UNIVERSIDAD AUTONOMA DE MADRID

Facultad de Ciencias
Departamento de Física Teórica

Measurement of the subjet multiplicity in neutral current DIS and determination of α_s at HERA

Medidas de la multiplicidad de subjets en procesos mediados por interacciones neutras en dispersión profundamente inelástica y determinación de α_s en HERA

Memoria del trabajo de iniciación a la investigación

Oscar González López

Director: Dr. Juan Terrón Cuadrado

25/08/2001

This work has been carried out in the frame of the ZEUS Collaboration
through the High Energy Physics group of the Universidad Autónoma de Madrid (Spain).

Este trabajo ha sido desarrollado en la Colaboración ZEUS
a través del grupo de Altas Energías de la Universidad Autónoma de Madrid

El autor ha sido financiado por el Ministerio de Educación, Cultura y Deporte.

Introduction

The study of the internal structure of the jets produced in high-energy processes gives direct insight into the transition from the partons taken part in the hard scattering and the hadrons that are observed in the detectors. Measurements of the jet internal structure provide very stringent tests on the predictions given by the theory and the phenomenological approaches used to describe the processes that drive this transition.

During the last years, the use of subjets to characterize the internal structure of the jets, particularly when the k_T-cluster algorithm is used to define the jets, has become more and more popular. This is due because it has been shown that it presents several advantages from the theoretical and experimental points of view.

This document presents a study of the subjet multiplicity for jets produced in neutral current interactions in ep collisions. The events are selected in the deep inelastic scattering regime, and the results are presented in terms of the mean subjet multiplicity, which has become the standard variable for the studies of subjets.

The measurements are corrected for detector effects and compared with the theoretical predictions in the frame of the Standard Model, namely with the predictions of Quantum Chromodynamics (QCD). After the comparison is made, a QCD analysis is performed and the value of the strong coupling constant is extracted. This is one of the fundamental parameters of the theory to describe the strong interaction and, hence, of the Standard Model. The determination of this coupling constant, made here for the first time using the internal structure of the jets at HERA, provides a new contribution to the world average value, which still presents a relatively large uncertainty.

The structure of this document is as follows. After describing the theoretical concepts related to the subject, in chapter 1, the experimental infrastructure is described in chapter 2. In chapter 3, the main aspects of physics simulation, phenomenological models and generation of simulated events are introduced. The complete description of the analysis is presented in chapter 4. This chapter contains the results of the analysis, which are discussed and compared with theory in chapter 5. In this chapter, the comparison between the measurements and the theory is used to extract the value of the strong coupling constant α_s. Finally, chapter 6 is devoted to the last remarks and a brief summary.

The results presented here, specifically in chapters 4 and 5, have been presented by the author in the DIS'2001 conference in Bologna [1]. Furthermore, they have been published by ZEUS as a contributed paper to the International Europhysics Conference on High Energy Physics, Budapest, Hungary (2001) [2].

Contents

List of Figures

List of Tables

Chapter 1

Theoretical introduction

At present the most successful way to describe the measurements in high energy experiments is summarized in the Standard Model. It describes interactions between the fundamental particles in terms of three kind of forces: the electromagnetic, the weak and the strong interaction. They are introduced as coming from fundamental gauge symmetries $U(1)_Y xSU(2)_L xSU(3)_C$ based on the solid grounds of Quantum Field Theory. It developed through the 1960's and 1970's by a fruitful combination of experimental and theoretical discoveries.

In this chapter a discussion about the Standard Model is presented. It includes how this model describes the formation of hadronic jets and its internal structure, which is the relevant part for the analysis.

1.1 Standard Model of particle interactions

The Standard Model is at present the last step in the developments done to understand the phenomena which we observe in the Universe. It can be considered the great thriumph of the discoveries made during the 20th Century in trying to explain the experimental results which were not described by the classical mechanics (developed in 16th and 17th Centuries) and electrodynamics (developed during 18th and 19th Centuries).

In the following sections we will describe how the different results were interpreted and related to complete a model which has been able to describe most of the experiments done during the last 25 years.

1.1.1 The internal structure of atoms

At the beginning of the 20th Century only two types of interaction were known: the electromagnetic interaction which was needed to describe the forces between electrically charged particles and the gravitation force, which was associated to the mass. The matter was understood by Dalton's picture that fundamental particles (atoms) exists and their combinations form everything we observe in Nature.

Different experimental results of those years were demonstrating that this simple picture had to be corrected: the atoms seemed not to be fundamental particles since some elements change into others and even one fundamental particle has been found inside the atom: the electron.

On the interaction side, although Faraday and Maxwell's vision of the electromagnetic fields looked successful, the energy and specially the electromagnetic one, displayed a behaviour easier to explain assuming that the energy presents not only wave- but particle-like behaviours.

After different experiments and models, a new theory was developed, Quantum Mechanics, and it turned very useful to describe the experiments designed to study the internal structure of the atoms. During the first three decades of the 20th Century, the theoretical and experimental progresses concluded with a picture where:

- At very small scales (atom size) matter cannot be described as fundamental particles in a classical way. Particles displays wave-like behaviours and have to be described using Quantum Mechanics.
- Matter and Energy cannot be understood in a separate way: The Theory of Relativity shows that energy can be converted into matter and viceversa. On the other hand, Quantum Mechanics shows that particles (matter) has to be described by fields and the interactions (energy) need to be described using *quanta of energy*, that is particles.
- The electrically neutral atom is formed by a nucleus which has a positive electric charge which define the chemical properties of the element. To compensate this charge the nucleus is surrounded with electrons. The structure of the electron cloud around the nucleus is described with Quantum Mechanics.
- The nucleus contains the majority of the mass of the atom. It contains a given number of protons and neutrons which are packed together to form the nucleus.

This picture of energy and matter is still today correct, although it was clear that it has to be completed because there must be some additional force to pack together protons and neutrons in the nucleus. At this point, two ways of study were followed: the first one tried to understand the interactions between electrons and the electromagnetic field. The second looked inside the nucleus to understand its internal structure.

1.1.2 The electromagnetic interaction and QED

Although Quantum Mechanics was very successful to describe the structure of the atoms, it was very clear that it was not enough, since it doesn't include the relativity effects which are needed for high energy processes. In this case high energy means in comparison with the electron mass, or the bounding energy of the electrons to the atom.

The first attempts to include the relativistic principles in the description of the particle behaviour were not very successful because strange solutions appear. One had to wait till Dirac's equation to describe the free electron. This equation can be written in the covariant form [3]

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0, \tag{1.1}$$

where m is the mass of the particle, ψ is the wave function to describe the quantum state where the particle is, and γ^{μ} are the four Dirac γ-matrices needed to recover the relativistic Klein-Gordon equation. For an electron, the γ-matrices has to be at least 4x4 matrices, so ψ is a four-component spinor[1]. Two of then were related to the two components of a "standard" electron

[1] An spinor is a mathematical vector but with a specific set of transformation properties.

of spin 1/2 and the other two, associated to negative energy states, were interpreted as the two components of a spin-1/2 particle with the same mass of the electron but positively charged (the positron). It should be noted that the spin is naturally introduced in the relativistic-quantum description of the electron.

The discovery of the positron was a sucess of Dirac's description of the relativistic free electron, so there were some attempts to modify the equation to include interactions. The inclusion of the interaction of the electron with a nucleus was very succesful when making the approximation to the non-relativistic regime. Some corrections predicted by Dirac's theory (for example related to the spin of the electron coupled to the magnetic field of the nucleus) has been observed experimentally [4].

On the other hand for a complete relativistic quantum description of the electron coupled to the electromagnetic field it was necessary to describe the electromagnetic field within the quantum mechanics framework. In the first attempts it was discovered that a lot of complications appear, mostly due to divergences in observables which were experimentally finite. It was needed to wait till the 40's and 50's to find the solution to all these problems.

The solution was the renormalization procedure: the divergences which appear in the calculation of observables are absorbed in some non-observable quantities which are present in the theoretical basis, i.e. contained in the Lagrangian. With this kind of mathematical treatments, one is able to use the Quantum Field Theory to obtain predictions and these ideas were successfully applied in Quantum ElectroDynamics (QED). In this theory the electromagnetic interaction is described by two quantum fields: one for the charged particles (the electron-positron field) and one for the electromagnetic (or photon) field. They are coupled in such a way that the variations in one field induces changes in the other.

This leads, in the classical limit, to Faraday's picture of interactions by means of fields created by the particles.

The validity of QED to describe the electromagnetic interaction has been confirmed during more than 50 years and at present this theory is among the best established ones in Physics. Its success confirms the validity of the Theory of Relativity and of Quantum Mechanics, which can be considered as the two great developments in Physics during the 20th Century.

1.1.3 The strong and weak interactions

After the discovery of the proton one question arose: how can the protons be packed together in the nucleus? The gravitational force was not strong enough to compensate the electrical forces which tends to push away particles with the same electric charge. The discovery of the neutron as a fundamental piece to build the atomic nucleus[2] showed that new forces are present in Nature and the atomic nucleus was the place where to study them. The stability of the neutron within the nucleus suggests that a new interaction is present between protons and neutrons. This force, called *strong force* due to the fact that it was stronger than the electromagnetic one, was present only when particles were very close.

With the development of new experimental devices, new particles were discovered and the analysis of their behaviour revealed that two different kinds of particles appear in Nature:

[2]The only nucleus that does not contain any neutron is the hydrogen nucleus, but it consists of only one proton... so it is not a "nucleus" but a proton.

- Leptons: which are particles which don't interact through the strong force. The electron is a lepton.

- Hadrons: which are particles which interact trough the strong force. The group is divided into mesons (which are bosons) and baryons (which are fermions). The proton is a hadron.

In both groups appear particles with and without electric charge so they are able to interact with the electromagnetic field. But, the study of some decays leads to introduce a new interaction: *the weak interaction*. It was named this way because its main property was that its effects were weaker than the effects induced by the electromagnetic interaction.

Although some principles were accepted (conservation laws, the existence of an antiparticle for each particle,...) no one knew how to describe the new interactions in a fundamental way. It took more than 35 years to understand the nature of these new forces and it concluded with the development of the Standard Model during the seventies.

1.1.4 Hadron spectroscopy and the Quark-Parton Model

After the discovery of the first hadrons, a new branch in fundamental physics appeared. It is *hadron spectroscopy* and its aim is to study the different kinds of hadrons and understand the different relations between them.

The first step in this new branch was the introduction by Heisenberg of the *isospin quantum number*. This quantum number sets the differences between a neutron and a proton, which seems otherwise to be identical particles regarding the nuclear interaction. Formally, a proton and a neutron can be considered as the same particle (a nucleon) but with different isospin quantum numbers.

The discovery of additional hadrons lead to use the isospin to classify them. In fact, it is noted that hadrons are classified in isospin families in such a way that the members are very similar.

The second classification of hadrons was made taking into account the spin-statistics connection for particles. In this way, hadrons can be fermions or bosons and they can have spin from 0 to 3/2. To differentiate baryons from mesons a new quantum number is introduced, the baryon number (B). It takes value 1 for baryons (which are hadrons following Fermi-Dirac statistical rules), -1 for antibaryons and 0 for mesons (which are hadrons following Bose-Einstein statistical rules). The conservation of the baryon number in Nature is the only explanation to avoid proton decay. The proton would be stable since is the baryon with smallest mass.

The discovery of new relations between hadrons needed the introdution of a new quantum number, the *strangeness* (S). It is needed to explain why some particles are only created in pairs when strong and electromagnetic interactions are present. The weak interaction does not conserve this quantum number, and that explains why the created pairs can decay to particles with S=0 via the weak interaction although the creation was done via the strong or electromagnetic force.

Using the strangeness and the baryon number, a new quantum number can be constructed which in fact it turns out to be more useful for the classification of hadrons. It is the "hypercharge" and it is defined by

$$\mathrm{Y} = \mathrm{B} + \mathrm{S}$$

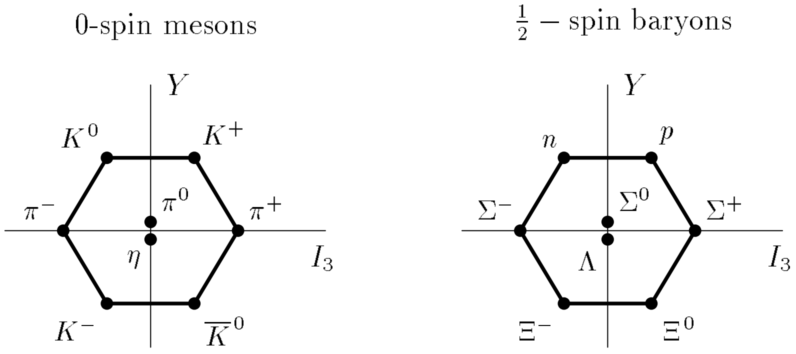


Figure 1.1: Two examples of the 'Eightfold Way' classification of hadrons. On the left some of the 0-spin ground-state mesons and on the right the $\frac{1}{2}$-spin ground-state baryons. The particles belonging to these groups are distributed in octets of SU(3). The plane containing these octets is defined, as mentioned in the text, by the hypercharge $Y\,(= B + S)$ and the third component of the isospin $I_3\,(= Q - Y/2)$.

which leads to another important relation introducing the electric charge (in units of the proton charge):

$$\mathrm{Q} = \mathrm{I}_3 + \frac{\mathrm{Y}}{2},$$

where I_3 is the value of the third component of isospin, which is the one which differentiates the particles belonging to the same isospin family.

The next important step in the studies on hadron spectroscopy was the introduction of the 'Eightfold Way', proposed independently by Ne'eman and Gell-Mann [5]. It consists of a classification scheme to group the observed baryons and mesons with the same spin according to their third isospin component (that is, the charge) and hypercharge (that is, the strangeness) using the SU(3) group symmetry. As shown in figure 1.1, these representations are very simmetric and the distribution of all the hadrons in this kind of schemes suggests that the origin of the hadronic structure is common to all these particles. In a similar way as with Mendeleev's table for chemical elements, some hadrons were not yet discovered when the 'Eightfold Way' appeared. The discovery of a predicted particle with Q=−1 and S=−3 in 1964 led to a wide acceptance of the 'Eightfold Way' for the classification of hadrons.

The nice way the hadrons are distributed in representations of the SU(3) group can be used as a hint to analyze its internal structure, at least in a formal way. It should be mentioned that it was already known that hadrons were not point-like particles. The first indication of this comes from the magnetic moment of the neutron: it is simpler to consider the neutron as a non-elementary particle than to explain how a neutral particle has a magnetic moment.

The second indication came from the *g factor* for the neutron and the proton. Although QED introduce corrections to the value of 2 predicted by Dirac's theory, the proton and the neutron have values too far away of 2, which is the expected value for 1/2-spin point-like particles, as the electron.

In 1964, Gell-Mann and Zweig independently proposed hadrons to be composed of elementary constituents, called *quarks* [6], providing the basis for the 'Eightfold Way'. The quarks, coming

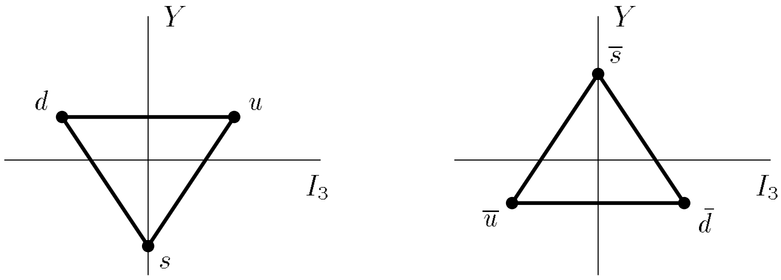


Figure 1.2: The multiplets for the quarks and the antiquarks as introduced in the original Quark model. With successive combinations of these particles, it is possible to reconstruct the various hadrons belonging to the different SU(3) multiplets.

in three flavours, up (u), down (d) and strange (s) and carrying spin $\frac{1}{2}$ were assigned fractional charges: $+\frac{2}{3}$ for the u and $-\frac{1}{3}$ for the d and s, in units of the proton's charge.

The quantum numbers of the three quarks (shown in table 1.1) are given in such a way that the quarks and the corresponding antiquarks are distributed in a fundamental multiplet of SU(3) (see Fig. 1.2). Combining the quarks, it is posible to reproduce the different representation of SU(3) which configures the 'Eightfold Way'. To do that, it is necessary to describe the baryons as formed by three quarks, the antibaryons by three antiquarks and the mesons by one quark and one antiquark. This quark model of the hadron internal structure simplifies the studies on hadron spectroscopy, where its use would be justified even in the case that quarks were not real particles[3].

Two strong points were against the quark model. The first one was that quarks had not been observed, and with a fractional electric charge its detection would have been very simple.

The second one was that the model needs an additional quantum number to explain some of the quark states. Some hadrons contain three quarks of the same flavour and the same spin component and since they are fermions, there must be another quantum number to be consistent with the Pauli Exclusion principle. This new quantum number was called *colour* [7] since it was used to construct a mnemonic rule for the particularly reduced quark association: *with three colours (or anticolours) or one colour and its relative anticolour it is posible to form colourless particles*. That means that colour is related to the quark world, although the observable hadrons are colourless... and only combination of three (anti)quarks or one quark and one antiquark are possible observable states.

The discovery in the last years of the 1960's that the structure functions of the nucleon scale with energy (see section 1.2.2 for a complete description) and that this result could be explained considering that the nucleon is formed by point-like particles with spin 1/2 leads to reconsider the validity of the quark model as a good description of the internal structure of the hadrons. This was the origin of the Quark-parton model which describes the proton as formed by free particles (the partons). This model does not contain interactions between the partons, which in fact are needed to explain the confinement of the quarks inside the colourless hadrons.

The different studies done during the begining of the seventies conclude with a model where the proton was formed by additional partons apart from the three quarks predicted in the Quark model as used in hadron spectroscopy. In the nucleon there were additional neutral partons,

[3] "Such particles [quarks] presumably are not real but we may use them in our field theory anyway" (Gell-Mann, 1964).

Quark	Q	B	S	Y	I_3
d	$-\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	$-\frac{1}{2}$
u	$\frac{2}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{2}$
s	$-\frac{1}{3}$	$\frac{1}{3}$	-1	$-\frac{2}{3}$	0

Table 1.1: Quantum numbers from the three quarks as introduced in the original SU(3) Quark model. The corresponding antiquarks have the same numbers but with opposite sign. As described in the text, the hypercharge is defined as $Y = B + S$ and the isospin is defined in terms of the electric charge $I_3 = Q - Y/2$.

which were called gluons, and from the interactions between the gluons and the quarks, new quarks appear. This leads to a separation of the partons inside the hadrons into three groups

- **The valence quarks**: they are the original quarks of the Quark model
- **The gluons**: the valence quarks can radiate gluons as a consequence of the strong force which keeps the partons together within the hadron. This gluon content has not contribution to the quantum numbers of the hadron and it does not affect the espectroscopic values of the original quark model.
- **The sea quarks**: the gluons can split into $q\bar{q}$ pairs. These $q\bar{q}$ pairs don't contribute to the quantum numbers of the hadron since those of the quark are compensated by those of the antiquak.

All these contributions configure the complete internal structure of the hadron, which is far from simple, is determined by the interaction between the partons and has to be added to the Quark-Parton model. The strong interaction between hadrons is related to the interactions between quarks and gluons. This means that the interaction between hadrons is not a fundamental interaction. It is only the indirect consequence of the interactions between the partons from the different hadrons. This explains why the interactions between hadrons (as the nuclear interaction) is only present at short distances: it is needed that the partons in one hadron "detect" the existence of the colour fields produced by the partons in the second hadron, and since the colour field is not produced by the whole hadron, only contributions resolving the internal structure of the hadron can take part in the hadron-hadron interaction.

Since the discovery of the c-quark (charm), the quark model has been considered as the best description of the internal structure of hadrons, at least from the point of view of hadron spectroscopy. During this time, quarks and gluons have become a fundamental part of the present knowledge on fundamental physics.

1.1.5 Quantum Chromodynamics

Following the way opened by QED and Yang-Mills theories, Quantum ChromoDynamics (QCD) was developed in the context of Quantum Field Theory starting with a global symmetry. This symmetry is based in the colour group $SU(3)_C$ and when the free-quark Lagrangian is imposed to preserve this symmetry locally, then a local non-abelian gauge theory of quarks and massless gluons appears. This theory can be described by the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{A}_{\alpha\beta}F^{\alpha\beta}_{A} + \sum_{flavours} \bar{q}_a \left[\gamma^{\mu} D_{\mu} - m\right]_{ab} q_b \;, \tag{1.2}$$

where $F^{A}_{\alpha\beta}$ is the field strength tensor defined in terms of the gluon field A^{A}_{α} by

$$F^{A}_{\alpha\beta} = \left[\partial_{\alpha} A^{A}_{\beta} - \partial_{\beta} A^{A}_{\alpha} - g f^{ABC} A^{B}_{\alpha} A^{C}_{\beta}\right] \;,$$

where indexes A, B, C run over the eight colour degrees of freedom of the gluon field. The last term in the expression for $F^{A}_{\alpha\beta}$ distinguishes QCD from QED since it gives rise to cubic and quartic gluon self-interactions; g is the coupling charge which determines the strength of the interaction and f^{ABC} are the structure constants of the SU$(3)_C$ group.

The sum in the lagrangian runs over the different flavours of quarks. The quarks fields (q_a) are in the triplet representation of the group (a=1, 2, 3). The covariant derivative D_{μ} takes the form

$$\left(D_{\mu}\right)_{ab} = \delta_{ab}\partial_{\alpha} + ig\left(t^{C} A^{C}_{\alpha}\right)_{ab} \tag{1.3}$$

where δ_{ab} is Kronecker's delta, and $\{t\}$ are matrices in the fundamental representations of SU(3). By convention, the normalization of the SU(N) matrices is chosen to be

$$Trace\{t^{A}t^{B}\} = \frac{1}{2}\delta^{AB}$$

With this choice, the colour matrices obey the following relations:

$$\sum_{A} t^{A}_{ab} t^{A}_{ba} = C_F \delta_{ab} \qquad C_F = \frac{N^2 - 1}{2N}$$

$$Trace\{T^{C}T^{D}\} = \sum_{A,B} f^{ABC} f^{ABD} = C_A \delta_{CD} \qquad C_A = N$$

where $\{T\}$ are matrices in the adjoint representation of SU(N). For the SU(3) group, these are the eight Gell-Mann matrices.

With all theses tools it is possible to describe the interactions between quarks and gluons. Using the perturbative approach, the theory is able to make predictions for those processes in which the strong coupling constant $\alpha_s = g^2/4\pi$ is small.

The validity of QCD to describe the strong force was experimentally confirmed with the discovery of the scaling violations of the structure functions and when the existence of the gluon was confirmed experimentally with the observation of three-jet events in e^+e^- annihilation. In this way, the colour interaction has been able to describe the results related to the strong interaction for more than 20 years and is at present one of the fundamental interactions included in the Standard Model.

1.1.6 The electroweak unification

If the undertanding of the strong interactions needed a very hard work to understand the hadron spectroscopy and the experimental results, even harder was the long way needed to complete a theory which included the weak interaction.

After the first steps introducing a new, undetected particle (the *neutrino*) and some rules for the calculation of cross sections and decay processes (known as *Fermi theory of the weak interaction*), it was demostrated in 1956 that the weak interaction violates parity and charge-conjugation symmetry. That is, the experiments would provide different results if all the particles are changed by their antiparticles. The same occurs if the space is mirror-reflected.

In the new theory to describe this property, the neutrino was introduced to have a well-defined and fixed helicity, which is only possible if these particles are massless. This characteristic of neutrinos obviously violates parity symmetry, since right-handed neutrino would be forbidden in Nature[4].

On the hadronic sector, interesting features were discovered in parallel. With the quark model description in hand, the weak interaction was shown to change the flavour of the quarks, similarly to the transformation of electrons into neutrinos and of neutrons into protons. Additionally, the weak interaction couples to quark states which are not the flavour-defined states. The interaction couples linear combinations of the physical quarks[5]. This mixing, which is understood as a rotation in the space containing the quark states, was first described by the Cabibbo angle. After the discovery of the third family of quarks, the 3x3 CKM matrix (Cabibbo, Kobayashi, Maskawa) describes the mixing in the quark sector [8].

The first successful description of the weak interaction based on a symmetry group was done by Glashow in 1961 [9]. The fundamental group was the SU(2) associated to the left-handed helicity of the particles. From the point of view of this group the particles are divided into left-handed doublets and right-handed singlets. For example, for the leptons of the first generation, the left-handed electron is associated to the neutrino in a doublet. The right-handed electron forms a singlet which do not couple to this interaction.

Of course, for a consistent description of this interaction the masses have to be neglected since $SU(2)_L$ cannot be a symmetry in nature if particles have mass. This is one of the difficulties of the theory.

Since an interaction based in $SU(2)_L$ is divided into a "charged part" (that is, exchanging the components of the doublet) and a "neutral part" (that is, leaving the doublets as they are), the model seemed to be inexact since neutral-current weak interactions had not been observed at that time. The discovery of this kind of interactions in 1973 confirmed the validity of this model.

In fact the complete model has to be slightly more complex since the observed neutral-current weak interaction displays a right-handed component, so there must be something else in the model. At present, the Standard Electroweak Model or GWS Model (Glashow, Salam and Weinberg) [9, 10] is based on the group $SU(2)_L$ x $U(1)_Y$. The first group, related to the left-handed particles, is described introducing a new quantum number, the *weak isospin*. The

[4] In fact, it is not forbidden that a right-handed neutrino exists. The theory only forbids this neutrino to interact with matter through the usual weak interaction.

[5] "Physical quarks" is meant for the particles with defined mass, which in the case of quarks are the ones with a given flavour.

second group is related to the *weak hypercharge*. The matrices $\{T\}$ are the representation of the $SU(2)_L$ algebra, while Y is the gauge charge associated to the $U(1)_Y$ group.

The weak isospin is formally similar to the hadronic isospin we described in section 1.1.4. It is associated to spin-like multiplets, in this case associated to $SU(2)_L$. The fermions are set in doublets (i.e., weak isospin $T=\frac{1}{2}$) in case of left-handed components, or singlets (i.e. $T = 0$) for the right-handed ones. The third component of the weak isospin (T^3) allows to differentiate the two components of the doublet. The weak hypercharge is then defined for each particle as [3]

$$Q = T^3 + \frac{Y}{2} \tag{1.4}$$

where Q is, as usual, the electric charge of the particle in units of the proton charge.

For this model, using Yang-Mills fields and local gauge invariance, the following Lagrangian is defined

$$\mathcal{L} = -\frac{1}{4}W_i^{\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + \bar{f}\big(i\gamma_\mu D^\mu\big)f \tag{1.5}$$

where f represent the fermionic fields and

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_W \epsilon^{ijk} W^i_\mu W^k_\nu$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$$D^\mu = \delta_{ij}\partial^\mu + ig_W\left(T \cdot W^\mu\right)_{ij} + iY\delta_{ij}g_Y B^\mu$$

The antisymmetric (Levi-Cività) tensor ϵ^{ijk} appears in the expressions because its components are the structure constants of SU(2) [11].

The Lagrangian constructed in this way evidently describes four massless vector bosons forming a singlet (B) and a triplet ($W^\pm, W^0$) under weak isospin transformations. The model cannot yet describe the real world because there is only one massless vector boson observed in Nature which can be considered as a candidate, the photon. The direct addition of explicit mass terms for three of the bosons would violate local gauge invariance and spoil the renormalizability of the theory.

On the other hand, since the weak isospin is only well-defined for fermions of a fixed helicity, mass terms for the fermionic field would break the symmetry and the local gauge invariance would be lost. But in Nature, the fermions are observed to have a mass which is only negligible for high energy process, so the model has to be completed to include mass terms in a natural way while preserving the properties of the theory.

