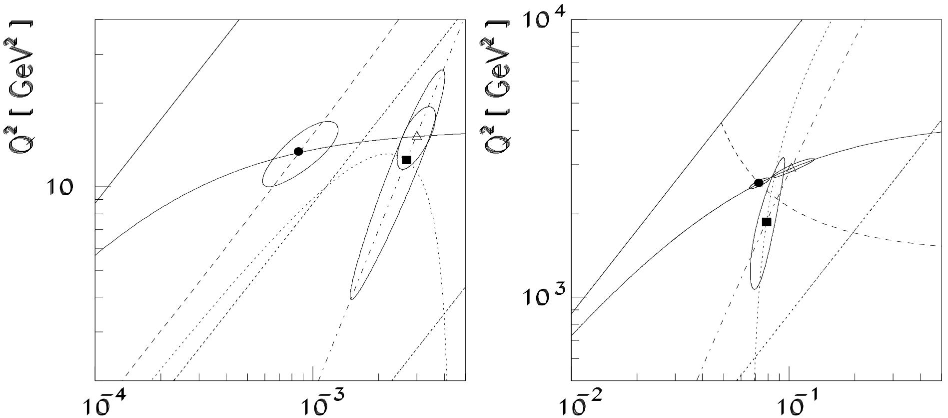


Figure 5.3: *For two events the error-ellipses are drawn in the (x, Q^2) phase space on the same logarithmic scale. The electron reconstruction is denoted with a black circle, Jacquet-Blondel with a black square and the double-angle method with an open triangle. The lines $y = 1$ (straight) and $y = 0.1$, $y = 0.01$ are drawn (dotted). The isolines in E_e (dashed), θ_e (straight), F (dotted) and γ (dashed-dotted) of the events are drawn.*

From this figure one concludes that the double angle method is powerful in this region of phase space since it has the smallest area covered by the error–ellipse. Overlap with the Jacquet–Blondel error-ellipse implies that the hadron energy measurement is consistent with the two angles. But the electron method does not overlap, which means that the electron energy is too low. Only after an addition of 3.4 GeV to the electron energy do the three ellipses overlap. The LUMI photon counter is silent, but as it has limited photon acceptance, this event might have radiated off an initial state photon (using formula 3.26 of chapter 3, one obtains $E_\gamma = 3.4$ GeV). An alternative explanation is an underestimate of the dead material in front of the CAL. We investigate the possibility and consequences of this explanation in the next section.

The second example is one of the highest Q^2 events recoreded in 1992, and is displayed in figure 5.2. This beautiful event has $E_e = 40.0$ GeV, $\theta_e = 77.9^o$, $F = 54.2$ GeV and $\gamma_h = 42.9^o$ with vertex position $\mathbf{z} = -5.8\ cm$. The error ellipses of the three reconstruction methods overlap, as is seen in the right plot of figure 5.3.

5.2 Energy deposits in the calorimeter

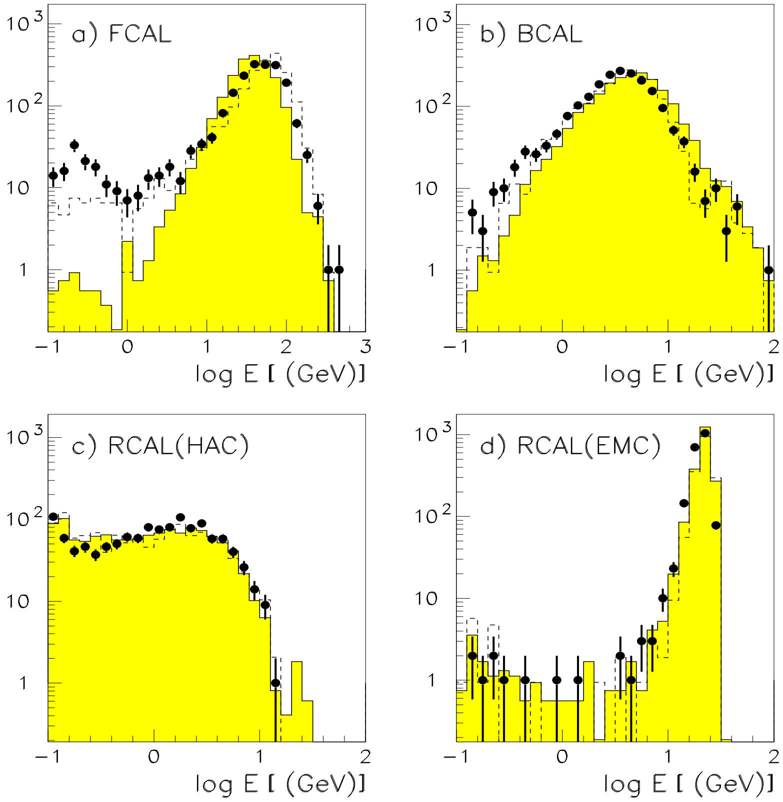


Figure 5.4: *Energy deposit in the calorimeter. The points are the data, the full histogram the* ARIADNE *DIS Monte Carlo and the dashed histogram the* HERWIG-SUE *Monte Carlo. Here we normalized the Monte Carlo to the number of data points. In (a) deposit in FCAL, (b) for BCAL, (c) for RCAL HAC section and (d) RCAL EMC section.*

The calorimeter plays an essential role in the analysis of the 2025 selected DIS events. In figure 5.4 we show the energy distributions in the FCAL, BCAL, RCAL-HAC and RCAL-EMC sections respectively. In all sections the cell–noise cut, 60 MeV for EMC and 110 MeV for HAC cells, is imposed on data as well as on the Monte Carlo.

The FCAL generally has large deposits, up to ∼ 300 GeV, that originate from the proton remnant. The precise energy flow of the proton remnant is described by the fragmentation process. Since in general the remnant carries large fractions of the proton beam energy the FCAL energy is very sensitive to the fragmentation process [47]. In the figure we compare two MC fragmentation models with the data. The ARIADNE Monte Carlo underestimates the FCAL high energy distribution slightly. The HERWIG Monte Carlo, with the soft underlying event (SUE) included, overestimates the high energy behaviour. Clearly the data seem to lie in the range of expectations and further detailed study is needed to tune the MC event generators with the high energy distribution of the FCAL. But, more interestingly, both models completely underestimate the events with very low energy in the FCAL. We will discuss this class of events, that is not described by the Monte Carlo's, in section 5.6.

The BCAL is hit most often by the current jet (besides a few electrons as we showed in figure 5.2), and the bulk of events deposit energies around a few GeV. The cell noise smoothly clouds the signal at very low energies. The difference of the two event generators for the BCAL energy distribution is less pronounced compared to the FCAL. This is explained by the fact that for hadron flows to the central direction of the BCAL, energy is contained. For these events the current jet energy is measured inclusively and depends only little on fragmentation.

Only very low-x events, at high values of y with the current jet in the direction of the RCAL, deposit hadronic energy in the RCAL-HAC. The RCAL-EMC has a characteristic peak around the electron beam energy, due to the impact of the scattered electron. The difference between the two fragmentation models is negligible in this region, but the position of the energy peak seems to be somewhat lower in data compared to Monte Carlo simulation (see also section 5.3.1).

We showed that the hadron energy deposits depend on the fragmentation processes, especially in the forward direction. For an inclusive measurement however, fragmentation should not affect the result too drastically. Indeed, the Jacquet-Blondel prescription to obtain the kinematic variables from inclusive measurements, show that the important quantities of the hadronic system are not energies themselves, but either transverse energies or differences of energy and longitudinal momentum ($\delta_{had} = \sum_h E_h - E_h \cos\theta_h$).

The scalar sum of the transverse energy, E_t, should be large for DIS events and well described by the Monte Carlo. We show in figure 5.5a the E_t distribution of the complete event. We compare with the ARIADNE fragmentation model for two different structure function parametrizations MTB1 and MTB2. The cross sections of those distributions is normalized to the luminosity of the data. We also show the distribution for the HERWIG-SUE, but now normalized to the number of data points. The E_t distribution of the selected DIS events is consistent with the DIS Monte Carlo prediction. The E_t typically has a value of a few tens of GeV's, with the tail up to $\sim$ 100 GeV. The E_t distribution shows only a small dependence to the fragmentation models.

The distribution of the vector sum of the transverse energy of the hadron system, P_t(hadron), is shown in figure 5.5b. The hadron system consists of all cells in the calorimeter that do not belong to the scattered beam electron. We make the same comparison with the Monte Carlo as in plot 5.5a.

We plot the polar diagram for the selected DIS events in figure 5.6. Compare

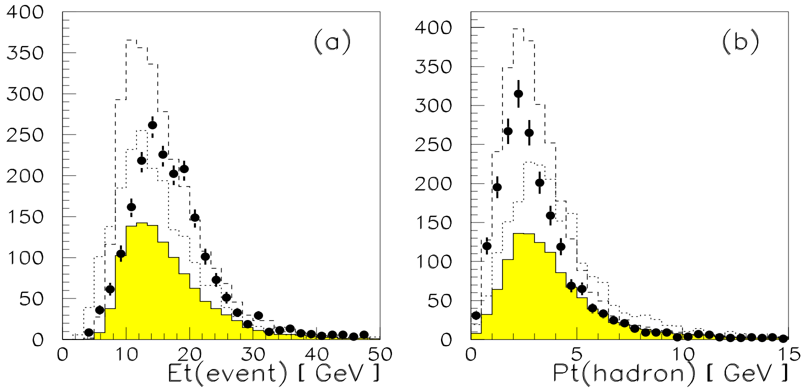


Figure 5.5: *In plot (a) the E_t distribution of the complete event. In plot (b) the P_t distribution of the hadron system. The full histogram is the $MTB1$ parametrization for the* ARIADNE *fragmentation. The dashed line for the $MTB2$ parametrization. The dotted line represents the* HERWIG *fragmentation, normalized to the number of data points.*

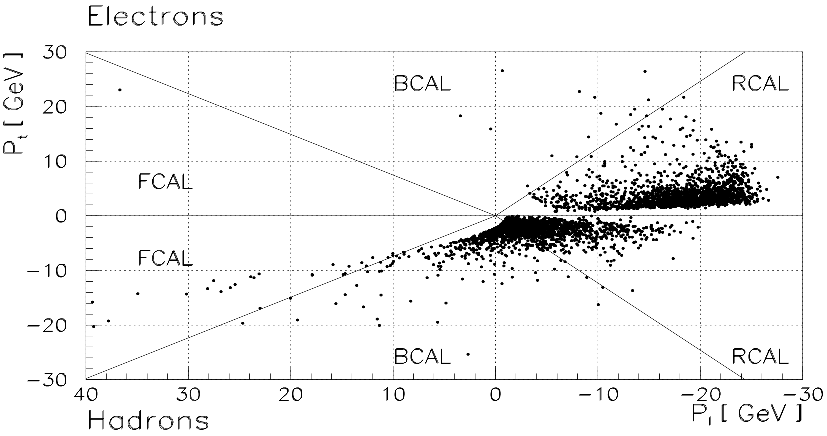


Figure 5.6: *The longitudinal and transverse momentum of the electron is plotted in the top half of this figure. The edges of the R-B- and FCAL are indicated. For the same events the momenta of the hadron system are plotted in the lower half. For the hadron system one has $P_t = F \sin\gamma_h$ and $P_l = F \cos\gamma_h$.*

this polar diagram with figure 3.4b, where we plotted the polar diagram with the isolines in x, y and Q^2. In both figures we use the same scale. In this figure, the edges of the CAL sections are indicated. As mentioned earlier, most of the electrons end up in the RCAL and only one electron is found in the FCAL. This is observed from the top half of the figure, the electron side. In the lower half of the figure we plot the P_l and P_t of the hadron system. We computed the longitudinal and transverse momentum of the hadron system, like we did in figure 3.4, as if one deals with two body kinematics (to resume: they are defined as $P_l = F \cos \gamma_h$ and $P_t = F \sin \gamma_h$, with F and γ_h reconstructed by inversion of the Jacquet-Blondel relations, as described in chapter 3). The distance of the points to the origin of the polar diagram is equivalent to the energy of the electron or hadron system. As one observes, the hadron system often has a smaller energy compared to that of the electron. This hadron energy is mostly deposited in either RCAL or BCAL. The requirement that $y_{_{JB}} > 0.04$ removes events with the hadronic energy-deposits close to the FCAL beam pipe.

With the exception of events with very low energy deposit in the FCAL, all distributions show the characteristics of DIS, as anticipated in the DIS MC.

5.3 The final state electron

The electron finder assigns a number of CAL cells to the identified electron. The total amount of energy deposit in those cells is defined to be the energy of the scattered electron. For a large part of the phase space the scattered electron energy differs only little from the incident electron beam energy, see for example in the isoline figure 3.5a. This effect results in the electron energy-peak. Since the propagator term of the cross section, $1/Q^4$, suppresses events with high values of Q^2, or large final state electron energy (a electron energy $E = 30$ GeV implies already that $Q^2 > 200$ GeV^2), the high energy end of the energy peak is, apart from detector resolution, sharply cut. On the other hand, the low energy tail of the electron energy represents DIS events with very low values of x. In this region the cross section predictions are uncertain due to the unknown proton structure function at small values of x. Hence the measured shape of the electron energy distribution discriminate between various structure function predictions.

In figure 5.7a we compare the energy distribution of the electron with MC predictions. Besides the normalization difference of the two structure function parametrizations MTB1 and MTB2, the high tails of the MC energy peak predictions overlap. The left, low energy, side of the energy peak however, differ in shape for the two structure function parametrization predictions. We therefore claim that the high tail of the energy peak provides a good test for the understanding of the energy scale. This is important because the distribution of the data shows a broadening and shift of the energy-peak that can not be explained by the structure function. We studied this effect extensively and it is believed to be due to an incorrect simulation of the configuration of the dead material in front of the

RCAL. It will not shift the impact position of the electron. The angle of the scattered electron, θ_e, is well described by the Monte Carlo, as is seen in figure 5.7b. We show the reconstructed impact position of the electron on the face of the RCAL in figure 5.7d. The fiducial cut of 16 *cm* on the impact position of the electron, symmetric in **x** and **y** position, is clearly visible.

The energy of the electron E can be determined alternatively via the double angle method in the following manner:

$$E[\theta,\gamma] = 2A\frac{\sin\gamma}{\sin\gamma + \sin\theta - \sin(\theta+\gamma)} = \frac{Q^2_{DA}}{2A(1+\cos\theta)} \tag{5.1}$$

For this method we have to rely on the determination of the hadron angle γ_h. In figure 5.7c we compare the calculated electron energy, $E[\theta,\gamma]$, with the Monte Carlo predictions. The $E[\theta,\gamma]$ is well described by the Monte Carlo and the shape of the distribution bears information of the structure of the proton.

5.3.1 The electron energy

We observed that the measured electron energy for the complete sample of selected DIS events is shifted and broadened compared to the MC predictions. This serious problem can be investigated further with the ratio

$$r[E] = \frac{E_e\,[\text{measured}]}{E\,[\theta,\gamma]}. \tag{5.2}$$

This ratio measures the ratio of the observed electron energy in the calorimeter and the calculated electron energy, using the double angle method. The ratio $r[E]$ is independent of the structure function and luminosity uncertainties and is a suitable variable to study the energy scale in the ZEUS detector. The ratio $r[E]$ generally is less than 1 since the dead material in front of the CAL degrades the energy measurement. This effect is present in data as well as in MC, but, as seen from figure 5.8a, much stronger in the data. In this figure we plot the ratio r as function of the azimuthal angle of the electron φ. The ratio $r[E]$ depends on φ for the data, whereas the MC shows almost no dependence on φ. This implies that the MC does not follow the data in parts of the detector. The scale 'problem', that is, the difference of the ratio $r[E]$ for data and MC prediction, is most pronounced at either small or large values of φ which translates in the positive **x**-position on the RCAL.

In figure 5.8b we inspect the ratio $r[E]$ on the face of the RCAL further. In this figure the RCAL is binned in areas of 5×5 cm^2. In each area the ratio $r[E]$ is determined for the data and MC. The hatching in the figure correspond to the mismatch between the ratio $r_{data}[E]$ for data and $r_{MC}[E]$ for the MC. We have defined the hatching as follows:

$$\text{dark hatching: } \frac{r_{data}[E]}{r_{MC}[E]} < 0.92 \qquad \text{light hatching: } \frac{r_{data}[E]}{r_{MC}[E]} > 1.08 \tag{5.3}$$

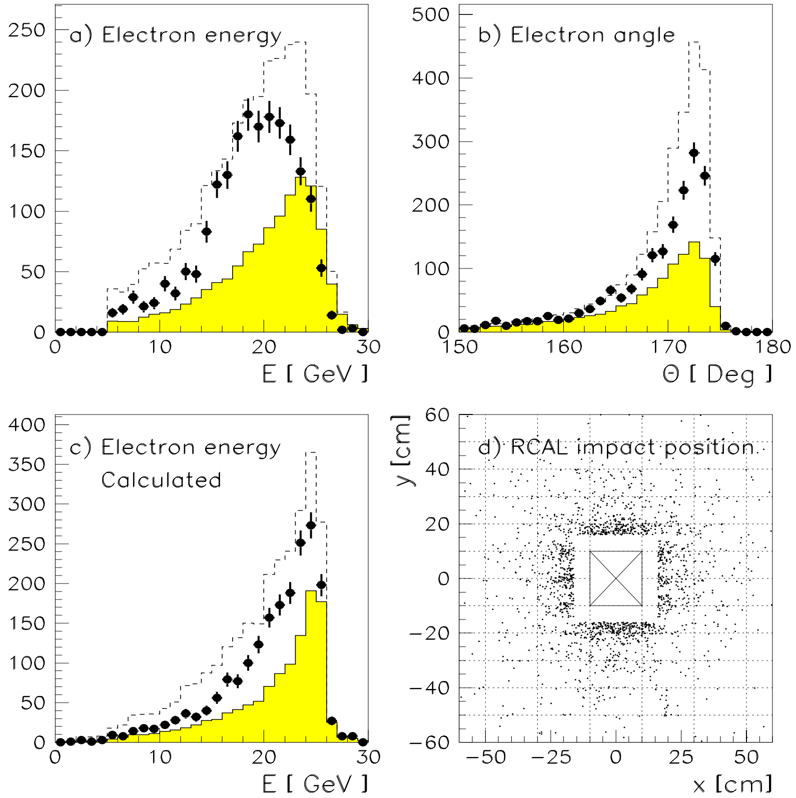


Figure 5.7: *Electron kinematics. The points are the data, the full histogram the DIS Monte Carlo (MTB1 parametrization), the dashed histogram the MTB2 parametrization. In (a) the electron energy, (b) the polar angle θ_{elec}. In (c) the energy calculated with the double angle, $E[\theta,\gamma]$, is plotted. The electron impact point in the RCAL is drawn in (d). The grid represent the EMC cell structure.*

$$\chi = \frac{|r_{data}[E] - r_{MC}[E]|}{\sqrt{\sigma^2_{stat}(r_{data}[E]) + \sigma^2_{stat}(r_{MC}[E])}} > 2.$$

The requirement that $\chi > 2$ ensures that the effect is not a statistical fluctuation. We have not implemented the fiducial cut on the impact position of the electron in this figure. Besides a few bins, the ratio $r_{data}[E]$ for the data is smaller compared to $r_{MC}[E]$ for MC on the face of the RCAL. The difference is not symmetric for positive and negative values of the x-position of the impact point and it is concentrated close

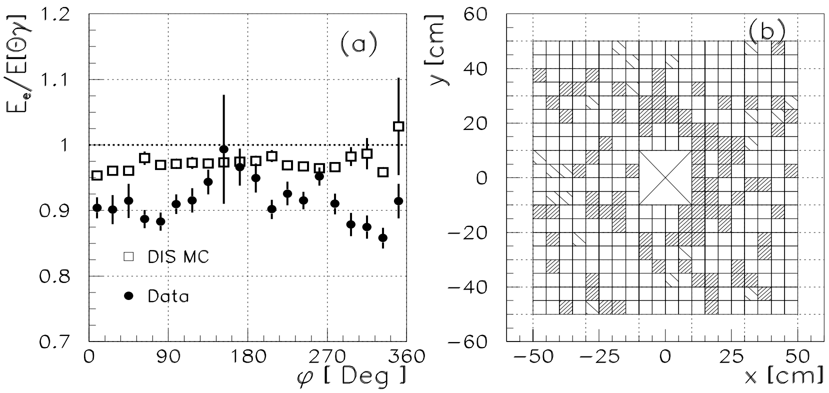


Figure 5.8: ***Comparison between the calculated electron energy $E[\theta, \gamma]$ and measured electron energy E_e. In the left plot the ratio $r = E[\theta, \gamma]/E_e$ as function of the polar angle φ is shown for Monte Carlo and data. In the right plot the face of the RCAL is binned (5×5 cm^2). For an explanation of the hatching, see text.***

to the RCAL beam pipe on the right side and above it. The shape seems a half spherical projection on the RCAL. This effect hints at the presence of material in front of the RCAL that is not simulated in the MC.

Generally, the electron produces a shower in the presence of dead material. This implies that the electron signal can deteriorate and broaden due to the dead material. The number of CAL cells that are assigned to the electron should increase in this case. In figure 5.9a we plot the distribution of the number of electron cells. In this plot the MC prediction is normalized to the number of data points. The number of CAL cells that is assigned to the electron is larger for the data compared to the MC prediction. In figure 5.9b we show the energy distribution of the electrons that have 4 or less cells assigned to it. The next figure, 5.9c, for the electrons that have 5 or more cells assigned. The distributions for the two structure functions predictions, MTB1 and MTB2, are normalized to the number of events in the data. For electrons of 4 or less cells, the shape is reasonably reporduced by the MC distributions. For electrons that have more than 4 cells however, the data show an excess of events with energy around 20 GeV.