This is done using the Higgs mechanism [12], based on the idea of "spontaneous symmetry breaking". A rough description of this idea is that the need to select between several identical options leads to a symmetry breaking. Of course, the way this selection is done can be completely at random, and in principle any other choice would be equally valid; it is the need to make this choice which leads to symmetry breaking, and not the choice itself.

In the Standard Model, a single complex doublet of scalar fields is introduced. These fields transform as a doublet under $SU(2)_L$. By assigning them a weak hypercharge of one half we fix their transformation properties under $U(1)_Y$. The inclusion of these fields in the Lagrangian after coupling them to the vector bosons makes a mass term for the scalar field to appear, but it is an imaginary mass. A real mass is made to appear transforming the fields into physical fields, but this transformation is not unique and a choice of parameters is needed. Here is where the "spontaneous symmetry breaking" takes place, because the transformation to the physical fields is not invariant under $SU(2)_L$ x $U(1)_Y$. A deeper study yields that three of the symmetries are spontaneously broken, and three of the massless bosons get a finite mass by means of the Goldstone theorem applied to this case.

One symmetry is preserved, which is related to the standard $U(1)_{em}$ group, and the photon is kept massless. The neutral currents of the original groups are now mixed and two new fields appear, being the relation

$$Z_\mu = W^0_\mu \cos\theta_W - B_\mu \sin\theta_W$$

$$A_\mu = W^0_\mu \sin\theta_W + B_\mu \cos\theta_W$$

where the second one is the standard electromagnetic field and θ_W the *weak angle* which dictates the mixing of the high-energy fields (given in the GSW symmetry group) with the low-energy ones (where the electromagnetic and the weak fields are separated).

The mass scale for this symmetry breaking is set to the mass of the weak bosons, which is of the order of 100 GeV. For energies well above this scale, the electromagnetic and the weak interactions are unified and the description of the electroweak force is done in terms of the weak isospin and hypercharge interactions. For energies close to that scale and smaller, the weak interaction separates from the electromagnetic one and massive bosons are needed to describe this low energy region. Since the degrees of freedom have to be conserved when going from high energy to low energy, the remaining degree of freedom of the Higgs complex fields has to be observed at low energies as a scalar field, since the other three degrees of freedom have been absorbed by the longitudinal helicity components of the massive fields.

This particle has not been yet observed[6], but it has to be included in the Standard Model to allow the gauge bosons to acquire a mass. Furthermore, the coupling between fermions and the Higgs field is the origin of the fermionic mass, as an additional consequence of the Higgs mechanism.

For a more complete description of the electroweak unification, the GSW model and the Higgs mechanism, see [3, 11].

1.1.7 Summary: The Standard Model and beyond

With the artifitial unification of the electroweak and strong interactions, we are left with the symmetry group $SU(3)_C$ x $SU(2)_L$ x $U(1)_Y$, which contains the description of all these fundamental forces. These forces are present in the interactions between the fundamental particles and the Standard Model is completely defined.

In the so-called matter sector, the particles are divided into three families or generations which are formally identical. The only difference between generations is the mass of the particles.

[6]Although some events from the LEP experiments in autumn 2000 suggest the presence of the Higgs boson at a mass around 115 GeV, the results have not been considered as a the discovery of this particle. Further experimental evidence is needed to confirm or rule out these results.

	Generations						Quantum numbers				
	1^{st}		2^{nd}		3^{rd}		Q	T	T_3	Y	colour
Quarks	u	u_L	c	c_L	t	t_L	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	Yes
	(1-5 MeV)	u_R	(1.5 GeV)	c_R	(170 GeV)	t_R		0	0	$\frac{4}{3}$	
	d	d_L	s	s_L	b	b_L	$-\frac{1}{3}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	Yes
	(3-9 MeV)	d_R	(75-170 MeV)	s_R	(4-5 GeV)	b_R		0	0	$-\frac{2}{3}$	
Leptons	ν_e	(L)	ν_μ	(L)	ν_τ	(L)	0	$\frac{1}{2}$	$\frac{1}{2}$	-1	No
	$(< 1$ eV$)$	$-$	$(< 0.2$ MeV$)$	$-$	$(< 18.2$ MeV$)$	$-$					
	e^-	e_L	μ^-	μ_L	τ^-	τ_L	-1	$\frac{1}{2}$	$-\frac{1}{2}$	-1	No
	(0.51 MeV)	e_R	(106 MeV)	μ_R	(1.8 GeV)	τ_R		0	0	-2	

Table 1.2: Quantum numbers and masses for all the fundamental fermions in the Standard Model. The only couplings which have been left out are the ones related to the weak neutral current. These are obtained as $C_V = T_3 - 2\sin^2 Q$ and $C_A = T_3$ for the vector and the axial part, respectively.

The first generation is the lightest one and contains the components of ordinary matter (i.e. the atoms). The formal structure of this generation is

$$\begin{pmatrix} \nu_e \\ e_L^- \end{pmatrix} \quad e_R^- \quad \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad u_R \quad d_R$$

In fact, the quark states with well-defined mass are not the same states that enter into the doublets and singlets of the SU$(2)_L$ group, as we described above. However, this is usually neglected in the presentation of the formal struture of the generations[7]. Additionally, all the antiparticles are included in the corresponding structure. It should be mentioned that the left-symmetry of the particles is translated into a right-symmetry of the antiparticles, due to the standard convention to define helicity states.

The second generation is identical but with the muon and the muonic neutrino in the lepton sector and the strange and charm quarks. For the third generation we are left with the tauon and its neutrino, the bottom and top quarks. No direct or indirect hint of additional generations has been found in the experimental analysis.

Table 1.2 shows the most important characteristics of the various particles. It should be noted that the colour charge is the same for all the quarks and for this reason is not included. Additionally, the antiparticles have the same quantum numbers but with the opposite sign.

On the interaction side, five different gauge bosons are present in the Standard Model. In principle, eight different gluons exists in Nature, but they can be simplified to one since colour charge is not observable. Table 1.3 shows the parameters associated to all the gauge bosons.

Since the Standard Model is founded on Quantum Field Theory, the processes it describes undergo quantum-mechanical effects. One example is that the neutral-current interaction via

[7] Obviously, the problem is that the three generations cannot be completely separated since the quark sector is mixed.

	Particle	Mass	Interaction
gluon	g	0	strong/colour $SU(3)_C$
photon	γ	0	electromagnetic $U(1)_{em}$
weak bosons	Z^0	91.19 GeV	weak, neutral
	$W^\pm$	80.42 GeV	weak, charged

$$\left.\begin{array}{l} \text{photon} \\ Z^0,\ W^\pm \end{array}\right\} \text{electroweak} \left\{ \begin{array}{l} B^0 \rightarrow U(1)_Y \\ \left.W^3 \right\} SU(3)_L \end{array}\right.$$

Table 1.3: Bosons associated to the fundamental interactions in the Standard Model. At very high energies the weak and the electromagnetic interactions are unified and in fact they are replaced by another two separate interactions, as described in the text.

the exchange of a neutral boson proceeds through all the possible bosons which can couple to the interacting particles. When an electron, for example, is scattered by a hadron, contributions from the photon and the Z^0 boson are present. Since the two contributions are added to compute the amplitude of the process, then the final observable will have not only two separate contributions, but also the interference term. This contribution is experimentally negliglible at energies smaller than the mass of the Z^0 boson, but it becomes very important when approaching that energy scale.

Apart from all these particles, the Standard Model includes an additional scalar boson, the Higgs particle, as we described before. This is the only particle which is left to be discovered from the complete spectrum of the Standard Model.

The success of the Standard Model to describe the experimental results during the last two decades is a great thriumph of theoretical physics in describing the fundamental interactions. However we cannot forget that this model still presents some problems and that gravity has to be included in the description of the fundamental forces. Apart from this, new questions and even some of the old questions still remain to be answered, so it is clear that the Standard Model is only a step in the understanding of Nature.

Some experimental results during the last years have shown that the Standard Model has to be completed. Among these, the mixing between neutrinos is the most important one. These results imply that neutrinos are massive and that is at present not included in the Standard Model.

Some studies have been done to construct models beyond the Standard Model. The most interesting ones unify all the known interactions at very high energies. The mechanism for "spontaneous symmetry breaking" is then applied to separate the interactions into the ones we observe at low energy.

However, at present the Standard Model seems to be enough to describe the experimental results. The presence of a well-established theory is used in the experimental side to determine the parameters of the theory, in addition to the tests which are performed to confirm the validity of the predictions in unexplored regions.

1.2 Deep Inelastic Processes

One of the most useful experiments to study the internal structure of hadrons is the scattering of leptons on hadrons, usually protons or nuclei. When the momentum transfers are very large, the process is said to be in the deep inelastic scattering (DIS) regime. In this case the target loses its identity completely and the resulting products are complicated multiparticle states whose study allows to gain insight into the internal structure of the initial target.

The following sections describe how the DIS processes have been used in the past as a fundamental tool to build the Standard Model. At the same time, the notation which will be used during the description of the analysis is introduced.

1.2.1 DIS processes

Usually, a DIS process is related to the scattering of a high-energetic lepton (l) on a hadron (h), which leads to the general reaction

$$l(k) + h(p) \to l'(k') + X(p') \; ,$$

where the initial state particles (l, h) are defined by the experiment and the final ones (l', X; where X means any state of particles) are given by the processes which have occured in the collision.

At HERA the initial state particles are electrons (or positrons) and protons, so we will concentrate on the related processes. Any other DIS process can be described in a very similar way.

When an electron is scattered on a proton, two major processes can occur depending on the boson which is exchanged in the collision (see figure 1.3). If this particle does not carry electric charge, the process is called *neutral current DIS*. If, on the other hand, it carries electric charge, then is called *charged current DIS*. To distinguish between the two process the only information we need is the type of the final-state lepton. If it's the same as the initial one, a neutral current event has happened. If not, the electric charge of the incoming lepton has been taken by the exchanged particle and the final-state particle cannot be the same as the initial one; the electron (positron) has transformed into a (anti-)neutrino via the exchange of a $W^{\pm}$ boson.

In the following, only neutral current processes will be considered, although most of the arguments can be applied to charged current studies.

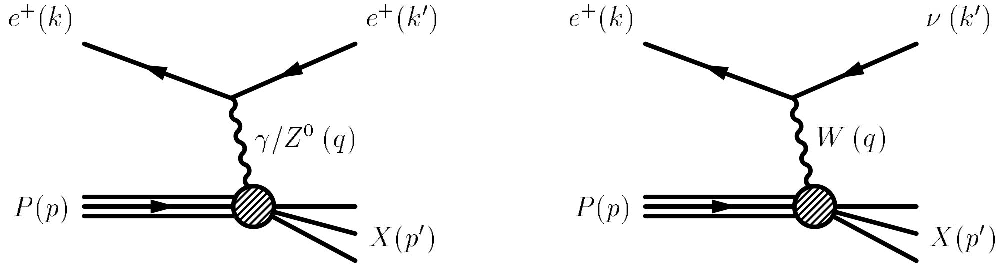


Figure 1.3: Feynman diagrams for the lowest order of neutral and charged current DIS interactions with a positron being scattered by a proton.

At a given center of mass energy ($\sqrt{s}$) the kinematics of a DIS event are completely described by two variables among the following three Lorentz-invariant quantities:

$$Q^2 = -q^2 = -(k - k')^2 \tag{1.6}$$

$$x = \frac{Q^2}{2\, p \cdot q} \tag{1.7}$$

$$y = \frac{p \cdot q}{p \cdot k} = \frac{Q^2}{s\, x} \text{ (ignoring masses)} \tag{1.8}$$

where the notation for the four-momenta is that of in figure 1.3. Q^2 is the negative square of the momentum transfer and specifies the virtuality of the exchange boson. If $Q^2 > 0$ the exchange boson can have longitudinal as well as transverse polarization. In this case Q^2 set the energy scale of the process. If $Q^2 = 0$ then the process takes place in the photoproduction regime, in which the proton collides with a real photon. The photoproduction regime is studied in a slightly different way within the context of DIS processes and we don't consider it further since the analysis is done in a region where Q^2 is well above 1 GeV^2.

x is the scaling variable introduced by Bjorken. In the Quark-Parton Model it is interpreted as the fraction of the proton momentum carried by the struck quark (see section 1.2.2).

It can be shown (see [3], for example) that at a fixed x, the exchange boson probes smaller distances with increasing Q^2 since the wavelength of the boson is inversely proportional to $\sqrt{Q^2}$.

Finally, in the proton rest frame y corresponds to the fractional energy transfer from the lepton to the proton.

In the single boson exchange approximation the cross section for deep inelastic ep scattering can be factorized into a leptonic tensor $L_{\mu\nu}$ and a hadronic tensor $W^{\mu\nu}$

$$d\sigma \sim L_{\mu\nu}\, W^{\mu\nu}$$

The leptonic tensor,which is symmetric in μ and ν, can be calculated exactly using the Standard Model. The ignorance of the structure of the proton and hence the details of the interaction at the hadronic vertex are parametrised in terms of the hadronic tensor. The most general form of this tensor, taking Lorentz-invariance and the symmetry of the leptonic tensor into account is [3]

$$W^{\mu\nu} = -W_1 g^{\mu\nu} + W_2 \frac{p^\mu p^\nu}{M^2} + iW_3 \epsilon^{\mu\nu\rho\sigma} p_\rho q_\sigma + W_4 \frac{q^\mu q^\nu}{M^2} + W_5 \frac{p^\mu q^\nu + q^\mu p^\nu}{M^2} ,$$

where M is the mass of the hadronic target, in our case the proton.

Imposing the conservation of the four-vector current yields

$$W_5 = -W_2 \frac{p \cdot q}{q^2} \quad \text{and} \quad W_4 = -W_5 W_2 + W_1 \frac{M^2}{q^2}$$

The three remaining functions depend on two independent Lorentz-invariant scalar variables, which can be chosen to be Bjorken x and Q^2. The dynamics of the strong interaction and related processes in the hadronic vertex are contained in these x and Q^2 dependencies. Nowadays, the notation given by the definitions

$$\begin{aligned} F_1(x,Q^2) &= M \cdot W_1(x,Q^2) \\ F_2(x,Q^2) &= \frac{Q^2}{2M\,x} \cdot W_2(x,Q^2) \\ xF_3(x,Q^2) &= \frac{Q^2}{2M} \cdot W_3(x,Q^2) \end{aligned}$$

is commonly preferred. The $\{F_i\}$ are the so-called *proton structure functions*. As $F_3(x,Q^2)$ describes the parity violation contribution this structure function is small for $Q^2 \ll {M_Z}^2$.

Using these definitions, the deep-inelastic $ep \to e + X$ scattering cross section can be written as

$$\frac{d^2\sigma(e^\pm p)}{dx\,dQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[y^2xF_1(x,Q^2) + (1-y)F_2(x,Q^2) \mp y(1-\frac{y}{2})xF_3(x,Q^2)\right] \tag{1.9}$$

or with the definition of $F_L = F_2 - 2xF_1$ as

$$\frac{d^2\sigma(e^\pm p)}{dx\,dQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[Y_+F_2(x,Q^2) - y^2F_L(x,Q^2) \mp Y_-xF_3(x,Q^2)\right], \tag{1.10}$$

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

In the single-photon exchange approximation (within QED), the structure function F_L is only related to the absorption cross section of longitudinally-polarised “virtual” photons. In the quark-parton model, this absorption process is forbidden and $F_L = 0$ (see section 1.2.2).

1.2.2 The Quark-parton model and Bjorken scaling

Already in 1968, Bjorken predicted that the structure functions would depend only on one dimensionless variable in the high energy limit $Q^2 \to \infty$ with finite x (the “Bjorken limit”). This behaviour, given by

$$F_i(x,Q^2) \to F_i(x), \tag{1.11}$$

is kwown as *scale invariance* or Bjorken scaling [13, 14]. As we mentioned earlier, the resolution power of the exchanged particle increases with Q^2. If the proton consists of point-like constituents, it turns out to be possible to observe Bjorken scaling using Q^2 finite since an increase in the resolution does not change the way a point is observed.

The experimental results from SLAC [15] showed the scaling behaviour very clearly (figure 1.4) for F_2.

According to Feynman's parton model [16], the proton is composed of free point-like constituents, called partons. In this model the deep-inelastic ep scattering cross section is the incoherent sum of quasi-elastic e-parton scattering, that is

$$\frac{d\sigma(e^\pm p)}{dx\,dQ^2} = \sum_i \frac{d\sigma}{dx\,dQ^2}\,(e^\pm - \text{parton-}i) \tag{1.12}$$

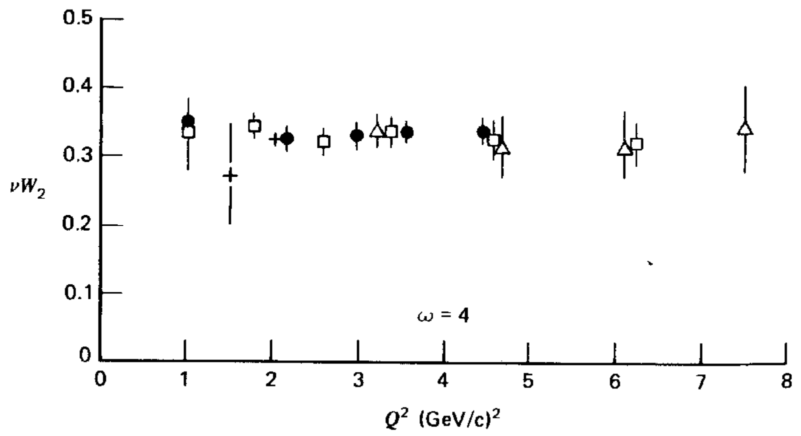


Figure 1.4: The structure function F_2 (here denoted νW_2 where $\nu = \frac{Q^2}{2M\,x}$ is the energy of the exchanged boson) is shown as a function of Q^2 for $\omega = \frac{1}{x} = 4$ as measured at SLAC. The scaling behaviour is evident.

In the proton's infinite momentum frame all transverse momenta are negliglible and the Bjorken scaling variable x receives a simple interpretation: x corresponds to the fractional longitudinal proton momentum carried by the struck parton, i.e. the parton which interacts with the boson. This can be verified neglecting the proton and parton mass. For the parton after colliding with the boson

$$(\xi p + q)^2 = m^2 \approx 0$$

where ξp is the four-momentum of the initial parton in the proton. Operating with this expression

$$0 = (\xi p)^2 - Q^2 + 2\xi(p \cdot q)$$

$$\text{thus: } \xi = \frac{Q^2}{2\, p \cdot q}$$

which is equal to the definition given above for x (Eq. (1.7)).

Taking Eq. (1.12) into account, the structure function F_i corresponds to the sum of the parton momentum distribution $x \cdot f_i(x)$ weighted with the square of the electric charge of the parton[8]

$$\begin{aligned} F_1(x) &= \frac{1}{2}\sum_i {e_i}^2 f_i(x) \\ F_2(x) &= \sum_i {e_i}^2 x f_i(x) = 2x\; F_1(x)\;, \end{aligned} \tag{1.13}$$

where the functions $\{x f_i(x)\}$ are called *parton distribution functions* or simply *parton densities*, and describe the probability of finding a parton of type i in the proton carrying a fraction x of the proton momentum.

The relation between F_1 and F_2 is known as the Callan-Gross relation [17]. It holds exactly only for $\frac{1}{2}$-spin partons which cannot couple to longitudinally polarised photons. The experimental

[8]Note that in this discussion, the single-photon exchange approximation is considered in order to simplify the expressions.

confimation of this relation allowed the identification of Feynman's partons with Gell-Mann's quarks and the model was called the Quark-Parton Model (QPM). It should be mentioned that the fractional charge of the partons was confirmed using neutrino-nucleon scattering and the postulated number of three valence quarks in the proton and neutron was experimentally confirmed using the Gross-Llewellyn-Smith sum rule[9]

$$\int_0^1 \frac{dx}{2}\left(F_3{}^{\nu p}(x) + F_3{}^{\nu n}(x)\right) = \int_0^1 dx(u_v + d_v) = 3 \tag{1.14}$$

which experimentally was found to be 3.2 ± 0.6 [18].

If the proton consisted only of charged quarks, their momentum would be expected to add up to the proton momentum, but experimentally it was shown that half of the momentum of the proton is carried by neutral partons. Direct evidence for the existence of these partons, called gluons, was provided at DESY via the observation of three-jet events in e^+e^- annihilation [19].

1.2.3 DIS and Quantum Chromodynamics

Although the Quark-Parton Model gives a good description of the experimental results for lepton-nucleon interaction at first order it cannot explain how the partons are bound together to form the nucleon or other hadrons if the quark structure of hadrons introduced by Gell-Mann and Zweig is considered. It is necessary to add a force between the quarks/partons in such a way they are bound to form the hadrons.

Thus, the naive QPM has to be completed to get the QCD theory where the picture shows not only the three quarks which determine the quantum numbers of the proton (valence quarks). The proton is formed by quarks interacting through gluons, and radiating gluons which can split into quark-antiquark pairs (the so-called sea quarks) or more gluons (see figure 1.5).

In the 1970's QCD was developed as the field theory governing the strong interaction between quarks and gluons and therefore being the basis for the interaction between hadrons including the nuclear forces. QCD is a non-abelian gauge theory based on the SU(3) symmetry group. Each quark has three possible "charges" related to the colour quantum number and the strong interaction can be mediated via the exchange of 8 different coloured gauge particles, called gluons. The gluon always carry a combination of colour and anti-colour and the emission of one gluon change the colour number of the quark (anticolour for the antiquarks). Since the symmetry group of the colour quantum number ($SU(3)_C$) is non-abelian, the gluons also carry a net colour charge themselves and therefore couple to each other. This is the main difference with QED[10], in which the photon is a neutral particle (and the group is the abelian $U(1)_{em}$).

The self-coupling of the gluon is the reason for the strong (QCD) coupling constant α_s to become large at small energies (large distances) and to decrease at high energies (small distances). This property of QCD is used to explain the observed behaviour of the strong interaction:

- **Asymptotic freedom**: at high energies (or small distances) the strong interaction proceeds via colour fields of reduced strength and the quarks and gluons behave as essentially free, non-interacting particles.

[9]This relation counts the number of valence quarks in the nucleon.

[10]In fact this is the only difference, since all the others are consequence of this non-abelian structure of QCD.

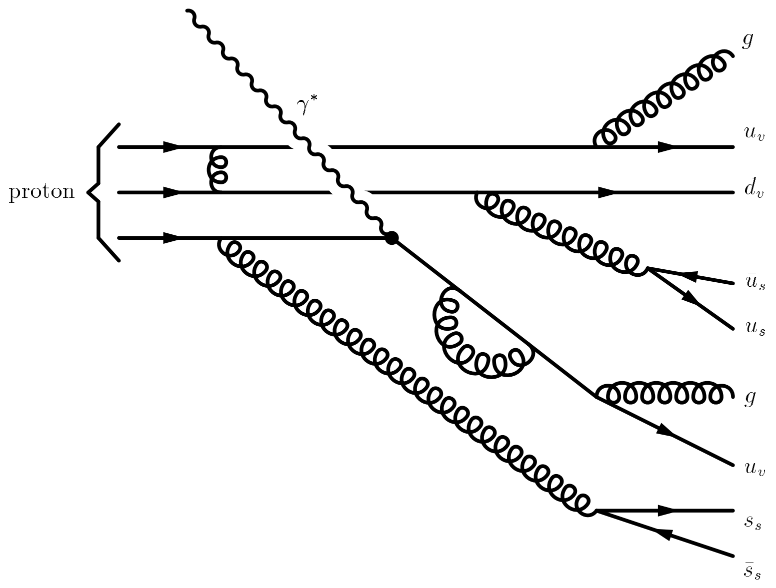


Figure 1.5: The structure of a proton interacting with a virtual photon. In addition to the three valence quarks defined in the quark model, the proton is made up of gluons and slow debris consisting of quark-antiquark pairs (the so-called sea quarks).

This property is needed to explain why in deep-inelastic scattering the hard interaction between the boson and the parton can be calculated with non-interacting partons during the time scale in which the interaction takes place. That is, this property explains the success of the naive QPM approach even considering that partons are not free in time-scales longer than that of the hard-interaction in deep-inelastic scattering.

As we will see in section 1.3, this property allows the application of perturbation theory to calculate scattering amplitudes.

- **Infrared slavery**: at low-energy (or large distances) the strength of the colour field is increasing (the potential behaves as $V(r) \sim \lambda r$) and in this way the quarks and gluons can never escape. If two interacting partons are separated, the energy of the field increases so much that it creates new interacting particles and at the end (that is, at long time scales) we are left with colourless hadrons containing the partons. This is the way QCD explains the **confinement** of the partons, which are not observed as free particles at long time scales.

  It should be noted that interacting partons can be separated and observed as different objets, although each of these partons will be confined in a different hadron. What is not observed as free particles are coloured partons, although the partons, as well as their quantum numbers (except the colour), can be studied experimentally.

This scale dependence of the strength of the interaction is translated into a dependence of α_s with energy, which is given by the so-called renormalization group equation and the β function [11]

$$\mu^2 \frac{d\alpha_s}{d\mu^2} = \beta(\alpha_s) = -b{\alpha_s}^2 \left(1 + b_1\alpha_s + \dots\right), \tag{1.15}$$

where n_f is the number of active quark flavours (whose mass is smaller than the energy scale μ). The coefficients are calculated to be

$$b = \frac{33 - 2n_f}{12\pi}$$
$$b_1 = \frac{153 - 19n_f}{2\pi\,(33 - 2n_f)}$$
$$\ldots$$

and α_s decreases with incresing the energy scale as long as $n_f < 17$. This is the first step for α_s to exhibit asymptotic freedom.

A solution of equation (1.15), to first order in an expansion in α_s, is

$$\alpha_s(\mu) = \frac{12\pi}{2\,(33 - 2n_f)\,\ln\left[\frac{\mu}{\Lambda_{QCD}}\right]} \tag{1.16}$$

which shows very clearly the asymptotic freedom property of QCD ($\alpha_s(\mu) \to 0$ for $\mu \to \infty$). The infrared slavery cannot be derived from this expression since it was obtained from a perturbative expansion valid when α_s is small enough. The QCD scale parameter (Λ_{QCD}) is the energy scale at which α_s becomes large and the perturbative approach is no longer valid.

Coming back to the structure functions of the proton, the inclusion of the gluon and the strong interaction dictated by QCD introduces some modifications of the conclusions derived in the previous section. The most obvious feature is that gluon radiation by the quarks results in a transverse momentum component on the partons. Consequently, quarks can couple to longitudinally polarised photons and the Callan-Gross relation is no longer satisfied exactly. The longitudinal structure function is non-zero and, due to its origin, is largely dependent on the gluon content of the proton and is therefore considered to be a good measure of the latter [11].

A more important consequence of the gluon content of the proton are the so-called *scaling violations* of the structure functions, which exhibit a logarithmic dependence on Q^2 at fixed x. In QCD, this is interpreted as follows: at low Q^2 the resolution of the exchanged boson is rather broad, so that only just the quark substructure of the proton is "seen". At high Q^2, however, a finer resolution is achieved and quark-antiquark pairs originating from radiated gluons can be resolved. This means that the history of a quark before it interacts with the boson becomes very important. It could radiate a gluon and thus, although the quark which is struck has momentum fraction[12] x, the quark originally had a larger momentum fraction $y > x$ (this kind of process is known as the QCD Compton process). Alternatively it may be that a gluon with momentum fraction y produced a $q\bar{q}$ pair and one of these became the struck quark of momentum fraction x (this process is known as Boson Gluon Fusion, BGF).

Although the QPM description is not exact, the analytical tools described in section 1.2.2 are still useful. Specifically, the parton distribution functions can be kept including a dependence on the scale at which the proton substructure is resolved, $xf_i(x) \to xf(x, Q^2)$.

Considering all these modifications, now we should take into account that the quark and gluon distribution functions $yf(y, Q^2)$ for all momentum fractions y such that $x < y < 1$ contribute

[11] Unfortunately, the measurement of F_L is experimentally very difficult, particularly at HERA [20].

[12] It should be noted that x cannot longer be related to the fraction of the proton momentum carried by the parton, but it is still a useful kinematic variable.

to the given reaction via QCD-Compton or BGF processes. At large x, where valence quarks dominate, the quark density and hence F_2 falls with Q^2 as a result of the gluon radiation, while at small x the amount of $q\bar{q}$ pairs which configures the partonic 'sea' increases, so that F_2 rises with Q^2. These scaling violations in conjunction with a strong rise at small x for fixed Q^2 which are clear predictions by QCD have been observed experimentally, specifically by the H1 and ZEUS experiments at HERA [21, 22], as can be seen in figure 1.6.

It should be noted that the presence of all these effects in the description of DIS processes is what allows the study of the strong interaction using the multihadronic final states in ep collisions.

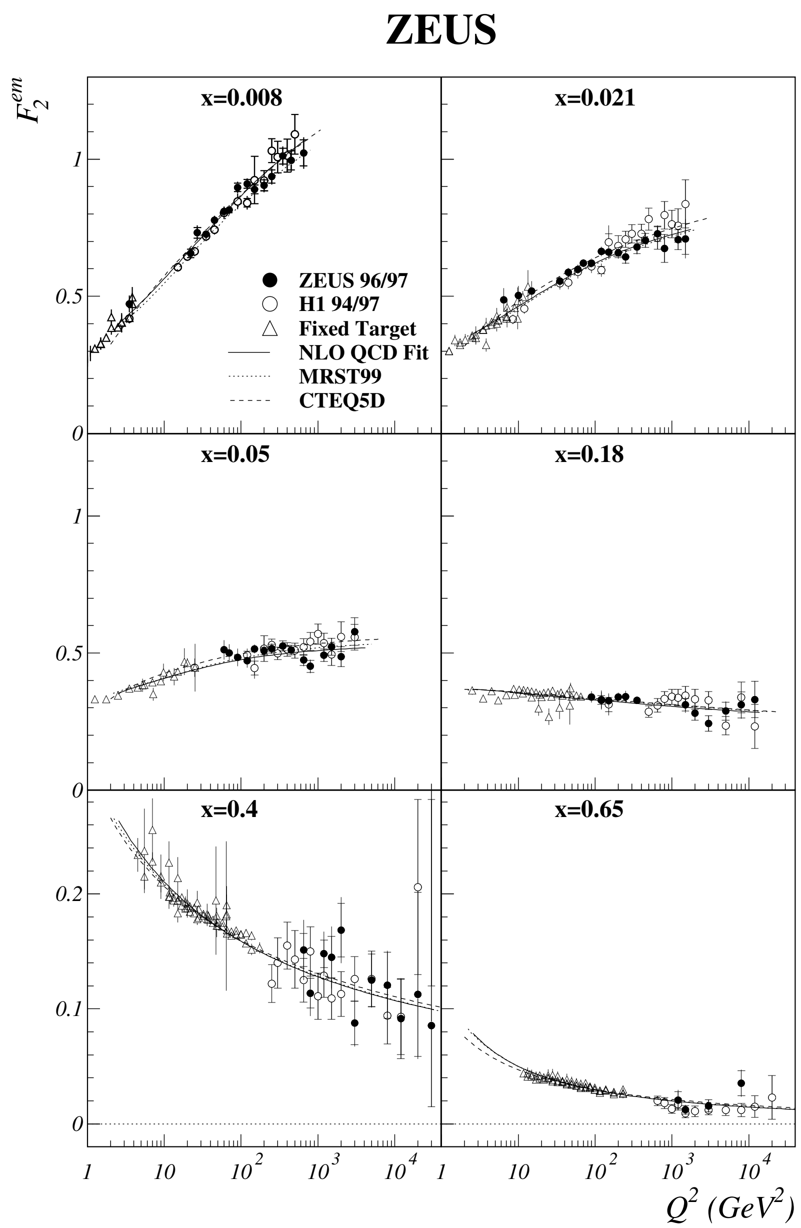


Figure 1.6: The results from ZEUS (solid points) for F_2^{em} versus Q^2, for six values of x, are compared with results from NMC, BCDMS, E665 (triangles) and H1 results (open symbols) [22]. A behaviour close to Bjorken scaling for intermediate x values and Q^2 not very small is observed. The scaling violations predicted by QCD are clearly observed for low x and high x values.

1.3 The perturbative QCD framework

As was already discussed in the previous section, QCD gives a very successful description of the strong interaction, containing in a very natural way the properties of asymptotic freedom and confinement. However, these properties make specially difficult to obtain predictions for the soft limit, where hadrons live and are observed in the experimental conditions. Nevertheless, the coupling constant can be kept small in some process and the usual perturbative approach can be applied giving rise to a predictive power never achieved before in hadron physics. This section gives a description of the way perturbative QCD (pQCD) is actually implemented in order to provide quantitative predictions.

1.3.1 The perturbative QCD approach

From the Lagrangian density for QCD different couplings appear in the theory. They can be related to the following processes:

- **A (anti)quark radiating a gluon** which is dictated by the quark-gluon coupling which, in turn, is formally equal to the coupling between an electric charge and the electromagnetic field. The same coupling is related to the process in which a **gluon splits into a quark-antiquark pair**.
- **A gluon splitting into more gluons** which is dictated by the self-interaction coupling of the gluon.

These processes are proportional to the strong coupling constant, so if the energy scale associated to the process is large enough, then the coupling is small and the interaction can be described by a perturbative expansion. This expansion allows the equations describing the processes to be solved in the limit where the strong coupling constant is small (*"hard processes"*). On the other hand, if the energy scales are small, i.e. the interacting partons become further apart, then the strong coupling constant becomes very large and the perturbative approach is no longer valid.

In order to be able to make predictions for the measurements performed in a collider, QCD should be studied in reactions in which the hard processes are dominant versus the non-perturbative ones. Usually the hard process occurs as a first step in the collision and afterwards, the soft, non-perturbative, process gives rise to the final-state hadrons (*hadronization*) from the partons taking part in the hard process. Perturbative QCD is applied to obtain predictions for the partons dynamics, and the results can be considered reliable if the hadronization process does not hide the properties of the partonic state. That is, if the experimental measurement can be related to the parton dynamics in some direct way.

Even in the region where perturbative QCD can be applied, additional difficulties are present in the calculations. In the same way as in QED, different types of divergences appear when the vacuum fluctuations (*UV divergences*) or radiation of partons with very small relative transverse momentum with respect to the original parton (*IR divergences*) are included. Both are dealt with by following the standard regularization procedures. This mechanism is very often called *renormalization* when it is applied to get rid of the UV divergences. To remove the IR divergences the regularization is made selecting observables which are insensitive to these particular divergences. This type of observable is called *infrared-safe* and usually is constructed in such a way that the result doesn't change in case an IR-divergent process can appear in the calculation.

The use of infrared-safe observables allows a complete separation of the hard process from the soft process in a formal way, and in this way perturbative QCD predictions can be obtained for those observables, which then avoid the divergences without any additional treatment.

When the interaction contains hadrons in the initial state, as in the case of deep-inelastic scattering, it is imposible to be fully insensitive to the soft phenomena (including IR divergences) occuring within the hadrons, or in the steps leading from the free, initial hadron to the hard interaction. In that case, a complete regularization mechanism is needed to separate (factorize) the theoretical description into a calculable and a soft parts. This mechanism is called *factorization* and is very similar to the renormalization procedure. We will describe both in the next sections.

1.3.2 Renormalization

The UV divergences arising from loop integrals associated to vacuum fluctuations are dealt with by renormalizing the bare parameters in the lagrangian, which are not observable, to the physical (observable) ones. The procedure is not unique, i.e. there is a renormalization scheme choice which is chosen depending on the properties of the parameters which are needed in the calculation. Of course, the physical results cannot depend on this choice, although a theoretical dependence is usually obtained in pQCD since the perturbative expansion is performed only to a given order. The renormalization scheme determines the values of the physical parameters when calculated to a given order in the pertubative approach. The $\overline{MS}$ renormalization scheme is the most widely used[13].

In the renormalization procedure a scale μ_R is needed to set the scale at which the renormalization is performed. At a fixed order in perturbation theory, a dependence on this scale is introduced in the predicted value of the observable. Of course, any physical observable should be invariant under variations of this scale. This is formally expressed using the renormalization group equation

$$\mu_R{}^2 \frac{d}{d\mu_R{}^2} \mathbf{\Gamma} = 0 \; , \tag{1.17}$$

where Γ is the observable we are interested in. In the perturbative approach, this equation has to be applied to the perturbative expansion of the observable, and in pQCD we have

$$\mathbf{\Gamma} = \mathbf{\Gamma_0} + \alpha_s \, \mathbf{\Gamma_1} + \alpha_s{}^2 \, \mathbf{\Gamma_2} + \cdots$$

and the renormalization group equation transforms into

$$\mu_R{}^2 \frac{d}{d\mu_R{}^2} \sum_{j=0}^{N} \alpha_s{}^j \mathbf{\Gamma_j} = \mathcal{O}(\alpha_s{}^{N+1}) \; , \tag{1.18}$$

that is the variation of the observable with the scale is given by terms which were not included in the perturbative expansion.

[13]Another common choice is the DIS scheme which has the property that at any order in the perturbative expansion the structure functions are computed as linear combinations of the quark distribution functions, as shown in Eq. (1.13). This property is not preserved in other renormalization schemes, including the $\overline{MS}$.

It is very common to use this relation to estimate the size of the theoretical uncertainty of the observable coming from the truncation of the perturbative expansion at order N in α_s. This estimation is obtained analyzing the variation of the pQCD theoretical prediction with the renormalization scale; that is by computing the prediction for different choices of this scale one can estimate the size of the truncated orders in the series.

1.3.3 Factorization

In order to separate the infrared-divergent phenomena within the hadrons in the initial state from the hard interaction calculable in pQCD, the soft processes are isolated and factorized out in the parton distribution functions (pdf's). The pdf's are not calculable in pQCD, but can be determined from the experimental results. The universality of the parton distribution functions allows to make predictions for a process after measuring the pdf's in another, so that the predictivity of the theory is not lost. The hard-scattering factorization has been proved to hold for a different number of processes and it is summarized in the following expression for a cross section (although it can be generalized to any obervable) for lepto-hadron scattering:

$$\sigma(P_1, P_h) = \sum_i \int d\xi\, f_i(\xi, {\mu_F}^2)\, \hat{\sigma}(P_1, \xi P_h, \alpha_s(\mu_R), \mu_R, \mu_F) \tag{1.19}$$

where P_1 is any parameter to characterize the final state; P_h is the four-momentum of the incoming hadron; the sum is over all the flavours, f_i are the parton distribution functions which have to be evaluated at the *factorization scale* μ_F; and $\hat{\sigma}$ is the lepton-parton cross section which is the term calculable by pQCD.

Thus, the factorization procedure introduces an additional scale and the prescription to be followed (factorization scheme) is not unique, although the physical results are independent on the scheme and on the factorization scale. In a similar fashion to the renormalization procedure, the parton distribution functions are regularized to bury IR divergences coming from configurations where soft and collinear partons are emitted.

This factorization scale is related to the scale that separates soft and hard processes occuring in the parton system which configures the initial-state hadron. In practice, this scale separates what occurs "inside" and "outside" the hadron, being the "outside" region where the hard process occurs and where the perturbative approach can be applied.

Perturbative QCD is not able to give the value of the parton distribution functions, but it is possible to obtain predictions about the variation with the scale that arises from the factorization procedure using the evolution equations (see section 1.3.4). The equations provide resummations of terms from the perturbative expansion and introduce a known dependence of the $\{f_i\}$ on the scale.

For some observables, like the ones used in hadron production, the theoretical predictions are not completely infrared-safe since the process transforming the final-state partons into the observable hadrons (*fragmentation*) is not free of new divergences. The kind of divergences that appear are the same as those removed using the factorization mechanism, but now related to the final state processes. When this kind of observables is under study, a new regularization process is needed, so as to absorbe all these divergences into the *fragmentation functions*, which again factorize in the expression to compute the observable. These functions are the equivalent of the parton densities, but now giving the information of the probability for a parton to transform into a given hadron with the corresponding momentum fraction and energy scale. Since in the analysis with jets, this kind of regularization is not needed and the observables are very independent of the fragmentation functions[14], we don't go deeper into this aspect of the factorization procedure for obtaining perturbative predictions in QCD.

The isolation of all divergences for any process in deep-inelastic scattering has left as inheritance the residual and non-physical dependence of the prediction on, at least, two scales. These dependencies characterize the precision of a calculation and if the changes produced by the scale variations are large, the next order(s) has to be calculated to gain accuracy and sometimes predictive power in the theoretical predictions. In some cases, some terms which appear at all orders are identified and resummed, improving in this way the accuracy of the predictions. An example of this resummation are the evolution of the pdf's given by the evolution equations or the calculation of the running of the strong coupling constant.

1.3.4 Evolution equations

The interactions between partons inside the hadrons lead to a dependence of the predictions on the partons with momentum fraction of the hadron larger than x, as we described above (see Section 1.2.3). This dependence can be introduced in the following way: the parton density associated to one parton type is given by the processes which directly or indirectly generate partons of that type. For example, the gluon content is partly given by the probability of the quarks to radiate a gluon.

As we already mentioned, perturbative QCD is not able to predict the values for the parton distribution functions $\{f_i(x,\mu^2)\} = \{q_i(x,\mu^2), g(x,\mu^2)\}$ but it can give predictions for its evolution. That is, given all the parton densities at a given scale for all x, pQCD can predict the values for each parton density at other scale.

To do this, pQCD allows the construction of a set of integro-differential equations known as "Altarelli-Parisi equations" [23], which can be formally written as follows, where the quark and gluon content of the proton is distinguished, although formally both follow the same equation:

$$
\begin{aligned}
\frac{dq_i(x,Q^2)}{d\ln Q^2} &= \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dy}{y}\left[q_i(y,Q^2)\mathcal{P}_{qq}(x/y) + g(y,Q^2)\mathcal{P}_{qg}(x/y)\right], \\
\frac{dg(x,Q^2)}{d\ln Q^2} &= \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dy}{y}\left[\sum_i \left(q_i(y,Q^2)\mathcal{P}_{gq}(x/y)\right) + g(y,Q^2)\mathcal{P}_{gg}(x/y)\right],
\end{aligned}
\tag{1.20}
$$

where the splitting function $\mathcal{P}_{kl}(z)$ represents the probability of a parton l emitting a parton k with momentum fraction z of the parent parton.

It is easy (see [3, 11]) to obtain the pQCD expressions for theses splitting functions at first order in α_s by simply analyzing the radiative process and the associated couplings for the radiative processes contributing to these evolution equations, which are shown in figure 1.7.

[14] In the studies of jets, the effects of fragmentation processes are estimated by studying the differences of applying the jet algorithm to the partons and to the hadrons. Of course, the differences are given by the different models of the hadronization process, which indirectly include the fragmentation functions.

The expressions associated to these diagrams lead to the results

$$
\begin{aligned}
\mathcal{P}_{qq}(z) &= \frac{4}{3}\,\frac{1+z^2}{1-z} \\
\mathcal{P}_{qg}(z) &= \frac{1}{2}\,\left[z^2+(1-z)^2\right] \\
\mathcal{P}_{gq}(z) &= \mathcal{P}_{qq}(1-z) = \frac{4}{3}\,\frac{1+(1-z)^2}{z} \\
\mathcal{P}_{gg}(z) &= 6\,\left[\frac{z}{1-z}+\frac{1-z}{z}+z(1-z)\right]
\end{aligned}
\qquad (1.21)
$$

In the context or perturbative QCD two types of terms can become large and hence important in the perturbative expansion in $\ln Q^2$ and in $\ln(1/x)$. Depending on which terms are considered on each case the evolution equations lead to different results and different regions of application.

DGLAP evolution

The equations describing the so-called DGLAP evolution (Dokshitzer, Gribov, Lipatov, Altarelli and Parisi) [23, 24] are formally derived in the leading logarithm approximation (LLA), where terms of the form ${\alpha_s}^n \cdot (\ln Q^2)^n$, which give the dominant contribution at large Q^2 and large x, are summed to all orders. In a field theory having asymptotic freedom such an approximation proves to be asymptotically exact. The amplitude for the inelastic *ep* scattering process can be obtained as the sum of ladder diagrams of consecutive gluon emissions (figure 1.8) thus losing, gradually, its longitudinal momentum. If the rungs of the gluon ladder are labelled 1 to n from the proton to the photon, the fraction of longitudinal momentum x_i carried by the rungs are ordered

$$x_1 > x_2 > x_3 > \ldots > x_n \ ,$$

while the transverse momenta of the emitted gluons increase strongly as going up the ladder

$${k_{t,1}}^2 \ll {k_{t,2}}^2 \ll \ldots \ll {k_{t,n}}^2 \ll Q^2 \ .$$

The solution of the DGLAP equations give the parton distributions as a function of x at any Q^2, provided their x dependence at an input scale ${Q_0}^2$ is known. The latter cannot at present be calculated, they are determined experimentally.

At small x, the $\ln(1/x)$ terms entering the cross section become important. In the moderate region, where $\alpha_s(Q^2)\ln Q^2 \ll 1$ and $\alpha_s(Q^2)\ln(1/x) \ll 1$, but $\alpha_s(Q^2)\ln Q^2 \ln(1/x) \sim 1$, the so-called Double Leading Logarithm Approximation (DLLA) is used. In this approach leading terms in $\ln(1/x)$ are summed when they are accompanied by a leading $\ln Q^2$.

In the solution of these equations, the gluon density presents a very steeply rise at small x, a rise which is faster than any power of $\ln(1/x)$, but slower than any power of x. However, over a limited x, Q^2 range, as in the case of HERA, it may mimic a power behaviour.

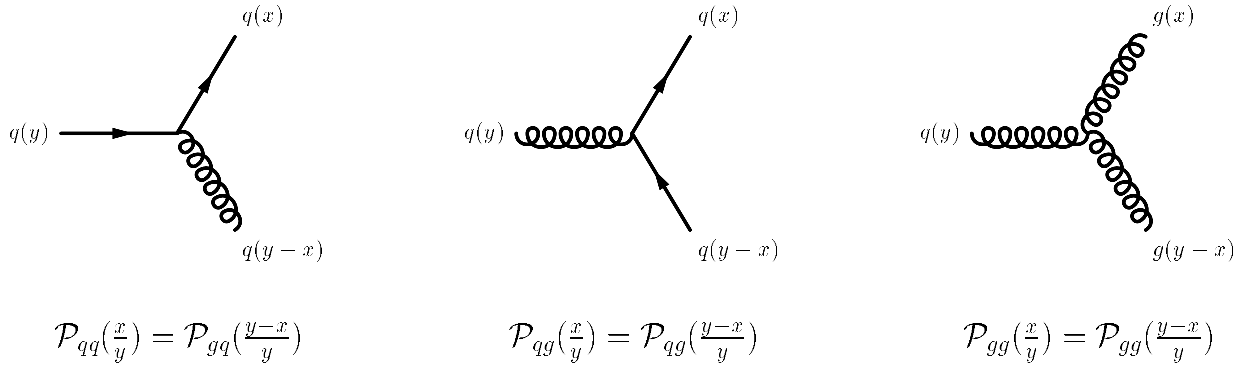


Figure 1.7: The splitting functions at lowest order in α_s. The diagram on the left shows the gluon radiation by a (anti)quark and the two diagrams on the right show the gluon splitting into two partons.

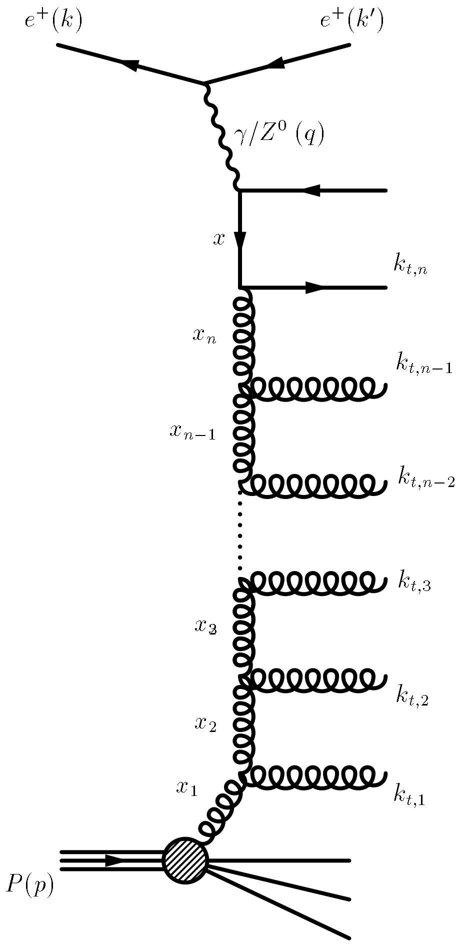


Figure 1.8: A ladder diagram initiated by a gluon in the proton. This kind of ladder illustrates how the parton evolution in the DGLAP equations is taken into account.

BFKL and other evolution equations

The steep behaviour of the gluon density and hence F_2 at low x has led to further developments in pQCD. In the low-x region the DLLA approach is not enough and it is also appropiate to sum contributions which are leading in $\ln(1/x)$ independent of $\ln Q^2$. This calculation has been done by Balitzki, Fadin, Kuraev and Lipatov [25], resulting into the BFKL evolution equations which are named after the authors. Summing $(\alpha_s \cdot (1/x))^n$ terms involves the evolution of a gluon distribution which is not integrated over k_t and the gluon ladder does not have to be ordered in k_t anymore (rather it involves a 'random walk in k_t').

The solution to the BFKL equation to leading order in $\ln(1/x)$ and fixed α_s has a very steep power law behaviour, which would violate unitarity in the limit $x \to 0$. However, the current form of the equations does not incorporate the running of α_s with Q^2 and the kernel has only been calculated to first order in $\ln(1/x)$. Furthermore cut-offs have to be introduced in the integral in the infrared and ultraviolet limit.

Due to all these problems and the difficulties of BFKL to deal with Q^2 evolution, progress has been made to achieve an 'unified' treatment of the x and Q^2 dependencies of the parton distributions and structure functions throughout the kinematic plane. The most important is the development of the so-called CCFM equation (Catani, Ciafaloni, Fiorani and Marchesini) [26]. It is based on the idea of coherent gluon radiation, which leads to angular ordering of gluon emission in the gluon ladder. The maximum angle of gluon emission is specified by an additional scale, which can be taken to be the Q^2 of the probing photon.

At small x the CCFM equation reproduces the BFKL behaviour, while at moderate x it reproduces the DGLAP equation for the integrated gluon density $q(x, Q^2)$ [27]. It should be mentioned that technically the CCFM equation involves similar considerations as the BFKL equation concerning the incorporation of the running of α_s and the UV and IR cut-offs.

Shadowing

The considerations so far have yielded a steeply rising gluon density at small x. If this were the asymptotic behaviour of the gluon density, the unitary bound would eventually be saturated or even violated.

On the other hand, all the evolution equations discussed above are linear integro-differential equations as they deal with the emission of gluon from quarks or the gluon splitting into quarks or additional gluons. However, as $x \to 0$ the gluon density becomes very large, and the self-coupling gluons may annihilate, or recombine into gluons. Such gluon *shadowing* or *screening process* may compete with the standard evolution and eventually saturate the gluon density.

A measurement at Q^2 probes a parton of transverse size $\sim 1/Q$. Assuming a homogeneous distribution of gluons in the proton, the $xg(x, Q^2)$ gluons occupy a transverse area of order $xg(x, Q^2)\frac{\pi}{Q^2}$ which increases with decreasing x. If this area gets comparable to the transverse area of the proton, shadowing effects are expected to emerge.

For a proton radius of 1 fm and $Q^2 \sim 10$ GeV^2, this limit is reached at $xg(x, Q^2) \sim 250$, which is well above the values found at HERA. Other models (like the scenario of 'hot spots' [28]) quote the relevant area to be smaller. In any case, no evidence for such an effect has so far been seen at HERA.

In order to take the recombination of gluons $gg \to g$ into account, Gribov, Levin and Ryskin added a quadratic correction to the evolution of the gluon distribution, yielding the so-called

GLR equation [29]. This equation can only be considered an approximation since possible interactions of gluons ladders before the recombination are not included and multi-ladder diagrams may also be important.

1.3.5 Modern parametrizations of parton densities

The measurement of structure functions serves two main purposes: to test QCD as the theory of the strong interactions and to extract the, as yet, uncalculable parton distribution functions and use them as input to obtain predictions for other processes. In this section the most important parametrizations of the proton parton densities are introduced. Most of them are obtained using the DGLAP evolution equations to evolve parton densities to the measured values of Q^2. The x dependence of the parton densities is parametrised from a starting scale ${Q_0}^2$. The parameters are determined by a χ^2 minimisation over data from structure functions in deep-inelastic e, μ or ν scattering, measurements of Drell-Yan production and W-asymmetry in $p\bar{p}$ collision as well as prompt photon production $pN \to \gamma X$.

Martin-Roberts-Stirling (MRS)

The MRS parametrization, obtained using the next-to-leading order DGLAP equations, uses the following functional form for the quark and gluon distributions at the starting scale ${Q_0}^2$ in the $\overline{MS}$ renormalization scheme

$$x f_i(x, {Q_0}^2) = A_i \cdot {x_i}^{\delta_i} (1-x)^{\eta_i} (1 + \epsilon_i \sqrt{x} + \gamma_i x) \,, \tag{1.22}$$

where f_i is a particular parton density and the set $\{A_i, \delta_i, \eta_i, \gamma_i\}$ are the parameters to be fitted. Not all the normalisation factors A_i are free parameters, but some are fixed from flavour or momentum sum rules. The charmed sea, assumed to be zero at the starting scale, is generated by gluon splitting as included in the DGLAP equations for massless partons.

The more recent parametrization of the parton densities coming from this group is the set donoted by MRST98 [30] (and the corrected version MRST99 [31]). It provides these parametrizations using very recent experimental measurements and the fits are done in different configurations, which allow the studies of the dependence on the α_s value or the uncentainty due to that of the gluon density for the perturbative QCD predictions. This kind of study has been performed in the analysis presented here and is explained in section 5.2.1.

CTEQ Collaboration parametrizations

The 'Coordinated Theoretical-Experimental Project on QCD' (CTEQ Collaboration) uses a very similar approach to the MRS group. For the quark distributions the same functional form as in equation (1.22) is chosen, whereas the gluon density is parametrised as

$$x g(x, {Q_0}^2) = A_0 x^{A_1} (1-x)^{A_2} (1 + A_3 x^{A_4}) \,. \tag{1.23}$$

CTEQ chooses a slightly different set of data for constraining the strange quark distribution. Furthermore, they do not suppress the charm distribution below the mass threshold but use the 'variable flavour technique'.

In 1996, the CTEQ Collaboration presented the set of distributions known as CTEQ4 [32]. This is the one used in the analysis for the pQCD predictions. There is a more recent parametrization by this collaboration, but the CTEQ4 has the advantage that fitted distribution of the parton densities are available for different values of α_s. These are needed for the determination of the strong coupling constant from the comparison of the measured values with the predictions, as it is described in section 5.2.1.

Glück-Reya-Vogt

Whereas the parton distributions of the MRS and CTEQ groups depend strongly on the input parametrizations at ${Q_0}^2$ this is, to a much lesser extent, the case for the GRV group [33]. Their basic idea was that at a very low starting scale (${Q_0}^2 = 0.34$ GeV^2) the nucleon only consists of constituent valence quarks. With increasing Q^2 the gluon and sea quarks would be generated dynamically from the valence quarks via the DGLAP evolution equations. As not all relevant data, in particular the prompt photon data, could be described adequately with this model, gluon and sea quarks had to be added at the starting scale. However, these distributions have a valence-like shape of the form

$$xg(x, {Q_0}^2) = Ax^{\alpha}(1-x)^{\beta} . \tag{1.24}$$

Charm is treated as a heavy quark. In this approach, there is no such a concept as charm parton distribution. Its contribution to F_2 comes from gluon splitting and it is calculated via the DGLAP evolution with massive-quark coefficient functions. The DGLAP evolution from the very low starting scale ${Q_0}^2$ ensures that the behaviour of the gluon and sea quark desities at small x for $Q^2 \gtrsim 1$ GeV^2, as described by GRV, is nearly a parameter free prediction of perturbative QCD.