In the next plot, 5.9d, we identify the cell with the highest energy deposit that belongs to the electron, E(cell,high). In this figure we plot the ratio of this cell to the total electron energy E[measured]. We observe that for this variable, which ranges between 0 and 1, the data agrees with the MC prediction. In the distribution of E(cell,high) itself however, we encounter the electron energy scale problem again, see 5.9e. The energy that is not contained in the highest cell, which we can write as

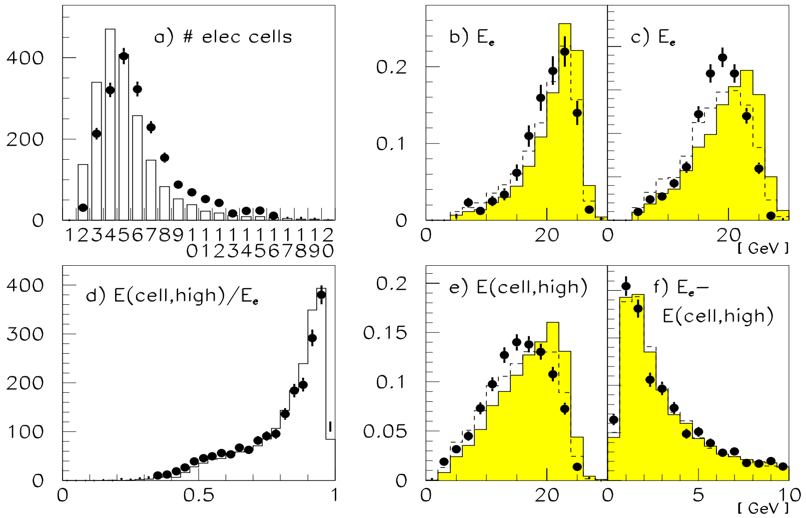


Figure 5.9: *In figure (a) we compare the number of cells assigned to the scattered electron with the MC prediction. In (b) we plot the electron energy distribution for electrons with less than 5 cells. In (c) the electrons with 5 or more cells are plotted. In plot (d) we show the ratio of the CAL cell energy that has the highest energy, E(cell,high), and the total electron energy. In (e) we plot the energy distribution of E(cell,high) itself. The last plot, (f), is the distribution of electron energy that is not contained in E(cell,high).*
The distributions in b), c), e) and f) are compared with the $MTB1$ (hatched histogram) and $MTB2$ (dashed histogram) MC predictions.

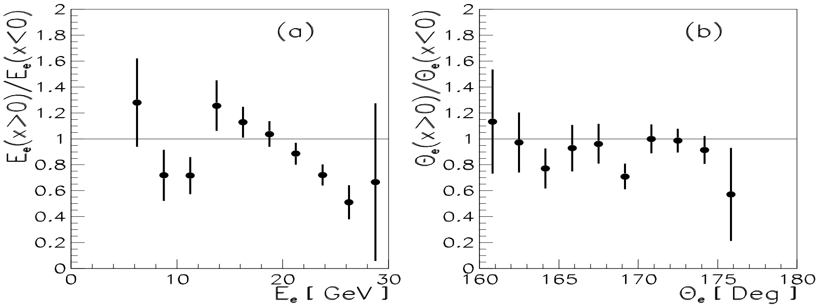


Figure 5.10: *In (a) the ratio of the electron energy-distributions with impact point in the right side of the CAL ($\mathbf{x}>0$ cm) and left side of the CAL ($\mathbf{x}<0$ cm). In (b) the same ratio for the electron angle θ distributions.*

$E_e - E(\text{cell,high})$, is plotted in figure 5.9f. This energy seems again to be reproduced by the MC.

We conclude that the electron shower in the data is broader than the MC prediction, as we observed from the number of cells that is assigned to it. Especially for electrons with a large number of cells the electron energy is degraded. The scale problem is more prominent in the cell that contains the highest energy. All this hints at an excess of dead material in front of the calorimeter in the data. However, we can draw no firm conclusion from these studies and we conclude that a precise understanding of the electron energy distribution is presently beyond our scope due to limited statistics.

A very important question is whether the angle of the electron is stable despite these problems of the electron energy scale. In figure 5.10a we plot the ratio of the electron energy distribution of the right side of the CAL, $\mathbf{x} > 0$, to the energy distribution of the left side, $\mathbf{x} < 0$. Also here we observe the shift of the data between the right and left distributions. The ratio of the θ_e distribution on the right to that on the left side of the CAL is plotted in figure 5.10b. This ratio is consistent with unity and therefore there are no major differences in the angle reconstruction between the right and left sides of the CAL.

5.4 The hadron system

In the previous section the hadron system was already heavily used to provide the hadron angle γ. The angle γ is calculated from the P_t and δ_{had} as described in section 3.3. A large fraction of the hadron energy is deposited in the rear direction, as is seen from figure 5.11a, where the γ distribution is shown. The low values of γ are rejected with the cut $y_{_{JB}} > 0.04$. The MC predictions are in accord with the observed distribution. The hadron energy F is displayed in the second plot. The tail of the data is somewhat lower than the MC whereas the peak is narrower. The mean hadron energy as a function of the hadron angle γ is plotted in figure 5.11c. The mean hadron energy is large in the forward and rear direction. We postpone a more detailed discussion about the hadron scale F to the next section.

One of the characteristics of the DIS events is that the hadron system balance the electron system in the transverse plane. In figure 5.11d we plot $\Delta\varphi = \varphi_{had} - \varphi_{elec}$, which indeed has a peak around $\Delta\varphi = 180^o$. In this plot the MC and data are normalized to one.

5.4.1 The hadron energy

We observed that the reconstructed hadron energy F does not match the MC prediction completely (figure 5.11b). Therefore the energy scale of the hadron system also needs to be investigated. Again we utilize the redundancy between the electron and hadron system. The hadron energy F (or the energy of the struck quark in the naive quark-parton model) can either be measured using the $(E - P_z)$ and the P_t of

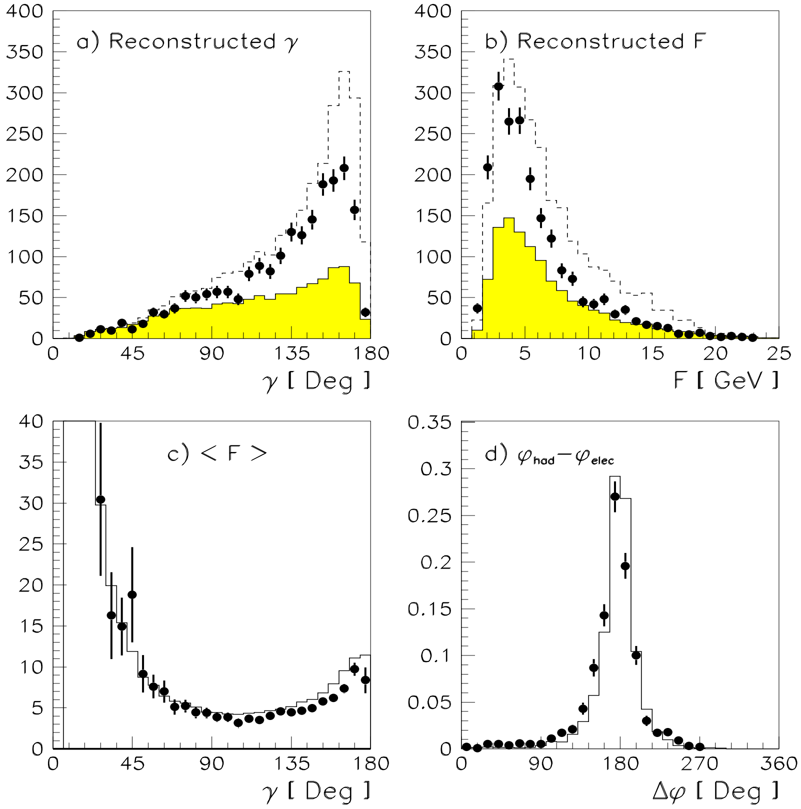


Figure 5.11: *The hadron kinematics. The points are the data, the full histogram the DIS Monte Carlo (MTB*1 *parametrization), the dashed histogram the MTB*2 *parametrization. In (a) the reconstructed hadron angle γ. In (b) the energy of the hadron system F. The mean hadronic energy as function of the hadronic angle is plotted in (c). The difference of the electron and hadron system in the azimuthal plane in (d).*

the hadron flow, as given in equation 3.5, or calculated with the two angles θ and γ, as follows:

$$F[\theta,\gamma] = 2A\frac{\sin\theta}{\sin\gamma+\sin\theta-\sin(\theta+\gamma)} = \frac{2A\ y_{DA}}{(1-\cos\gamma)} \tag{5.4}$$

In chapter 3 we showed that the determination of the angle γ is insensitive to the hadron energy in first order. Hence we can use the ratio, that we used before to

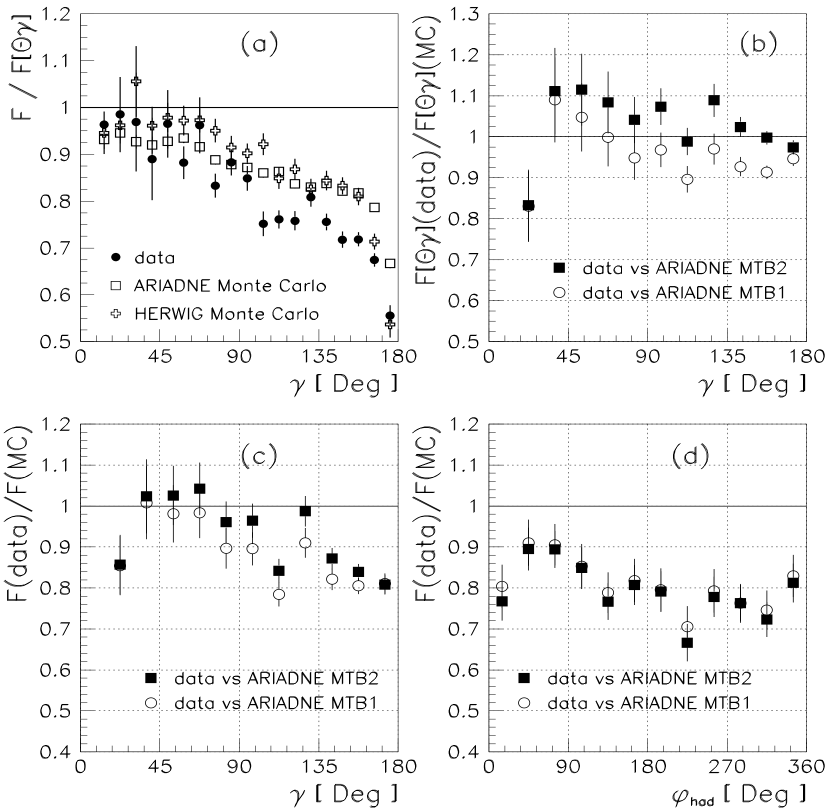


Figure 5.12: *A hadron scale study. In figure (a) we plot the ratio $F_{meas}/F\,[\theta,\gamma]$ as function of γ for the data,* ARIADNE *and* HERWIG *MC. In (b) the ratio $F_{\theta,\gamma}(\mathrm{data})/F_{\theta,\gamma}(MC)$. The structure function parametrization on this ratio is shown explicitly. In (c) we plot the ratio $F_{measured}(\mathrm{data})/F_{measured}(MC)$. In (d) the same ratio as a function of the azimuthal angle φ_{had}. In plot (d) we required that $\gamma > 130^o$.*

investigate the electron energy problem, for the hadron scale as well. We therefore define:

$$r[F] = \frac{F\,[\text{measured}]}{F\,[\theta,\gamma]} = \frac{y_{JB}}{y_{DA}}. \tag{5.5}$$

The ratio $r[F]$ is, like the ratio $r[E]$, not unity due to detector effects. Again the ratio is independent of the structure function parametrization. It is instructive to plot the ratio $r[F]$ as function of γ in figure 5.12a. To investigate the dependence on the fragmentation of this ratio, we included the HERWIG MC in this plot as well.

The two MC fragmentation models show that $r[F] < 1$, and that they match reasonably, as we expect for an inclusive measurement. A difference of the ratio between the two models is observed in the very rear direction where the jet approaches the RCAL beam-pipe. In this region the fragmentation model becomes important, where it simulates the energy-losses in the beam-pipe. The striking fact however, is that the ratio $r[F]$ is lower in the data, especially for the hadron flow in the rear direction.

This mismatch between the MC and the data can have its origin in either the determination of the hadron angle γ or the measured energy F. It is important to disentangle these effects. Therefore we inspect two alternatively defined ratio's, that depend either on the hadron angle γ or the hadron energy F only. The first one is the ratio $r[F_{\theta,\gamma}] = F[\theta, \gamma](\text{data})/F[\theta, \gamma](\text{MC})$ which is a function of the angles γ and θ only. Deviation from unity of this ratio implies that the MC expectation differs from the data, and hence that there is a mismatch between the MC and data in the determination of γ. The disadvantage of this ratio however, is that it depends on the input structure function. We therefore plot in figure 5.12b the ratio for the two extreme structure function predictions MT$B1$ and MT$B2$. The figure shows that the ratio is unity within the margin that is given by the structure functions. We conclude from this plot that there is no problem in the determination of the angle γ.

The second ratio, $r[F_{measured}] = F[\text{measured}](\text{data})/\ F[\text{measured}](\text{MC})$ depends on the hadron energy only. We plot this ratio in figure 5.12c. Also this ratio depends on the input structure functions, hence we plot the ratio for the two extreme structure function predictions. The ratio is less than one for hadron angles in the rear direction and this effect cannot be explained by the structure function. The plot therefore hints at a similar scale problem for the hadrons. In figure 5.12d we plot this ratio $r[F_{measured}]$ as function of the hadron azimuthal angle φ_{had}. In this plot we required $\gamma > 130^\circ$, to select hadron energy deposits in the RCAL. We do not observe a strong dependence of the ratio on the azimuth angle φ_{had}, as we did for the electron energy.

5.5 Reconstruction of kinematic variables

In the previous sections we have argued that there is some discrepancy between the energy scale for electrons and hadrons between the MC predictions and data. The resolution of the angles however show no such discrepancy and therefore we expect that the reconstruction method of x and Q^2 that is insensitive to the energy scale and depends on reconstructed angles only, has the smallest systematic uncertainty. Together with the fact that the double angle method is accurate for the complete phase space and almost completely insensitive to photon radiation on the lepton side (see section 1.6), we choose it as the nominal (x, Q^2) reconstruction method. The distributions for the reconstructed $y_{_{DA}}$, $x_{_{DA}}$ and $Q^2_{_{DA}}$ are shown in figure 5.13a,b

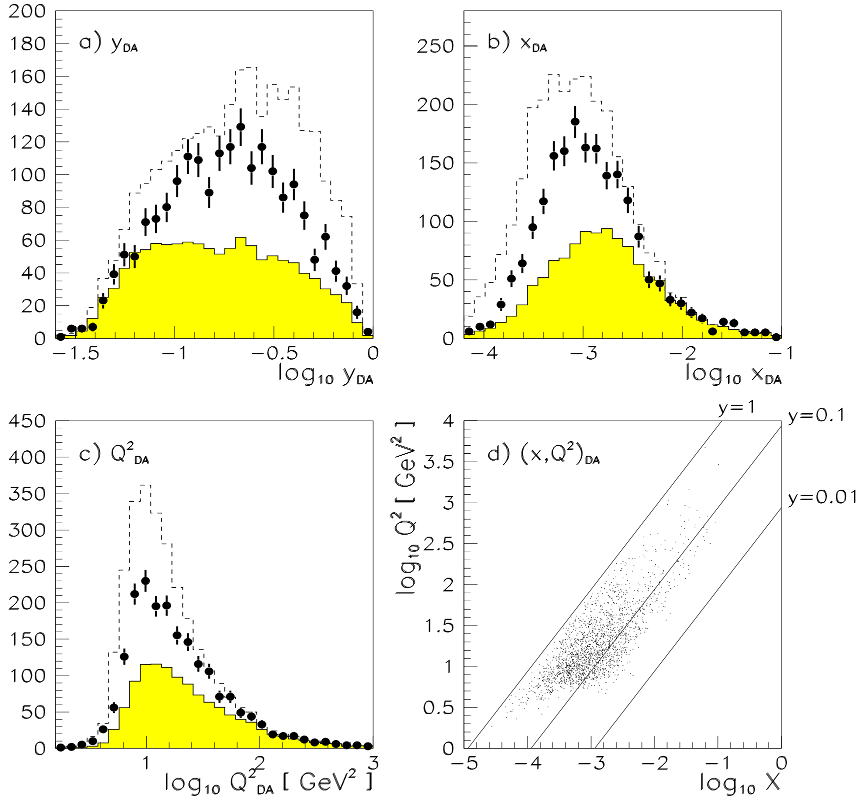


Figure 5.13: *The kinematic variables reconstructed with the double angle method. In (a) the y_{DA} distribution. In (b) the reconstruction of x_{DA} and in (c) the distribution of Q^2_{DA}. The data is compared with the* HERACLES+ARIADNE *DIS MC with the MTB1 parametrization (full histogram) and the MTB2 parametrization (dashed histogram). The events in the (x, Q^2) phase space are drawn in plot (d). The isolines $y = 1$, $y = 0.1$ and $y = 0.01$ are drawn.*

and c. Figure (5.13d) shows the distribution of the 2025 DIS events in the (x, Q^2) phase space.

The difference in the two extreme structure function parametrization vanishes at high values of x ($x > 0.01$, (plot 5.13b). In this part of phase space the data is in good agreement, in normalization and shape, with the MC predictions. It is clear that the number of events at low-x, and correspondingly low-Q^2, distinguishes

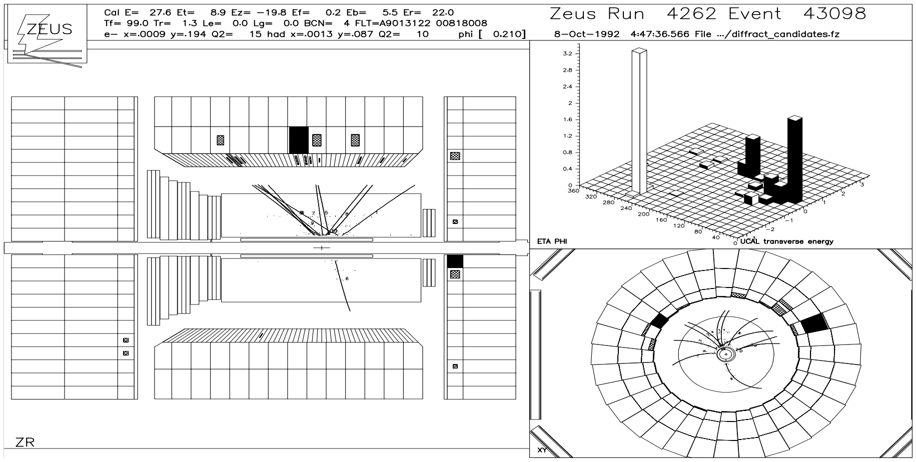


Figure 5.14: *An example of a diffractive-like event. The ZR-projection, XY-projection and the transverse CAL energy are plotted. The proton remnant around the FCAL beam-pipe is absent. The reconstructed kinematics are* $Q^2_{DA} = 17$ *GeV*^{2} *and* $x_{DA} = 0.0018$.

between the various structure function predictions.

The Q^2 distribution has a characteristic cut-off around 10 GeV^2. This corresponds to the low trigger acceptance for electrons in the vicinity of the RCAL beam-pipe. At around $Q^2 = 10$ GeV^2 the acceptance becomes close to unity. The high tail of the Q^2 distribution shows a fall-off that is almost completely due to the photon propagator term $1/Q^4$ with a small effect from logarithmic Q^2 dependence of the structure function itself.

5.6 Events with a large rapidity gap

In the data we observe a substantial number of events with low energy deposit in the FCAL which are not observed in the two MC predictions (section 5.2). In figure 5.14 we show an example of such an event. The event is clear DIS, but it lacks any energy deposit of the proton remnant.

These events can be separated from standard DIS events by a cut on the maximum pseudorapidity η_{max}. Pseudorapidity is defined for each hadronic cluster or condensate as

$$\eta = -\ln[\tan(\theta_h/2)]. \tag{5.6}$$

In this definition, θ_h is the polar angle of the hadronic cluster that has at least an energy of 400 MeV. The pseudorapidity ranges from -3.4 in the rear direction at the edge of the RCAL beam pipe to +3.8 at the edge of the FCAL beam pipe. The measured pseudorapidity however, can exceed these limits due to the fact that the

Figure 5.15: *In the left plot we show the* η_{max} *distribution of the data compared with the MC prediction. The edges of the R- B- and FCAL are indicated. The data show an excess of events that have a low value of* η_{max}*, that is, events with a large rapidity gap. The right plot shows the correlation between the visible invariant mass* M_X *and the calculated invariant mass* W_{DA} *for* $\eta_{max} > 1.5$ *and* $\eta_{max} < 1.5$*. For this last sample the visible invariant mass remains low.*

center of the cluster or condensate is reconstructed in the beam-pipe. We define the maximum pseudorapidity of an event, η_{max}, as the pseudorapidity of the hadronic cluster closest to the proton beam axis in the forward direction. In the left plot of figure 5.15 we show the distribution of η_{max} for the DIS events and the MC expectation. The data show a large number of events with a low value of η_{max}, which is not expected by the MC. This new type of events possess a large rapidity gap between the FCAL beam pipe and the hadronic cluster closest to it. They have been reported in reference [67].