1.4 Jet physics

High-energy quarks and gluons are not directly observed due to the confinement of colour charge in QCD. Instead they are revealed through the appearance of *jets*, collimated flows of hadrons, that can be reconstructed from the observed particles, and thus relate the final jets of hadrons to the original partons emerging from the interaction. To make a quantitative comparison between theory and experiment, one must go beyond a qualitative definition and use a precise algorithm for defining a jet. This algorithm has to be able to specify a jet configuration unambigously both in theoretical calculations and in experiments.

1.4.1 Jet search algorithm

Jet algorithms are not uniquely defined and the experimental results depend on the criteria used to define a jet. However, a well designed jet algorithm must have some basic properties: it must be easy to apply to the measured hadronic final state and infrared-safe so that the cross sections can be calculated order-by-order in perturbation theory. In perturbative QCD there is a collinear divergence when any of two massless partons are parallel. In the calculation of the total cross section this divergence is cancelled by the contribution from the virtual correction to the equivalent process, with the two partons replaced by their sum. However, for this cancellation to take place also in the jet calculation, it is necessary to ensure that a collinear pair of particles are treated identically as a single particle with their combined momenta. From the experimental point of view, the equivalent problem is the fact that parallel particles go into the same calorimeter cell and cannot be resolved. Likewise the requirement of infrared safety, i.e. insensitivity to emission of low energy particles, is necessary in perturbative calculations to avoid the soft divergence and in the experiment to avoid bias from the threshold trigger of the calorimeter cells and the background noise.

A clear difficulty in choosing a jet definition is that the latter most likely depends on the kind of high nergy scattering under consideration. Although the basic hard scattering processes studied in the different types of collisions can be described within the same theoretical framework, the overall event structure is quite different. In e^+e^- annihilation into hadrons, the initial state is purely leptonic and the entire final state can be thought of as arising from the virtual boson which creates a $q\bar{q}$ pair. Therefore, all the hadrons in the final state are associated with the hard scattering process. In contrast, in hadron (lepton)-hadron collisions, there are different contributions to the overall final state: only a fraction of the final-state hadrons are associated with the hard scattering process, and the remainder form part of the beam remnants or come from eventual *soft* (i.e. small transverse momentum) interactions of the remaining partons in the incident hadrons. Moreover, since only a fraction of the initial energy of the incoming hadron takes part in the hard process, the angular distribution of the final-state particles could be affected by the boost in the longitudinal direction.

As a general rule, any jet definition should satisfy the conditions of being

i) collinear and infrared safe (see [34] for more details);

ii) easy to apply in experimental analysis to the hadronic final state and to the order-by-order calculations in pQCD to the partons in the final state;

iii) little sensitive to hadronization corrections. Having small hadronization corrections entails a closer correspondence between the measurements from final state hadrons and the parton dynamics.

In the case of hadron-hadron and lepto-hadron collisions, the jet definition should also fulfill the additional requirements of being

iv) able to factorize the collinear singularities arising from the initial state hadron into universal parton densities; that is, the algorithm has to conserve the factorization properties of pQCD;

v) not strongly affected by contamination from hadron remnants or additional soft processes;

vi) longitudinally invariant. That is, the definition of distances to define a jet as a "collimated set of particles" has to be done in terms of quantities which are invariant under longitudinal boosts[15]

One of the first attempts to define jet cross sections in perturbative theory was made by Sterman and Weinberg for e^+e^- collisions [35]. They showed how to define cross sections free of singularities by integrating over finite ranges of energy and angle. The cross section for the process $e^+e^- \to q\bar{q}g$ has the form

$$\frac{1}{\sigma}\frac{d^2\sigma}{dx_1\,dx_2} = C_F\frac{\alpha_s}{2\pi}\frac{{x_1}^2+{x_2}^2}{(1-x_1)(1-x_2)}\ ,$$

where x_1, x_2 are the energy fractions of the final-state quark and antiquark ($x_i = 2 * E_i/\sqrt{s}$). This cross section has a singularity when one or both x_i are close to one. this corresponds to

[15] "Longitudinal" means along the initial hadron(s) direction.

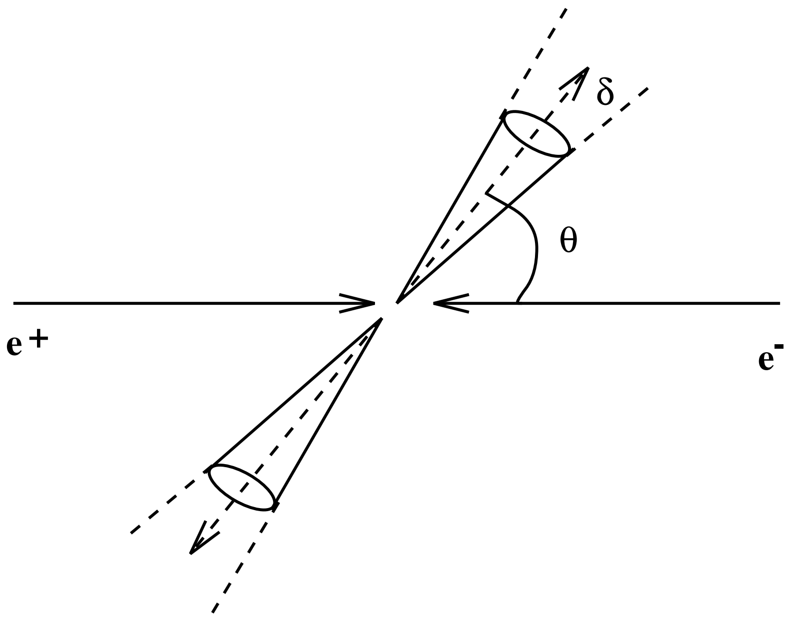


Figure 1.9: Sterman-Weinberg jet algorithm for e^+e^- interactions. The jets are studied by considering the energy distribution in the event adding up the energy within two opposite cones whose central axis forms an angle θ with the axis containing the incoming particles.

the collinear and infrared singularities in the gluon emission. For this process, Sterman and Weinberg introduced the quantity $\sigma_2(s, \theta, c, \delta)$, which is the two-jet cross section, defined for all events where a fraction $(1-c)$ or more of the total energy $\sqrt{s}$ is emitted within two opposite cones of half-angle δ, making an angle θ with the beam axis see figure 1.9. The total cross section for this process at first order in α_s can be written as an incoherent sum of two and three-jet cross section $\sigma = \sigma_2 + \sigma_3$. The fraction of two-jet events is then written as

$$\frac{\sigma_2}{\sigma} = 1 - 8C_F\frac{\alpha_s}{2\pi}\left[(\ln\frac{1}{\delta})\left[\ln(\frac{1}{2c}-1) - \frac{3}{4} + 3c\right] + \frac{\pi^2}{12} - \frac{7}{16} - c + \frac{3}{2}c^2 + \mathcal{O}(\delta^2 \ln c)\right]$$

which is free of any soft divergences as far as δ and c take reasonable and finite values.

At higher orders in perturbation theory, the final state can consists of more than three jets. It turned out that this algorithm is not well-suited to analyze these multijets events, and new algorithms were developed to improve the measurements of jet cross sections and the comparison with the theoretical calculations.

At present there are essentially two classes of algorithms in use: cone-type algorithms and cluster-type algorithms of the type first introduced by the JADE collaboration [36]. The cone-type algorithms are essentially modifications of the original Sterman-Weinberg algorithm to adapt it at the process under study and to partially avoid the difficulties of using this algorithm in multijet final states.

Generally, the cone-type algorithms define the jets using fixed geometrical structures which are positioned in the angular space occupied by the particles in such a way that the energy or the transverse energy is maximized. To determine an algorithm of this kind, the only requirements are the geometrical definition of the 'cone' (usually is a definition of a circle in the angular space with a given radius) and the criterion to be followed if two or more of these 'cones' overlap.

The cluster-type algorithms [37] are characterized by successively finding pairs of particles that are 'nearby' in phase-space and merging them together to form new 'pseudoparticles' which are those considered in the next iteration (instead of the two original ones). To determine this kind of algorithm, a definition of distance is needed to decide what 'nearby' particles means in a quantitative way. As an additional requirement, usually some scale is needed to stop the iterative procedure and define the final jets from the pseudoparticles.

It should be noted that in the discussion of jet algorithms, the word 'particles' is applied to any set of four-momenta. It could be the partons in the theoretical calculation, or the energy deposits detected experimentally in a calorimeter.

Several algorithms of the two types are at present available and currently used. Some of these algorithms, as used in e^+e^-, $p\bar{p}$ and ep at HERA are presented below.

The cone algorithm

In hadronic colliders is still very common the use of cone-type algorithms. The most used is that developed for $p\bar{p}$ colliders by the UA1 collaboration [38] and modifications to it. In this algorithm, all the particles are characterized by their transverse energy ($E_T = E \cdot \sin\theta$, where θ is measured with respect to the initial direction of the hadrons defined as z-axis), their azimuthal angle (ϕ) and their pseudorapidity ($\eta = -\ln(\tan(\theta/2))$)[16].

The particles with transverse energy above a threshold value are considered as *initiators*. For each of them a cone is defined taking a radius of R in the $\eta - \phi$ plane. All the particles whose distance to the cone axis is smaller than R

$$\sqrt{(\eta_{cone} - \eta_i)^2 + (\phi_{cone} - \phi_i)^2} < R$$

are included in the cone.

Then, the cone position is redefined using the particles included in the cone via the expressions

$$E_{T,cone} = \sum_i E_{T,i} \quad \eta_{cone} = \frac{1}{E_{T,cone}} \sum_i E_{T,i}\,\eta_i \quad \phi_{cone} = \frac{1}{E_{T,cone}} \sum_i E_{T,i}\,\phi_i$$

and the algorithm is repeated till the position of the cone converges. At that moment the jet is defined using the center of the cone.

Since two or more cones may overlap, a criterion is needed to decide if two cones should be merged into one or if the "overlapping" particles are distributed between the two different jets. After this criterion has been applied, we are left with a number of jets. An additional cut on E_T allows to obtain the final sample of jets.

It should be noted that this additional cut on E_T is needed to separate jets coming from the hard process and the hadrons coming from the remnants. At the same time, this cut allows the reduction of the dependence on the threshold to select the initiators.

As a final remark, the structure of this algorithm permits particles to be left unassigned to the jets. This means that the algorithm is "inclusive" and it is the main property for its use in hadron colliders: the particles which form the remnant must not be considered as parts of the jets and should not be taken into account.

[16] The pseudorapidity is the "longitudinally invariant" polar angle. It is very easy to show that a Lorentz transformation along the z-axis only changes the pseudorapidity by an additive constant. As a consequence, differences in pseudorapidity are invariant under such a transformation. This makes this variable very useful to define jet algorithms for hadron colliders.

The JADE algorithm

The JADE algorithm [36] was the first cluster-type algorithm. It is very common in studies of jets in e^+e^- colliders in the original way or via some of the modifications which have been developed since it was presented.

The jet reconstruction procedure starts by calculating the distance parameter m_{ij} for all pairs of particles according to the expression

$$m_{ij}{}^2 = 2E_iE_j(1-\cos\theta_{ij}) ,$$

where E_i (E_j) is the energies of particle i (j), and θ_{ij} is the angle between particles i and j. The variable m_{ij} is the invariant mass of one particle decaying into i and j under the approximation that both particles i and j are massless. The pair with the smallest m_{ij} is combined into a cluster by the addition of their four-momenta. The procedure is repeated until all remaining pairs have an invariante mass exceeding a preselected cut-off value, $m_{ij}{}^2 > y_{cut}M^2$, determined by a resolution parameter y_{cut} and a reference mass M. The final clusters represent the jets of the event.

As a difference with respect to the cone algorithm described above, in this algorithm all the particles are assigned to the jets. This property is usually denoted as *exclusiveness*, and it should be noted that this type of algorithm is only useful when all the hadronic particles are coming from the hard interaction, i.e. no remnant is present.

The k_T-cluster algorithm

A two-step procedure, similar to that of the original JADE algorithm, is used in the k_T-cluster algorithm. Although it can be applied with slightly modifications to every kind of experiment, the DIS version [37] uses the first step to perform the preclustering of hadrons into a beam-jet and final state macro jets. The second step aims at resolving jet structures within the system of macro jets.

In the preclustering, a variable $k_{T,ij}$ is calculated for all pairs of particles i and j according to the expression

$$k_{T,ij}^2 = 2\,(1-cos\theta_{ij})\,min\left\{E_i^2, E_j^2\right\} ,$$

where θ_{ij} is the angle between the momenta vectors of particles i and j and E_i and E_j the corresponding energies. In addition, a parameter $k_{T,iP}$, related to the distance of particle i to the beam-jet, is calculated for each particle i according to

$$k_{T,iP}^2 = 2\,(1-cos\theta_{iP})\,E_i^2 ,$$

where θ_{iP} is the angle between i and the incoming proton direction.

At this stage a scale E_T is introduced and the iterative procedure starts till the smallest of all the $\{k_{T,ij}, k_{T,iP}\}$ is greater than E_T. If this is not the case the two particles related to that smallest value are combined into a new cluster and a new iteration is required. Any particle combined with the proton remnant (the smallest is $k_{T,iP}$) is considered as part of the spectator jet and is not included in the next iteration. If the two particles combined are "real", then a *recombination scheme* has to be introduced to define the four-momentum of the new cluster, which is considered in the following iterations.

In the second step, all the particles which where not assigned to the spectator jet are considered. The process is similar to that used in the first step, but now the scale can be selected to be smaller in such a way as to resolve more jets. This second parameter sets the scale to resolve the jets. The first one (E_T) is the scale which separates the hard process(es) with respect to the soft ones.

The longitudinally-invariant k_T-cluster algorithm

The original k_T-cluster algorithm can be modified to define a longitudinally invariant algorithm whose advantages for jet studies in DIS have made it to become the most extended algorithm in HERA physics. This algorithm is the one used in the analysis presented here and make use of the longitudinally invariants E_T, η and ϕ, which were defined in the description of the cone algorithm.

The algorithm proceeds according to the following steps [39]:

1. For every pair of particles, define a closeness between the two particles d_{ij}

$$d_{ij} = min\left\{E_{T,i}, E_{T,j}\right\}^2 \left[(\eta_i - \eta_j)^2 + (\phi_i + \phi_j)^2\right] ,$$

   where it should be noted that for small distances in $\eta - \phi$, d_{ij} represents the relative transverse momentum squared between particles i and j.

2. For every particle, define a closeness to the beam direction d_i

$$d_i = E_{T,i}^2 \cdot R^2 ,$$

   where R is a parameter of the algorithm which plays the rôle of a jet radius in the $\eta - \phi$ plane. It is usually set to one, which is the preferred value from the theoretical point of view as it treats initial and final state radiation on equal footing.

3. The smallest value of all $\{d_{ij}, d_i\}$ is considered. If this is one of the d_{ij} then the two particles are merged into a new one following the recombination scheme. If the minimum is one of the d_i, then that cluster is considered as a *protojet* and no longer considered for clustering.

4. The procedure is repeated till no remaining cluster is left and all the particles have been assigned to protojets. It should be noted that at each iteration, one particle is removed, so the number of iterations is always equal to the original final-state particles.

5. From the sample of protojets, the final jets are selecting imposing a cut on E_T which sets the scale to distinguish the hard and the soft processes.

Although the recombination scheme is not fixed by the algorithm, the most natural definition is to recombine the particles using the following expressions (where the two particles i, j are merged into a new one k):

$$E_{T,k} = E_{T,i} + E_{T,j} \ , \ \ \eta_k = \frac{1}{E_{T,k}}\Big(\eta_i\, E_{T,i} + \eta_j\, E_{T,j}\Big) \ , \ \ \phi_k = \frac{1}{E_{T,k}}\Big(\phi_i\, E_{T,i} + \phi_j\, E_{T,j}\Big) \ . \qquad (1.25)$$

In this way the jet quantities are defined according to the Snowmass convention [40], in the same way as it was introduced above for the cone algorithm.

Although this algorithm does not impose a given geometry to the jets, the parameter R can be related to the radius of the cone algorithm. The definition of the k_T-cluster algorithm has the property that the jets are far away at least a distance in $\eta - \phi$ equal to R (if not, the algorithm would have merged them). On the other hand any particle in a radius R in the $\eta - \phi$ plane around the jet axis is included in the jet (since the protojets are only selected from clusters which have no other clusters at a distance smaller than R). That is the reason why the jets defined with this algorithm are phenomenologically similar to those defined by the standard cone algorithm described above.

However, the invariant k_T algorithm is less influenced by soft particles than the cone algorithm, which results in smaller hadronization and detector corrections. Qualitatively one can say that the longitudinally invariant k_T algorithm pays more attention to the core of the jet and only merges neighbouring particles if they are close enough, whereas the cone algorithm, in order to maximize the jet transverse energy, pulls in as much neighbouring energy as possible.

1.4.2 The internal structure of jets

The internal structure of quark and gluon jets, which in the frame of pQCD is ascribed to the emission of additional partons according to the SU(3) colour factors (as discussed in sections 1.1.5 and 1.3.4), is an exclusive event property mainly ruled by the pattern of parton radiation in the process under study. Furthermore, the analysis of the internal structure gives insight into the process steering the transition from the partons to the observable hadrons. Thus, the comparison with the theoretical prediction is a stringent test of the validity of perturbative QCD as the theory describing the strong interaction.

In this section the physical observables which are commonly used to study the internal structure of the jets are presented. Special stress is laid on the definition of the subjet multiplicity, which has been measured in the analysis presented here.

The most common way of resolving the internal jet structure, inspired by the cone-type algorithms, is to measure the energy distribution around the jet axis direction. The physical observable used is the so-called *jet shape*, which is related to the energy distribution around the jet axis. Close to this axis, the jet shape is dominated by collinear gluon emission, which can be calculated perturbatively. However, the cluster-type algorithms naturally suggest an alternative, additional way of analyzing the internal structure of a jet that is much closer to the partonic picture of how that structure arises: by resolving 'sub-objects' (called *subjets*) within a jet making use of a resolution variable. The physical observable used in this case is the so-called *subjet multiplicity*, which is the number of subjets within a given jet.

These quantities are usually studied in a statistical way and, therefore, the average values over a sample of jets are considered in the measurements.

The formal description of the quantities will be done as it is usually done in HERA physics. Definitions for other type of processes could be sligthly different although the underlying notions are the same.

The jet shape

The *integrated jet shape* $\psi(r, R)$ is defined as the average fraction of the jet's transverse energy that lies in an inner cone of radius r, in the $\eta - \phi$ plane, concentric with the jets axis (as shown in the drawing on the left of Fig. 1.10),

$$\psi(r, R) = \frac{1}{N_{jets}} \sum_{jets} \frac{E_{T,jet}(r)}{E_{T,jet}} ,$$

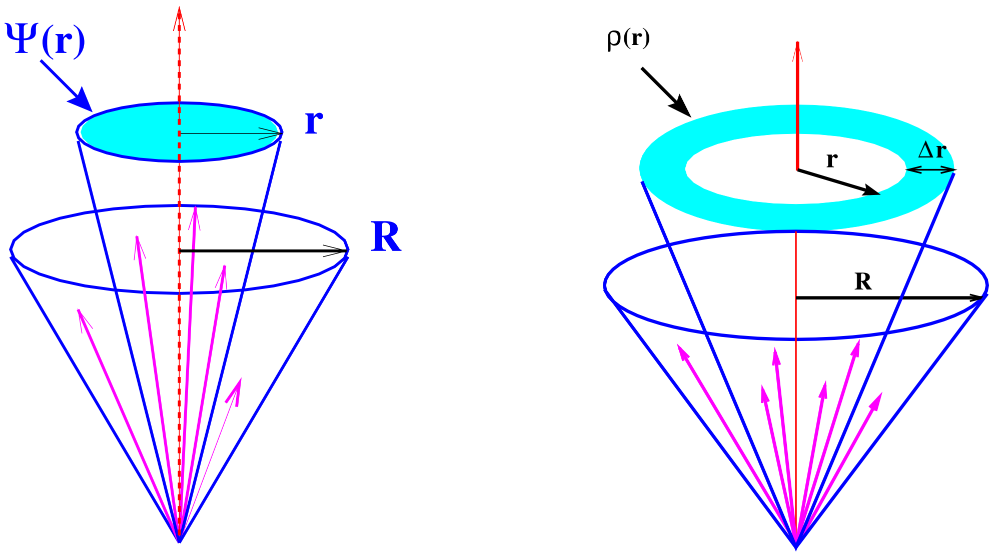


Figure 1.10: Schematic representation of the integrated (left) and differential (left) jet shape.

where $E_{T,jet}(r)$ is the jet's transverse energy within the inner cone of radius r, N_{jet} is the number of jets in the sample, and $E_{T,jet}$ is the entire transverse energy of the jet.

The jet shape can be also analyzed in terms of the *differential jet shape*, i.e. the average fraction of the jet's transvers energy that lies in an angular annulus concentric with the jet axis (see drawing on the right of Fig.1.10),

$$\rho(r) = \frac{d\psi}{dr}(r, R) = \frac{1}{N_{jets}} \sum_{jets} \frac{E_{T,jet}(r + \Delta r) - E_{T,jet}(r)}{E_{T,jet}} \ .$$

Narrower jets are characterized by larger values of $\psi(r, R)$, larger values of $\rho(r)$ at small r and smaller values of $\rho(r)$ at large r.

The subjet multiplicity

After identifying a jet of a given transverse energy $E_{T,jet}$, the longitudinally invariant k_T-cluster algorithm (see section 1.4.1) is rerun over only those particles that were assigned to the jet. This second clustering is stopped when all values of d_{ij} satisfy the condition:

$$d_{ij} > y_{cut} \cdot E^2_{T,jet} \ ,$$

i. e. when all the internal relative transverse momenta are greater than $\sqrt{y_{cut}} \cdot E_{T,jet}$ and the number of sub-structures inside the jet is counted (n_{sbj}. For a resolution scale $y_{cut} \sim 1$ the jet always consists of only one subjet (i.e. the jet itself). This can be formally expressed as

$$\lim_{y_{cut} \to 1} n_{sbj} = 1.$$

When the y_{cut} approachs zero every hadron is eventually considered as a separate subjet. Thus, varying the resolution scale y_{cut} allows a detailed study of the transition of the colour charged fields (partons) into colourless final-state hadrons final states[17].

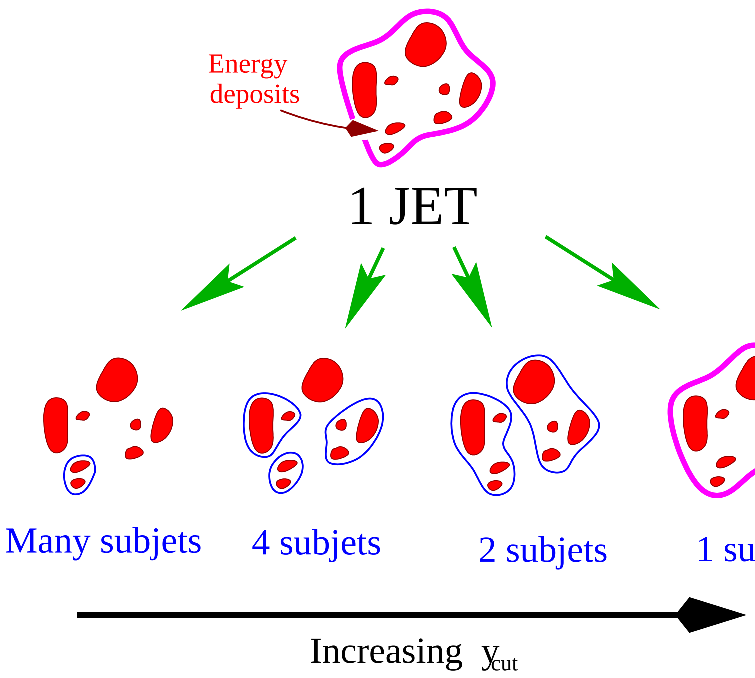


Figure 1.11: Schematic representation of the subjet structure within a jet. Depending on the value of the resolution scale (y_{cut}) the number of resolved subjets is different. For very small values, any energy deposit is defined as a subjet. In the limit where the resolution scale is close to 1, no internal structure is resolved.

Figure 1.11 shows a graphical representation of the meaning of subjet structure and how the definition presented above is related to the internal structure of the jets.

Since subjets are physical objets, at the same level as the original jets, they can be used to make very different kinds of analysis. The simplest way to study the internal structure of the jets using these subjets is characterizing that structure with the number of subjets resolved at the given resolution scale.

For statistical studies, the *mean subjet multiplicity* is defined as the average number of subjets in the sample of jets considered,

$$\langle n_{sbj} \rangle (y_{cut}) = \frac{1}{N_{jet}} \sum_{jets} n_{sbj}^{jet}(y_{cut})$$

and this is the quantity used in the analysis described here.

1.4.3 Quark- and gluon-initiated jets

In the frame of pQCD, the average number of hadrons in a jet initiated by a parton at a given scale can be calculated using the fragmentation functions and the evolution equations applied to that functions [34]. Since the sppliting functions are determined by the colour charge of the initial parton, it is clear that the multiplicity inside a jet will be dictated by the type of parton which initiates the jet.

[17] From the experimental point of view, very small values of y_{cut} cannot be considered since the detector effects would be very large. On the theoretical side, a very small resolution scale introduces large hadronization corrections and the model dependence of the parton radiation and fragmentation becomes very large.

To make a quantitative study of the differences between quarks and gluons, it is needed that pQCD gives a good description of the observable. For this, the scale should be relatively large and infrared-safe quantities have to be defined for a direct comparison of theory and measurements. The observables defined above, specially the mean subjet multiplicity, can be used to study these differences and to compare with the predictions of pQCD.

One of these predictions is that the asymptotic behaviour (i.e. at very large scales in comparison with the one related to the fragmentation process) of the averaged multiplicity of any type of objet in quark and gluon-initiated jets can be expresssed as [11]

$$\langle n \rangle_{quarks} \sim \frac{C_F}{C_A} \langle n \rangle_{gluons} \propto \exp\left[\frac{1}{b}\sqrt{\frac{2C_A}{\pi \alpha_s(t)}}\right] \tag{1.26}$$

This predicts that the multiplicity distribution should become larger for gluon-initiated jets than for quark jets by the factor C_A/C_F and the distribution of particles inside the jets should become wider by the square root of that factor.

These predictions are qualitatively correct, and it has been experimentally observed that gluon jets are broader than quark jets. However, the ratio of the two multiplicities is quite below the asymptotic value, although it is well described by the phenomenological models [41]

For the jet shape, a broader jet means that the distribution of energy is less concentrated in a small radius, so the qualitative prediction from QCD is

$$\psi_{quarks}(r, R) > \psi_{gluon}(r, R)$$

and the differential jet shape will be larger for quark jets at small radius and smaller at large radius than in gluon jets, as we described before.

For the mean subjet multiplicity, it is expected that the expression (1.26) holds for $\langle n \rangle = \langle n_{sbj} \rangle$, so the gluon jets are predicted to display a larger $\langle n_{sbj} \rangle$ than in quark jets.

1.4.4 Determination of α_s from the internal structure of jets

When the non-perturbative effects are small, i.e. at large energy scales, the internal structure of the jets is expected to be dominated by "hard" parton radiation and perturbative QCD should be able to give predictions that can be compared to the measurements.

The calculations are done by applying the same definition for jets as for hadrons, and in this way, if a jet contains two or more partons, an internal structure can be resolved. This case is only posible when computing the perturbative expansion at least to the next order to the first one needed for the definition of jets.