We divide our DIS sample into one with $\eta_{max} < 1.5$ and one with $\eta_{max} > 1.5$. With the requirement that $Q^2_{DA} > 10$ GeV^2, in the data 89 of 1441 events have $\eta_{max} < 1.5$ while only 4 ± 1 events are expected from MC simulations. The background in the data is estimated at 7 ± 3 events. We present in the right plot of figure 5.15 the correlation between the visible invariant mass of the hadronic system measured in the detector, M_X (see for the determination of M_X section 3.4.3), and the invariant mass W_{DA} of the $\gamma^\star p$ system, calculated with the double angle method. The events with a large rapidity gap, i.e. $\eta_{max} < 1.5$, have a small visible invariant mass M_X, typically smaller than 10 GeV, compared to W_{DA} which ranges up to 260 GeV. All other characteristics of DIS are similar for the events with a large rapidity gap.

The fraction of events is roughly independent of the measured Q^2_{DA}, which implies that the events are produced via a leading twist mechanism. Also, for large values of W_{DA} the contribution of the events with a large rapidity gap is constant. These features are expected from diffractive scattering.

In normal DIS, the colour flow of the hard interaction implies that the proton breaks up, as indicated in figure 5.16a. The invariant mass of the $\gamma^\star p$ system correlates with the visible mass in this case. In diffractive scattering the absence of the proton remnant in DIS is explained by the exchange of a colourless object between the incident proton and the hard scattering. The proton (or neutron for charged colourless objects) disappears in the forward beam pipe undetected. The visible invariant mass equals the mass of system that consists of the $\gamma^\star$ and the colourless object, and is small compared to the invariant mass of the $\gamma^\star p$ system.

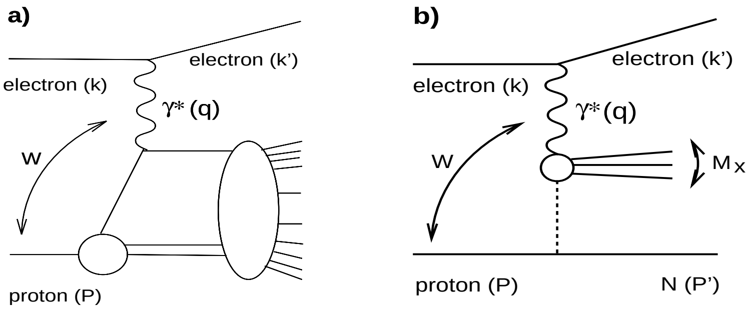


Figure 5.16: *The mechanism of 'normal' DIS, figure (a), versus the diffractive scattering (b).*

Unfortunately, in 1992 the Leading Proton Spectrometer (LPS) was not installed. With this device, it will be possible to detect protons in the very forward direction and hence confirm the suggested production mechanism of this class of events.

The diffractive events are DIS events which contribute to the cross section and have to be included in the determination of F_2. With a diffractive Monte Carlo event generator [68] we checked that the acceptance and detector smearing of kinematic variables of this class of events is similar to standard DIS events. This gives us confidence that we can use the same acceptance corrections, obtained with the HERACLES+ARIADNE generated events, for the diffractive events as well.

5.7 Summary

After the selection filters are applied, the sample consists of 2025 events. The first concern is to convince ourselves that these events are DIS. In this chapter we showed two typical examples in which the characteristics of DIS, the final state electron and the jet-like hadron flow, are present. We compared elementary distributions such as energy flow in the different CAL sections, with the MC predictions. The distributions are in accord with DIS, except for an excess of events with very low FCAL energy. This new type of events 'with a large rapidity gap' are characterized by the absence of the target remnant in the detector and are consistent with diffractive scattering.

For the reconstruction of the kinematic variables, understanding of the distributions of $\{E, \theta, F, \gamma\}$ is important. We found the two angles θ and γ in accord with the MC predictions. This was checked using the redundancy of the final state electron and hadron flow. A slight discrepancy between data and MC was observed in the distributions of E and F. For the electron energy a region on the RCAL face can be distinguished where the scale problem is most pronounced. The origin of these problems is not well understood, but is believed to be due to an incorrect configuration of dead material in front of the RCAL in the MC simulation.

The use of the double angle method to reconstruct the kinematic variables circumvents the energy scale problems. We show the distributions of x, y and Q^2 for data and MC. From these distributions it is clear that they contain information on the structure of the proton.

Chapter 6

Extraction of F_2

In this chapter we present the extraction of the structure function F_2. The chapter is organized as follows: Since the measured cross section includes the longitudinal structure function F_L, and is not corrected for QED radiation and detector smearing, we first summarize the unfolding of the structure function F_2 for these effects. We then choose bins in x and Q^2, in which F_2 will be determined. We discuss the the acceptances and photo-production background in these bins. The actual measurement of the structure function F_2 is followed by a detailed discussion on the systematic error. The final results on the measurement of F_2 are listed and presented graphically.

6.1 Unfolding of F_2

The cross-section for inclusive deep-inelastic neutral current ep scattering is written in terms of the three structure functions F_2, F_L and F_3 (see also section 1.1.2). The structure function F_3 incorporates parity violation and is produced by the weak interaction only. Since xF_3 as compared to the structure function F_2 reaches the percent level only for $Q^2 \sim 1000$ GeV^2, we can safely neglect the exchange of the weak gauge boson Z^0 at the low values of Q^2 covered by the data sample, and consider $\gamma^\star$ exchange only. With this assumption the differential cross section becomes a sum of contributions from F_2 and F_L. It is convenient to split the NC differential cross section in two terms containing these structure functions, and define:

$$\sigma(F_2) = \frac{2\pi\alpha^2}{xQ^4} Y_+ \, F_2 \quad ; \quad \sigma(F_L) = \frac{2\pi\alpha^2}{xQ^4} Y_+ \left[-\frac{y^2}{Y_+} F_L \right] \tag{6.1}$$

with $Y_+ = 1 + (1-y)^2$. The double differential cross section for neutral current DIS is then written as

$$\frac{d^2\sigma_{NC}}{dx\, dQ^2} = \frac{2\pi\alpha^2}{xQ^4} Y_+ \left[F_2 - \frac{y^2}{Y_+} F_L \right] = \sigma(F_2) + \sigma(F_L) \tag{6.2}$$

With the reconstruction of the kinematic variables of the events and the measured luminosity, one can measure the visible, uncorrected double differential cross section for the data, $d^2\sigma(data)/dxdQ^2$. The aim of the unfolding is to determine from this quantity the function $F_2(x, Q^2)$. Therefore the cross section needs to be corrected for the contribution of F_L, the detector smearing and acceptance, and QED radiation from the electron.

We determine the correction for the longitudinal structure function from the next-to-leading QCD prediction of F_L. Detector smearing, acceptance and QED radiation are corrected using the Monte Carlo sample.

6.1.1 Contribution from $d\sigma(F_L)$

The longitudinal structure function can only be disentangled from F_2 experimentally, when DIS events with various center of mass values s are available[1]. It is then possible to determine the triple differential cross section $d^3\sigma/dxdQ^2dy$ as a function of y at fixed values of both x and Q^2. The constant term of this function is proportional to $F_2(x, Q^2)$, while a term containing y^2/Y_+ is proportional to $F_L(x, Q^2)$. It is interesting to note that it is possible to obtain different values of s even without changing the beam energies. The events that radiated collinear photons from the lepton in the initial state, effectively lost electron energy before the collision. The collision that follows is then equivalent to an ep collision with degraded electron beam energy and hence has a lower value of the center of mass energy s. One can therefore argue that radiative events can be used to measure the longitudinal structure function [70].

In this analysis we correct for the longitudinal structure function, without measuring it. To this end we construct the longitudinal structure function from a parametrization of the parton densities in the part of phase space under consideration. In the absence of weak interactions, F_L is given in next to leading order QCD in the DIS scheme as [8]:

$$F_L(x, Q^2) = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dz}{z} \left(\frac{x}{z}\right)^2 \left[\frac{8}{3}F_2(z, Q^2) + 4\sum_f e_f^2 \left(1 - \frac{x}{z}\right) zG(z, Q^2)\right] \tag{6.3}$$

with $G(z, Q^2)$ the gluon distribution. At low x the term with the gluon distribution is the dominant term. We neglect contribution to F_L stemming from target mass corrections and higher twist, since both these effects are proportional with $1/Q^2$ and decrease with increasing Q^2.

We present F_L as a function of x for various values of Q^2 in figure 6.1. For this figure we used the MRS$D-'$ parametrization for F_2 and the gluon distribution. In these plots we draw F_L and F_2, as well as the combination of the two as they appear in the cross section, $F_2 - (y^2/Y_+)F_L$. The effect of F_L is small for to two

---

[1] DIS events with fixed x and Q^2 but different values of s in effect have different values of y, because $Q^2 = sxy$.

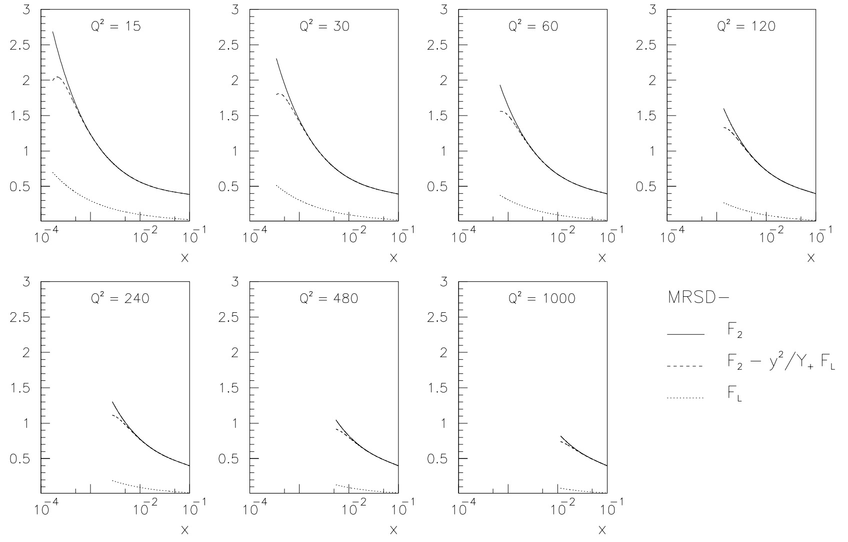


Figure 6.1: *The effect of F_L on F_2. The solid line represents the F_2 structure function as a function of x for the various Q^2 bins, according to the MRSD–′ parametrization. The dotted line is the QCD contribution to F_L. The dashed line represents the definitive correction of F_L to the cross section, which we write as $[K(F_L)]^{-1} F_2 \equiv F_2 - y^2/Y_+ F_L$. Note that at high values of y the correction becomes significant.*

reasons. Firstly, F_L vanishes in leading order QCD, as expressed by the Callan-Gross relation. So the leading contribution to F_L originates from second order QCD, and secondly the term y^2/Y_+ damps the contribution of F_L at low values of y. Only at high values of y does the contribution of F_L becomes appreciable.

We introduce a multiplicative factor $K(F_L)$ such that F_2 is written in terms of the cross section as:

$$K(F_L)\frac{d^2\sigma}{dx dQ^2} = \sigma(F_2) \quad ; \quad K(F_L) = \left(1 + \frac{\sigma(F_L)}{\sigma(F_2)}\right)^{-1} \tag{6.4}$$

In the expression for $K(F_L)$ only the ratio of $\sigma(F_L)$ and $\sigma(F_2)$ appears. Since both the parametrizations of F_2 and F_L are obtained via evolution of parton densities with the Altarelli-Parisi equations (and hence the evolution of the gluon distribution is coupled to the evolution of F_2), the contribution of $\sigma(F_L)$ to $\sigma(F_2)$ for moderate and high values of Q^2 depends only weakly on the parametrization that is used. This justifies the fact that we use a parametrization, which we take to be MRS$D-'$, to obtain $K(F_L)$ and correct with this factor the measured cross section $d\sigma/dxdQ^2$

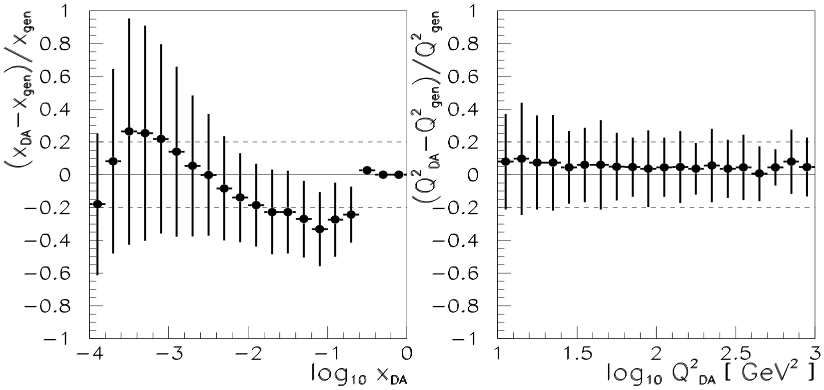


Figure 6.2: *The resolution and bias of $(x,Q^2)_{DA}$ after all selection cuts are applied. We plot $(x_{DA}-x_{gen})/x_{gen}$ as function of x_{DA} and $(Q^2_{DA}-Q^2_{gen})/Q^2_{gen}$ as function of Q^2_{DA}. The resolution is denoted by the error bar.*

for the contribution of F_L. The calculated factor $K(F_L)$ is sufficiently accurate even if the measured F_2 does not coincide with the MRS$D-'$ parametrization completely.

The correction of F_L in a finite bin $\Delta x \Delta Q^2$ is obtained by the integration of $d\sigma(F_L)$ and $d\sigma(F_2)$ over the bin. Hence for each bin we calculate the correction factor $K(\Delta F_L)$ as:

$$K(\Delta F_L) = \left(1 - \frac{\int_{\Delta x \Delta Q^2} \sigma(F_L)\, dx dQ^2}{\int_{\Delta x \Delta Q^2} \sigma(F_2)\, dx dQ^2}\right)^{-1} . \tag{6.5}$$

The value of $K(\Delta F_L)$ for the bins $(\Delta x, \Delta Q^2)$ that we choose are listed in the table with the final results on the measurement of F_2, table 6.7.

6.2 Defining bins in phase space

We determine the structure function in rectangular bins Δx and ΔQ^2. In order to reduce the systematic error on F_2, it is important to restrict the measurement of F_2 to a region of (x, Q^2) that is well measured and where the correction factors are uniform and not too large. In this section we define the bins, introduce the correction factors and label bins for which these criteria are fulfilled.

6.2.1 Defining the bins

The bin widths Δx and ΔQ^2 were chosen in such a way that a large fraction of the events remain in the bin when detector smearing is applied. Hence the bin widths are chosen to be larger than the resolutions of x and Q^2:

$$\begin{aligned} \Delta x &\geq \sigma(x) \\ \Delta Q^2 &\geq \sigma(Q^2). \end{aligned} \tag{6.6}$$

The resolution and systematic shift of x_{DA} and Q^2_{DA} after the selection cuts is shown in figure 6.2. The values in the plot correspond to the relative systematic shift and the error bar represents the mean resolution. We show the lines of ±20%. From this plot we conclude that if we select a bin width of 100%, equation 6.6 is satisfied. Although the resolution in Q^2 allows for a finer bin grid, the bins in Q^2 are chosen logarithmically equidistant to account for the lower statistics at high Q^2 values. In x the bins are large at high values of x, where due to the small event selection acceptance the statistics are also low. In this region of high x, previous measurements of F_2 are available and hence the structure function is known. A ZEUS data point in this region provides a check of the normalization of the data.

The bin edges that we chose are listed in table 6.1. We draw the bins in the phase space, together with the data points, in figure 6.3. The structure functions will be determined at the bin centers $(x, Q^2)_{cent}$.

x bin edge	x_{cent}	Q^2 bin edge	Q^2_{cent}
0.00010	0.00017	5	7.5
0.00030	0.00042	10	15
0.00060	0.00085	20	30
0.00120	0.00170	40	60
0.00240	0.00490	80	120
0.01000	0.03162	160	240
0.10000	0.31623	320	480
1.00000	-	640	1000
		1280	-

Table 6.1: *Definition of the bins in* (x, Q^2).

6.2.2 Acceptances

In this section we define several quantities that contain information about each bin (we commence each definition with the symbol ¶). These items quantify the effect of event selection and migration of events due to radiation and reconstruction errors.

¶Number of events in the bin dN: As one expects this is the number of events in a bin that have the x and Q^2 value within the limits of the particular bin. For Monte Carlo events in the bin, denoted by $dN(\mathrm{MC})$, a further subdivision is made for generated versus reconstructed kinematics, denoted by $dN(true)$ and

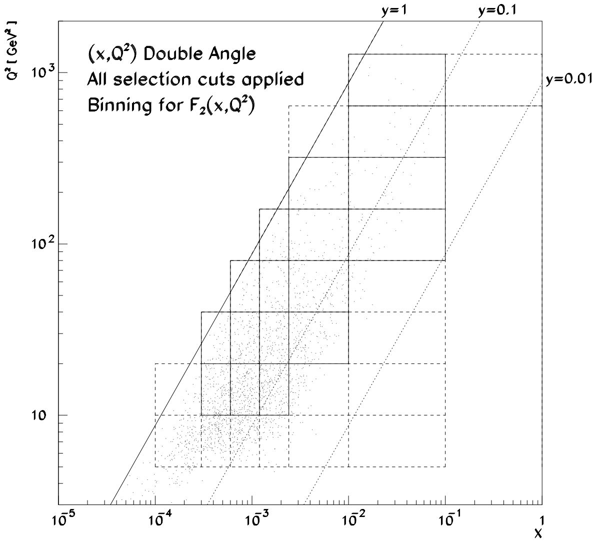


Figure 6.3: *The binning for the F_2 measurement in the (x, Q^2) phase space. The bins drawn with the dashed lines, will not survive the criteria on acceptance and the F_2 will be determined for the remaining bins only. We also plot the data.*

$dN(reco)$ respectively. Also the number of events with and without selection cuts, denoted by $dN(+cuts)$ and $dN(-cuts)$, generated with radiative corrections or Born level only ($dN(RAD)$ and $dN(Born)$) and with a particular structure function parametrization SF ($dN(SF)$) are indicated. The number of events in the bin for the data is denoted by $dN(data)$.

¶**Bin area:** The area $\Delta x \Delta Q^2$ is calculated in units of GeV^2 with the kinematical constraint $y < 1$.

¶**Acceptance due to selection:** The acceptance due to selection, Acc(sel), is defined as the fraction of the generated number of DIS Monte Carlo events that pass the selection cuts:

$$\mathrm{Acc(sel)} := \frac{dN(MC, true, +cuts, RAD, SF)}{dN(MC, true, -cuts, RAD, SF)} \tag{6.7}$$

This quantity gives the fraction of events that is lost by applying the selection cuts on the Monte Carlo.

¶**Acceptance for migration:** The acceptance for bin migration, due to smearing of the kinematic variables, Acc(mig), is defined as the ratio of the number of reconstructed and generated events in the bin that pass the selection cuts:

$$\mathrm{Acc(mig)} := \frac{dN(MC, reco, +cuts, RAD, SF)}{dN(MC, true, +cuts, RAD, SF)} \tag{6.8}$$

Deviation from unity of this quantity is due to bin to bin migration. Both detector smearing and photon radiation are at the origin of this migration.

¶**Total acceptance:** The total acceptance of a bin, Acc(tot), is defined as the ratio of reconstructed events that pass selection and the number of generated events in that bin:

$$\mathrm{Acc(tot)} := \frac{dN(MC, reco, +cuts, RAD, SF)}{dN(MC, true, -cuts, RAD, SF)} = \mathrm{Acc(sel)} \cdot \mathrm{Acc(mig)} \tag{6.9}$$

The total acceptance is the full correction factor for detector smearing and selection acceptance.

¶**Radiative correction:** Although the event generator includes emission of photons, virtual corrections also change the cross section in the bin. This effect cannot be traced event by event and hence a bin dependent multiplicative factor, Radcor, is introduced to describe this effect:

$$\mathrm{Radcor} := \frac{\mathcal{L}_{RAD}}{\mathcal{L}_{Born}} \cdot \frac{dN(MC, true, -cuts, Born, SF)}{dN(MC, true, -cuts, RAD, SF)} \tag{6.10}$$

where the luminosity $\mathcal{L}$ for Monte Carlo events equals $N_{tot}/\sigma(tot)$. In order to obtain the numerator of equation 6.10 we generated a sample of $100k$ HERACLES DIS events, with the contribution to the cross section from real and virtual QED radiation turned off. With this definition, the combination 'Radcor/Acc(tot)' is the ratio between the measured and corrected Born cross section in the bin.

¶**Stability:** The sensitivity of the acceptance and radiative corrections to the input structure function is quantified with the 'stability' (Stab). The 'stability' is defined to be the ratio of the correction factors for the cross section in the bin, for two different structure functions:

$$\mathrm{Stab}(SF, SF') := \frac{\mathrm{Radcor}(SF)/\mathrm{Acc}(\mathrm{tot}, SF)}{\mathrm{Radcor}(SF')/\mathrm{Acc}(\mathrm{tot}, SF')} \tag{6.11}$$

Large deviations from unity of this quantity indicate that the correction factors depend strongly on the input structure function.