The idea is that to define a jet we need an order in the perturbative expansion which contains at least one parton to be taken as a jet after imposing the selection cuts, for example on the transverse energy. On the other hand, although one parton is enough to define the jets, the minimum requirement for the predictions of the internal structure is two partons in the same jet. This two-parton configuration represents the lowest order contribution to the internal structure. To compute the pQCD predictions to higher orders, configurations containing more partons are needed.

As for every prediction of pQCD, the internal structure will depend on the value of the strong coupling constant, which enters in the calculation at first order via the splitting of one parton into two which still belong to the same jet, but are resolved when the internal structure of the jet is taken into account.

For a meaningful determination of α_s from the measured values of the internal structure, at least two orders of the pQCD prediction are needed. This limits the kind of process where this determination can be done, since in most of them the calculation of those orders is still not available. This is the case in hadron-hadron collisions.

On the other hand, in deep inelastic scattering the production of jets is at first order described by the QPM diagram, which does not include any contribution from QCD. That is, the first order for describing jet production is given by order 0 in α_s. To describe the internal structure of the jets, the order α_s^2 is required, and these terms are known.

Thus, all the tools which are needed for the determination of α_s from the internal structure of jets are available for jets produced in deep inelastic scatering. Since neutral-current processes have larger cross sections than charged-current ones and the analysis displays some advantages from the experimental and theoretical points of view, the most natural choice is to determine the strong coupling constant using jets produced in neutral-current deep-inelastic scattering events.

This is the goal of the analysis presented here, which uses the mean subjet multiplicity. A complete description of the next-to-leading order (NLO) QCD calculations is given in section 5.2.

Chapter 2

Experimental setup

In this chapter a brief description of the HERA ep collider and ZEUS detector is presented, specifically of the components that are more relevant for the analysis.

2.1 The HERA collider

The Hadron Electron Ring Accelerator (HERA) [42] located at the DESY (Deutsches Elektronen-SYnchrotron) laboratory in Hamburg (Germany) is the first electron (positron)-proton collider in the world. The tunnel in which it resides was built 10 to 25 meters under the earth surface and it has a circumference of 6.3 Km. The storage ring consists of four circular arcs (with a radius of 779 m) and four straight segments (each 360 meters long) as can be seen in figure 2.1. Four interactions points (IP) are placed in the straight segments, where the experiments H1, ZEUS, HERMES and HERA-B are located.

One of the remarkable features of HERA, which distinguishes it from a 'classical' collider, is the asymmetry between the two beam energies. HERA was designed to collide electron (positron) and protons with nominal energies of 30 GeV and 820 GeV respectively, resulting in a center-of-mass energy of $\sqrt{s} \sim 300$ GeV.

Beam injection is performed by a procedure involving several preaccelerators (see figure 2.2). Electrons are initially accelerated up to 200 MeV in the electron linear accelerator LINAC II, and then in the DESY II synchrotron up to an energy of 7.5 GeV. Afterwards, the electrons are injected into PETRA[1], where their energy is increased up to 14 GeV, before they are finally transferred to HERA. The protons start as H^- ions in the proton LINAC and are accelerated up to 50 MeV. After acceleration in DESY III (to 7.5 GeV), the electrons are stripped off and the protons are injected in PETRA, where their energy increases up to 40 GeV, before injection into HERA.

While the high momentum of the proton beam requires superconducting magnets, the electrons are controlled using conventional magnets. The HERA proton ring consists of 422 main dipoles delivering a bending field of 4.68 Teslas and 244 main quadrupoles. Standard cells of 47 m length combining 4 dipoles, 4 quadrupoles, 4 sextupoles and correction magnets are installed in the arcs of the proton ring and are cooled down to 4.2 °K. The conventional electron ring consists of 456 main dipoles of 0.164 Teslas and 605 main dipoles grouped in 12 m long magnet modules

[1]PETRA (Positronen-Elektronen Tandem Ring Anlage) is a previously existing e^+e^- collider at DESY.

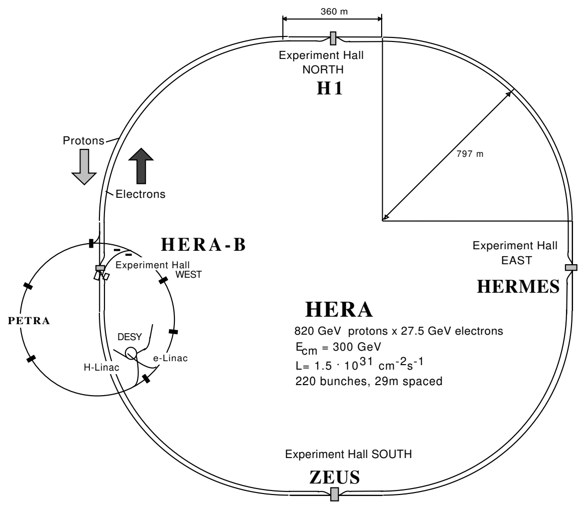


Figure 2.1: Layout of the HERA collider.

which contain one dipole, one quadrupole, one or two sextupoles and several correction dipoles. The energy loss due to synchrotron radiation is compensated by 500 MHz copper cavities installed in the straight segments of HERA. In addition, 16 superconducting 4-cell cavities with an electric field gradient of 5 MVolts/m are used to accelerate the electrons from 14 GeV (the energy they have when they are injected from PETRA to HERA) to the maximum energy. The electrons and protons are stored in separated bunches with a distance of 28,8 m between two successive bunches. This distance corresponds to a bunch crossing time of 96 ns (see figure 2.3). In order to maximize the luminosity, up to 210 bunches of particles can be stored for each beam. Unpaired bunches with only protons or electrons as well as completely empty bunches are used to estimate the number of background events coming from interactions with the residual beam gas in the storage ring, events produced by cosmics rays or other "non-ep" events.

Close to the interaction region, protons and electrons have to run in the same orbit in order to collide at zero angle. The distance of about 81 cm between the electron and the protons rings has to be traversed. A total of 123 guiding magnets are used to deflect the protons until they run in the same vacuum pipe as the electrons. After passing the interaction point they are brought back to the level of the proton ring.

From the '94 running period it was decided to accelerate positrons instead of electrons in order to achieve longer lepton-beam lifetimes. Interactions with almost stationary positive charged particles, most likely originated from the ion getter pumps of the HERA vacuum system, tend to defocus the eletron beam. For a positron beam the pile-up of the ions is avoided due to the repulsive electromagnetic force between the positrons and the positively charged ions.

During some part of the '98 and '99 running periods the machine accelerated electrons From that moment till September 2000, positrons were collided with protons. Since then, HERA and the experiments were involved in a long shutdown to improve the performance and increase the luminosity. From 1998 the proton beam has an energy of 920 GeV, which yields to an increase of 5% in the center of mass energy.

In addition to the available center-of-mass energy, another important quantity that characterizes the performance of a collider is the luminosity. This quantity determines the rate at which interactions occur. From the luminosity (L) and the cross section for a given process (σ) the

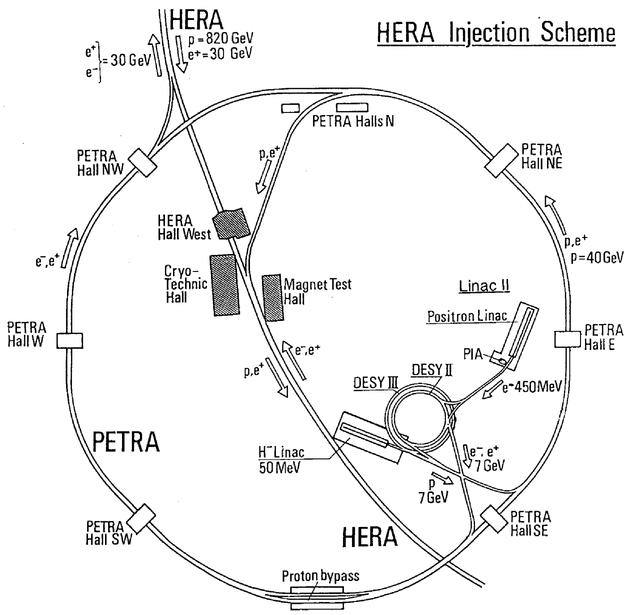


Figure 2.2: Scheme of the injection facilities at DESY.

number of events observed is

$$N = L \cdot \sigma \tag{2.1}$$

For an specific collider the luminosity L is controlled by the parameters of the collider: it depends of the number of bunches, the number of particles per bunch, the properties of the bunches and the crossing time.

Figure 2.4 shows the luminosity delivered by HERA during the several running periods from 1994 up to the end of the 2000 running.

2.2 The ZEUS detector

ZEUS is one of the currently four experiments at the HERA collider. It is a large multipurpose detector designed to study the wide spectrum of HERA physics. The ZEUS Collaboration consists of around 450 physicists belonging to more than 50 institutes from 12 different nations. The detector is located 30 m underground in the South Hall of the HERA tunnel outside of the DESY site and has a size of 12 m x 11 m x 20 m and weighs 3600 tons. A detailed description of the ZEUS detector can be found in [43].

Cross sections through the ZEUS detector are shown in figures 2.5 and 2.6. The ZEUS geometry is described by a right-handed coordinate system with its origin at the nominal interaction point. The z-axis is pointing along the incoming proton direction and the x-axis horizontally towards the center of HERA. A polar coordinate system (r, θ, ϕ) is very commonly used, where θ is the polar angle w.r.t. the positive z-axis and ϕ the azimuthal angle w.r.t. the x-axis.

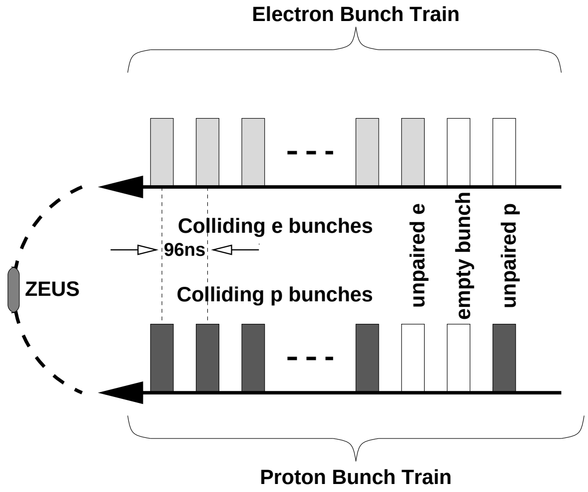


Figure 2.3: Bunch distribution for electrons and protons.

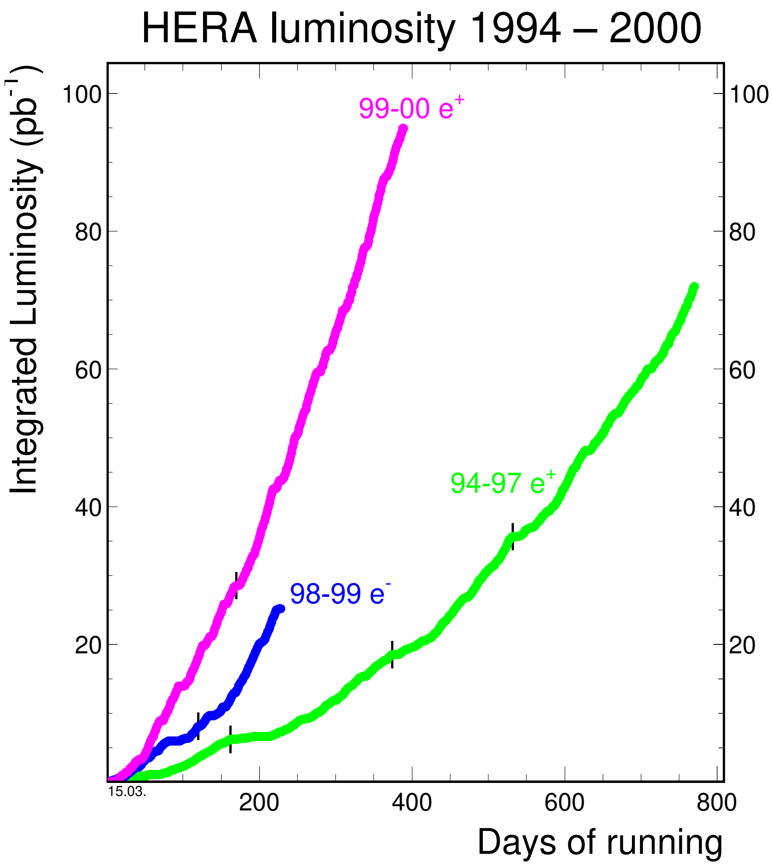


Figure 2.4: Accumulated luminosity delivered by HERA during RUN I divided in periods with equal lepton charge and proton beam energy.

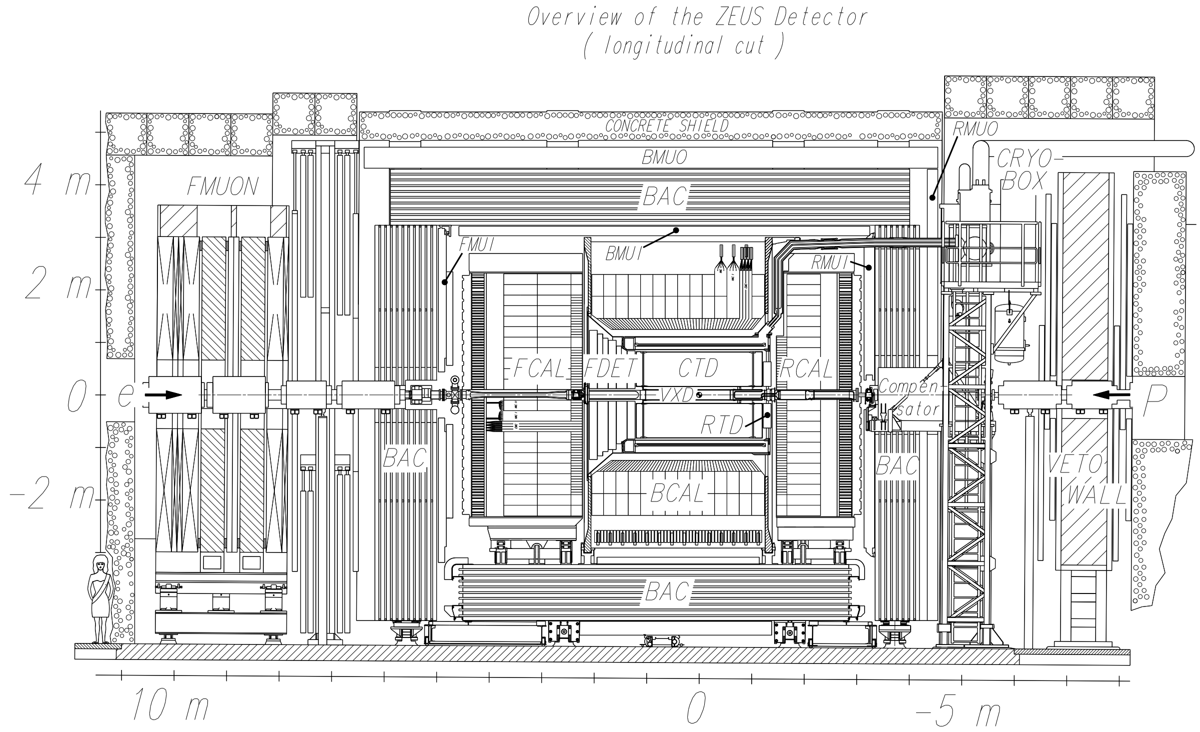


Figure 2.5: Longitudinal cut of the ZEUS detector. A person is drawn in the lower-left corner of the picture for reference of the dimensions. The vertex detector (VXD) was extracted before 1996 running. For a brief description of the most important components see section 2.2.

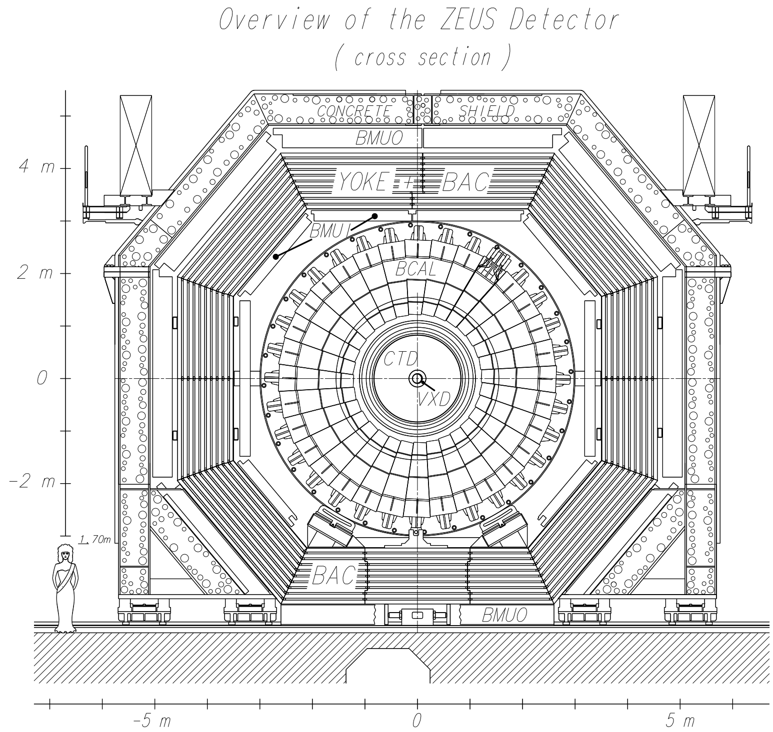


Figure 2.6: Cross section of the ZEUS detector perpendicular to the beam direction. Only the barrel components are shown in the picture.

Starting from the interaction point (see figure 2.5) ZEUS consist[2] of traking detectors for charged particles. They are divided into central (CTD), forward (FTD) and rear (RTD) traking detectors. The tracking chambers are surrounded by a high-resolution Uranium-Scintillator Calorimeter (UCAL). The UCAL is divided into three sections: the FCAL in the forward region, the RCAL in the rear region and the BCAL, a barrel section surrounding the central region. The small-angle rear track detector (SRTD) is situated behind the RTD and covers the face of the RCAL to a radius of $\sim$ 34 cm around the center of the beam-pipe hole. At a longitudinal depth of 3 radiation lengths in the RCAL the Hadron-Electron Separator (HES) is installed. It consists of 3x3 cm^2 silicon-diodes and provides a better discrimination power between electromagnetic and hadronic showers originating from low energetic particles ($<$ 5 GeV). Between the UCAL and the CTD there is a superconducting solenoid. The coil has an overall length of 2.46 m and an inner diameter of 1.91 m, and produces a magnetic field of 1.43 Tesla. A high-field solenoid (5 Tesla) situated behind the RCAL compensates the effect of the main solenoid on the electron beam. The axial field produced by the solenoid is used to measure the momentum of the charged particles. A second low-resolution calorimeter (BAC) was constructed to measure the energy leakage out of the main calorimeter. The BAC is made of 7.3 cm thick iron plates and also serves as the return yoke for the magnetic flux from the solenoid. In order to detect muons and measure their momenta, muon chambers are located between the BAC and the CAL (inner muon chambers), as well as after the iron return yoke (outer muon chambers). The muon detector, as the UCAL, is divided into three parts, forward (FMUI, FMUO), rear (RMUI, RMUO), and barrel (BMUI, BMUO). In the rear part of the main detector, at 7.5 m upstream of the interaction point, an iron-scintillator VETOWALL is used to reject beam-related background. A small lead-scintillator counter, located around the beam pipe at $z = -3.15$ m, monitors the synchrotron radiation accompanying the beams and the timing and longitudinal structure of the proton and electron bunches. At $\sim$ 20-90 m downstream in the proton-beam direction the Leading-Proton Spectrometer (LPS) and the Forward-Neutron Calorimeter (FNC) detect protons and neutrons scattered through small angles, respectively. The Proton-Remnant Tagger (PRT) is a lead-scintillator counter located at $z = 5.1$ m around the beam pipe and covers the pseudorapity[3] range $4.3 < \eta < 5.8$. In the electron-beam direction, two small lead-scintillator calorimeters are installed at $z = -34$ m and $z = -104$ m for the measurement of outgoing electrons and photons, respectively, at low-scattering angles. They are used for the determination of the luminosity and also to detect $Q^2 \sim 0$ events (photoproduction regime) and radiative events. The short time interval of 96 ns between bunch crossing at HERA results in a nominal rate of 10 MHz. ZEUS uses a three-level trigger system to reduce the rate to a few hertzs (see section 2.2.6).

In the next sections a brief description is given of the components which were relevant to the analysis presented here.

2.2.1 The Central Traking Detector (CTD)

The CTD [44] is a cylindrical wire chamber with nine superlayers of eight sense wire layers each, as shown in figure 2.7, covering the polar-angle range of $15^o < \theta < 164^o$. The inner radius of the chamber is 16.2 cm, whilst its outer radius is 85 cm. The longitudinal length is 241 cm. Five of the superlayers have wires parallel to the chamber axis, four layers are stereo wich have

[2] The description given here is the one related to the 96 and 97 running periods, which are the ones used in the analysis.

[3] The pseudorapidity is defined as $\eta = -\ln(\tan\theta/2)$, as described in section 1.4.1.

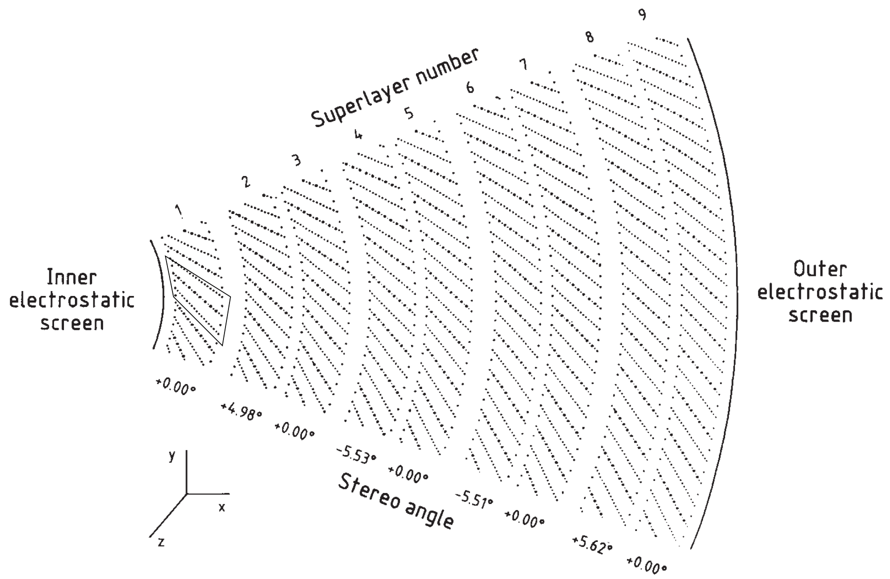


Figure 2.7: Layout of a CTD octant. The wires of the four superlayers with even numbers are slightly tilted with respect to the beam axis.

wires with a small angle (5^o or 7^o) with respect to the beam, and are used to determine the z-coordinate of the hits.

The CTD is filled with a gas mixture of argon, carbon dioxide and ethane (in the ratio 83:5:12) which is bubbled through ethanol (see [45] for more details). This mixture has been chosen on the grounds of safety and detector lifetime [46], although a pure argon-ethane mixture (50:50) would provide a better resolution and less noise.

For trigger purposes, the three inner axial layers are additionally equipped with a z-by-timing system which determines the z-position of the hit from the difference in arrival times of a pulse at both ends of the chamber.

After the final offline reconstruction, the $r - \phi$ hit resolution is $\sim$200 μm (depending slightly on the polar angle θ). The resolution in z using the axial and stereo superlayers is $\sim$2 mm and when using z-by-timing is 4 cm. The relative transverse-momentum resolution of the CTD is $\sigma(p_T)/p_T = 0.0058 p_T \oplus 0.0065 \oplus 0.0014/p_T$ for particles transversing all 9 superlayers and with p_T in GeV. The actual numbers in this parametrization have been obtained running the detector simulation (tuned with data) on generated tracks coming from the $D^* \to K\pi\pi$ channel [47]. The first term corresponds to the resolution of the hit positions, the second term to smearing from multiple scattering within the CTD and the last term to multiple scattering before the particle enters into the CTD [48].

In the forward region the acceptance of the CTD is extended by the FTD, a system of three planar drift chambers. Between the three FTD chambers, the modules of a Transition Radiation Detector (TRD), designed to enhance hadron-electron separation in the forward region, are positioned. FTD and TRD together form the Forward Detector FDET. In the rear direction, a single chamber (RTD) of identical design of those from the FTD extends the CTD acceptance to polar angles between 160^o and 170^o.

2.2.2 The Uranium-Scintillator Calorimeter (UCAL)

In high-energy experiments the energy of the particles and jets is measured with calorimeters, which are the central components of the large experiments at present storage rings. One of the essential properties of a calorimeter is its sensitivity to both charged and neutral particles.

2.2.2.1 Principles of Calorimetry

Before giving the description of the ZEUS Calorimeter, a brief introduction to theoretical aspects of calorimetry is presented here.

Passage of radiation trough matter

Particle reactions which occur when radiation penetrates matter are the basis of particle detection devices. The interactions depend on the type of radiation, its energy and the penetrated material. A calorimeter absorbs an incident particle completely and transforms a part of its energy into a measurable signal. The penetrated detection material can suffer radiation damage which may produce a change of the detection characteristics. Therefore, the knowledge of the particle reactions and effects produced by these processes on the material is of fundamental importance when designing a detector.

The passage of particles through matter with kinetic energies from several keV upwards is mainly characterized by the following processes [49]:

- nuclear reactions,
- elastic or inelastic scattering from nuclei,
- inelastic collisions with the atomic electrons of the material,
- emission of Cherenkov radiation and
- bremsstrahlung.

Particle reactions resulting from the radiationg passage through matter significantly depend on the type of radiation. In case of neutrons, for example, the most important type of interaction is collision with nuclei. Because of the short range of the associated force these elastic or inelastic reactions are much rarer than the Coulomb interactions of charged particles in matter. The main energy loss of heavy charged particles in matter (i.e. particles heavier than an electron) occurs in inelastic collisions with the atoms causing an ionization or excitation of the latter. The energy transferred in each collision is a small fraction of the total kinetic energy of the particles. The energy deposition dE/dx can be described by the Bethe-Bloch formula [50]. During its passage though matter charged particles can also produce Cherenkov radiation, which arises when the velocity of the particle is larger than the speed of light in the same medium.

The energy loss of electrons (and positrons) is similar to heavy charged particles, but besides the collisional energy loss an additional effect appears; it is the electromagnetic radiation due to the electron being scattered in the electric field of the nucleus (bremsstrahlung). For kinetic energies of a few MeV or less the energy loss is dominated by ionization. As the energy is increased up to a critical energy, the energy loss due to bremsstrahlung is comparable to ionization loss. Above this energy, bremsstrahlung processes dominate the energy loss of the electron in matter.

The interaction of photons in matter, specially high energy ones (γ- or X-rays), is mainly described by three types of processes: the photoelectric effect, Compton scattering and pair production. At higher photon energies, pair production ($\gamma \rightarrow e^+/e^-$ in presence of a partner, usually an atom) dominates the energy loss in matter completely.

Electromagnetic and hadronic showers

During the passage in matter electrons (and positrons) of high energy mainly produce bremsstrahlung photons, which initiate pair production if their energy is bigger than[4] 1.022 MeV. The combined effect of bremsstrahlung emission by electrons and positrons and pair production by high energy photons forms a cascade of photons, electrons and positrons. This **electromagnetic shower** continues until the energy of the particles drops below the *critical energy* E_c defined as the energy at which the energy loss of an electron due to bremsstrahlung is equal to the energy loss through ionization. Then collisional energy loss dominates the interactions and no further electron-positron pairs are produced. Thus, the propagation of the shower is halted [51]. The propagation of an electromagnetic shower converting bremsstrahlung into pair produced electrons/positrons and vice versa doubles the particle number N and bisects their average energy after each radiation length X_o. We can estimate the maximum number of particles produced in a shower using the relation

$$N \approx \frac{E_o}{E_c},$$

whereas the depth of the shower can be estimated as:

$$X_{max} \approx X_0 \, \log_2 N \approx X_0 \, \log_2 \frac{E_o}{E_c}. \tag{2.2}$$

An estimation of the depth D_{98} required to contain 98% of an electromagnetic shower produced by a 25 GeV electron leads to $D_{98} = 22X_0$ [52].

The critical energy decreases with increasing atomic number of the absorbing material. Therefore, heavy materials such as lead or tungsten are preferred absorbing materials inside calorimeters.

A **hadronic shower** in matter is initiated by interactions between the incident hadronic particles and nuclei of the absorbing material. Similar to an electromagnetic shower, but accompanied by a large variety of different interactions, a cascade of shower particles is produced. The hadronic shower propagation is dominated by inelastic scattering from nuclei with possible subsequent ejections of atomic particles. The nuclei are left in excited states which later may decay by emitting a photon or by means of other forms of emission as vaporization, nucleus fission and neutral pions. Some of these particles and also the decay of π^0 into two photons produce electromagnetic showers inside the complete hadronic showers.

In practice, the energy deposition of a pure hadronic shower depends strongly on the energy of the incident hadron. A fraction of its energy is deposited due to ionization. The rest is lost by nuclear interactions and production of neutrinos. The consideration of the absorption length encountered for shower products allows the simulation of hadronic cascades [53].

[4] 1.022 MeV is twice the mass of the electron, which is the particle with smaller mass which couples directly to the photon due to its electric charge.

The length scale appropriate for hadronic showers is the nuclear interaction length λ_{int} [54] defined as $\lambda = A/\mathrm{N}_A\sigma_i$, where N_A is the Avogadro number and σ_i denotes the inelastic cross section. A good approximation of λ_{int} is given by [52]:

$$\lambda_{int} \propto 35 \cdot \frac{A^{\frac{1}{3}}}{\rho} \ \ [\mathrm{cm}]$$

where A is the number of protons and neutrons in the nuclei and ρ the density of the material in units [g/cm^3].

As reference, the depth required to contain 95% of the hadronic shower energy is given by the parametrization [55]

$$D_{95} \sim 0.2 \ln E + 2.5 E^{0.3} + 0.7$$

Acording to this, a shower of a 300 GeV hadron can be contained in a calorimeter with a depth of $7.1\,\lambda_{int}$.

In conclusion, hadronic showers have much larger longitudinal and transverse dimensions than electromagnetic ones. As a reference, in a lead-based calorimeter the hadronic shower is roughly five times deeper and wider than the electromagnetic one. This difference is very commonly used to identify electrons, positrons and photons using the calorimeters in present high-energy physics experiments.

Calorimeter design

A calorimeter based on scintillating materials is one of the prime devices for energy measurements in high energy physics. One important condition for a correct energy measurement is the complete energy absorption of incident particles within the calorimeter and a partial transformation of the energy into fluorescense light. The light is transformed into measurable electrical pulses by photo-detectors. Several posibilities exist to configure a calorimeter. A homogeneous calorimeter consists of one uniform material which is used for both energy dissipation and measurement(e.g. lead-glass calorimeter). A sampling calorimeter is based on alternating layers of passive absorbing material and active detector material. For the absorption of particles mostly a material is used with high density ρ and atomic number such as lead, tungsten or iron. Plastic scintillators, liquid argon or silicon diodes as active materials offer possibilities for detection.

Independently from any particular configuration a calorimeter has to fulfill some characteristics. It should be sensitive to both charged and neutral particles. The calorimeter has to absorb the total energy of the incident particles. The penetration depth of an incident particle into the absorbing material is proportional to the logarithm of the initial energy E_o. If the measurement uncertainties are dominated by statistical processes depending on the number of incident particles, the energy resolution σ/E decreases with increasing energy E like [56]

$$\frac{\sigma(E)}{E} \propto \frac{1}{\sqrt{E}}$$

while the depth D required to stop incoming particles increases only logarithmically with the particle energy (as in Eq. (2.2)).

In sampling calorimeters a significant fraction of energy remains in the absorbing material due to the production of fragments of nuclei and does not contribute to the measurable signal. At the same incident energy the ratio of the measurable energies of electromagnetic to hadronic showers, the so-called *sampling fraction* is larger than unity, i.e. $e/h > 1$.

In order to reduce the differences from the energy measurement of electromagnetic and hadronic showers, the energy of the latter which remains in the absorbing material has to be compensated. For this purpose the ZEUS calorimeter was constructed using depleted uranium (DU) as an absorbing material of a thickness of 3.3 mm (equivalent to one radiation length X_0 in that material) along with plastic scintillators as active material of 2.6 mm (equivalent to $0.006 \cdot X_0$) per plate. A hadronic shower traversing the DU plates produces fast neutrons whose kinetic energies are proportional to the energy loss due to nuclear reaction of the shower. These neutrons lose most of their energy by scattering elastically off the free protons of the scintillator and hardly due to any scattering off the heavy uranium nuclei. The deposited energy of the neutrons compensates the energy loss of the hadronic shower and achieves a sampling fraction close to 1 [57]. It should be mentioned that a compensating calorimeter can also be obtained with lead, provided that the layers of active material and passive material have the proper thickness relative to each other [58].

2.2.2.2 The ZEUS Calorimeter

The high-resolution ZEUS calorimeter is a sampling calorimeter with equal response to electrons (positrons and photons) and hadrons ($e/h = 1$) using absorber plates of depleted uranium and plastic scintillator layers as active plates. The depleted uranium plates have a composition of 98.1% U^{238}, 1.7% Nb and less than 0.2% U^{235}. The optical readout is performed via plastic wavelength shifters, lightguides and photomultipliers. In order to obtain equal calorimeter response to electron (positrons and photons) and hadrons, 3.3 mm thick uranium plates corresponding to one radiation length alternate with 2.6 mm thick scintillator plates. For a detailed description see [59] and [60]. The main features of the high-resolution calorimeter are:

- hermeticity over a large solid angle (99.7% of the solid angle is covered);
- energy resolution for hadrons of $\sigma(E)/E = 35\%/\sqrt{E} \oplus 2\%$;
- energy resolution for electrons of $\sigma(E)/E = 18\%/\sqrt{(E)} \oplus 2\%$;
- calibration of the absolute energy scale to 1%;
- precise angular resolution for particles ($\leq 10\%$ mrad);
- longitudinal segmentation for hadron-electron separation;
- short signal-processing time at the nano-second level.

The solenoid and the tracking chambers are completely enclosed by the uranium calorimeter, except in the region around the beam-pipe. The UCAL is constructed in three main parts, the hollow cylindrical central part of BCAL with a polar angle acceptance of $36^o < \theta < 129.1^o$ and the two end caps FCAL and RCAL, which cover the polar angle of $2.2^o < \theta < 39.9^o$ and $128^o < \theta < 176.5^o$, respectively (see figure 2.8 and table 2.1). The total depth of the three different calorimeter sections takes into account the Lorentz boost and reflects the maximum energy deposition of scattered particles. The FCAL has a maximum depth of 7.1 interaction length (λ_{int}), while BCAL has $5.3\lambda_{int}$ and RCAL $4\lambda_{int}$. Each section is divided into modules and further subdivided into towers which provide granularity for position reconstruction.

The three calorimeter components have similar structures and are subdivided longitudinally into an electromagnetic calorimeter (FEMC, BEMC, REMC) with a depth of $1\lambda_{int}$, which is

Part of the UCAL	Polar angle	Pseudorapidity
FCAL (forward)	$2.2^o < \theta < 39.9^o$	$4.0 > \eta > 1.0$
BCAL (barrel)	$36.7^o < \theta < 129.1^o$	$1.1 > \eta > -0.74$
RCAL (rear)	$129.1^o < \theta < 176.5^o$	$-0.74 > \eta > -3.5$

Table 2.1: CAL sections and the angular ranges covered by them. The polar angle and the pseudorapidity ranges are calculated with respect to the nominal interaction point. The regions of the FCAL and RCAL which are behind the BCAL are not considered.

sufficient to fully contain the electromagnetic showers, and a hadronic calorimeter (HAC). In FCAL and BCAL the HAC section is divided into two subsections HAC1 and HAC2, while the RCAL has only one HAC section.

The whole calorimeter has a modular structure. The FCAL and RCAL calorimeters consist of 23 modules which follow the same construction principles, as shown in figure 2.9, where an isometric view of a FCAL module is presented. The FCAL and RCAL modules have the same width of 20 cm and have a height vaying from 2.2 m to 4.6 m so that roughly a clindrical structure with a radius of $\sim$2.3 m can be built. The depth varies from $7.1\,\lambda_{int}$ in the central region to $5.6\lambda_{int}$ in the outer horizontal regions for FCAL. For RCAL the depth is smaller ($\sim 4\,\lambda_{int}$) since momentum conservation does not allow very high energy particles going in the rear direction of the detector.

The centre module of the FCAL and RCAL calorimeter are split into a separate upper and lower part leaving a hole in between to allow the beam-pipe crossing through the detector. The horizontal segmentation is determined by the width of the modules. The transverse segmentation depends on the height of the wavelenght shifters which collect the scintillator light. Each longitudinal section (EMC, HAC1, HAC2) is read out in both sides by wavelegth shifters. For FCAL the segmentation of the EMC section is 5x20 cm^2, and for RCAL 10x20 cm^2. These sections are called EMC cells. The hadronic cells of FCAL and RCAL have a segmentation of 20x20 cm^2.

The BCAL modules [43] present a similar distribution than in FCAL, but the geometrical design is slightly different to adapt to the barrel-shape requirements of this calorimeter part. Each module consists of 14 towers parallel to the z-axis. The 32 modules complete the circle around the z-axis containing the interaction point. In the BCAL, each tower is similar to one tower in FCAL except that the depth is around $5\,\lambda_{int}$. The first (last) tower in each module contains only 3 (2) electromagnetic cells. To increase the efficiency for detecting electromagnetic showers, the first and the last towers have their electromagnetic cells oriented towards the interaction point.

Due to the optical read out combined with photomultipliers a fast readout processing can be performed which allows the determination of the arrival times of incoming particles at the nanosecond level. The timing resolution for each calorimeter cell is $\sigma_t = 1.5/\sqrt{E} \oplus 0.5$ ns, where E (GeV) is the energy deposited in the cell [61]. The timing information from the calorimeter is useful to remove both beam-gas and cosmic ray related backgrounds. The time $t = 0$ ns is defined to be the time at which the particles originating from ep collinsions at the interaction point arrive at the calorimeter.

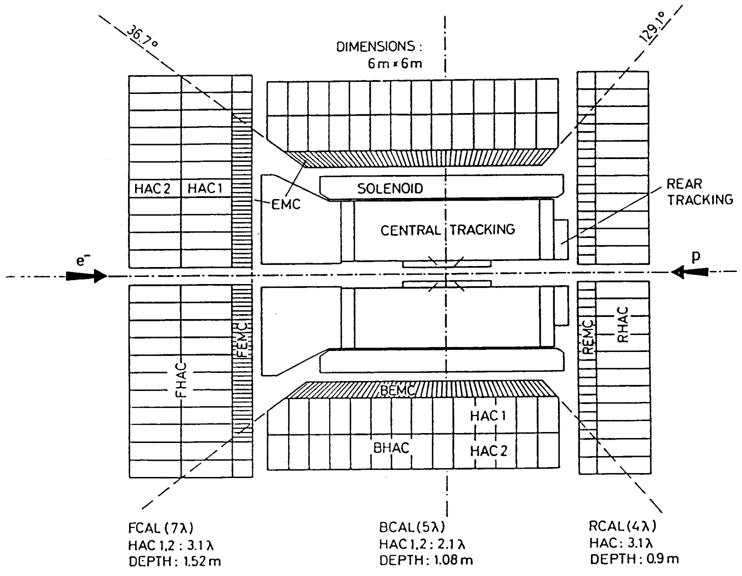


Figure 2.8: Layout of the ZEUS high resolution uranium-scintillator calorimeter. The picture shows the three sections of the UCAL and the polar-angle acceptance of each section.

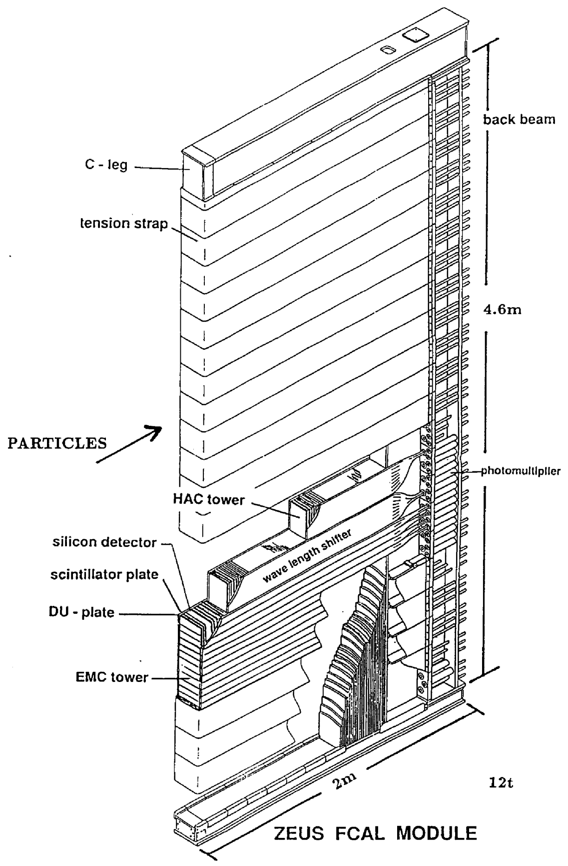


Figure 2.9: Internal structure of an FCAL module. The RCAL modules are very similar but they contain only two EMC cells instead of the four in FCAL and BCAL. One module in BCAL follows the same distribution as in FCAL but its geometry is adapted to the barrel requirements (see section 2.2.2.2).

When particles from beam-gas interactions which occur behind the RCAL, deposit energy in the RCAL, this time is negative. Therefore, a cut on the RCAL time can remove beam-gas events. In addition, the "up-down" time difference[5] should be zero for deposits related to an ep collision. However, for cosmic rays, this time difference is greater than 10 ns. Therefore, cosmic rays events can also be removed with the calorimeter time information. The vertex position along the the z-axis can also be measured using calorimeter time information. The time t at which energy is deposited in the FCAL near the beam-pipe from an ep collision will depend on the vertex position along the z-axis (z_{vtx}) according to the equation $t = -2z_{vtx}/c$, where c is the speed of light. The correlation between the z_{vtx} measured using the tracking detectors and the FCAL timing information can also be used to remove beam-gas interactions.

The readout of the calorimeter cells is fast enough to avoid mixing the information of two successive bunch crossings, but since the trigger decission about each event is slower (Sec. 2.2.6), the information has to be stored waiting for it. This is done by means of a pipelined readout of the calorimeter [62]. Each of the channels stores the information in a system of capacitors which delays the transmission of the read values during the time the trigger system takes a decision about the event. This system can store up to a maximum of 58 events, which means that the information is kept $\sim 5.6\mu$s after the event occurs. This time delay is enough for taking a fairly good decision about the usefulness of the event.

The calibration of the calorimeter is performed using several redundant tools, which are described in [60]. The main calibration source is the use of the natural uranium radioactivity, the so-called uranium noise (UNO), which produces a low background current in the photomultiplier. This current is statistically very stable and the deviations from the expected value allow the detection of problems in the operation of the photomultiplier. On the other hand, to calibrate the electronic system reading the information for each channel, injectors of charge are used to simulate the signal coming from the photomultiplier. Since the quantity of charge is known, the returned value given by the complete readout system is used to calibrate the effects of the electronics, after substraction of the noise contribution. All these tools provides a stable diagnostic tool for monitoring and calibration of the calorimeter.

2.2.3 The Backing Calorimeter and the Muon Chambers

The main purpose of the Backing Calorimeter (BAC) is to measure the energy leakage out of the main calorimeter. In addition, it is used to detect muons from ep interactions and cosmics. The BAC totally covers the UCAL using as absorber material the 7.3 cm thick plates of the iron yoke. The iron yoke for the return of the magnetic field has the shape of an octogonal prism which is closed by endcaps on both sides. The yoke can be magnetized by copper coils, in order to create a magnetic field for an independent measurement of the momentum of muons in the barrel muon chambers. It consists of slabs of iron separated by 3.7 cm gaps instrumented with proportional tubes.

The Muon Chambers are located inside and outside of the iron yoke, and measures tracks penetrating the calorimeter and cosmics rays. This detector is constructed in three main modules: the forward (FMUON), barrel (BMUON) and the rear region (RMUON); more components are placed in the forward region in order to cover the higher particle flux. By measuring the momenta of reconstructed track segments, and comparing this measurement with track segments

[5] Defined as the difference between the time at which energy is deposited in cells at the top and at the bottom of the BCAL.

in the CTD, prompt muons can be detected. Each barrel chamber consists of 2 double layers of limited streamer tubes, with wires oriented along the beam axis. The position resolution is better than 1 mm. The barrel muon chambers accept muon tracks at $\theta > 34^o$. In order to reach the muon chamber, muons must transverse the UCAL and BAC. As a minimum ionizing particle, the muon typically deposits ~2 GeV in the UCAL and this is the energy needed for the muon to be detected in the inner muon chambers. In crossing the iron yoke between the inner and the outer muon chambers, the muon loses approximately 1 GeV of energy.

The forward direction is equipped with special muon chambers, which give an independent measurement of the muon momentum up to 100 GeV and down to very small angles. The detector consists of a toroidally magnetized iron region, interleaved with sections of drift chambers, limited streamer tubes and time of flight counters. The outer diameter of the toroids is 6 m and the average magnetic field inside is 1.7 T.

2.2.4 Luminosity measurement

Any process that is to be used to monitor the luminosity seen by an experiment must fulfil the requeriments that its cross section must be fairly large, so that a short measurement time is needed for a reasonably accurate estimate of the instantaneous luminosity, and that its cross section must be very well known, i.e. calculable and be short of uncertainties.

For HERA, the ep-bremsstrahlung process

$$e \, + \, p \, \rightarrow \, e' \, + \gamma \, + \, p$$

is used for the luminosity measurement. For the experimental measurement of this process, the outgoing electron and photon could be measured in coincidence in the so-called luminosity monitor (LUMI), with $E_{e'} + E_\gamma = E_e$, i.e. the scattered electron and photon energies must add up to the electron beam energy. The cross section for this processes is given semi-classically by the Bethke-Heitler formula [63]:

$$\frac{d\sigma_{bh}}{dE_\gamma} = 4\alpha r_c^2 \, \frac{E'_e}{E_\gamma \, E_e} \left(\frac{E_e}{E'_e} + \frac{E'_e}{E_e} - \frac{2}{3} \right) \left(\ln \frac{4E_p \, E_e \, E'_e}{M \, m \, E_\gamma} - \frac{1}{2} \right), \tag{2.3}$$

where E_γ is the photon energy, E_e, E'_e are the energies of the initial and final electrons, E_p is the proton energy, M (m) is the proton (electron) mass, α is the fine structure constant and r_c is the classical electron radius.

It has been shown that, within the experimental conditions, this cross section agrees excellently with QED calculations [64]. The radiative corrections to this process have been calculated and amount to -0.3% within the measurable area of phase space. The ZEUS luminosity monitor consists of two separate detectors, one for the photon branch and one for the electron branch. The layout of the luminosity monitor components is shown in figure 2.10.

Photons emerging from the bremsstrahlung process at scattering angles $\theta_\gamma \leq 0.5$ mrad leave the beam pipe through a window located at a distance of 92.5 m from the interaction point, and are measured in a lead-scintillator electromagnetic calorimeter at 107 m from the interaction point. In front of the photon calorimeter, a 2 radiation length carbon filter absorbs the large flux of synchrotron radiation photons, which typically have low energy ($\sim$ keV), whereas the ep-bremsstrahlung photons used for the luminosity measurement have energies well above 1 GeV. A Cerenkov counter placed between the filter and the calorimeter detects events in which the

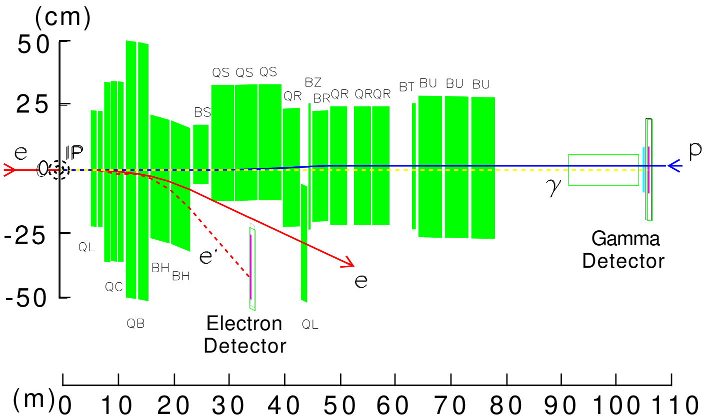


Figure 2.10: Layout of the ZEUS luminosity monitor.

photon has been converted in the filter in an e^+e^- pair. The acceptance for bremsstrahlung photons is about 98%, independent of energy.

Electrons scattered at angles $\theta_{e'} \leq 6$ mrad, and with energies $0.2E_e \leq E_{e'} \leq 0.9E_e$ are deflected by the beam magnets away from the nominal electron beam orbit. They leave the electron beam line through a window at 27.5 m from the interaction point, and hit the electron calorimeter located at 34.7 m from the interaction point. The acceptance is measured to be above 70% and flat for scattered electrons with energies in the range $0.35E_e \leq E_{e'} \leq 0.65E_e$. This is the region used in the luminosity measurement. In addition, the electron calorimeter of the LUMI can be used to tag electrons scattered at very small angles. In this case, no coincidence with the photon calorimeter is required. This electron-tagger mode covers the Q^2 range between 10^{-7} and $2 \cdot 10^{-2}$ GeV^2/c^2, although the location of this calorimeter is such that only a fraction of the electrons are tagged in this region of Q^2.

The luminosity value is easily obtained from the observed ep-bremsstrahlung rate R_{ep} by $L = R_{ep}/\sigma_{obs}$ where σ_{obs} is the ep-bremsstrahlung cross section corrected for detector inefficiencies and acceptances. The uncertainty in σ_{obs} is small and, therefore, the error in the luminosity measurement is largely dominated by the uncertainty in R_{ep}. The luminosity measurement suffers from a number of background processes. By requiring a coincidence between a high-energy electron and photon, this background is dominated by interactions of electrons with the remaining gas in the beam pipe. The most problematic contribution is the beam-gas bremsstrahlung process

$$e\,Z \;\rightarrow e'\,Z\,\gamma$$

which is impossible to distinguish from the process we want to measure and its contribution to the total counting rate of the luminosity monitor has to be estimated by using the pilot electron bunches. Let the fully measured counting rate of the luminosity detector be R_{tot}, the rate of the pilot bunches be R_{pilot}, the total current in the electron ring be I_{tot}, and the current in the pilot bunches be I_{pilot}, the actual ep-bremsstrahlung rate for the luminosity measurement can be estimated as:

$$R_{ep} = R_{tot} - R_{pilot} \cdot \frac{I_{tot}}{I_{pilot}},$$

from where the luminosity can be extracted using the Bethe-Heitler expression corrected for the acceptance of the LUMI detector. For the determination of the luminosity photons with

Period	HERA delivered (pb^{-1})	ZEUS on-tape(pb^{-1})	ZEUS physics (pb^{-1})	ZEUS physics (10^6events)
e^- 93-94	2.17	0.99	0.82 ($\pm$1.5%)	8.34
e^+ 94-97	70.92	51.2	48.32 ($\pm$1.1%)	80.58
e^- 98-99	25.2	17.78	16.67($\pm$1.8%)	23.57
e^+ 99-00	94.95	73.37	66.04($\pm$2.25%)	66.76

Table 2.2: Luminosity collected by ZEUS in 1993-2000. The difference between "on-tape" and "physics" comes from the rejection of data considered not useful for analysis by the offline data quality monitoring due to problems in the detector.

energies in the range 10-16 GeV are used. However, these cut values are changed for systematic checks.

The ZEUS luminosity available for physics analysis is listed in Table 2.2 for the 1993-2000 running periods. In these years the average ZEUS efficiency was around 70%.

2.2.5 Other components

Veto Wall

The veto wall (VETO) is an iron wall equipped with two layers of scintillator on both sides and placed 7 m away from the interaction point in the electron direction. Its dimensions are 800 cm x 760 cm x 87 cm, with a square hole of 95x95 cm^2 in the center for the beam pipe and magnets. Its main purpose is to shield the detector against particles from the proton beam halo and to give a veto signal for the main detector components.

C5 Counter

The C5 Counter is an scintillator counter which is located near the collimator C5, about 3 m away from the interaction point in the electron direction, very close to the beam (4 cm - 10 cm). Its main purpose is to provide an accurate time measurement for the backgroud in order to veto events which are not in the time window of the crossing of electron and proton bunches. From this timing information it is possible to reduce the background rate by an order of magnitude.

The Small-angle Rear Track Detector

The SRTD is installed in the region around the RCAL beam pipe [65]. The active part of the SRTD consists of arrays of scintillator strips grouped in 4 quadrants, 24x24 cm^2. Each quadrant contains two layers of scintillator strips arranged in orthogonal directions. The SRTD improves the measurement of charged particles, mainly small-angle scattered electrons, which is essential for an accurate reconstruction of the event kinematics at low Q^2.

2.2.6 Trigger and Data Acquisition Systems

At HERA one of the challenges for the data acquisition system is the bunch crossing time of 96 ns, clearly too short to perform a full data readout or to make a trigger decision. The purpose of the ZEUS trigger system is to distinguish interesting physics events from the background events. The total interaction rate, which is dominated by background from upstream interactions of the proton beam with residual gas in the beam pipe, is of the order of 10-100 kHZ while the rate of ep physics events in the detector is of the order of 3-5 Hz. In addition, other background sources like electron-gas collisions, halo muons and cosmic rays are also present. Background suppression is achieved by a sophisticated three-level trigger system [66], where each successive level has more time available to take more complicated trigger decissions. A schematic diagram of the ZEUS trigger and data acquisition systems is presented in figure 2.11.

The *First Level Trigger* (FLT) is a hardware trigger, designed to reduce the input rate below 1 kHz. Each component of the ZEUS detector has its own FLT, which stores the data in a pipeline and makes a trigger decision within 2 μs after the bunch crossing. The decisions from all the FLT parts are collected by the Global First Level Trigger (GFLT), which decides whether to accept or reject the event, and returns this decision to the readouts of the different components within 4.6 μs, corresponding to 46 bunch crossings. The FLT operates only on a small subset of detector data in order to achieve such rate in the decision. At this level most of the beam-gas and beam-halo events are eliminated.