¶**Purity:** Due to bin to bin migration, some of the events are lost in the bin, while other events migrate into the bin. Whereas the acceptance due to migration, Acc(mig), is sensitive to the sum of these two effects only (after the selection), purity is defined to be the ratio of the number of events that remain in the bin when the smearing is applied to the total number of events in that bin:

$$\text{Purity} := \frac{dN(MC, reco, -cuts, RAD, SF; remained)}{dN(MC, reco, -cuts, RAD, SF)} \tag{6.12}$$

Purity is by definition less than one and its value depends strongly on the bin size. A low purity implies that the bin is unstable and correspondingly the systematic error will be large.

In table 6.2, the above defined quantities are listed for the DIS Monte Carlo. The events were distributed in phase space according to the MRS$D-'$ parametrization. The subscript $+cuts$ indicate that the nominal selection cuts are applied, which are listed in section 4.9 and $reco$ means that the kinematics are reconstructed with the double angle method. We quote the 'stability' of the two extreme input structure function parametrizations, MT$B1$ and MT$B2$. The number of data events $dN(data)$ is listed in the last entry.

To reduce the systematic uncertainty on F_2 we required that the total acceptance ranges between:

$$0.5 \leq \text{Acc(tot)} \leq 2 \tag{6.13}$$

This requirement implies that the structure functions are not determined for the dashed lined bins in figure 6.3. From table 6.2 we conclude that especially the acceptance due to selection, Acc(sel), is low for the rejected bins. At low values of Q^2 the selection acceptance is dominated by the fiducial cut of the electron impact point around the RCAL beam pipe and by the calorimeter First Level Trigger (CAL FLT). A source of systematic error is the uncertainty on the value of the hard-wired thresholds that are used in the CAL FLT. This uncertainty is negligible for events with $Q^2 > 10$ GeV^2, where the CAL FLT acceptance is very close to unity.

Also bins at low values of y are rejected. The main reason for the low value of the acceptance is the requirement $y_{_{JB}} > 0.04$.

The radiative corrections in the selected bins are below $\sim 10\%$. The purity is not very high for the bins with low Q^2 or high y values, but reaches more than 50% otherwise. Although we do not use the values of the purity in the determination of the systematic error, the values correlate with the quoted systematic uncertainty, which are listed in table 6.7 at the end of this chapter.

Q^2	x	$\Delta x \Delta Q^2$	Acc (sel)	Acc (mig)	Acc (tot)	Rad cor	Pur %	Stab	dN (data)
7.5	0.00017	0.001	0.39	0.73	0.29	0.92	24	0.96	139
	0.00042	0.002	0.38	1.20	0.46	0.85	39	1.13	194
	0.00085	0.003	0.29	1.25	0.36	0.90	42	1.07	139
	0.00170	0.006	0.20	1.17	0.23	0.91	45	1.04	60
	0.00490	0.038	0.03	0.53	0.02	0.93	41	1.06	9
	0.03162	0.450	-	-	0.00	0.97	-	-	0
15	0.00017	0.003	0.34	0.74	0.25	0.93	12	1.04	25
	0.00042	0.003	0.76	1.11	0.85	0.91	37	1.11	165
	0.00085	0.006	0.82	1.21	0.99	0.89	50	1.09	204
	0.00170	0.012	0.79	1.41	1.11	0.93	59	1.07	175
	0.00490	0.076	0.33	0.86	0.28	0.91	66	1.09	66
	0.03162	0.900	-	-	-	1.03	-	-	0
30	0.00042	0.005	0.52	1.02	0.53	0.99	26	1.04	42
	0.00085	0.012	0.89	1.11	0.99	0.99	50	1.05	119
	0.00170	0.024	0.97	1.20	1.16	0.96	63	1.04	119
	0.00490	0.152	0.79	1.09	0.86	1.00	75	1.02	125
	0.03162	1.800	0.07	0.08	0.01	0.96	47	0.99	1
60	0.00085	0.020	0.68	1.24	0.84	1.00	42	1.09	27
	0.00170	0.048	0.95	1.04	0.98	0.93	57	1.01	48
	0.00490	0.304	0.96	1.21	1.16	0.92	81	0.99	106
	0.03162	3.600	0.34	0.54	0.18	1.00	64	1.02	15
120	0.00170	0.078	0.68	1.21	0.83	0.75	65	1.10	22
	0.00490	0.608	0.96	1.12	1.07	0.92	81	1.02	50
	0.03162	7.200	0.66	0.90	0.60	0.94	78	1.02	31
240	0.00490	1.146	0.87	1.31	1.14	0.94	78	1.08	19
	0.03162	14.400	0.88	1.17	1.03	1.20	82	0.99	24
480	0.00490	1.440	0.83	1.20	0.99	0.86	69	1.21	2
	0.03162	28.800	0.90	1.16	1.04	0.93	74	0.98	18
1000	0.03162	56.648	0.92	1.39	1.28	0.89	88	0.99	8
	0.31623	576.000	0.89	0.58	0.51	0.67	43	1.00	0

Table 6.2: *This table contains quantities that are determined from a sample of 50k generated Monte Carlo DIS events. The bin area is given in the third column. The acceptances, 'radcor' and 'purity' are determined with the MRSD$-'$ parametrization, double angle reconstruction,* EF1 *electron finder and standard cuts. We plot Stab$(MTB2, MTB1)$ as the most extreme case. The number of events for the data is listed in the last column.*

6.3 Background subtraction and energy scale correction

For a correct measurement of the structure function F_2, we have to subtract the remaining photoproduction background from the sample. In contrast to the estimate of the background using the photoproduction Monte Carlo PYTHIA, as described in section 4.8, we determined the photoproduction background from the data itself. Therefore we investigated the δ distribution for data and Monte Carlo *without* the selection cut $\delta > 35$ GeV, in each (x, Q^2) bin. From the shape of δ as determined from the Monte Carlo, the contribution of DIS events in the δ distribution of the data was fitted. The background contribution of photoproduction events to the distribution, characterized by lower values of δ, was fitted in the data sample as well such that contributions from the two event types could be disentangled.

In this method, we took care of the energy scale problem. Since the value of δ is very sensitive to energy deposits in the rear direction, the δ distribution of the data is somewhat broader and shifted with respect to the Monte Carlo distribution. The fraction of events that is removed with the cut $\delta > 35$ GeV can therefore be different for the data and Monte Carlo. We correct for this effect.

We showed the δ-distribution for Monte Carlo DIS events in figure 4.4, which peaks around $2A$ and has a 'radiative tail' towards lower values of δ. We fit for each bin in $(\Delta x, \Delta Q^2)$ the δ distribution to functional forms $\Delta(\delta)$, which contains six free parameters. These are the widths $\sigma_1(MC)$ and $\sigma_2(MC)$ of a Gaussian shape on either side of a central value δ_0^{MC} with two normalization factors C_1 and C_2. We determined the central value δ_0^{MC} with the fit as well. For the lower side of the distribution an exponential form, parametrized by α, is added to account for the radiative tail of the δ distribution:

$$\Delta_{DIS}^{MC}(\delta) = \begin{cases} C_1 e^{-(\delta-\delta_0^{MC})^2/2\sigma_1^2(MC)} & \delta > \delta_0^{MC} \\ C_2 e^{-\alpha(\delta-\delta_0^{MC})} + & \\ \quad (C_1 - C_2) e^{-(\delta-\delta_0^{MC})^2/2(\sigma_1^2(MC)+\sigma_2^2(MC))} & \delta \leq \delta_0^{MC} \end{cases} \tag{6.14}$$

The functional form of $\Delta_{DIS}^{MC}(\delta)$ is chosen purely phenomenologically and it is used because it describes the asymmetric Monte Carlo shape of δ very well.

The contribution to the δ distribution from DIS events in the data has, apart from a different overall normalization, the same functional form $\Delta_{DIS}^{data}(\delta)$. The width and the central value of the data distribution, $\sigma_0(data)$ and δ_0^{data}, are allowed to differ from the Monte Carlo distribution. In this way we take into account the broadening and degrading of the data δ distribution as compared to the Monte Carlo distribution in the fit. In addition to the functional form $\Delta_{DIS}^{data}(\delta)$, a piece that describes the photoproduction background is added, Δ_{BG}^{data}. The shape of the photoproduction as a function of δ for Monte Carlo events is also drawn in figure 4.4. We parametrized this background by a simple Gaussian. The Gaussian shape of the background part $\Delta_{BG}^{data}(\delta)$ was found to give an acceptable fit for the minimum bias photoproduction events that pass the selection cuts. Hence to fit the δ distribution of the data we used:

$$\begin{aligned}
\Delta^{data}_{DIS}(\delta) &= \begin{cases} C_3\left[C_1 e^{-(\delta-\delta_0^{data})^2/2\sigma_0^2(data)}\right] & \delta > \delta_0^{data} \\ C_3\Big[C_2 e^{-\alpha(\delta-\delta_0^{data})} + \\ \quad (C_1 - C_2) e^{-(\delta-\delta_0^{data})^2/2(\sigma_0^2(data)+\sigma_2^2(MC))}\Big] & \delta \leq \delta_0^{data} \end{cases} \qquad (6.15) \\
\Delta^{data}_{BG}(\delta) &= C_4 e^{-(\delta-\delta_0^{BG})^2/2\sigma_3^2(BG)} \\
\Delta^{data}(\delta) &= \Delta^{data}_{DIS}(\delta) + \Delta^{data}_{BG}(\delta)
\end{aligned}$$

The central value of δ for the background, δ_0^{BG}, is constrained to be less than 32 GeV and $5 < \sigma_3(BG) < 9$ GeV.

In figure 6.4 we present the results of the fit procedure. The δ distribution of the data is shown for $16 < \delta < 60$ GeV in each $(\Delta x, \Delta Q^2)$ bin. The distributions $\Delta^{data}_{DIS}(\delta)$ and $\Delta^{data}_{BG}(\delta)$, obtained from the fit, are shown as histograms with different hatching styles. We conclude that the requirement for the nominal selection, $\delta > 35$ GeV, is chosen perfectly. In all bins this cut optimally divides the DIS events and photoproduction background and hence the correction is small. From figure 6.4 we conclude that at low values of y the photoproduction background is almost absent, but pronounced at high values of y.

The fits describe the data well, and the corrected number of events in the bins was estimated as:

¶**Corrected number of events in bin $dN(data, corr)$:** The number of events in the bin corrected for photoproduction and energy scale problems, is determined from the fit results. The contribution of the background (photoproduction), $dN(data, BG)$, is calculated as

$$dN(data, BG) = \int_{35\text{ GeV}}^{60\text{ GeV}} d\delta\ \Delta^{data}_{BG}(\delta). \qquad (6.16)$$

The correction for the energy mismatch between the data and the Monte Carlo, characterized by the difference of the central value of δ for Monte Carlo, δ_0^{MC}, and data, δ_0^{data}, is given by:

$$dN(data, lost) = \int_{35-(\delta_0^{MC}-\delta_0^{data})}^{35\text{ GeV}} d\delta\ \Delta^{data}_{DIS}(\delta). \qquad (6.17)$$

The number of events in the bin was also corrected for the electron-gas background. We estimated this fraction from the number of events in the bin stemming from the electron pilot bunch, $dN(data, e\text{-pilot})$, see section 4.8. The corrected number of events in the bin, $dN(data, corr)$ is consequently given by

$$\begin{aligned} dN(data, corr) &= dN(data) \qquad (6.18) \\ &\quad - dN(data, BG) + dN(data, lost) - 9.73 \cdot dN(data, e\text{-pilot}). \end{aligned}$$

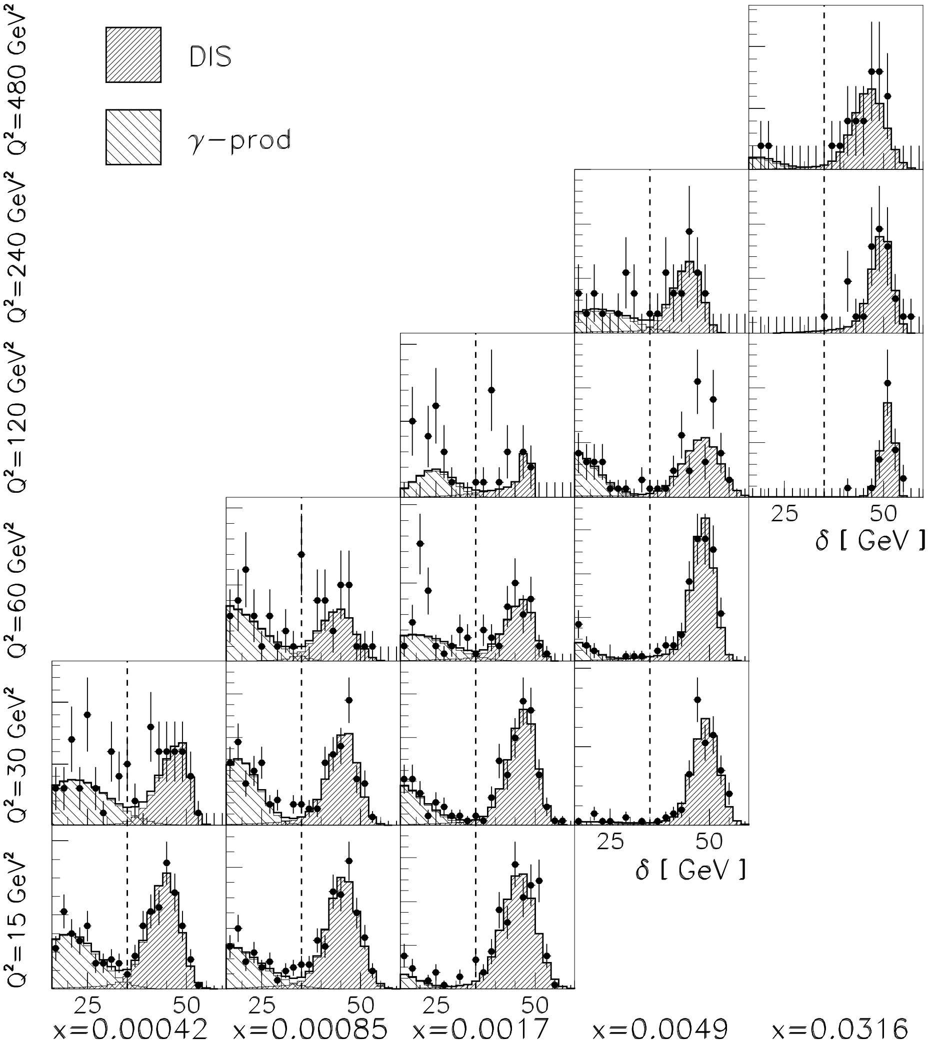


Figure 6.4: *The δ distribution in the (x, Q^2) bins. The δ ranges from 16 to 60 GeV in each bin, and the cut of $\delta = 35$ GeV is represented by the dashed line. The data is shown with the fitted contributions from DIS and photoproduction. The histogram with the light hatching represents the fitted photoproduction background, the dark hatching the contribution of DIS events.*

		$\delta > 35$ GeV			$\delta > 33$ GeV			$\delta > 40$ GeV		
Q^2	x	dN (BG)	dN $(lost)$	dN $(corr)$	dN (BG)	dN $(lost)$	R_c (33)	dN (BG)	dN $(lost)$	R_c (40)
15	0.00042	5.3	3.6	154	9.9	2.2	1.01	0.8	17.7	0.97
	0.00085	2.1	4.0	206	3.6	2.9	1.03	0.4	18.0	0.98
	0.00170	0.0	2.9	178	0.0	2.0	1.01	0.0	13.6	0.99
30	0.00042	2.8	0.1	39	4.3	0.0	1.04	0.8	0.1	1.00
	0.00085	0.1	2.4	121	0.3	1.8	1.01	0.0	9.5	1.00
	0.00170	0.1	1.7	121	0.3	1.3	0.99	0.0	6.4	0.99
	0.00490	0.1	0.7	126	0.2	0.6	1.00	0.0	1.4	1.00
60	0.00085	0.6	0.5	27	1.0	0.3	1.16	0.1	2.1	0.89
	0.00170	1.4	0.8	47	2.2	0.6	1.04	0.4	3.0	0.94
	0.00490	0.0	1.2	107	0.0	1.0	1.00	0.0	2.1	0.94
120	0.00170	0.2	0.1	22	0.5	0.1	0.97	0.0	0.0	0.57
	0.00490	0.0	0.4	50	0.1	0.3	1.03	0.0	1.5	0.99
	0.03162	0.0	0.4	31	0.0	0.0	1.00	0.0	0.1	1.01
240	0.00490	0.0	0.0	19	1.2	0.0	0.96	0.0	0.4	0.78
	0.03162	0.0	0.0	24	0.0	0.0	1.03	0.0	0.0	1.02
480	0.03162	0.0	0.0	18	0.0	0.0	0.96	0.0	1.3	1.02
1000	0.03162	0.0	0.0	8	0.0	0.0	0.99	0.0	0.0	0.90

Table 6.3: *Estimate of photoproduction background and energy mismatch in the bins for three different cut values on* δ. *For the three cut values the* $dN(data, BG)$ *and* $dN(data, lost)$ *are listed. For* $\delta > 35$ *GeV the corrected number of events in the bins are listed. The ratios* $R_c(33)$ *and* $R_c(40)$ *are listed as well. For an explanation, see text.*

with the factor 9.73 obtained from the reweighting of the pilot to the non-pilot electron beam current. Only one electron-pilot event was found in the bins, which was in bin $(x = 0.00042, Q^2 = 15)$.

In table 6.3 we list the number $dN(data, BG)$, $dN(data, lost)$ and $dN(data, corr)$ for each bin. The estimated background obtained with this fit is in good agreement with the estimated background as obtained with the PYTHIA photoproduction Monte Carlo, described in section 4.8. Correction for the energy mismatch, $dN(data, lost)$, is appreciable in bins where the δ distribution is broad and shifted. This happens at low values of Q^2, or small electron scattering angles θ_e, where the scale problem is most pronounced due to dead material in front of the RCAL.

Since photoproduction is characterized by low values of δ, the amount increases when the cut value on δ is lowered. The value for $dN(data, lost)$ decreases in this case since we move away from the peak $2A$ of the δ distribution. We show the values of these quantities for $\delta > 33$ GeV in table 6.3. The situation is reversed when we increase the cut on δ to $\delta > 40$ GeV.

To investigate whether the fits are reasonable, we introduce the ratio $R_c(\delta_c)$ as:

$$R_c(\delta_c) = \frac{dN(data, corr, \delta > 35)}{dN(data, corr, \delta > \delta_c)} \cdot \frac{dN(MC, reco, +cuts, RAD, SF, \delta > \delta_c)}{dN(MC, reco, +cuts, RAD, SF, \delta > 35)}, \quad (6.19)$$

and claim that, apart from statistical fluctuations, R_c is independent of the cut value on δ, (denoted by δ_c), in the case that the fits describe the background and energy mismatch perfectly. For then the fits correct for these effects when the value δ_c is changed, and the fraction of DIS events in the Monte Carlo that is rejected with this new cut $\delta > \delta_c$ equals the fraction of data events that is rejected. Deviations from unity of the ratio $R_c(\delta_c)$ signals our ignorance of the background events, especially when $\delta_c < 35$ GeV, and energy mismatch when $\delta_c > 35$ GeV.

We list the values for $R_c(33)$ and $R_c(40)$ in table 6.3 as well. The uncertainty of the photoproduction is typically a few percent and reaches 16% for bin ($x = 0.00085, Q^2 = 120$). This is due to a large statistical fluctuation in that bin, see figure 6.4. The uncertainty of the fit due to the energy scale mismatch is also typically a few percent, but reached, due to a statistical fluctuation, 43% for bin ($x = 0.0017, Q^2 = 120$). We note that the highest Q^2 bin has too few events for the fit procedure.

6.4 Iterative method to obtain F_2

The Monte Carlo DIS sample provides the correction factors for event migration and acceptance used for the F_2 measurement. The relative event migration and radiation depend on the distribution of the events in phase space, hence the correction factors depend on the input structure function. Therefore the Monte Carlo input structure function, F_2^{SF}, should coincide with the unknown, measured and corrected, structure function F_2^{meas}. This is the reason that in principle it is necessary to obtain the corrected measured F_2^{meas} via an iterative approach. Note, however, that the the propagator term of the cross section is dominant for the dependence of the correction factors on the event migration and therefore the correction factors are expected to depend only little on the structure function.

To perform the iteration we compare the number of events of the data and Monte Carlo simulation in each bin, and define the ratio $d\rho$ as:

$$d\rho = \frac{\mathcal{L}_{RAD}}{\mathcal{L}_{data}} \frac{dN(data, corr)}{dN(MC, reco, +cuts, RAD, SF)} \quad (6.20)$$

Deviations of this ratio $d\rho$ from a flat distribution in the (x, Q^2) phase space probe the difference between the input structure function of the simulation and the one which has to be obtained from the data. The normalization is left free to account for the luminosity measurement uncertainty. In the iteration, the information contained in $d\rho$ is used to modify the input structure function F_2^{SF} of the simulation. This implies that the distributions of the true $(x, Q^2)_h$ (the non-smeared (x, Q^2) from the hadron vertex) are altered. This change is reflected in the distribution of events with respect to the reconstructed $(x, Q^2)_{DA}$ as well (in equation 6.20 parametrized with the label SF), and hence the ratio $d\rho$ changes. The iteration stops if the ratio $d\rho$ is constant and just a normalization factor. The measured structure function F_2^{meas}, corrected for radiation and detector effects, equals the input structure function of the last iteration after taking into account the normalization $d\rho^{last\ iter}$.