If the event is accepted by the FLT, the data are transferred to the *Second Level Trigger* (SLT), which is software-based and runs on a network of transputers. It is designed to reduce the rate below 100 Hz. For the SLT almost all the event information and the full dynamic range are used. The Global Second Level Trigger (GSLT) receives all of the SLT information from the individual components as well as the trigger information from the GFLT. The SLT produces a decision within 3 ms after the corresponding *ep* interaction. If the event is accepted by the GSLT, all the detector components send their data to the Event Builder, which collects the information to reconstruct a complete event. The event is then passed to the *Third Level*

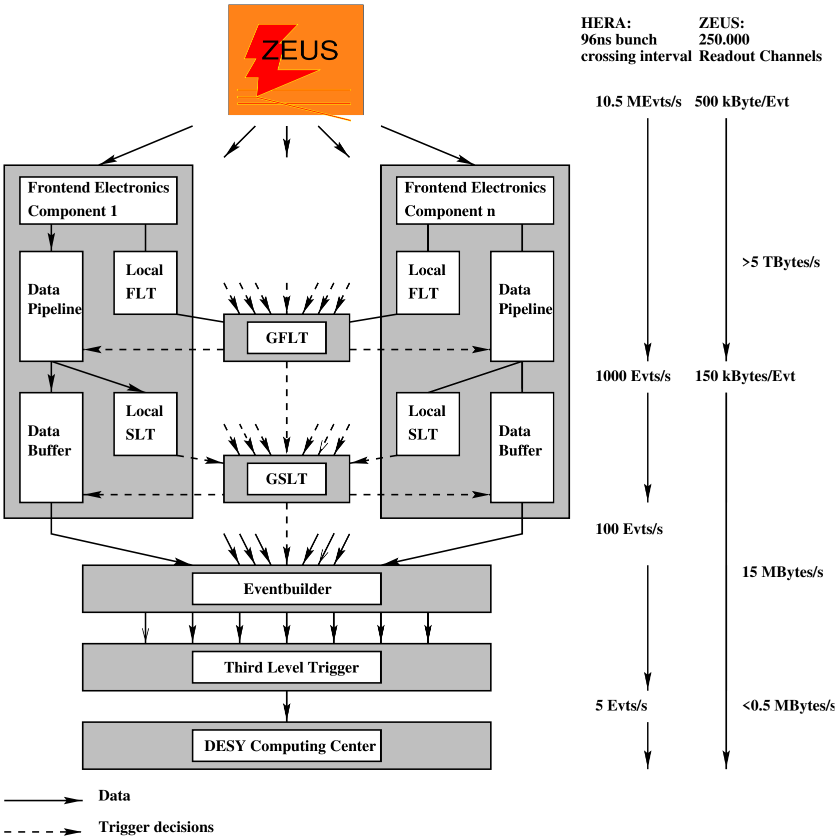


Figure 2.11: Scheme of the ZEUS data adquisition system. The effects of the trigger selection on lowering the rate (in events per second and in data size) is shown in the lines parallel to the decision diagram.

Trigger (TLT) which is software-based and runs a reduced version of the offline data analysis code. It is designed to reduce the rate to a few Hz. Events accepted by the TLT are written to tape for further analysis.

2.2.7 Offline and detector simulation

The data previously stored on tape is reconstructed with the ZEPHYR package. During the reconstruction a preselection logic based on very soft, process oriented, requirements is performed. The results of this preselection are DST bits, which are stored in the header of the event file. Only the header is read for events which do not fulfill the required DST bit logic. This allows to run more efficiently over large amounts of data samples.

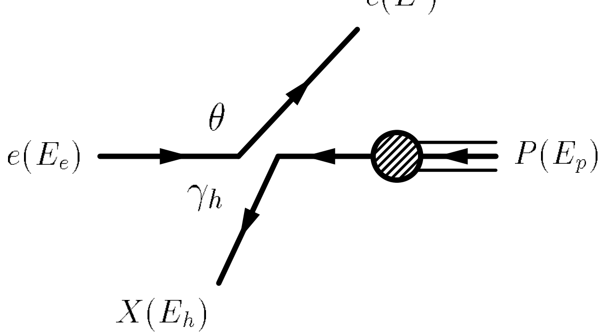
During the reconstruction procedure the information of the different components is reanalyzed applying corrections given by the data quality monitoring and the calibration of the different channels on each component. At that point the whole detector information is available and can be used for the reconstruction of information which is obtained indirectly from the detector. For example the tracking reconstruction using the different hits in the tracking chambers or the reconstruction of the vertex position using the reconstructed tracks.

After this point the data are written on disk and is available for the final physics analyses [67].

The Monte Carlo (MC) detector simulation is performed with the MOZART program (based in GEANT [68]). It allows the application of the detector response to simulated final states. After passing through the simulation the MC samples supply the detector information in the same format as for the data. This information is input for the trigger simulation (ZGANA) and afterwards the simulated events are reconstructed using ZEPHYR as for the real data.

2.3 Reconstruction of the kinematic variables

A precise reconstruction of x and Q^2 from the final state is of vital importance for the description of DIS events. The kinematic reconstruction based on quantities measurable in the ZEUS detector using several methods is described in this section. In deep inelastic scattering, the kinematics of the event must be reconstructed from the measured final state. At first order, the final state in a neutral current DIS event consists of the scattered lepton and the scattered quark



where E_e, E' and θ are the initial energy, final energy and scattered angle of the lepton, respectively. E_p, E_h and γ_h are the beam proton energy, the final hadronic energy and the polar angle of the final hadronic system respectively. The kinematics of the event can be completely known from any pair of the variables E', E_h, θ and γ_h. In the following, some methods are considered and they are classified depending on the information that was used: either the electron variables, the hadronic final state information or a combination of the scattered lepton and hadronic system variables.

2.3.1 Reconstruction using the scattered electron variables

In neutral-current DIS there is an electron in the final state. In fixed target experiments, x and Q^2 are reconstructed from the angle and energy of the final-state electron. Using the definition of the kinematic variables given in the previous chapter, the following expressions are obtained:

$$\begin{aligned} Q^2_{ele} &= 2E_e E' \left(1+\cos\theta\right) \\ y_{ele} &= 1 - \frac{E'}{2E_e}\left(1-\cos\theta\right) \\ x_{ele} &= \frac{Q^2_{ele}}{s\, y_{ele}} \end{aligned} \tag{2.4}$$

It is noted that the following relation holds

$$Q^2_{ele} = \frac{(p_{T,ele})^2}{1-y_{ele}}\,, \tag{2.5}$$

where $p_{T,ele} = E'\sin\theta$ (in the massless aproximation) is the transverse momentum of the scattered electron. The contours of constant energy or scattering angle in the (x, Q^2) plane accesible to HERA are displayed in figure 2.12. It is seen that at low values of Q^2 the electron scattering angle is large (measured with respect to the proton direction) and it hits the rear part of the UCAL. Occassionally the electron is deflected to the forward part of the detector. The scattering angle increases with Q^2, and is largely independent of x, except at high values of y.

The very low-Q^2 region, which is not covered by the nominal vertex data due to the beam hole in the ZEUS calorimeter, is indicated by the grey-shading in the right-plot of figure 2.12.

From the expressions:

$$\begin{aligned} (\Delta Q^2/Q^2)_{E'} &= \Delta E'/E' \\ (\Delta Q^2/Q^2)_{\theta} &= \tan(\theta/2)\,\Delta\theta \\ (\Delta x/x)_{E'} &= \frac{1}{y}\,\Delta E'/E' \\ (\Delta x/x)_{\theta} &= (x\frac{E_p}{E'_e} - 1)\tan(\theta/2)\,\Delta\theta \end{aligned}$$

it is seen that the resolution in Q^2 as obtained by the electron variables is very good (except for very small scattering angles, that is, high Q^2 values). However, the resolution in x becomes very poor at low y due to the limited energy resolution.

2.3.2 Reconstruction using the hadronic system

Since ZEUS is an almost 4π detector, it is possible to use a completely independent reconstruction method based on the hadronic final state. This method is mainly used when the electron variables are not reconstructed since the final lepton is not detected. For neutral-current DIS events this method allows to extend the measurable region in the (x, Q^2) plane to that in which the electron scapes through the beam pipe hole. For charged-current DIS events, where the final-state neutrino leaves the detector unobserved, this method provides the only way to reconstruct the kinematics.

From the scattered-quark energy E_h and angle with respect to the proton direction γ_h, the kinematics can be reconstructed using the following expressions:

$$\begin{aligned} y &= E_h \, \frac{1-\cos\gamma_h}{2\,E_p} \\ Q^2 &= {E_h}^2 \, \frac{\sin^2\gamma_h}{1-y} \\ x &= \frac{Q^2}{s\,y} \end{aligned} \tag{2.6}$$

However it is not possible to measure directly E_h and γ_h. One possibility could be to look for jets in the final state and associate one (or more) of them with the struck quark. Since jets could be originated from soft and hard QCD processes, their use obscures the direct use of E_h and γ_h. Instead, the hadronic energy flow as measured in the cells of a segmented calorimeter can be used without explicit jet finding. Jacquet and Blondel [69] proposed such a method, based on an appropiate weighting of the energy deposits in the cells. The idea is that the hadronic final state consists of a component that has taken part in the hard scattering (which therefore contains the information on E_h and γ_h) and a component that was only an spectator during the interaction (the target remnant). The kinematics can be accurately reconstructed from the hadronic energy flow if the latter is suppressed. The spectator component is characterized by a boost in the incoming proton direction. The Jacquet-Blondel method uses the following expressions:

$$\begin{aligned} y_{JB} &= \frac{1}{2\,E_e} \sum_i (E_i - p_{z,i}) \\ {Q^2}_{JB} &= \frac{1}{1-y_{JB}} \, (\sum_i p_{T,i})^2 \\ x_{JB} &= \frac{{Q^2}_{JB}}{s\,y_{JB}}\,, \end{aligned} \tag{2.7}$$

where the sums run over the hadronic particles (or energy deposits) in the final state. The influence of the target remnant and any hadron lost in the forward beam-pipe hole, characterized by low p_T and low $E - p_z$, is suppressed in the above expressions. The problem with the use of the hadronic energy flow lies in the limited hadronic energy resolution of a calorimeter.

Figure 2.13 shows the contours of constant current-jet energy and angle in the (x, Q^2) plane. In contrast to the electron method, the effect on the resolution in x from the measurement errors of the jet energy is small at high x. However, the forward beam-pipe hole at $\lesssim 2.2^o$ limits the reach of this reconstruction method to $y > 10^{-3}$. At very low x and high y a similar limit is set by the rear beam hole (see grey-shaded regions in the right-plot of figure 2.13).

2.3.3 Reconstruction from mixed variables

The methods which combine information from the electron and the hadronic final state are known as *mixed methods*. Taking into account the advantages of both the electron and the Jacquet-Blodel methods, it is straightforward to design the following expressions:

$$
\begin{aligned}
y_{mix} &= y_{JB} \\
{Q^2}_{mix} &= {Q^2}_{ele} \\
x_{mix} &= \frac{{Q^2}_{mix}}{s\, y_{mix}}
\end{aligned}
$$

The Double-Angle method

In addition, it has been proved useful to use θ and γ_h in the event kinematics reconstruction (the "double angle" (DA) method) [70], which exploits the advantage that angles are measured with better accuracy than energies, and do not suffer (to first order) from calorimeter energy calibration uncertainties. The expressions to reconstruct the kinematic variables in terms of these two angles are the following:

$$
\begin{aligned}
y_{DA} &= \frac{\sin\theta\,(1-\cos\gamma_h)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h+\theta)} \\
{Q^2}_{DA} &= 4{E_e}^2\,\frac{\sin\gamma_h\,(1+cos\theta)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h+\theta)} \\
x_{DA} &= \frac{{Q^2}_{DA}}{s\, y_{DA}} = \frac{E_e}{E_p}\;\frac{\sin\gamma_h + \sin\theta + \sin(\gamma_h+\theta)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h+\theta)}
\end{aligned}
\tag{2.8}
$$

The difficulty in the DA method appears in the determination of γ_h. It has been suggested to determine γ_h from the Jacquet-Blondel variables,

$$
\cos\gamma_h = 1 - \frac{8{y_{JB}}^2{E_e}^2}{{Q^2}_{JB}(1-y_{JB}) + 4{y_{JB}}^2{E_e}^2}\;,
$$

which is completely equivalent to (in terms of the four-momentum of the final particles or energy deposits)

$$
\cos\gamma_h = \frac{(\sum p_x)^2 + (\sum p_y)^2 - (\sum E - p_z)^2}{(\sum p_x)^2 + (\sum p_y)^2 + (\sum E - p_z)^2}\;, \tag{2.9}
$$

where the sums run over all the particles (energy deposits) associated to the hadronic final state. It should be noted that when the transverse momentum of the spectator jet and the jet masses are neglected, this expression is equivalent to the angle of the current jet defined as

$$
\cos\gamma_h = \frac{\sum E_i\;\cos\gamma_{h,i}}{\sum E_i}\;,
$$

where the sum now runs over all the particles in the jet.

The Σ method

The Σ method [71] combines the measurements of energies and angles of the electron and hadronic system. Making use of the relation

$$\sum_i (E_{h,i} - p_{z,i}) + E'(1 - \cos\theta) = 2E_e \ ,$$

which holds exactly for a hermetic and ideal detector, a new expression for y can be derived

$$y_\Sigma = \frac{\sum_i (E_{h,i} - p_{z,i})}{\sum_i (E_{h,i} - p_{z,i}) + E'(1 - \cos\theta)} \ ,$$

which gives y at the hard interaction vertex even if an initial state radiation (ISR) photon has been emitted. Using the expression

$$Q^2{}_\Sigma = \frac{(E' \sin\theta)^2}{1 - y_\Sigma}$$

it is possible to determine Q^2 independently of ISR assumming that the initial-state radiated photon is collinear to the beam line and does not carry transverse momentum. The value of x_Σ is calculated from the last two expressions using the usual relation:

$$x_\Sigma = Q^2{}_\Sigma / (s \ y_\Sigma)$$

The e-Σ method

Considering the good resolution for $Q^2{}_{ele}$ it is possible to use it in combination with y_Σ to make a kinematics reconstruction which is less sensitive to ISR photons. The variable x is reconstructed from these two variables using the usual relation:

$$x_\Sigma = Q^2{}_{ele} / (s \ y_\Sigma)$$

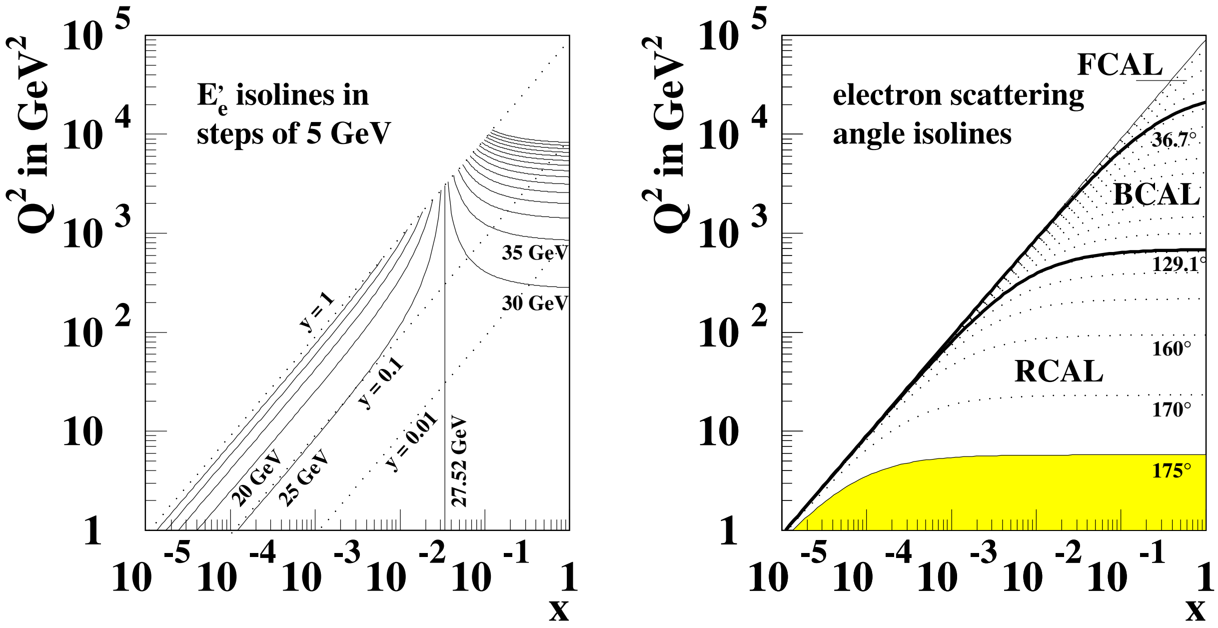


Figure 2.12: Contours of constant electron energy (left) and scattering angle (right) in the (x, Q^2) plane. The shaded region shows the angles for which the scattered electron escapes through the rear beam-pipe hole. In that case the electron is not detected in the main calorimeter.

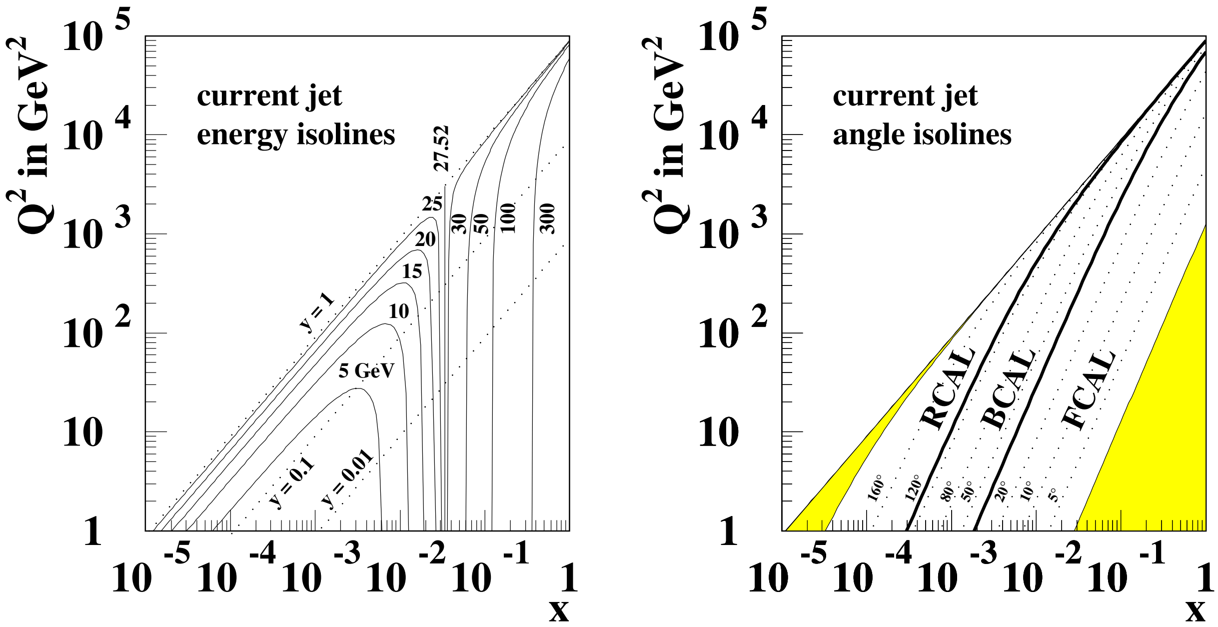


Figure 2.13: Contours of constant current jet energy (left) and hadronic angle (right) obtained in the QPM approach in the (x, Q^2) plane. The shaded regions show the angular ranges in which the scattered quark escapes through the forward or rear beam-pipe holes.

Chapter 3

Event simulation

In order to understand how the detector effects modify the observation of the final state in a high energy process is necessary to study these effects by means of a detector simulation. This simulation is applied to final state particles which have been generated according to the physical process under study. The generation of *Monte Carlo (MC) events* has become a fundamental tool in high energy experiments for the understanding of the complicated detectors.

On the other hand, since the measurements are done for hadrons but pQCD only gives predictions for partons, it is necessary to incorporate some phenomenological models to transform the partons into hadrons. These models are not only useful to obtain the simulated events at hadron level which are passed through the detector simulation. They can also be used to estimate the corrections in going from partons to hadrons; these corrections are needed to make a direct comparison with the theory. Furthermore, the comparison of different models allows the estimation of the uncertainty for these corrections.

In this chapter, a description of the most important aspects of Monte Carlo event generation is presented, particularly for the models and generators which were used in the analysis

3.1 Global description

As already mentioned in section 2.2.7, generated events are passed through the detector and trigger simulation to study the response of the detector. The output of this simulation has the same format as the real data, but it also contains the information about the generated event. This information consists of tables containing the list of particles which are created in the process under study with a Monte Carlo event generator.

In the case of *ep* scattering the generation of simulated events relies in the phenomenological approaches to QCD to describe the different processes which appear in the partonic level. The theoretical justification of the way the MC event generators apply the different steps in the simulation, lies in the factorization theorem for hard processes [72]. That is the event generators utilize the fact that a scattering process with a hard scale can be factorized into separate stages (see figure 3.1):

1. **hard sub-process**: it involves a hard scale and can therefore be calculated in a fixed order perturbative expansion. The nature of this interaction determines the main characteristics of the event.

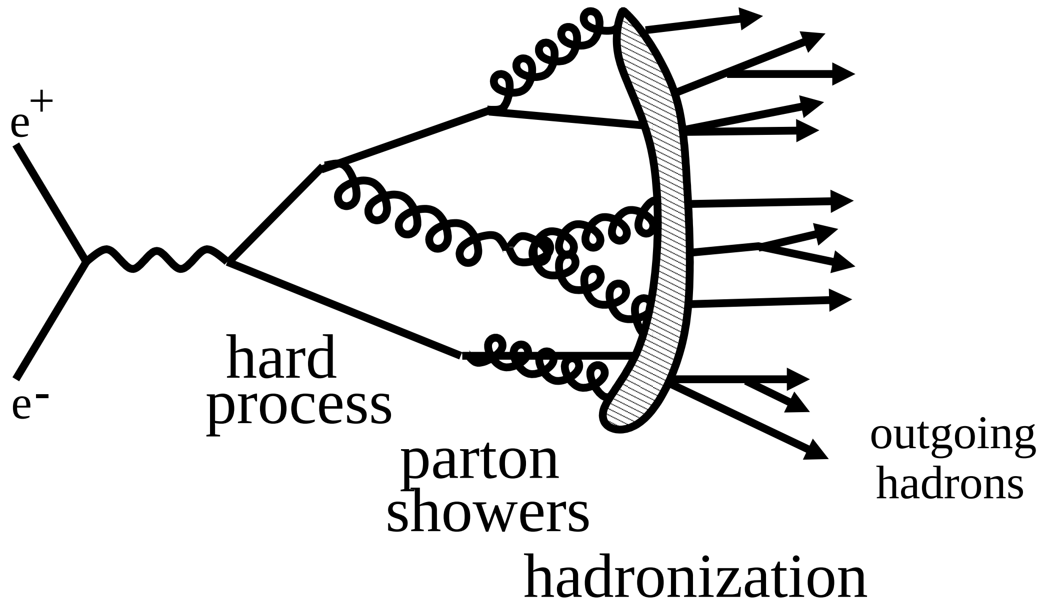


Figure 3.1: Representation of the different steps in the generation of an $e^+e^- \to$ hadrons event. The hard scattering is followed by the parton showering and the hadronization processes. It should be noted that the different steps are artificially separated in the generation.

2. **initial- and final-state radiation**: in every process that contains coloured and/or charged objets in the initial- or final-state, additional partons and/or photons may be radiated, which can give rise to large corrections to the overall topology of the event. A very common approach to the modelling of these radiation processes is the use of an arbitrary number of branchings of one particle into two by using the corresponding splitting functions (see figure 1.7 for the case of QCD radiation). This generates a shower of particles which configures the partonic final state in the case of QCD radiation. This parton radiation tends to simulate the higher-orders contribution to the cross section and the inclusion of partly-known higher-order terms is standard. It is done in a similar way as described for the evolution equations, although the parton radiation is clearly dominated by the approximation considered.

   In the case of QED radiation, the showers generate a set of particles which will appear in the final state and may have an important contribution to the global properties of the event, although it does not affect the processes which tranform partons into hadrons.

   Note that initial-state QCD radiation leads to the evolution of the parton distribution functions, as it was described in section 1.3.

3. **hadronization (fragmentation)**: it is the process by which the coloured partons are transformed into colourless hadrons. This non-perturbative process has yet to be understood from first principles and at present only phenomenological models are available to generate the hadronic final state starting from the partons.

4. **beam remnant fragmentation**: in the scattering process, the algorithm of the initial-state radiation reconstructs in each beam particle the shower initiator by backwards evolution from the hard sub-process. This shower initiator takes only some fraction of the total beam energy, leaving behind a beam remnant that takes the rest. If the shower initiator is coloured, so is the remnant. Being colour-connected to the hard interaction, the beam

remnant forms part of the same fragmenting system and needs to be reconstruted and connected to the rest of the event. In addition, in collisions where the two incoming beam particles have a composite nature (e.g. hadron-hadron interaction) there is the additional posibility that several parton pairs undergo separate hard or semihard scattering, usually referrerd to as 'multipartonic interactions'. At present there is very little understanding of these mechanisms and the event generators have to resort to phenomenological models and parametrizations from existing data to simulate these effects.

Following the scheme above, it should be clear that for the generation of simulated events, the two most important steps from the point of view of modellization are parton radiation and hadronization. The radiation of photons does not present the difficulties of parton radiation, mostly because it is easy to simulate than the complicated and unobservable processes which transforms the hard partons into observable hadrons.

For this reason, in the following we will concentrate on the simulation of parton radiation and in the models to generate the hadrons from a given set of partons.

3.2 Multipartonic production

The description of the hadronic final state in high energy processes requires the calculation of multiple partons emissions in QCD. For these higher-order QCD processes it is usually not possible to perform exact matrix element calculations (i.e. to obtain pQCD predictions) and one has to use aproximation schemes. Two different approaches will be described here and used in the analysis: the Parton Shower (PS) approach [73], based on the QCD parton-branching processes, and the Coulour Dipole Model (CDM) [74], in which the gluons are radiated from colour dipoles between the separated quarks, antiquarks and gluons originating from the initial hard process. Both models are able to describe the measurements in e^+e^- annihilation into hadrons [75]. However, the situation in deep inelastic scattering is more complicated, as mentioned before. In the following sections, a brief description of each model applied to DIS proceses is presented.

3.2.1 The Parton Shower approach

In DIS, the quark struck by the electroweak boson can emit partons both before and after the interaction giving rise to initial- and final-state parton showers, respectively; it is schematically represented in figure 3.2.

A parton close to the mass-shell in the incoming nucleon can initiate a parton cascade where in each branching one parton becomes increasingly off-shell with a space-like virtuality ($m^2 < 0$) and the other is on-shell or has a time-like virtuality ($m^2 > 0$). This initial-state space-like shower results in a space-like quark which finally interacts with the electroweak boson that turns it into an outgoing quark which is either on shell or has a time-like virtuality. After the interaction, the different partons with time-like virtualities generate a parton shower where the off-shell mass is reduced by branching into daughter partons with decreasing virtualities and decreasing opening angles. This shower continues until all partons are essentially on-shell.

The general behaviour of initial- and final-state showers are similar since they are both based on the branching processes as described by the DGLAP evolution equations (Sec. 1.3.4) in the leading-logarithm approximation of perturbative QCD.

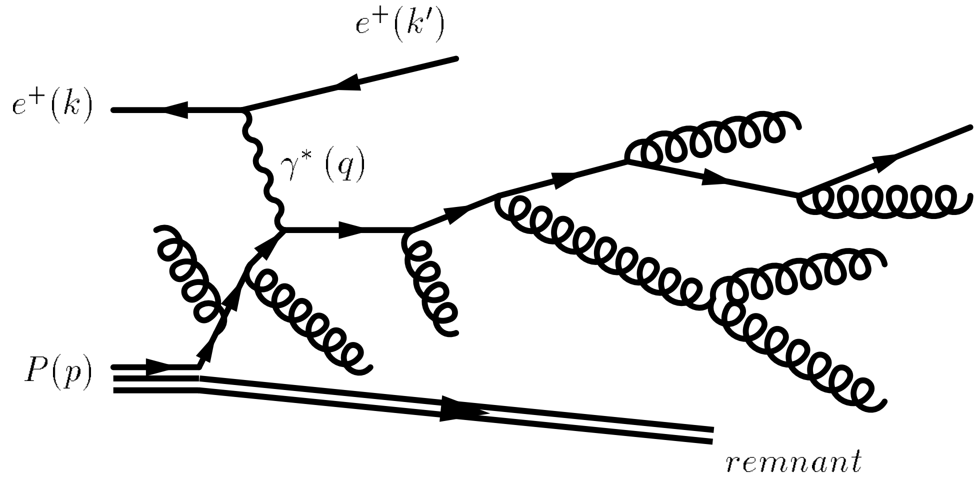


Figure 3.2: Schematic representation of the parton shower approach to the simulation of the parton cascade. The partons are generated by succesive splitting processes.