We first determine the deviation of the ratio $d\rho$ from a flat distribution in (x, Q^2) phase space:

$$\chi^2 = \sum_{bins} \left(\frac{d\rho - \langle\rho\rangle}{\epsilon_\rho} \right)^2 \tag{6.21}$$

where ϵ_ρ represents the statistical fluctuation of the ratio $d\rho$, and $< \rho >$ the mean value of $d\rho$ over the bins. Secondly, we fitted the ratio $d\rho$ to a function, which we chose as

$$f_1(x, Q^2) = A(1 + B \log^C x) \left[1 + D \log x \log(Q^2/5\mathrm{GeV}^2)\right] \tag{6.22}$$

with A, B, C and D free parameters. This function is used to weight the true $(x, Q^2)_h$ events in the Monte Carlo simulation. The ratio of the data to the reweighted Monte Carlo events is determined again, yielding a new ratio $d\rho^{2nd\ iter}$, and the χ^2 was re-evaluated. The new ratio was again fitted to the same functional form as f_1, denoted by f_2. The Monte Carlo events were then reweighted with $f_2 \cdot f_1$.

The procedure was continued until the change of χ^2 between two iterations was less than one. The measured structure function becomes:

¶**The measured structure function F_2^{meas}** is defined as

$$F_2^{meas}(x_{cent}, Q^2_{cent}) = d\rho^{last\ iter} \cdot (f_1 \cdot f_2 \cdots) \cdot K(\Delta F_L) F_2^{SF}(x_{cent}, Q^2_{cent}) \tag{6.23}$$

Note that we included the correction for the longitudinal structure function as well.

In table 6.2 we listed the sensitivity of the correction factors to the two extreme input structure functions, Stab($MTB1, MTB2$). The values for the stability are typically on the percent level, and below 10% for the bins in which we determine F_2. The variation of F_2 due to the iteration is smaller than these numbers, and is negligible with the present statistics of the data. If we use only one iteration step to obtain F_2^{meas}, equation 6.23 is reduced to

$$F_2^{meas}(x_c, Q^2_c) = \frac{\mathcal{L}_{RAD}}{\mathcal{L}_{data}} \frac{dN(data, corr)}{dN(MC, reco, +cuts, RAD, SF)} K(\Delta F_L) F_2^{SF}(x_c, Q^2_c) \tag{6.24}$$

This is of course only valid if the ratio does not vary appreciably over the bin.

6.5 Systematic uncertainties of F_2

The systematic uncertainty on the nominal value of the extracted structure function has been studied extensively. In this section we present the result of this study. First, as a systematic check, we investigated the variation of F_2 when we altered the analysis slightly. Secondly we discuss the possible sources of systematic uncertainties on the measurement of F_2 and assign errors to each of these sources in order to obtain a realistic systematic error on the measured F_2 for each bin.

6.5.1 Systematic checks

The study of the systematic checks was performed by extracting the structure function many times under somewhat different conditions [50]. The values for the selection cuts were modified, the input structure function parametrizations were chosen differently, the EF2 electron finder was used, etcetera. The variations on the extracted F_2 indicate the sensitivity of each bin to these modifications.

For the nominal extraction of F_2 all standard selection cuts were applied, the EF1 electron finding algorithm was used and the kinematics was reconstructed with the double angle method. The Monte Carlo events were distributed according to the MRS$D-'$ parametrization. The checks we made were:

1 Nominal extraction of F_2. No iteration, F_L correction or background subtraction.
2 The value of the cut on δ was lowered from 35 GeV to 33 GeV.
3 The value of the cut on δ was increased from 35 GeV to 40 GeV.
4 The identified electron has an minimum energy of 7 GeV in stead of 5 GeV.
5 The box cut was discarded completely.
6 The value of the cut on y_{JB} was lowered to $y_{JB} > 0.03$
7 The value of the cut on y_{JB} was raised to $y_{JB} > 0.05$
8 The EF2 electron finding algorithm was used.
9 The analysis was performed with electrons with an impact position $\mathbf{x} < 0\ cm$.
10 The same for the other side of the calorimeter, $\mathbf{x} > 0\ cm$.
11 The Monte Carlo electron energy was scaled with a factor 0.9.
12 The Monte Carlo cell energies which do not belong to the electron were scaled by $1.1 - 0.001\theta_H$, θ_H in degrees.
13 Only those events were used that have the event vertex position $\mathbf{z} > 0\ cm$.
14 The events were used that have the event vertex position $\mathbf{z} < 0\ cm$.
15 The event vertex was shifted by $+3\ cm$.
16 The event vertex was shifted by $-3\ cm$.
17 The Monte Carlo events were reweighted to the MTB1 parametrization.
18 The Monte Carlo events were reweighted to the MTB2 parametrization.
19 The Monte Carlo events were reweighted to the MRS$D0'$ parametrization.
20 In stead of the double angle reconstruction method, the $y\theta$ method was used.
21 The electron method was used for reconstruction of kinematic variables, with $y_{elec} > 0.05$.

Q^2 (GeV^2)	15			30			
x ($\cdot 10^{-4}$)	4.2	8.5	17	4.2	8.5	17	49
1. Nominal	1.27	1.07	0.78	1.67	1.58	1.08	0.66
Stat error	9	8	9	20	13	13	11
2. $\delta > 33$	2	3	1	12	3	-1	1
3. $\delta > 40$	-15	-10	-7	-10	-9	-5	-2
4. $E_e > 7$	3	0	0	-2	1	0	0
5. No box cut	0	-6	0	0	-1	1	1
6. $y_{JB} > 0.03$	0	1	1	0	0	0	4
7. $y_{JB} > 0.05$	0	-1	-6	0	0	0	0
8. EF2 finder	19	0	2	14	-8	1	5
9. x-pos > 0	-3	-2	-8	-16	-7	-4	-2
10. x-pos < 0	3	3	8	13	7	4	2
11. $0.9 \cdot E_e$	2	2	3	1	2	2	2
12. E_h	9	1	1	21	8	2	-5
13. $\mathbf{z} > 0.$	-9	-6	4	18	2	-24	-5
14. $\mathbf{z} < 0.$	16	8	-3	-3	4	19	5
15. $\mathbf{z} + 3\ cm$	2	1	1	13	5	7	2
16. $\mathbf{z} - 3\ cm$	-3	-3	0	-5	-5	-5	2
17. MTB1 par	11	9	9	9	9	8	6
18. MTB2 par	-2	-2	0	0	1	1	1
19. MRS$D0'$ par	8	7	5	4	5	4	3
20.$(x, Q^2)_{y\theta}$	-34	-4	13	-39	-38	0	26
21. $(x, Q^2)_{elec}$	27	-11	-31	11	25	-7	-34

Table 6.4: *Relative systematic variation of F_2 in percents for the low Q^2 bins.*

In tables 6.4 and 6.5 we present the results of these systematic checks. In these tables we list the values for the 'nominal' F_2's with the statistical error in percents. The percentage variation of this F_2 for each systematic check is listed thereafter. The results are also displayed in figure 6.5. In this figure each histogram represents a bin in which we draw the extracted structure function with its statistical error for 17 systematic checks.

In these variations of the analysis we did not perform the photoproduction background subtraction and energy scale correction, and we did not consider the modification of the cross section due to the longitudinal structure function. We also did not extract the structure function iteratively. The purpose of these checks is to investigate the maximum size of these variations and the sensitivity of F_2 to the various selection cuts and kinematic reconstruction uncertainties. We briefly discuss the effects of these variations.

Changing the cut value on δ (2 & 3) has appreciable effects on F_2. As discussed before at length, this is due to photoproduction events and the energy scale mismatch between the Monte Carlo and data events. Without correction, F_2 varies

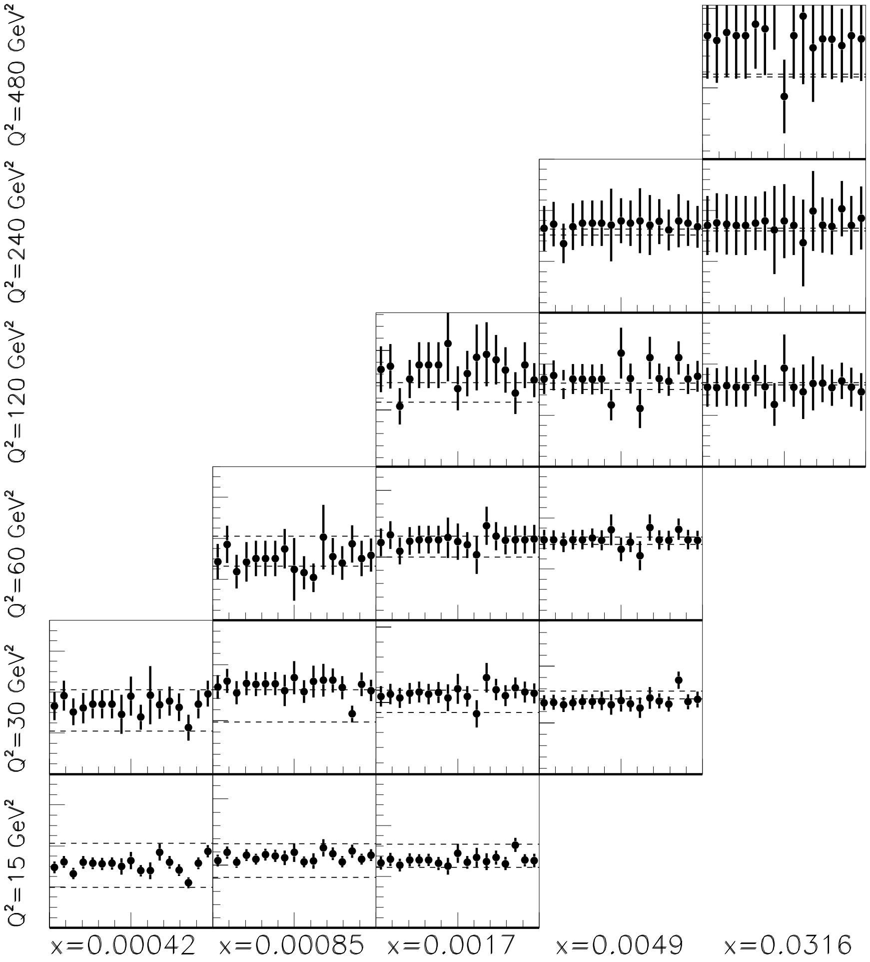


Figure 6.5: *Systematics study on $F_2(x,Q^2)$ in which each histogram represents a bin in the (x,Q^2) phase space. We plot the F_2 with the statistical error only under different conditions. Variations of the points indicate the instability of the extraction of F_2. This figure is the pictorial version of the tables 6.4 and 6.5.*
We plot in each bin the MRSD0′ (lower dashed line) and the MRSD−′ (upper dashed line) parametrizations. The normalization of the histograms is taken with respect to the MRSD−′ parametrization.

Q^2 (GeV^2)	60			120			240		480	1000
x ($\cdot 10^{-4}$)	8.5	17	49	17	49	316	49	316	316	316
1. Nominal	1.22	1.22	0.83	1.66	0.96	0.51	1.01	0.57	0.82	0.56
Stat error	30	17	12	23	17	22	27	32	33	34
2. $\delta > 33$	22	5	0	-1	3	0	-1	3	-4	-1
3. $\delta > 40$	-20	-13	-3	-40	-6	1	-21	2	2	-9
4. $E_e > 7$	-1	-2	0	-14	0	0	-4	0	0	0
5. No box cut	0	0	0	0	0	0	0	0	0	0
6. $y_{JB} > 0.03$	0	0	2	0	0	10	0	2	8	0
7. $y_{JB} > 0.05$	0	0	-1	0	0	-	0	5	5	0
8. EF2 finder	6	-1	0	-12	4	-6	-12	8	-7	-29
9. x-pos > 0	11	2	12	27	-27	-20	-2	-5	51	28
10. x-pos < 0	-13	-2	-12	-27	26	22	2	5	-47	-27
11. $0.9 \cdot E_e$	1	2	1	3	1	0	0	0	4	2
12. E_h	13	4	3	13	0	-8	9	4	10	7
13. $\mathbf{z} > 0.$	-32	-17	-19	13	-30	-6	1	-18	14	27
14. $\mathbf{z} < 0.$	37	15	14	7	21	5	-1	15	-9	-18
15. $\mathbf{z} + 3\ cm$	-1	1-	4	13	4	2	4	2	2	2
16. $\mathbf{z} - 3\ cm$	3	-4	-6	-3	-2	-2	0	-2	-4	-2
17. MTB1 par	11	7	1	11	4	5	11	2	-1	0
18. MTB2 par	0	2	0	1	1	0	1	0	-2	-1
19. MRS$D0'$ par	6	3	1	5	3	0	9	2	1	1
20.$(x, Q^2)_{y\theta}$	16	-9	8	-41	13	7	4	15	-17	-10
21. $(x, Q^2)_{elec}$	36	-2	12	-11	-15	17	1	7	-1	-29

Table 6.5: *Relative systematic variation of $F_2(x, Q^2)$ in percents for the high Q^2 bins.*

predominantly at high values of y. Apart from a bin with large statistical fluctuation, the deviations of F_2 including the corrections from the fits become much smaller. Also requiring a minimum electron energy $E_e > 7$ GeV (4) affects F_2 predominantly at high values of y. It is an alternative way to observe the effect of photoproduction events, since they are characterized by low -and wrongly- identified electron energies. This variation of F_2 is correlated, although smaller, to the variation of F_2 due to $\delta > 33$ GeV.

Discarding the box cut (5) has a small effect in the bins with lowest Q^2. As Q^2 increases, the electron angles θ become smaller and fall outside the fiducial area.

The effect of calorimeter noise is most pronounced in the determination of the kinematic variables at low values of y. By changing the requirement on y_{JB} (6 & 7), we investigate the effect of calorimeter noise on F_2. The variation of F_2 is small and, as expected, F_2 varies only at low values of y_{DA}.

Using the alternative electron finder EF2 (8) is also a measure for the photoproduction background. In chapter 4 we argued that this algorithm picks up more photoproduction events, and hence the F_2 increases at high values of y. The differences in F_2 due to the different electron finders reach 19% in the low Q^2 bins.

We also divided the sample into parts. First we considered the right and left side of the calorimeter separately by selecting events with electron impact position $\mathbf{x} > 0\ cm$ and $\mathbf{x} < 0\ cm$ respectively (9 & 10), to investigate asymmetries of the detector. The effect of these selections on F_2 is rather large, but this is mainly due to statistical fluctuations. We also scaled the Monte Carlo electron energy by 0.9 (11), without changing the data electron energy. Now the variation of F_2 did not exceed 4%. Since the scale problem for hadrons was most pronounced at high values of γ, we introduced a function, which depends on the polar angle (12), to scale each calorimeter cell that did not belong to the electron. We chose the function of the form $1.1 - 0.001\theta_H$, θ_H the polar angle of the cell in degrees. Hence calorimeter cells in the rear direction were scaled down by a factor 0.92. In this way we implicitly distorted the value for the Monte Carlo hadron angle γ, as the result of the scale problem for the hadron side. Its effect on F_2 reaches 21%.

We also divided the sample with positive and negative event vertices (13 & 14). Also here we encountered the problem that the statistical fluctuations are large. Hence we further investigated the uncertainty on the event vertices by changing the $\mathbf{z}$ vertex position of the Monte Carlo events by $\pm 3\ cm$ without changing the data vertices (15 & 16; note that the resolution of the event vertex from the tracking devices is $4.5\ cm$ in $\mathbf{z}$). The effect on F_2 is typically larger than a percent, and reaches 13%.

With checks (17, 18 & 19) we determined the size of the variation of F_2 due to different Monte Carlo input distributions. This is done by reweighting the Monte Carlo events to the extreme MTB1 and MTB2 parametrizations and the MRS$D0'$ parametrization. The effect is largest when the low MTB1 parametrization is used as input Monte Carlo distribution, and reaches 11%.

For completeness we also measured F_2 with the $y\theta$ (20) and the electron method (21) for reconstruction of the kinematic variables. For the electron method we included the requirement $y_{elec} > 0.05$ to reject the badly reconstructed events. Since we applied no energy scale corrections, these methods show large deviations of F_2 with respect to the double angle method, and can not be trusted. However, once the energy problems have been controlled, these methods will provide a nice way to investigate the systematic uncertainty on F_2, for the electron method especially at high values of y and for the $y\theta$ method at low values of y.

6.5.2 Determination of the systematic error

In order to obtain the total systematic uncertainty on the measurement of F_2 in each bin, we classified the largest possible uncertainties in five categories. These are: the understanding of the photoproduction background, electron and hadron energy scale, the electron and hadron flow position, the CAL noise, and the effects of QED radiation and different input structure function parametrizations.

For each category we quote a systematic uncertainty separately. The categories are chosen such that the uncertainties are independent. To obtain the total system-

Q^2	x	back-ground	energy scale	position	CAL noise	SF input	total
15	0.00042	1	9	3	0	8	12
	0.00085	3	2	3	1	7	8
	0.00170	1	3	1	6	5	8
30	0.00042	4	21	13	0	4	25
	0.00085	1	8	5	0	5	11
	0.00170	1	2	7	0	4	8
	0.00490	0	5	2	4	3	7
60	0.00085	16	13	3	0	6	22
	0.00170	4	6	4	0	3	9
	0.00490	0	6	6	2	1	9
120	0.00170	3	43	13	0	5	45
	0.00490	3	1	4	0	3	6
	0.03162	0	8	2	10	0	13
240	0.00490	4	22	4	0	9	24
	0.03162	3	4	2	5	2	8
480	0.03162	4	10	4	8	1	14
1000	0.03162	1	10	2	0	1	10

Table 6.6: *The estimated systematic error in percent for the different categories. The final systematic error on the determination of F_2 is listed in the last column.*

atic error for each bin, we subsequently added the errors of the five categories in quadrature.

Photoproduction In the final analysis of the structure function we subtract the photoproduction background as determined from the fits (see figure 6.4). We estimated the uncertainty on the corrected F_2 by shifting the value on δ to $\delta > 33$ GeV. We presented the results as the ratio $R_c(33)$ in table 6.3.

Energy scale We also corrected for the energy scale mismatch between the data and the Monte Carlo with the fits, and presented the result as the ratio $R_c(40)$ of table 6.3. For the total uncertainty of the energy scale problem we took a conservative attitude and quote the maximum deviation of F_2 for this ratio $R_c(40)$, the electron scale (11) and the hadron scale (12) checks.

Electron and hadron position The uncertainties in the angles of the final state electron and hadron flow were obtained by shifting the vertex by 3 *cm* (13 & 14) and discarding the box cut (5). Uncertainty on the hadron angle γ was parametrized also by the hadron scale function of the previous category (12). Since we cannot make a correction for this effect, we quote the systematic uncertainty to be the largest percentage deviation of these checks from the nominal extraction of F_2 (third column of table 6.6).

Calorimeter noise The effect of the calorimeter cell noise was investigated by varying the cut value on y_{JB} (6 & 7). Also here we list the largest deviation of these checks with the nominal extraction as the systematic uncertainty on F_2 due to calorimeter noise (fourth column of table 6.6).

Structure function input Via the iteration procedure we correct for the Monte Carlo input structure function parametrization to obtain the correction factors. We quote for the systematic uncertainty due to the unknown input structure function the variation on F_2 if we take the parametrization MRS$D0'$ (19) as input and do not perform the iteration. The measured F_2 data points fall between the MRS$D0'$ and MRS$D-'$ parametrizations (see table 6.7) and therefore this uncertainty is estimated conservatively.

The overall systematic uncertainty on the final determination of F_2 is listed in the last column of table 6.6.

6.6 Final results on F_2

The final results for the F_2 measurement are presented in table 6.7. In the table we include the number of events that are observed in the bin, $dN(data)$, and the number of events in the bin after photoproduction background subtraction and and energy scale mismatch correction, $dN(corr)$. We list the correction for the longitudinal structure function, $K(\Delta F_L)$, in the fifth column. The structure function F_2, obtained from the ZEUS 1992 fall data, is listed thereafter, with the statistical and systematic uncertainties quoted separately. The overall uncertainty due to the luminosity measurement, 5%, is not included.

We compare the measured structure function in table 6.7 with the parametrizations according to MRS$D0'$ and MRS$D-'$, in the next two columns.

Additionally, we selected the events with a large rapidity gap (see section 5.6), with the requirement $\eta_{max} < 1.5$. The number of these 'diffractive' events in the bins are listed in the column dN_{diff}. We assumed identical smearing of the kinematic variables for these events as well as identical fractional background in the bins. Hence we obtained the contribution to F_2 from the diffractive events by scaling the measured F_2 with the fraction of diffractive events to $dN(data)$. This contribution to the structure function, with the statistical error, is listed in the last column of table 6.7.

In figure 6.6 we present the measured structure function F_2 with the MRS$D0'$ and MRS$D-'$ parametrizations, as a function of x in bins of Q^2. In the figure the statistical and systematic error are added in quadrature but the normalization

Q^2 (GeV^2)	x	dN $(data)$	dN $(corr)$	$K(\Delta F_L)$	$F_2 \pm$ stat $\pm$ sys (ZEUS 1992)	F_2 MRS$D0'$	F_2 MRS$D-'$	dN $_{diff}$	$F_2(diffract)$ $\pm$stat
15	0.00042	165	154	1.032	$1.22 \pm 0.12 \pm 0.15$	0.89	1.79	17	0.13 ± 0.04
	0.00085	204	206	1.006	$1.09 \pm 0.09 \pm 0.09$	0.81	1.32	16	0.09 ± 0.03
	0.00170	175	178	1.001	$0.80 \pm 0.07 \pm 0.07$	0.73	1.00	7	0.03 ± 0.01
30	0.00042	42	39	1.119	$1.75 \pm 0.34 \pm 0.44$	1.09	2.10	6	0.25 ± 0.10
	0.00085	119	121	1.027	$1.66 \pm 0.21 \pm 0.18$	0.97	1.54	11	0.15 ± 0.05
	0.00170	119	121	1.005	$1.10 \pm 0.14 \pm 0.09$	0.86	1.15	5	0.05 ± 0.02
	0.00490	125	126	1.001	$0.66 \pm 0.07 \pm 0.05$	0.71	0.78	3	0.02 ± 0.01
60	0.00085	27	27	1.101	$1.34 \pm 0.36 \pm 0.29$	1.14	1.76	6	0.30 ± 0.12
	0.00170	48	47	1.023	$1.24 \pm 0.21 \pm 0.11$	0.99	1.31	2	0.05 ± 0.04
	0.00490	106	107	1.003	$0.84 \pm 0.10 \pm 0.07$	0.79	0.86	3	0.02 ± 0.01
120	0.00170	22	22	1.086	$1.79 \pm 0.38 \pm 0.81$	1.13	1.46	3	0.24 ± 0.14
	0.00490	50	50	1.011	$0.98 \pm 0.16 \pm 0.06$	0.87	0.94	0	-
	0.03162	31	31	1.000	$0.51 \pm 0.11 \pm 0.07$	0.53	0.52	0	-
240	0.00490	19	19	1.040	$1.05 \pm 0.27 \pm 0.26$	0.95	1.02	0	-
	0.03162	24	24	1.001	$0.57 \pm 0.18 \pm 0.04$	0.55	0.54	0	-
480	0.03162	18	18	1.004	$0.82 \pm 0.27 \pm 0.12$	0.57	0.55	0	-
1000	0.03162	8	8	1.016	$0.57 \pm 0.19 \pm 0.06$	0.59	0.57	0	-

Table 6.7: *Final result for the F_2 structure function measurement extracted from the fall 1992 ZEUS data.*

uncertainty is not included. The same for figure 6.7, where we plot F_2 as function of x for the diffractive events.

In figure 6.8 we plot F_2 as function of Q^2. Note that the F_2 values are multiplied by different factors (shown in parenthesis) for the different x bins. For comparison we included the NMC data points [18].

6.7 Summary

In this chapter we presented the measurement of the structure function F_2 from the 'fall 1992' ZEUS data sample. In this period the integrated luminosity was measured to be 24.7 nb^{-1}.

The results on the measurement of F_2 were preceded by a discussion on the longitudinal structure function F_L, for which we introduced the correction factor $K(\Delta F_L)$. We then defined rectangular bins in x and Q^2 and introduced various types of correction factors. These factors were designed such that the effects of event selection, event migration and radiative effects were disentangled. The values for these correction factors were obtained with the Monte Carlo sample. A requirement on the total acceptance, Acc(tot), was used to select bins in which the correction factors are small and uniform. Seventeen bins remained, ranging in x from 0.00042 to 0.03162 and in Q^2 from 15 GeV^2 to 1000 GeV^2. The remaining photo-production background in the bins was then estimated with a fit procedure on the variable δ. The same fits were used to correct for the energy scale problem, which we encountered in chapter 5. We discussed the iterative method to obtain F_2.

For the determination of the systematic uncertainty on the extracted F_2 we modified the analysis many times to investigate the stability of F_2. We then defined five categories and quoted the systematic uncertainties for each of those. The total systematic uncertainty was obtained by adding the uncertainty of each category in quadrature.

In table 6.7 we presented the final results of the analysis. Graphical representations of this table are shown in figures 6.6, 6.7 and 6.8.

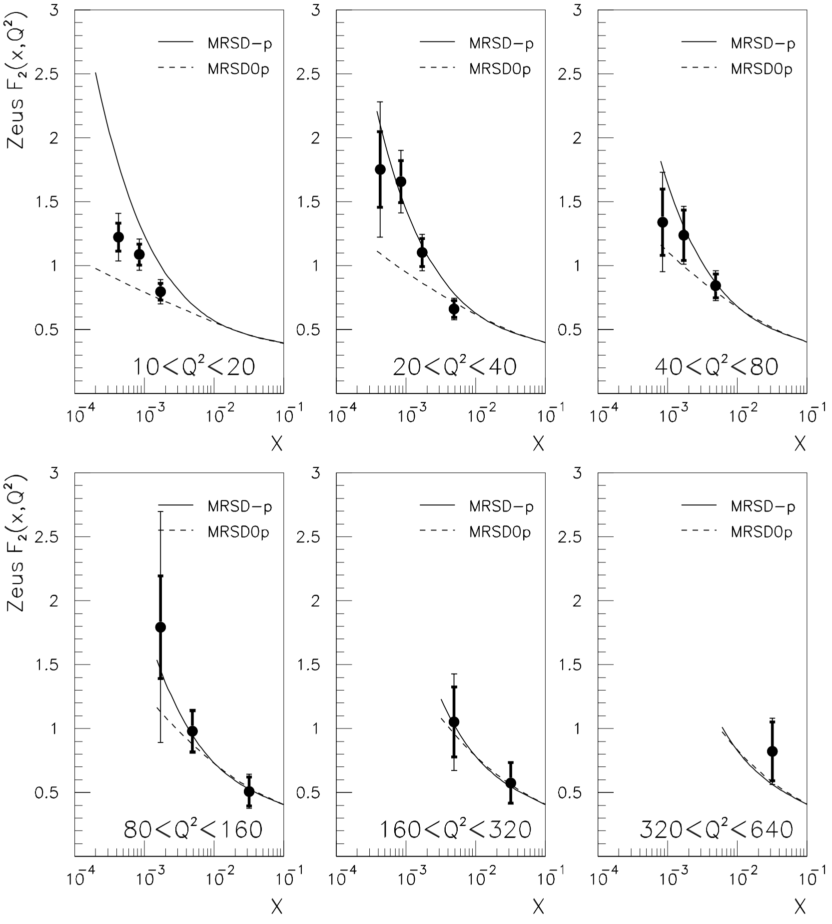


Figure 6.6: *The measured F_2 as a function of x in different Q^2 bins. The statistical error is represented by the inner error bar. The systematic error, represented by the outer error bar, has been added in quadrature. The overall normalization uncertainty of 5% is not shown.*

The curves are the F_2 parametrization of MRSD$-'$ (full line) and MRSD0$'$ (dashed line).

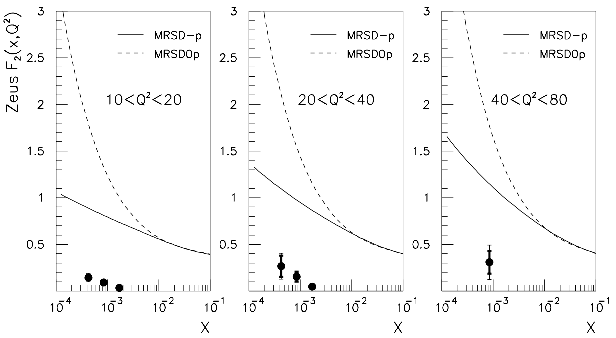


Figure 6.7: *The measured F_2 as function of x for the diffractive events.*

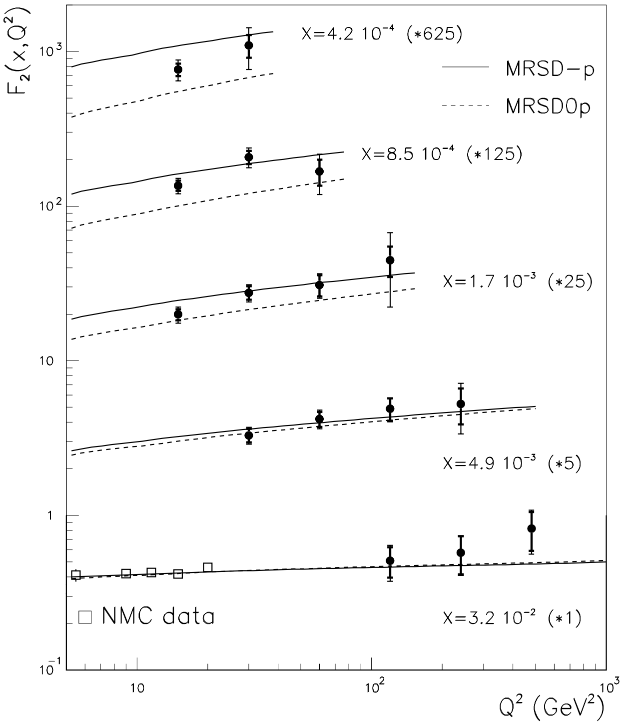


Figure 6.8: *The measured F_2 as function of Q^2. To separate the points of different x bins, the F_2 for each bin is multiplied by a factor, quoted between parenthesis. The MRSD$-'$ and MRSD$0'$ parametrizations are drawn. The open squares are the NMC data points.*

Chapter 7

Implications of the F_2 measurement

With the new measurement of F_2, obtained with the ZEUS data, a previously unexplored region in phase space is entered. Although the statistical and systematic errors on the F_2 are large, the measurement irrefutably contains new information on the structure of the proton. In this chapter we compare the F_2 of the ZEUS data with the results of the H1 collaboration. We briefly discuss the comparison with various parametrizations of F_2.

To appraise the new measurement in more detail, we develop a new parametrization of F_2. We obtained this parametrization from leading order QCD evolution of parton distributions, fitted to the NMC data and the new ZEUS results. The results of the fit show that the gluon and sea quark distributions behave singularly as x tends towards small values.

7.1 The F_2 measurements compared

We presented the result of the F_2 measurement for the 'fall 1992' ZEUS data in table 6.7 of the previous chapter. Our measurement is comparable with the results of the H1 collaboration [71]. In figure 7.1 we plot our results and the results of the H1 collaboration for F_2 as function of x in the two lowest bins of Q^2. The H1 results are somewhat higher, but with the present statistics this difference is not significant. In the same figure we show the previously published values of F_2 from the NMC collaboration.

It is very interesting to note that the F_2 is rising at low values of x. Independent of any parametrization of F_2, the new DIS results from the HERA collider show an increase of the structure function F_2 at low values of x for fixed Q^2. This increase of the structure function is not only visible when one compares our measurement with results of previous fixed target experiments, but is also visible from the ZEUS

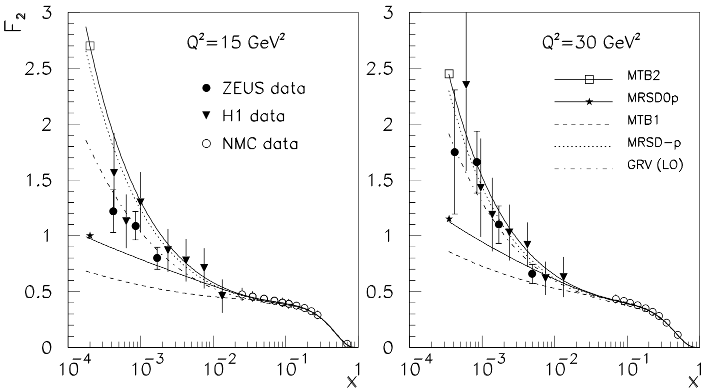


Figure 7.1: *Comparison of the F_2 measurement for the ZEUS, H1 and NMC collaborations. The F_2 is plotted as function of x for the two lowest Q^2 bins. We draw the MRSD0′ (straight line), MRSD−′ (dash-dotted line), MTB1 (dashed line), MTB2 (straight line) parametrizations, obtained with a next-to-leading order QCD fit, and the GRV parametrization in leading order QCD (dotted line).*

results themselves. For example, the F_2 for the lowest x value in the bin $Q^2 = 30$ GeV^2 ($x = 0.00042$) is approximately 2.5 times as high as F_2 for the largest x value ($x = 0.00490$). This rise of F_2 is consistent with the measurements of the H1 collaboration.

Especially for the lowest three Q^2 bins, where the parametrizations diverge, our measurements of F_2 contain new information on the structure of the proton. Already with the present statistics we observe that various parametrizations of the proton are excluded by our result. For example, the prediction according to the MRS$D0'$ parametrization is too low and does not describe the data consistently. The same conclusion holds for the MT$B1$ parametrization, which is even lower. The parametrizations of MRS$D-'$ and MT$B2$ seem to overshoot the data but, especially for $Q^2 > 20$ GeV^2, are in reasonable agreement with the data. The parametrization of GRV is available both for a leading order (LO) and next-to-leading order (NLO) QCD fit. Whereas the other parametrizations are drawn for the NLO QCD fit, in figure 7.1 we draw the GRV parametrization in LO, for reasons that will become clear. It provides a reasonable description of the new data.

As a function of Q^2, the measured data points show the expected scaling violations. The statistics of the data are too low and the scaling violations are too weak to observe possible differences in F_2 as a function of Q^2 between the present data and the parametrizations.

7.2 Leading order QCD analysis

The new ZEUS results provide structure function measurements in a hitherto inaccessible region of low x. We therefore investigate the implication of these measurements on the gluon and sea quark distributions. To this end, we performed a first order QCD fit to the NMC F_2^p data and ZEUS results on the F_2 measurement.

7.2.1 QCD evolution

We performed a LO QCD evolution for six parton distribution functions (PDF's). To improve numerical convergence at low values of x we evolved x times the PDF's.

We determined the evolution of the valence quark distributions, $x(u_V + d_V)$ and xd_V, separately. These are non-singlet distributions for which the evolution in Q^2 is independent of the gluon distribution. Since the singlet strange sea is known to be roughly half $\bar{u}$ and $\bar{d}$ sea [21], we assumed $\bar{s} = \frac{1}{2}\bar{u} = \frac{1}{2}\bar{d}$, and we did not allow the $\bar{d}$ to be different from $\bar{u}$ (hence the Gottfried-Sum rule equals $\frac{1}{3}$ by construction).

For the $x(u_V + d_V)$, xd_V, xS and xg parton distribution functions at $Q^2 = Q_0^2 =$ 4 GeV^2 we used the functional forms of the MRS parametrizations:

$$
\begin{aligned}
x\left[u_V(x) + d_V(x)\right] &= A_{ud} x^{\eta_1} (1-x)^{\eta_2} (1 + \epsilon_{ud}\sqrt{x} + \gamma_{ud} x) && (7.1)\\
xd_V(x) &= A_d x^{\eta_3} (1-x)^{\eta_4} (1 + \epsilon_d\sqrt{x} + \gamma_d x) && (7.2)\\
xS(x) &= A_S x^{\delta_G} (1-x)^{\eta_S} (1 + \epsilon_S\sqrt{x} + \gamma_S x) && (7.3)\\
xG(x) &= A_g x^{\delta_G} (1-x)^{\eta_G} (1 + \gamma_g x) && (7.4)
\end{aligned}
$$

The heavy quarks, charm and bottom, are set to zero at $Q^2 = 4$ GeV^2 and are generated dynamically. Note that we used the same leading $x \to 0$ behaviour for the sea quarks and the gluon distribution at $Q^2 = Q_0^2$, parametrized by δ_G. This reflects the assumption that for low values of x the sea is generated by the gluon distribution already at $Q^2 = Q_0^2$.

In first order the Gribov-Lipatov-Altarelli-Parisi equations for the above defined PDF's can be written as (with $t = \ln(Q^2/\Lambda^2)$):

$$
\begin{aligned}
d_t\left\{x(u_V + d_V), xd_V\right\} &= \frac{\alpha_s}{2\pi}\left[P_{qq} \otimes \left\{x(u_V + d_V), xd_V\right\}\right] && (7.5)\\
d_t xS &= \frac{\alpha_s}{2\pi}\left[P_{qq} \otimes xS + 6P_{qG} \otimes xG\right] &&\\
d_t\left\{xc, xb\right\} &= \frac{\alpha_s}{2\pi}\left[P^h_{qq}(Q^2) \otimes \left\{xc, xb\right\} + 2P^h_{qG}(Q^2) \otimes xG\right] &&\\
d_t xG &= \frac{\alpha_s}{2\pi}\left[P_{GG} \otimes xg + P_{Gq} \otimes x(u_V + d_V + S + c + b)\right] &&
\end{aligned}
$$

with $d_t = d/dt$ and the convolution integral '$\otimes$' defined in equation 1.17. For the splitting functions P_{qq}, P_{qG}, P_{Gq} and P_{GG} we used the first order expressions, with the 'plus' prescription to circumvent the divergencies at $z = 1$ in the convolution integrals. With $y = x/z$ the evolution for the non-singlet distributions $\Delta = \{x(u_V + d_V), xd_V\}$, singlet $\Sigma = xS$ and the gluon $g = xG$ become

$$
d_t\Delta(x,t) = \frac{\alpha_s}{2\pi}\left\{\int_x^1 dz \tfrac{4}{3}\left(\tfrac{(1+z^2)\Delta(y,t) - 2\Delta(x,t)}{1-z}\right) + \left[2 + \tfrac{8}{3}\ln(1-x)\right]\Delta(x,t)\right\}
$$

$$\begin{aligned}
d_t\Sigma(x,t) &= \frac{\alpha_s}{2\pi}\Big\{\int_x^1 dz \tfrac{4}{3}\left(\tfrac{(1+z^2)\Sigma(y,t)-2\Sigma(x,t)}{1-z} + 3\left[z^2+(1-z)^2\right]g(y,t)\right) \\
&\quad + \left[2+\tfrac{8}{3}\ln(1-x)\right]\Sigma(x,t)\Big\} \\
d_tg(x,t) &= \frac{\alpha_s}{2\pi}\Big\{\int_x^1 dz\left(\tfrac{6[zg(y,t)-g(x,t)]}{1-z} + \tfrac{6(1-z)(1+z^2)g(y,t)}{z} + \tfrac{4}{3}\tfrac{1+(1-z)^2}{z}\cdot\right. \\
&\quad \left.\textstyle\sum_f y\left[q_f(y,t)+\bar{q}_f(y,t)\right]\right) + \left[\tfrac{11}{2}-\tfrac{n_f}{3}+6\ln(1-x)\right]g(x,t)\Big\}
\end{aligned}$$

with n_f the number of flavours that participate in the evolution at the particular value for t. The GLAP-equations for the heavy quarks charm and bottom, described by the heavy quark splitting functions $P^h_{qq}(Q^2)$ and $P^h_{qG}(Q^2)$, are given according to the prescription of Glück, Hoffman and Reya [73]. This prescription takes into account the quark mass thresholds such that the heavy quarks participate in the evolution only for positive β^2, defined as

$$\beta^2 = 1 - 4zM_q^2/Q^2(1-z) \tag{7.6}$$

Effectively this means that only for $W^2 > (2M_q)^2$ are these quarks generated dynamically.

For the running coupling constant of the strong interactions, $\alpha_s(Q^2)$, we adopt the first order QCD expression with a prescription to achieve a smooth crossing of the bottom quark mass threshold:

$$\frac{1}{\alpha_s(Q^2)} = \frac{25}{12\pi}\ln(Q^2/\Lambda_4^2) - \frac{1}{6\pi}\theta(Q^2-16M_b^2)\ln(Q^2/16M_b^2) \tag{7.7}$$

with $\theta(x)$ the step-function. The value for Λ_4 is given in first order with the number of flavours $n_f = 4$.

With the PDF's evolved to all values of Q^2, the electromagnetic part of the structure function $F_2^{\gamma\gamma}$ becomes:

$$F_2^{evo} = \sum_f e_f^2\left[xq_f + x\bar{q}_f\right] = \frac{4}{9}x(u_V+d_V) - \frac{1}{3}xd_V + \frac{11}{45}xS + \frac{4}{9}xc + \frac{1}{9}xb \tag{7.8}$$

We compare this F_2^{evo} with the experimental ZEUS and NMC data.

We used a N-point grid in x with power P to perform the integrations in x. To take into account the measurements at small values of x, we start the integration at $x_{min} = 10^{-4}$. The grid at point x_i is defined as

$$x_i = x_{min} + (1-x_{min})\cdot\left(\frac{i-1}{N-1}\right)^P \tag{7.9}$$

We typically used $N = 40$, $P = 3$ to reach sufficient accuracy at both very low and high values of x. We checked that our results are stable against these numbers. Simpson integration of a function over this grid, used to perform the convolution integrals, is accurate up to $\mathcal{O}(\Delta h^3)$ with $\Delta h = 1/(N-1)$.

We performed the integration in t on a linear grid with the simple Euler integration rule.

$$q(t+\Delta t, x) = q(t,x) + \Delta t\, d_t q(t,x) \quad ; \quad \Delta t = \frac{1}{N_t - 1}\ln(Q^2_{max}/Q^2_0) \tag{7.10}$$

We typically used $\Delta t = 0.04$. We calculated the distributions at all gridpoints (t_i, x_i), and used a linear interpolation of the distributions on the grid to obtain the PDF's and F_2^{evo} for continuous points (x, Q^2).

We made several consistency checks for this evolution. For example, we checked that the number of valence quarks is independent of Q^2. We also evolved the PDF's from $Q^2 = Q_0^2$ according to the LO EHLQ parametrizations [56], and found no significant difference between our results and the literature.

7.2.2 Fit to the data

We compared the computed structure function F_2^{evo} with the measurements of the structure function F_2 for the ZEUS and NMC data, by defining the χ^2 function as

$$\chi^2 = \sum_i \frac{\left(F_2^{evo} - F_{2\,i}^{data}\right)^2}{\sigma_i^2(stat) + \sigma_i^2(sys)} \tag{7.11}$$

We used the MINUIT [72] program with the MIGRAD option to find the minimum χ^2 by varying the parameters at $Q^2 = Q_0^2$ GeV^2.

In principle all parameters of equations 7.1–7.4, together with the value of Λ_4, can be varied in the minimization of χ^2. But there are a number of constraints on the PDF's at Q_0^2 that restrict this freedom.

- The normalizations A_{ud} and A_d are fixed by requiring the correct number of valence quarks in the proton:

$$\int_0^1 dx\,(u_V + d_V) = 3 \quad ; \quad \int_0^1 dx\, d_V = 1 \tag{7.12}$$

- The normalization A_G is fixed by the requirement that the sum of momentum fractions carried by the PDF's equals one:

$$\int_0^1 dx\, x(u_V + d_V + S + c + b) + xG = 1 \tag{7.13}$$

- The PDF's should be positive in the interval $x \in \{0, 1\}$, which we implemented by requiring that either both ϵ and γ are positive or

$$\epsilon < 0 \rightarrow \gamma > \frac{\epsilon^2}{4} \quad ; \quad \gamma < 0 \rightarrow \epsilon > -(1+\gamma) \tag{7.14}$$

  for the $(u_V + d_V)$, d_V and S distributions. The parameter γ_G is required to be $\gamma_G > -1$.

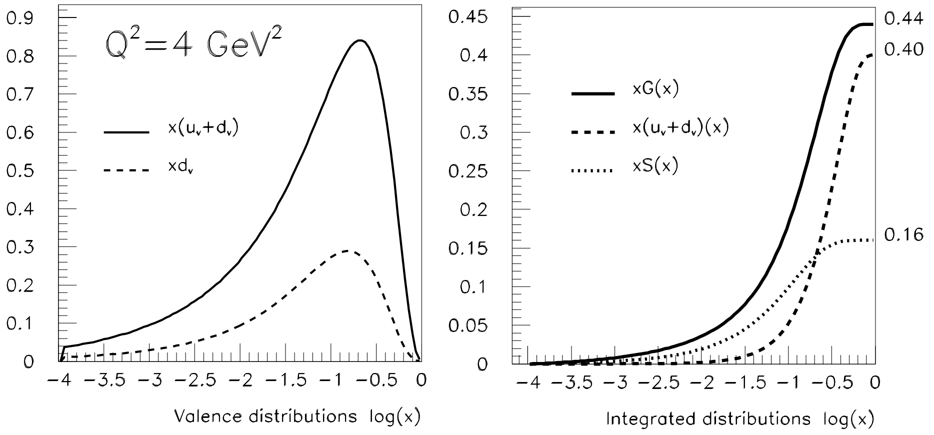


Figure 7.2: ***On the left plot we show the valence distributions at $Q^2 = Q_0^2 = 4$ GeV^2. On the right plot we show the integrated distributions from zero as a function of the integration maximum, for x times the gluon, valence and sea distributions. The values at $x = 1$ correspond to the total momentum carried by the distributions at $Q^2 = Q_0^2$.***

In principle we are left with 16 free parameters in the minimization of χ^2. These are the QCD scale Λ_4, eight parameters that describe the singlet distributions, four parameters for the sea distribution, and three parameters for the gluon.

The value for Λ_4 is fixed in the present analysis to $\Lambda_4 = 230$ MeV. Although the slopes of F_2 as a function of Q^2 from the ZEUS data points show the expected scaling violations, they are not measured very accurately with the present statistics, and hence contain no new information on the value of Λ_4. The sensitivity to Λ_4, without correlation to the gluon distribution, is for the evolution in Q^2 mainly at high values of x. By fixing the value of Λ_4, we remove the correlation of Λ_4 with the gluon distribution at low values of x.

We did not include hadron cross-section data, other than the NMC results. In the left plot of figure 7.2 we draw the valence $x(u_V + d_V)$ and xd_V distributions of the MRS parametrization in the DIS scheme. From this plot we conclude that the contribution of the valence distributions to the structure function F_2 is small in the region of x covered by the new ZEUS data points. Hence the ZEUS data will provide no new information on the valence distributions. Therefore we took the MRS results for the valence distributions $x(u_V + d_V)$ and xd_V at Q_0^2. This means

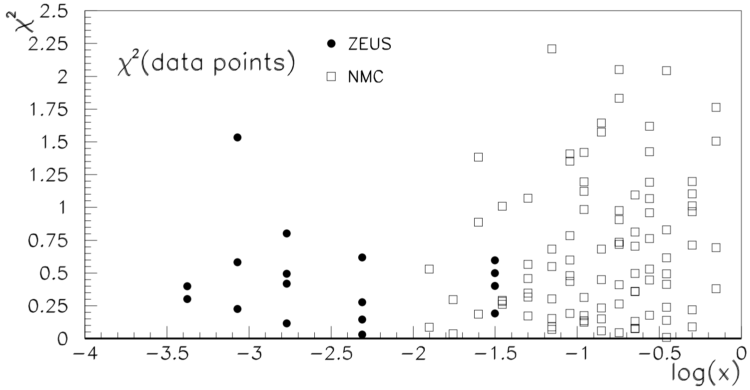


Figure 7.3: *The* χ^2 *for each data point, plotted as function of* x*. The ZEUS data with black circles and NMC data with open squares.*

that we fixed all eight parameters of the valence distributions in the fit procedure.

It was found previously [24] that the total sea distribution at $Q^2 = Q_0^2 = 4$ GeV^2 carries approximately 16% of the proton momentum. In the right plot of figure 7.2 we show the integrated momentum fraction carried by the gluon, the valence quarks and the sea quarks at $Q^2 = Q_0^2$, as a function of the upper boundary of the integration. Hence at a particular value for x_0, the plot shows the momentum fraction carried by the distributions with $x < x_0$. From this figure we conclude that the sea carries only a few percent of the proton momentum fraction for values of $x_0 < 0.01$. Since this fraction is small, the ZEUS data points contain almost no information on the normalization of the sea distribution. Therefore we fixed the parameter A_s in the present analysis, such that $\int_0^1 dx\; xS(x, Q_0^2) = 0.16$.

Finally, we note that the parameter η_G only affects the gluon distribution at high values of x. Since again this is not addressed by the present data, we fixed this parameters in the fit to the value of the MRS distributions as well.

This leaves the five parameters δ_G, γ_G, γ_s, ϵ_S and η_s as the free parameters in the fit, with constraint 7.14 for the sea distribution. The most important parameter, which describes the gluon and sea distribution at very low values of x, is δ_G.

7.2.3 Results

The fit procedure minimizes the χ^2 for 17 ZEUS and 92 NMC points, all with $Q^2 > Q_0^2 = 4$ GeV^2. The question whether the diffractive type of events should be included in the fit procedure of the PDF's is not answered yet. Since a theoretical explanation of the diffractive type of events is presently lacking, we include the contribution of this type of events to F_2 and fit the total inclusive F_2, as listed in table 6.7.

Typically the minimization of MINUIT converged in approximately 500 iterations. Note that for each iteration the GLAP equations over the full phase space in (x, Q^2) had to be calculated. After the fit procedure was completed, it was found that

$$\begin{array}{lll} \text{NMC: 92 data points} & \sum \chi^2 = 70.7 & <\chi^2>= 0.77 \\ \text{ZEUS: 17 data points} & \sum \chi^2 = 5.8 & <\chi^2>= 0.34 \end{array} \tag{7.15}$$

The values for $<\chi^2>$ for the NMC data is somewhat low because we included the systematic error in quadrature. This is also true for the ZEUS data, and the fit procedure indicates that the systematic error for the ZEUS data may be somewhat overestimated. The systematic error will decrease in future measurements of F_2 at ZEUS, when higher statistics are available and systematic effects are better understood. In figure 7.3 we plot the χ^2 value for each data point as function of x. We observe no structure in the value of χ^2 as a function of x, which indicates that the fit matches the data uniformly as a function of x.

	A_{ud}	η_1	η_2	ϵ_{ud}	γ_{ud}
$x(u_V + d_V)$	1.78	0.43	3.11	0.92	2.35

	A_d	η_3	η_4	ϵ_d	γ_d
xd_V	0.50	0.43	3.50	4.64	-3.25

	A_{ud}	δ_G	η_S	ϵ_S	γ_S
xS	0.11	-0.43*	8.34*	7.68*	13.9*

	A_G	δ_G	η_G		γ_G
xG	0.34	-0.43*	5.10	-	15.0*

Table 7.1: *The parameters after the fit procedure is completed. Note that we fitted only the parameters that are marked with an asterix.*

In table 7.1 we present the parameters at $Q^2 = 4$ GeV^2 that minimize the χ^2 distribution for the ZEUS and NMC data. The parameters that are left free in the fit procedure are marked with an asterix (*). In figure 7.4 we present the comparison of the fit result with the data, both as a function of x and of Q^2. In the figure we compare with the parametrizations of MRS$D0'$, MRS$D-'$ (NLO) and GRV (LO).

We find that the gluon distribution is almost singular, with a value of $\delta_G = -0.43$. We present the gluon distribution for two different values of Q^2 in figure 7.5. The gray band is the maximum variation of the gluon distribution due to the ZEUS luminosity uncertainty of 5%.

To set a scale, we draw in the same plots the gluon distribution according to the MRS$D0'$, MRS$D-'$ and GRV parametrizations. An important remark should be made when these distributions are compared with the results of the fit. We

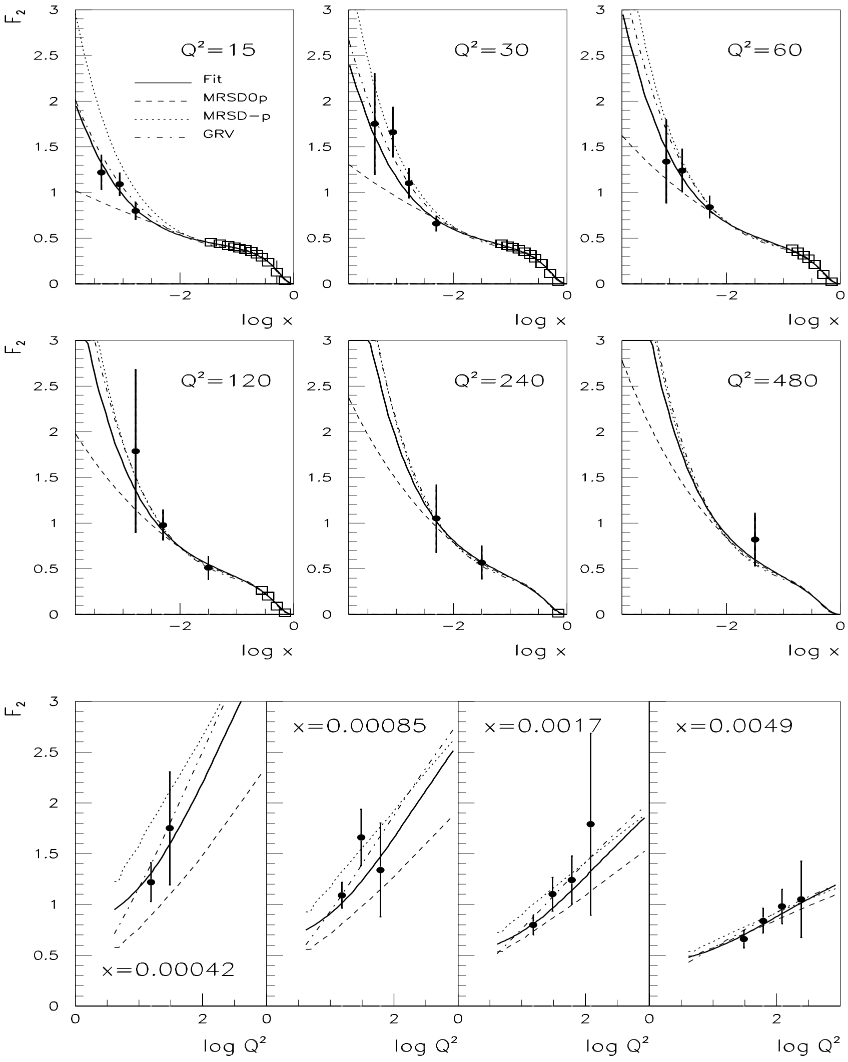


Figure 7.4: *The fitted structure function* $F_2(x, Q^2)$. *The ZEUS (black circles) and NMC (open squares) data points are drawn, together with the NLO MRSD0′, MRSD−′ and LO GRV parametrizations. In the upper six plots we show the* F_2 *as a function of* x *for the various* Q^2 *bins. In the lower four plots we show* F_2 *as function of* Q^2.

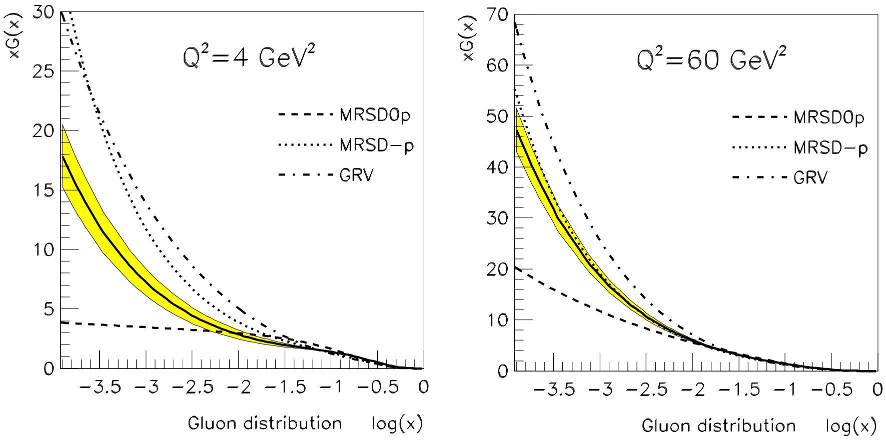


Figure 7.5: *The gluon distributions at two values for Q^2. The gray band represents the variation of the gluon distribution due to the uncertainty in the ZEUS luminosity measurement (±5%). The gluon distributions from the MRS (NLO QCD in DIS scheme) and GRV (LO QCD) parametrizations are drawn also. Note the warning in the text for comparison of the MRS parametrization with the fit results.*

are forced to draw the MRS parametrizations in NLO QCD evolution, whereas the present fit is in LO. In the figure, we draw the NLO distributions in the DIS scheme, since in this scheme the expression for F_2 in terms of the PDF's is equal to the LO expression. But the GLAP evolution equations are not identical for LO and NLO. There is no problem in comparing the fit-results with the LO GRV parametrization.

In figure 7.6 we present the result of the fit for the sea distribution. Here also we indicate the maximum variation of $xS(x)$ due to the uncertainty of the ZEUS luminosity measurement. We make the same warning as before in comparing the fit with the MRS parametrizations.

An important issue is the uncertainty of the parameter that governs the low x behaviour of the gluon and sea distributions, δ_G. The minimum χ^2 is obtained for $\delta_G = -0.43$, but there are large correlations of this parameter with the remaining four fit parameters. In order to investigate the variation of χ^2 as a function of δ_G, we repeated the fit procedure with δ_G set to a range of values from -0.6 to zero. For each value of δ_G we fit the remaining four parameters to estimate the effect of the correlations. In figure 7.7 we present the result of this scan in δ_G. The black circles

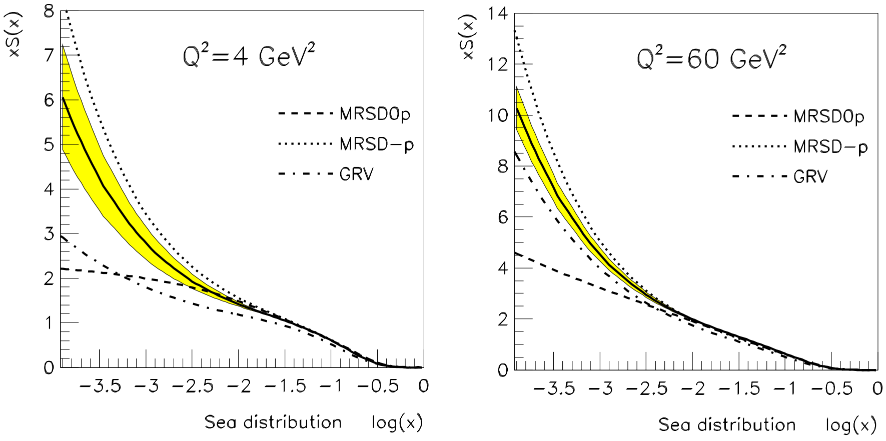


Figure 7.6: *The sea distributions at two values for Q^2, with the gray band representing the luminosity uncertainty. The sea distributions from the MRS (NLO QCD in DIS scheme) and GRV (LO QCD) distributions are included. Note the warning in the text for comparison of these results.*

represent the variation of δ_G with the ZEUS and NMC data, the open squares for the NMC data only. We indicate a variation of 1 in χ^2, which represents the one standard deviation statistical error. It is very interesting to observe that the NMC data alone hardly contains information on δ_G, and this parameter ranges from zero to -0.35 within $\Delta\chi^2 = 1$. The complete range in δ_G from -0.5 to zero is covered within $\Delta\chi^2 = 4$. When the ZEUS data points are included, the one sigma standard deviation on the value of δ_G covers the interval $-0.5 < \delta_G < -0.3$. The variation of δ_G within $\Delta\chi^2 = 4$, including the ZEUS data, is in the range $-0.525 < \delta_G < -0.2$. A future challenge is to limit this allowed range of δ_G with more statistics and better understanding of the systematic errors on F_2.

At a certain value of Q^2 the proton is composed of valence quarks, light and heavy sea quarks and gluons. The value of F_2 is then the charge weighted sum of the valence and sea quark distribution. The change in F_2 has three contributions:

1. The change of the valence quark distribution, which becomes softer due to gluon radiation. This leads to a negative slope contribution to F_2 at high x, and a positive slope contribution at low values of x.

2. The change of the sea quark distribution, which also becomes softer through

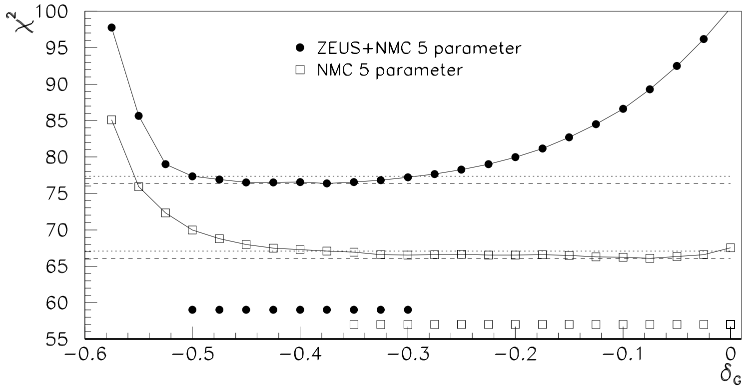


Figure 7.7: *Scan of the parameter δ_g, for the NMC data only (open squares) and NMC+ZEUS data (black circles). For each value of δ_G a fit for the remaining parameters, η_S, γ_S, ϵ_S and γ_G was performed. We draw the lines for the minimum χ^2_{min} (hatched lines) for the two data sets. The dotted line is for $\chi^2_{min}+1$. The values of δ_G that fall inside $\Delta\chi^2 = 1$ are marked above the x-axis.*

gluon radiation. At high x the slope contribution to F_2 is negative. At low x the slope contribution will depend on the exact shape of the sea distribution. If the sea distribution is flat as $x \longrightarrow 0$, the slope contribution of the sea quarks to F_2 vanishes as $x \longrightarrow 0$. A singular sea quark distribution yields a negative slope contribution as $x \longrightarrow 0$.

3. The splitting of gluons into $q\bar{q}$ pairs. As new quarks are created this contribution to the slope of F_2 is always positive. The exact value of the slope depends on the shape of the gluon distribution.

In figure 7.8 we plot the slope of the structure function, $dF_2/d\ln Q^2$, as a function of x for $Q^2 = Q_0^2$. In the left plot we show the slope for the fitted distributions, in the right plot the slope for $\delta_G = 0$, which is not the result of the fit. Figure 7.8 shows that independent of the shape of the gluon, the slope of F_2 is completely dominated by contribution (3), below $x \sim 0.01$. It therefore follows that the slope of F_2 at a fixed value of Q^2 gives a direct measure of the gluon distribution at low values of x. It should also be noted that process (1) and (2), and the splitting $g \longrightarrow gg$, also enhance the gluon distribution, and so process (3) becomes even stronger at higher values of Q^2. This is the prime reason that the Q^2 evolution of F_2 at low values of x is not linear in $\ln Q^2$ as can be observed in figure 7.4.

The second reason that the slope of F_2 increases as Q^2 increases, is the fact that the heavy quarks are produced dynamically for $Q^2 > Q_0^2$, via process (3). In

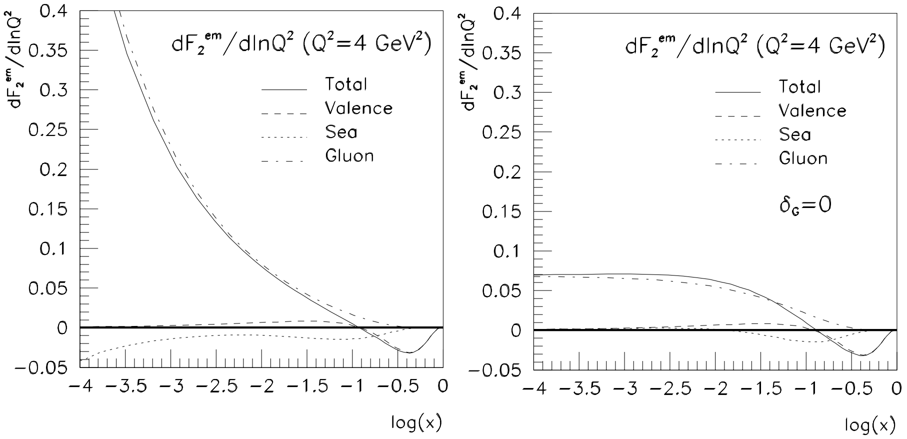


Figure 7.8: *The slope of F_2 in Q^2, at the start value $Q^2 = Q_0^2$, as function of x. The contributions to the slope from the valence (1), sea (2) and gluon (3) distributions are drawn separately. On the left plot we show the slope for the fitted structure function F_2, on the right plot for $\delta_G = 0$, which is not the result of the fit procedure.*

figure 7.9 we present the evolution in Q^2 for the charm and bottom distributions. Note that in the GLAP equations we included for the charm and bottom distributions the heavy quark splitting functions, which realize a smooth mass threshold in Q^2. At very high values of Q^2 the masses of the charm and bottom can be neglected and the evolution in Q^2 becomes comparable to the LO GRV parametrization.

7.3 Discussion

The measured structure function F_2 at ZEUS from the 1992 fall data is steeply rising at low values of x. This behaviour is measured by the H1 collaboration as well.

In order to investigate the implications of this measurement on the parton distribution functions, we performed the evolution of the PDF's in Q^2 according to the GLAP equations, in leading order perturbation theory. The parametrization of the PDF's at the start value Q_0^2 are copied from the MRS distributions. We defined a χ^2 function, which is minimized by varying the parameters at $Q^2 = Q_0^2$. Besides the ZEUS data, the NMC data are included as well to describe the high x behaviour

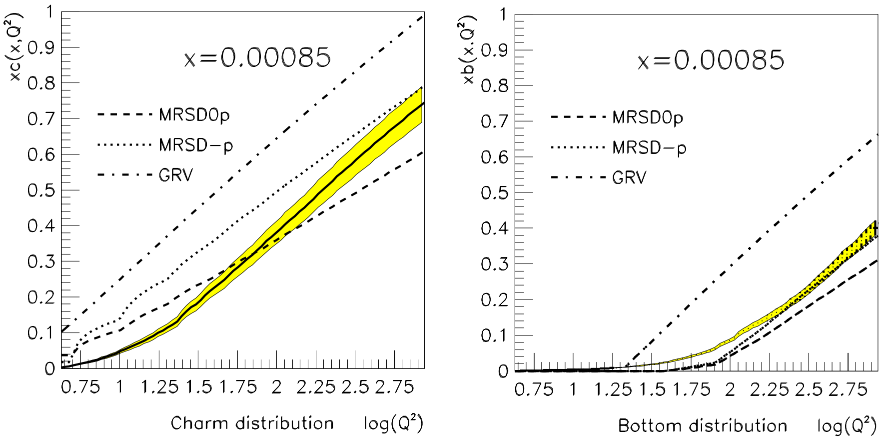


Figure 7.9: ***The charm and bottom distribution at** $x = 0.00085$ **as a function of** Q^2. **Note that we required** $xc(x, Q_0^2) = xb(x, Q_0^2) = 0$ **and included a smooth mass threshold. For high values of** Q^2 **the evolution of the GRV and our parametrizations become equal.***

of F_2. Since we did not include other hadron cross section measurements, we fixed the valence distribution in the fit procedure. We also fixed the normalization of the sea and the high x behaviour of the gluon distribution. In total we were left with five free parameters in the fit procedure.

The result of the fit describes the DIS ZEUS and NMC data satisfactorily. The ZEUS data is fitted with a steeply rising gluon and sea distribution at low values of x. The parameter of particular interest which describes this low x singularity, δ_G, was found to be $-0.5 < \delta_G < -0.3$ within one standard deviation. Although this region is large, the present measurements of F_2 indicate that the gluon distribution is not flat as $x \longrightarrow 0$. A more accurate determination of δ_G is foreseen in the near future, when more ZEUS DIS data become available. A scan of this variable showed that the NMC data alone contains hardly any information on this parameter.

The five parameters show large correlations in the fit procedure. MINUIT minimizes χ^2 for $\delta_G = -0.43$, but the scan of δ_G showed that the one standard deviation statistical error on this parameter is in the range $-0.5 < \delta_G < -0.3$. The variations of the presented gluon and sea distributions in figures 7.5 and 7.6 do not incorporate this large range in δ. The true variation of the gluon and sea distributions with the present data points is much larger due to the parameter correlations. The parametrization of the gluon and sea quarks depend on the specific incorporation of the heavy quarks, charm and bottom, in the GLAP equation. The parametrizations also change when a next-to-leading order QCD evolution is performed. More statistics are needed for an accurate determination of the value of δ_G.

Nevertheless it is interesting to compare our result with the parametrization of GRV in LO (see also section 1.4.3). For low values of x, at $Q^2 = Q_0^2$, the sea distribution of GRV is lower, whereas the gluon density is larger than in our fit. This results in a steeper Q^2 evolution of F_2 with GRV, because the gluon density determines the slope of F_2. Also, the evolution of the charm and bottom distribution starts at $Q^2 = Q_0^2$ in our parametrization at zero, whereas in the GRV parametrization the charm already contributes at $Q^2 = Q_0^2 = 4$ GeV^2, see figure 7.9. Therefore the increase of F_2 as function of Q^2 is smaller in our fit-result as compared to GRV at low values of Q^2. Ultimately, at high values of Q^2, the slope of F_2 for the GRV parametrization and our fit become comparable, as we observe in figures 7.4.

The gluon distribution is the main source of the evolution in Q^2 of the structure function at low x. Its singular behaviour indicates that the gluon density in the proton is enormous at very low values of x. The fit results showed, however, that the present data is described satisfactorily by the GLAP evolution equations. When more statistics are available, recombination of the gluons, which is allowed by QCD but is not included in the GLAP equations, or specific signatures of the BFKL equation may appear to play a role at these high densities. These new and unexplored processes may become visible with the center-of-mass energy that HERA makes available and will inevitably provide exciting times in this field of high energy physics.

Appendix: Reconstruction methods

Energy and momentum conservation on the quark level read in our notation:

Energy conservation	$xP + A = E + F$
Longitudinal momentum conservation	$xP - A = E\cos\theta + F\cos\gamma,$
Transverse momentum conservation	$E\sin\theta = F\sin\gamma$

where we neglected the rest-masses of the quarks and leptons as is done throughout this thesis.
Now we turn to the definition of Q^2 and y (where we write x in terms of Q^2 and y). The most easy and standard way is to define these quantities in terms of E and θ using the final state lepton (where $\mathbf{k}$, $\mathbf{k}'$ are the four-vectors of the initial respectively final state lepton and $\mathbf{P}$ the four-vector of the proton):

$$\begin{aligned} Q^2 &= -(\mathbf{k}-\mathbf{k}')^2 &&= 2AE(1+\cos\theta) \\ y &= \frac{\mathbf{P}\cdot\mathbf{q}}{\mathbf{P}\cdot\mathbf{k}} &&= 1-\frac{E}{2A}(1-\cos\theta) \\ x &= \frac{Q^2}{2\mathbf{P}\cdot\mathbf{q}} &&= \frac{Q^2}{sy} \end{aligned}$$

It is possible to determine (x, Q^2) as functions of any combination of two variables out of the set (E, F, θ, γ) [54]. It is convenient to compute x from Q^2 and y so we only have to consider the formulae for Q^2 and y. We get:

A1
$$\begin{aligned} Q^2[E,\theta] &= 2AE(1+\cos\theta) \\ y[E,\theta] &= 1-\frac{E}{2A}(1-\cos\theta) \end{aligned}$$

A2
$$\begin{aligned} Q^2[E,\gamma] &= 4A^2(y[E,\gamma]-1)+4AE \\ y[E,\gamma] &= \frac{1}{2}\left\{\left(1-\frac{E}{2A}\right)(1-\cos\gamma)\pm\frac{E}{A}\sqrt{\frac{1}{4}(1-\cos\gamma)^2+\frac{A}{E}\left(1-\frac{A}{E}\right)\sin^2\gamma}\right\} \end{aligned}$$

A3
$$\begin{aligned} Q^2[\theta,\gamma] &= 4A^2\frac{\sin\gamma(1+\cos\theta)}{\sin\gamma+\sin\theta-\sin(\theta+\gamma)} \\ y[\theta,\gamma] &= \frac{\sin\theta(1-\cos\gamma)}{\sin\gamma+\sin\theta-\sin(\gamma+\theta)} \end{aligned}$$

A4
$$\begin{aligned} Q^2[E,F] &= \frac{4A^2(A-F)}{E+F-2A}+4EA \\ y[E,F] &= \frac{E-A}{E+F-2A} \end{aligned}$$

A5
$$\begin{aligned} Q^2[F,\theta] &= 4A^2(1-y[F,\theta])\left(\frac{1+\cos\theta}{1-\cos\theta}\right) \\ y[F,\theta] &= 1-\frac{1}{4A}(1-\cos\theta)\left\{(2A-F)\pm\sqrt{(2A-F)^2-\frac{8A(A-F)}{1-\cos\theta}}\right\} \end{aligned}$$

A6
$$\begin{aligned} Q^2[F,\gamma] &= \frac{F^2\sin^2\gamma}{1-y[F,\gamma]} \\ y[F,\gamma] &= \frac{F(1-\cos\gamma)}{2A} \end{aligned}$$

Apart from the familiar formulae A1 and A6, where formula A6 is the Jacquet-Blondel method, we derived four new methods for reconstructing x, y and Q^2. Note that in method A2 and A6 one obtains 2 solutions. These double solutions with physically allowed values only occur for $x < x_0 = A/P$.

References

[1] I. Aitchinson and A. Hey, *Gauge theories in particle physics*, A. Hilger LTD (1989).
P. Collins and A. Martin, *Hadron interactions*, A. Hilger LTD (1984).
A. Martin and F. Halzen, *Quarks and Leptons*, J. Wiley (1984).
T. Muta, *Foundations of QCD*, World Scientific (1987).

[2] R. Ellis and W. Stirling, *QCD and collider physics*, FERMILAB-CONF-90/164-T.
Proceedings of CTEQ *Summer school on QCD analysis and phenomenology*, Fermilab Pub 93/094(1993).
W. van Neerven, *QCD corrections to deep-inelastic structure functions*, Private Comm. (1991).

[3] M. Gell Mann and Y. Ne'eman, *The eightfold way*, Benjamin (1964).

[4] R. Taylor, H. Kendall and J. Friedman, Rev. Mod. Phys. **63**(1991).

[5] R. Brandelik et al., *Evidence for planar events in e^+e^- annihilation at high energies*, Phys. Lett. **B86**(1979)243.

[6] G. Ingelman and R. Rückl, *Quark momentum distributions from ep collisions: strategies and prospects*, DESY 89-025.
J. Blümlein et al., *Testing QCD scaling violations in the HERA energy range*, DESY-89-101.

[7] C. Callan and D. Gross, Phys. Rev. Lett. **22**(1969)156.

[8] A. Martin et al., RAL-87-052
A.M. Cooper-Sarkar et al., *Measurement of F_L and the small-x gluon density of the proton*, HERA Workshop Proceedings, ed. R. Peccei, 1987.

[9] J. Björken, Phys. Rev. **179**(1969)1547.
R. Feynman, Phys. Rev. Lett. **23**(1969)1415.
J. Björken and D. Paschos, Phys. Rev. **185**(1969)1975.

[10] V. Gribov and L. Lipatov, Sov. J. Nucl. Phys. **15**(1972)78.
G. Altarelli and G. Parisi, Nucl. Phys **B126**(1977)298.
G. Altarelli, *Partons in quantum chromodynamics*, Phys. Rep. **81**(1982)1.

[11] Ya. Balitski and L. Lipatov, Sov. J. Nucl. Phys. **28**(1978)822.
E. Kuraev, L. Lipatov and V. Fadin, Sov. Phys. JETP **45**(1977)199.

[12] A. Ali et al., *Small-x behaviour of DIS in QCD*, DESY topical meeting, Nucl. Phys. B(Proc. Suppl.)**18C**(1990).
J. Bartels, *Structure functions at small x: new physics at HERA*, Part. World **2**(1991)46.

[13] A. Mueller, *Parton distributions at very small x values*, Nucl. Phys. **B18C**(1990)125.

[14] L. Gribov et al., *Semihard processes in QCD*, Phys. Rep **100**(1983)1.
L. Gribov et al., Nucl. Phys **B188**(1981)155.

[15] M. Bonesini et al., WA70 collaboration, Z. Phys. **C38**(1988)371.

[16] C. Brown et al., E605 collaboration, Phys. Rev. Lett. **63**(1989)2637.

[17] A. Whitlow et al., SLAC, Phys. Lett. **250B**(1990)193.

[18] P. Amaudruz et al., NMC collaboration, Phys. Lett. **B295**(1992)159.

[19] J. Aubert et al., EMC collaboration, Nucl. Phys. **B259**(1985)189.
A. Benvenuti et al., BCDMS collaboration, Phys. Lett. **B223**(1989)485.
B. Adeva et al., SMC collaboration, CERN-PPE-93-206.

[20] J. Berge et al., CDHSW collaboration, Z. Phys. **C49**(1990)187.
D. Allasia et al., WA25 collaboration, Phys. Lett. **B135**(1984)231.
K. Varvell et al., WA59 collaboration, Z. Phys. **C36**(1987)1.

[21] C. Foudas et al., CCFR collaboration, Phys. Rev. Lett. **64**(1990)1207.

[22] R. Roberts et al., *A compilation of structure functions in deep-inelastic scattering*, J. Phys. **G 17**(1991)D1.

[23] J. Owens and W. Tung, *Parton distributions functions of hadrons*, FERMILAB-PUB-92/59-T.

[24] A. D. Martin et al., *New information on the parton distributions*, Phys. Rev. D**47**(1993)867.
A. D. Martin et al., *Parton distributions updated*, Phys. Lett. **B306** (1993)145.

[25] J. G. Morfin and W. K. Tung, Z. Phys. C **52**(1991)13.

[26] J. Botts et al., *CTEQ parton distributions and flavour dependence of the sea quarks*, Phys. Lett. **B304**(1993)159.

[27] M. Glück, E. Reya and A. Vogt, Nucl. Phys. **B(Proc. Suppl.)18C**(1990)49
M. Glück, E. Reya and A. Vogt, . Z. Phys. **C53**(1992)127.

[28] H. Spiesberger et al., *Radiative corrections at HERA*, Proc. of the Workshop on Physics at HERA vol. 2 (1992)798.

[29] B. Badelek et al., *Report of the working group on radiative corrections at HERA*, J. Phys. G: Nucl. Part. Phys. **19**(1993)1671.

[30] H. Plothow-Besh, PDFLIB, Proceedings of the 3^{rd} workshop on detector and event simulation in high energy physics, Amsterdam 1991.
S. de Jong, SFLIB, Amsterdam 1991.

[31] A. Kwiatkowski et al., HERACLES, Proc. of the Workshop on Physics at HERA vol. 3 (1992)1294.

[32] G. Ingelman, LEPTO, Proc. of the Workshop on Physics at HERA vol. 3 (1992)1366.

[33] L. Lönnblad et al., ARIADNE, Proc. of the Workshop on Physics at HERA vol. 3 (1992)1440.

[34] B. Webber, HERWIG, Proc. of the Workshop on Physics at HERA vol. 3 (1992)1354.

[35] T. Sjöstrand, the JETSET 7 program and manual.
T. Sjöstrand, Computer Phys. Comm. **39**(1986)347.
T. Sjöstrand et al., Computer Phys. Comm. **43**(1987)367.

[36] R. G. Roberts, *The structure of the proton*, Cambridge University Press, 1990.

[37] A. D. Martin et al., *Present status of parton distributions*, J. Phys. G: Nucl. Part. Phys. **19**(1993)1429.

[38] H. Bethe and W. Heitler, Proc. Roy. Soc. **A146**(1934)83.

[39] M. van der Horst, PhD thesis, University of Amsterdam, 1990

[40] A. de Roeck and M. Klein, *First measurements at HERA of deep inelastic scattering at low-x*, DESY 93-014.

[41] M. Derrick et al., Nucl. Instr. Meth. A **309**(1991)77;
A. Andresen et al., Nucl. Instr. Meth. A **309**(1991)101

[42] M. Derrick et al., (ZEUS BCAL group), Nucl. Instr. and Meth. **A309**(1991)
I. Gialas, Proc. Second Intern. Conf. on Calorimetry for high energy physics (Capri 1991) p.162.

[43] J. Krüger, Habilitationsschrift, DESY F35-92-02 (1992).

[44] J. Straver, *Design, Construction and Beam tests of the high resolution Uranium scintillator calorimeter for ZEUS*, PhD thesis, University of Amsterdam, March 1991.

[45] U. Behrens et al., *Test of the ZEUS forward calorimeter prototype*, Nucl. Instr. and Meth. **A289**(1990)115.
A. Andresen et al., *Construction and beam test of the ZEUS forward and rear calorimeter*, Nucl. Inst. Meth **A309**(1991)101.

[46] W. Smith et al., *The ZEUS trigger system*, ZEUS-note 89-084
L. W. Wiggers, *Monte Carlo study of dead-time in the ZEUS DAQ system*, ZEUS-note 87-056

[47] M. Derrick et al., *Observation of two-jet production in deep-inelastic scattering at HERA*, Phys. Lett. **B306**(1993)158.
M. Derrick et al., *Hadronic energy distributions in deep-inelastic electron-proton scattering*, Z.Phys.**C.59**(1993)231.

[48] M. Derrick et al., *A measurement of $\sigma_t(\gamma P)$ at $\sqrt{s} = 210$ GeV*, Phys. Lett. **B293**(1992)465.

[49] M. Derrick et al., *Initial study of deep inelastic scattering with ZEUS at HERA*, Phys. Lett. **B303**(1993)183.

[50] M. Derrick et al., *Measurement of the proton structure function F_2 in ep scattering at HERA*, Phys. Lett. **B316**(1993)412.

[51] J. Feltesse, *Measurement of inclusive differential cross sections*, Proc. of the HERA workshop, (1987)33.

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

[53] S. Bentvelsen, J. Engelen and P. Kooijman, *Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA*, Proc. Workshop on physics at HERA vol. 1(1992)23.

[54] S. Bentvelsen, *Optimal use of kinematical variables for x-Q^2 reconstruction in hard NC events*, ZEUS note 91-14(1991).
S. Bentvelsen and P. Kooijman, *Reconstruction of x and Q^2 and the effects of initial state radiation*, ZEUS note 91-57(1991).

[55] K. Hoeger, *Measurement of x, y, Q^2 in neutral current events*, Proc. Workshop on physics at HERA vol. 1(1992)43.

[56] E. Eichten, I. Hinchliffe, K. Lane, C. Quigg, Rev of Mod Physics 56(1984)579.
E. Eichten, I. Hinchliffe, K. Lane, C. Quigg, Rev of Mod Physics 58(1986)1065.

[57] The ZEUS detector, Technical Proposal, Hamburg, March 1986.

[58] The ZEUS detector, Status Report 1993, DESY.
The ZEUS detector, Status Report 1989, DESY.

[59] S. de Jong, *The ZEUS second level calorimeter trigger: physics simulation, design and implementation*, PhD thesis, University of Amsterdam, 1991.

[60] H. Uijterwaal, *The global second level trigger for ZEUS*, PhD Thesis, University of Amsterdam, 1992.

[61] H. van der Lugt, *The data-acquisition and second level trigger system for the ZEUS calorimeter*, PhD Thesis, University of Amsterdam, 1993.

[62] S. Fisher and P. Palazzi, *The* ADAMO *data system*, Programmers manual, Version 3.2 (1992).

[63] A. Abramowicz et al., *A determination of F_2 with the 1992 data*, ZEUS-note 93-078(1993).

[64] W. Dorth, LAZE, internal ZEUS note 91-069(1991).

[65] A. Dake, GAZE, internal ZEUS note 91-006(1992).

[66] G. Wolf, *First results from HERA*, DESY-92-190 (1992).

[67] M. Derrick et al., *Observation of events with a large rapidity gap in deep inelastic ep scattering at HERA*, Phys. Lett. **B315**(1993)481.

[68] K. Nikolaev, B. Zakharov, *Pomeron structure function and diffractive dissociation of virtual photons in perturbative QCD*, Z. Phys. **C53**(1992)331.

[69] P. de Jong, *The measurement of the hadronic energy flow and jet production with the ZEUS calorimeter in DIS at HERA*, PhD Thesis, University of Amsterdam, 1992.

[70] M. Krasny et al., *Determination of the longitudinal structure function at HERA from radiative events*, Proc. Workshop on physics at HERA vol. 1(1992)171.

[71] I. Abt et al., *Measurement of the proton structure function F_2 in the low-x region at HERA*, Nucl. Phys. **B407**(1993)515.

[72] CERN-D506, MINUIT function minimization and error analysis, 1992.

[73] M. Glück et al, *Scaling violations and the gluon distribution of the nucleon*, Z. Phys. **C13**(1982)119.