The final-state radiation is analogous to parton radiation in $e^+e^- \to q\bar{q}$. The evolution is based on the Sudakov form factor [76], which expresses the probability that a parton does not branch between some initial maximum virtuality and some minimum value. From this, one can find the mass of the decaying parton, the energy fraction in the branching and the flavour of the daughter partons. The process is iterated with a reduced virtuality until all parton vitualities are below some cut-off m_0^2 (around 1 GeV^2). Coherence in soft gluon emission is taken into account through angular ordering, by decreasing opening angles in subsequent branches [77].

Initial-state radiation is handled within the backwards-evolution scheme, where the shower is constructed from the hard electroweak interaction backwards with decreasing virtualities down to the on-shell partons in the incoming nucleon. This is a more complicated process since the nucleon parton densities must be taken into account. When combining the initial- and final-state radiation to construct the complete event, special precautions have to be taken in order to preserve energy-momentum conservation and keep the definition of the kinematic variables, in particular that of Bjorken x.

The amount and hardness of the initial- and final-state radiation depends on the degree of off-shellness of the struck quark just before and after the interaction with the electroweak gauge boson. These virtualities are chosen, based on a Sudakov form factor, between m_0^2 and a maximum to be set by the energy or momentum transfer scale in the process.

The parton-shower approach has some shortcomings due to its approximate nature. The separation of initial- and final-state parton emission implies the neglect of interference terms between the two and it is not gauge invariant. The use of a leading-order approximation means that the emission of the partons that are soft or close to the direction of the emitting partons should be well described, while the emission of hard partons at large angles could be mistreated. Therefore, the rate of events with additional hard partons that give rise to separate jets will not be well described. These problems were partially solved in e^+e^- by correcting ths first emission in the parton shower to agree with the result from exact pQCD first order calculations. In DIS, this matching is more difficult due to the more complicated matrix elements. Starting with the matrix elements, the hard emission is generated and extra, but softer emissions are then added using the parton-shower approach. The singular region of the matrix elements is avoided by a cut-off in the invariant mass of every parton pair.

3.2.2 The Colour Dipole Model

The Colour Dipole Model (CDM) is based on the fact that a gluon emitted from a $q\bar{q}$ pair in an e^+e^- collision can be treated as a radiation from the colour dipole between the q and the $\bar{q}$, and that to a good approximation the emission of a second softer gluon can be treated as a radiation from two independent dipoles, one between the q and the first gluon and the second between the same gluon and the $\bar{q}$. This approach can be generalized and the emission of a third gluon is given by three independent dipoles, and so on (see figure 3.3).

For DIS, the CDM does not divide the parton radiation process into an initial and final state ones, as is made in the Parton Shower approach discussed above. Instead it is assumed that all the radiation can be described by the colour dipole formed between the struck quark and the proton remnant. In $e^+e^- \to q\bar{q}$ both the q and the $\bar{q}$ are treated as point-like particles, but in DIS only the struck quark can be considered point-like while the proton remnant is an extended object. It is a well known fact that emission of small wavelengths from an extended antenna is suppressed and that for an antenna with a certain transverse size, effectively only a part proportional to the transverse wavelength $\lambda \propto 1/p_T$ is actually participating in the emission. In the model, this feature is implemented so that only a fraction

$$a = (\mu/p_T)^\alpha$$

of the proton remnant takes part in the emission of a gluon with transverse momentum p_T, where μ describes the size and α the dimension of the proton remnant. In this way, the parton radiation is only dependent of the energy associated to the hadronic system in the hard interaction.

The CDM model correctly takes into account the QCD coherence effects included in the parton-shower approach by imposing angular ordering in the final-state emissions. The colour coherence effects due to the extension of the proton remnant included in the CDM has however no correspondence in the Parton Shower approach, where the remnant is considered as a non-radiating spectator.

The emission of partons from dipoles only describes the transverse momentum and rapidity distributionsof the emitted parton in the rest frame of the emitting dipole. To describe com-

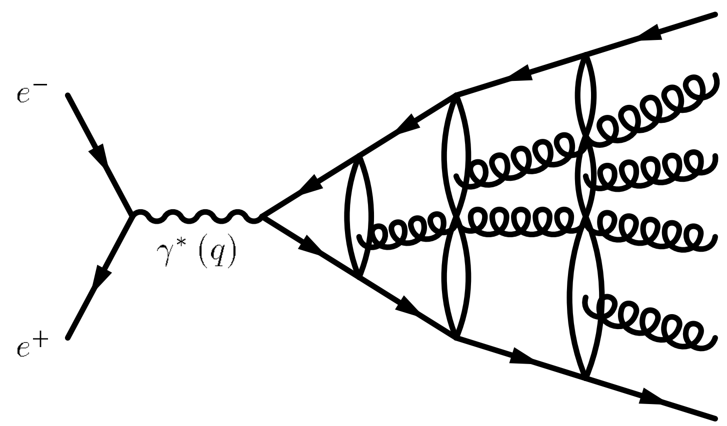


Figure 3.3: Schematic representation of the Colour Dipole Model for the simulation of the parton cascade of an $e^+e^- \to q\,\bar{q}$ event. The parton radiation is generated by colour dipoles defined between the partons present in each step.

pletely an emission, there are two more degrees of freedom to fix. One is the azimuthal angle. The other describes the distribution of transverse recoil among the emitting partons. This is also different in DIS as compared to $e^+e^- \rightarrow hadrons$. Since only a fraction of the hadron remnant takes part in the emission, only that part should be allowed to recoil. This is realized by introducing an extra "recoil" gluon taking a fraction a of the remnant energy. The recoil is then distributed between the recoil gluon and the other emitting parton.

Although gluon emission in the CDM is close to the form of the exact matrix element, the processes initiated by gluons (i.e. the Boson Gluon Fusion process) are not included at all. Therefore, a matching procedure has been introduced for the first emission in a DIS event. In this procedure, the initial dipole between the struck quark and the hadron remnant can either emit a gluon according to the $\mathcal{O}(\alpha_s)$ matrix element or emit the antipartner of the struck quark according to the BGF matrix element. The selection of the process is, as usual, made by the Sudakov form-factor prescription. In the case of BGF, there are then two dipoles, connecting each of the struck quarks with the hadron remnant, which continue radiating independently.

3.3 Hadronization models

After the parton shower has finished, the final state consists of a set of partons with virtualities of the order of the cut-off scale m_0^2. From this point, the low momentum-transfer (long-distance) regime, in which non-perturbative effects cannot be neglected, has to be considered to simulate the hadronization step.

One general approach to hadronization, based on the observation that perturbation theory seems to work well to rather low scales, is the hypothesis of *local parton-hadron duality*, where it is supposed that the flow of momentum and quantum numbers at the hadron level is established at the parton level. Hence, the flavour of the quark initiating a jet should be found in a hadron near the jet axis. The extent to which the hadron flow deviates from the parton flow reflects the smearing due to hadron formation. This smearing is related to energy scales of order Λ_{QCD} so its effects will be negligible if the energy scale associated to the partonic interaction is very large in comparison with Λ_{QCD}, which is of the order of 200-300 MeV.

Several phenomenological models have been constructed to simulate the formation of hadrons. Here three different models will be described: the independent fragmentation model, the string or Lund model and the cluster fragmentation model. The Lund model is the one which was used to generate the Monte Carlo samples which were used in the analysis.

3.3.1 The independent fragmentation model

The simplest scheme, suggested by Field and Feynman [78], for generating distributions of hadrons from those of partons is to assume that each parton fragments independently. The fragmentating quark q creates a colour field in which a new light $q'\bar{q}'$ pair is produced. A meson $q\bar{q}'$ is formed with a fraction z of the energy of the quark q. The leftover quark q', with energy fraction $(1-z)$, is fragmented in the same way, and so on until the leftover energy falls below some cut-off. For gluon fragmentation, the gluon is first split into a quark-antiquark pair, either assigning all the gluon's momentum to one or the other with equal probability, so that the gluon behaves as a quark of random flavour, or using the $g \rightarrow q\bar{q}$ splitting function.

Such a prescription contains an arbitrary function $f(z)$ plus other input parameters for the flavour/spin/mass options at each step. These parameters must be determined from the comparison with experimental data. The details of the procedure are, however, not fixed. One can identify z with the energy fraction or with the $E + p_L$ fraction instead. Alternatively, for example, the $q\bar{q}'$ pairs formed can be identified with low-mass hadronic cluster instead of single mesons.

3.3.2 The string model

The assumption of linear confinement provides the starting point for the string model [79]. As a q and $\bar{q}$ move apart from their common vertex, the physical picture is that of a colour vortex line being stretched between both partons. The transverse dimensions of the tube are of the order of the hadronic scale (~ 1 fm). If the tube is assumed to be uniform along its length, this produces a confinement picture with a linearly rising potential. In order to obtain a Lorentz covariant and causal description of the energy flow due to this linear confinement, a massless relativistic string with no transverse degree of freedom is needed. From hadron spectroscopy, the string constant, i.e. the amount of energy per unit length, is of the order of $k \sim 1$ GeV/fm. As q and $\bar{q}$ move apart the potential energy stored inside the string increases, and the string may break by the production of a new $q'\bar{q}'$ pair, so that the system splits into two colour singlet systems $q\bar{q}'$ and $q'\bar{q}$, each of them with a string evolving in an independent way (see drawing on the right of figure 3.4). If the invariant mass of either of these strings is large enough, further breakings may occur.

In the Lund string model [80], the string break-up process is assumed to proceed until only on-mass-shell hadrons remain, each hadron corresponding to a small piece of the string. In order to generate the quark-antiquark pairs, the Lund model invokes the idea of quantum mechanical tunnelling. This leads to a flavour independent gaussian spectrum of the p_T of the pairs and a suppression of heavy-quark production. A tunnelling mechanism can also be used to explain the production of baryons. In the simplest approach a diquark in a colour antitriplet state is just treated like an ordinary antiquark, such that a string can break either into a quark-antiquark or into an antidiquark-diquark pair.

If several partons are moving apart from a common origin, the string structure between them becomes more complicated. For a $q\bar{q}g$ event in e^+e^- collisions, a string is stretched from the q end to the $\bar{q}$ end via the gluon. The string will fragment along its length. To first approximation there is one fragmenting string piece between the quark and the gluon and a second one between the gluon and the antiquark. When considered in detail, the string motion and fragmentation is more complicated, with the appearance of additional string regions during the time evolution of the system. These corrections are specially important for soft and collinear gluons, since they have to provide a smooth transition between events where such radiation took place and events in which it did not. In this way, the string fragmentation model is *infrared safe* with respect to soft or collinear gluon emission.

3.3.3 The cluster fragmentation model

An important property of the branching process is the preconfinement of colour [81]. This suggests a class of *cluster hadronization models*, in which colour-singlet cluster of partons form after the perturbative phase of jet development and then decay into the observed hadrons (see

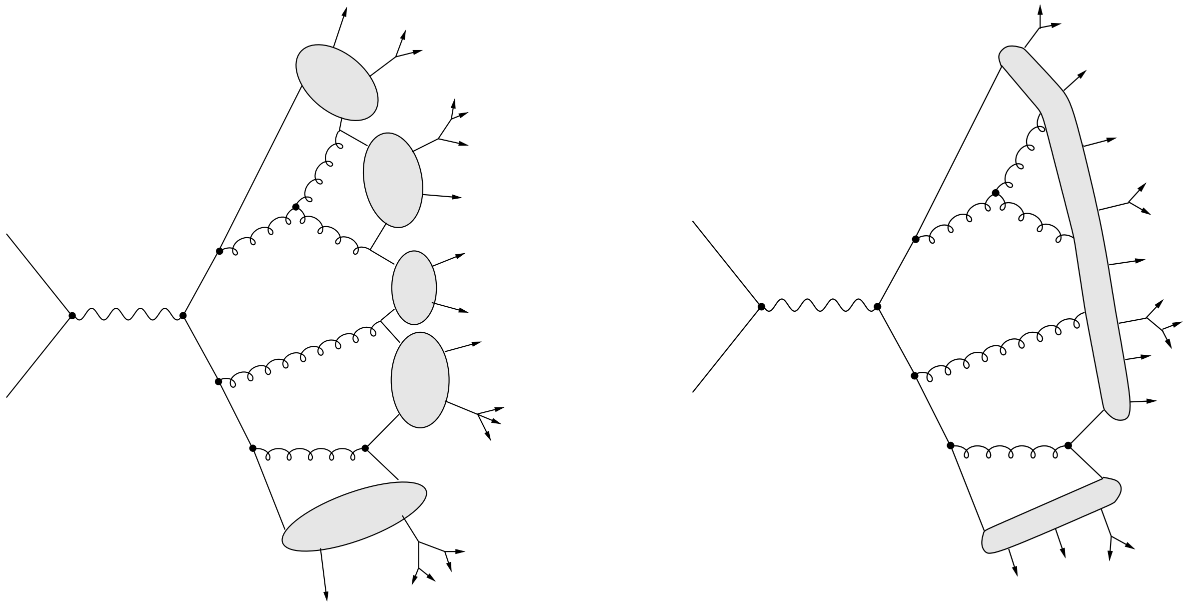


Figure 3.4: Schematic picture showing a parton shower followed by cluster hadronization (left) and Lund-string hadronization (right) for $e^+e^- \to$ hadrons.

drawing on the left of figure 3.4). The simplest way for colour-singlet clusters to form after parton branching is made via non-perturbative splitting of gluons into $q\bar{q}$ pairs [82].

Neighbouring quarks and antiquarks can then combine into singlets. The mass of the resulting clusters is typically two or three times m_0. If a low value of m_0 is used (of the order of 1 GeV) most clusters have masses of up to a few GeV and it is reasonable to consider them as superpositions of resonances. The hadronic energy and transverse momentum distributions for final states in $e^+e^- \to hadrons$ are quite well described without including additional adjustable fragmentation functions.

3.4 Event generators for neutral current DIS

In this section the main programs that are used for the simulation of events in neutral current deep inelastic scattering are presented. Obviously the event generator used in each case depends on the kind of events which are under study. Since the complete list of generator is very large, we concentrate here on those most commonly used for inclusive studies in DIS. These generators are those used in the analysis presented later on.

The LEPTO Monte Carlo generator

The first step of the event generation is, as mentioned above, the simulation of the hard scattering process. This can be done using the LEPTO program [83], which is based on the leading-order electroweak cross sections for the underlying parton level scattering, and includes QCD corrections using exact first order matrix elements.

Since LEPTO is able to calculate lepton-hadron scattering using the parton level, a set of PDF's has to be provided for the calculation. This is done using the standard library PDFLIB [84], which contains the information for the most important and used parametrizations of the proton parton densities.

Additionally, since LEPTO cannot be used to obtain radiative corrections, the events are generated using this program interfaced to HERACLES [85] via DJANGO [86]. The HERACLES program includes first-order electroweak corrections and the generation of events using this interface allows to calculate the size of the radiative corrections on one side. On the other hand, since these radiative corrections are present in the real data, the addition of them to the calculation should improve the similarity between the real events and the simulated events. This improvement is sometimes very important to understand the detector effects.

After the hard scattering process has been computed, the final products are considered as the input for the generation of the parton cascade. In LEPTO it is possible to select between the Parton Shower approach (in the Matrix-Element plus Parton Shower (MEPS) option of the program) and the Colour Dipole Model, as it is described below.

Once the parton radiation is completed, the hadronization is performed with the Lund hadronization model [79] as implemented in JETSET [87]. This program allows the selection of the independent hadronization model, but its use is not recommended with LEPTO since it does not provide a smooth transition from configurations with radiation of soft partons and those without it [83]. After the fragmentation process, the final-state hadrons are provided and the simulated event is ready to be used for the extraction of predictions to be compared to the data, or for the detector simulation as it is described above.

The MEPS and ARIADNE options for parton radiation

The parton cascade can be simulated following different phenomenological models, as it is mentioned in section 3.2. The most standard ones can be used to evolve the hard partons provided by the first step of the LEPTO program, as mentioned in the previous section.

The parton model approach is included in LEPTO as the MEPS option, which allows to calculate the predictions at parton level using the exact matrix element for the lowest orders and including higher orders in the leading logarithm parton shower approach, i.e. using the DGLAP equations. The simulation of the initial- and final-state parton radiation is done in the standard way described in section 3.2.1. However, it should be noted that the calculations to first order in α_s are done using the exact matrix element and this implies that the hard parton emission is better implemented than in a model where all the parton radiation is calculated using the Parton Shower approach, which is only a good model for soft radiative processes.

A switch allows the user to implement the parton radiation using the Colour Dipole Model and including the boson-gluon fusion process, as implemented in the ARIADNE program [88]. This program is not a complete event generator, and only generates the QCD cascade. It has to be interfaced to other programs which handle the hard interaction (as LEPTO) and the subsequent hadronization and particle decays.

Since the two models can be used with the same program for the calculation of the hard process, the comparison of two samples only differing on the choice of parton cascade allows the extraction of the differences between the models. This is often used to estimate the uncertanty in the prediction for the parton cascade.

The HERWIG Monte Carlo generator

The program HERWIG[1] [89] is a general multipurpose particle physics event generator including

[1]HERWIG stands for Hadron Emission Reactions With Interfering Gluons.

the simulation of a complete set of interaction processes in one package. Among the different kind of interactions, it is possible to select the simulation of neutral current DIS events. This can be done selecting the lowest order (QPM process) or the order-α_s processes. In both cases the higher orders are approximated with a parton cascade generated using the parton shower approach, as it was described above.

After the parton radiation, the fragmentation process is simulated using the cluster fragmentation model which is performed in different steps:

- All the outgoing gluons are split non-perturbatively into quark-antiquark pairs (and optionally diquark-antidiquark pairs).
- For the quark and diquark states, a colour line can be followed, in the planar approximation, from each quark to an antiquark or diquark with which it can form a colour-singlet cluster.
- These clusters satisfy the preconfinement property and are then fragmented into hadrons with a mechanism which globally is like the one described in section 3.3.3. Special actions are taken when the cluster has small mass.

The philosophy of this program is to describe the perturbative-calculable parts in as much detail as possible, so it assumes that the observable will be dominated by the perturbative part of the calculations. Following this approach, simple models are taken to describe the non-perturbative hadronization and beam remnant components, which results into a very few tunable parameters and a high predictive power.

On the other hand, the cluster fragmentation model usually gives a poorer description of the data than the Lund string model. This feature makes it not so useful for obtaining the detector effects at the hadron level by using the detector simulation. HERWIG is still very commonly used in the different analysis of DIS at HERA because it allows the estimation of the uncertainty coming from the hadronization process by comparing the prediction of the two different phenomenological models.

Chapter 4

Description of the analysis

The analysis was performed using the data taken with the ZEUS detector in the 96 and 97 running periods. During those years HERA collided protons of energy $E_p = 820$ GeV and positrons of energy $E_e = 27.5$ GeV. The integrated luminosity available for physics analyses is 38.4 ± 0.6 pb^{-1}. For the Monte Carlo, two independent samples were used, which differ in the way the parton radiation is simulated. In one, called "ARIADNE", the CDM has been used. In the called "LEPTO-MEPS", the parton shower approach simulates the parton cascade. The samples have been generated using standard parameters within the ZEUS Collaboration and no further tuning was made to improve the comparison with the data at detector level.

In this chapter, the complete experimental analysis is described. It shows the different requirements used to select the events in the data and simulated events.

4.1 Event selection

The first step in the analysis consists of selecting the events which are under study. The initial sample include all the events taken with the detector and trigger after the data have been processed. Depending of the kind of analysis, the selection requires different event properties. The effects of these requirements on the measurements are analyzed using simulated events, as is described in section 4.2.

4.1.1 Identification of the scattered positron

The main signature of neutral current DIS events is the presence of a scattered positron in the final state. A correct and efficient identification of the positron and a precise reconstruction of its position and energy are of vital importance for the analysis of this type of events. The different topology of electromagnetic and hadronic showers is the property used to identify the positron in the UCAL. In fact, using this property, it is very easy to identify an electromagnetic shower as the one associated to the scattered positron when it has sufficiently large energy and is well isolated.

In the kinematic regions where the positron is expected to have small energy, the identification of the electromagnetic shower is difficult because it is very similar to a hadronic shower of similar energy. Additionally, the segmentation of the calorimeter sets further limits on the energy values which allow to differentiate between a hadronic and an electromagnetic shower.

If the hadronic final state deposits energy very close to the positron, it is difficult to recognize the energy deposit as coming from an electromagnetic shower since the deposit is bigger than the shower initiated by the positron alone. The additional hadronic shower masks the positron deposit which no longer looks like an electromagnetic one.

Apart from the difficulties coming from positrons which are taken as hadronic deposits, others coming from particles looking like positrons are also present. Photons in the final state produce energy deposits which are very similar to those of positrons. Neutral pions are very frequently identified as positrons because their very fast decay into two photons ($\pi^0 \rightarrow \gamma\gamma$) is detected in the calorimeter as only one electromagnetic shower if the pion has sufficiently large energy[1]. Other particles present similar decays to photons, but they are not so frequent due to their larger masses. In any case, all these fake positron signals in the calorimeter give rise to background, which has to be suppressed.

There are several parameters which can be combined to extract the information to identify positron candidates.

The complexity of the possible ways to obtain a combined, but simple answer made the problem appropriate for being solved by a neural network. In this analysis the neural network SINISTRA [90] is used for the identification of the scattered positron. SINISTRA is based on islands of calorimeter cells, where an island is a group of cells which are merged following a predefined criterion: each cell is merged with the cell with the highest energy of all cells which are contiguous to it. The procedure is made in such a way that the biggest islands are formed by 3x3 calorimeter towers.

The input variables are the energy registered in the corresponding island and the neural network projects the information into one output variable $\mathcal{P}$, which is interpreted as the probability that the island is the scattered positron ($\mathcal{P} \sim 1$) or of hadronic origin ($\mathcal{P} \sim 0$).

Furthermore, it is possible to find a position for each island, using the different cells and weighting their respective positions with the cell energy deposit. Using the energy and position, the four-momentum of the island is reconstructed. This four-momentum is that associated to the positron candidate if the island is accepted as the final candidate.

In the end, the different islands have an associated probability of being the scattered positron. The ones with largest probability are taken as good candidates. Efficiency and purity studies of the positron candidates have shown that a good selection of candidates is made by considering only candidates with $\mathcal{P} > 0.9$ and that the probability given by SINISTRA is reliable if the island has an energy larger than 10 GeV.

This sets the standard way to select the positron candidate to the following selection method, which is the one used in the analysis:

i) The candidate with largest probability is taken.

ii) If the candidate energy as measured in the calorimeter is larger than 10 GeV and the probability $\mathcal{P}$ is larger than 0.9, it is considered as a good candidate for the scattered positron.

[1] The neutral pion is taken as a positron candidate if it has an energy larger than the ones which allow to differentiate a hadronic shower from an electromagnetic one. If the energy of the π^0 is much larger than its own mass, then the two photons are very collimated and behave in the detector like one single photon.

iii) If the candidate does not fulfill the requirements in ii), then the candidate is rejected and the event is discarded, i.e. no longer considered as neutral current DIS event.

iv) If the value of y_{ele} (see Eq. (2.4) in section 2.3.1) associated to the candidate is larger than 0.95, the event is discarded. This requirement reduces the background due to photoproduction events.

To improve the purity and efficiency of the positron identification, two additional cuts are imposed **only if the candidate has a polar angle in the region** $20^o < \theta_e < 140^o$. They are related to the size of the electromagnetic shower and to the isolation of the candidate. These cuts reduced the chance that a hadronic jet or part of a jet is misidentified as an electromagnetic deposit and associated to the positron candidate.

v) The energy distribution of the electron is study as a function of the distance in the $\eta - \phi$ plane to the candidate's direction, given by the four-momentum:

$$\Delta_{ele} = \sqrt{(\eta - \eta_e)^2 + (\phi - \phi_e)^2}$$

If the energy inside a cone of 0.3 in this plane (i.e. $\Delta_{ele} < 0.3$) is less than 90% of the total energy of the candidate, the candidate is rejected and the event discarded. This cut rejects candidates whose shower is bigger than the usual size.

vi) The same kind of study is done for the energy not associated to the candidate. In this case, if the summed energy not associated to the candidate in a cone of 0.7 around the candidate's direction is larger than 10% of the candidate's energy, the candidate is rejected and the event discarded. This cut allows to reject candidates which are not well isolated, and are frequently fake candidates (fake positrons[2]).

In some analyses, the positron candidate is required to have an associated track if the polar angle of the candidate is in the central region. This requirement, which can be applied in different ways, prevents to consider a photon as a candidate for the scattered positron. The difficulty in this case is that the track-reconstruction efficiency is slightly different in real data and in simulated events, and this method, although it provides a very clean sample of candidates (i.e. high purity), are problematic when correcting with Monte Carlo samples. For this reason, this kind of requirements is not considered in this analysis.

4.1.2 Selection of the neutral current DIS event sample

After the positron candidate has been accepted, the event is a good candidate to be a neutral current DIS event. Additional requirements are imposed, some of them to override the trigger requirements since it is obvious that at the first and second trigger level it is not possible to perform a sophisticated identification of the scatttered positron as that provided offline by an electron finder like SINISTRA.

The complete set of cuts used in the analysis to select the final sample of events is listed below. The cuts are grouped for a clearer description.

[2]Note that "fake positron" may refer to a real positron but not the scattered one. For example, an electron/positron coming from the semileptonic decay of a heavy quark. This kind of candidates are usually close to a hadronic deposit and will be not isolated at all.

Selection of events, energy scale and noise

The first step in the selection of the sample consists of removing the events which are not "good enough" for physics analysis by applying a cut on the events for which the detector conditions are considered as "bad". In this analysis, the standard selection used in the ZEUS analyses is done. Since this selection only depends on detector conditions, it is only applied to the real data.

During the data selection, the energies of the calorimeter cells are scaled to improve the accuracy of the energy scale. These scale factors have been calculated by comparison of real and simulated events and yield a better comparison between the real data and simulated events. The factors applied in this analysis are the following [91]:

- For the FCAL, a factor of 1.04 is applied to the EMC cells. A factor of 0.95 is applied to the HAC cells.
- In the BCAL, a factor of 1.04 is applied to the EMC cells. A factor of 1.08 is applied to the HAC cells.
- In the RCAL, a global factor of 1.025 is applied to all cells.

These factors are only applied to the data.

The noise per calorimeter cell due to the natural radiactivity of the uranium has approximately a Gaussian shape, centered around 0, with $\sigma \sim 18$ MeV in EMC cells and $\sigma \sim 27$ MeV in HAC cells. This noise has to be suppressed for a better reconstruction of the event. A noise suppresion cut of 60 MeV for EMC cells and 110 MeV for HAC cells is therefore applied to each calorimeter cell for the final selection. A cut on the relative imbalance between the energies measured by the two photomultipliers (left, E_l and right, E_r) of the cell is also applied if the total energy of the cell is smaller than 1 GeV:

$$\left|\frac{E_l - E_r}{E_l + E_r}\right| > 0.49 + \frac{0.03\,\mathrm{GeV}}{E_{cell}}$$

where $E_{cell} = E_l + E_r$ is the total energy of the cell.

Some calorimeter cells are identified as "noisy" during certain running periods, i.e. they have a significantly higher mean energy or "firing" frequency compared to other cells. This is usually due to malfunctions in the electronics reading the information of the photomultiplier. A list of noisy cells is made for each year and cells identified as "noisy" are not included in the final reconstruction if their energy is less than three sigma above the noise level.

To further reduce the noise level a cut of 100 (150) MeV is applied to isolated EMC (HAC) cells. All these cuts are applied also to the Monte Carlo generated events. It should be noted that the detector simulation includes the simulation of the noise due to the uranium radioactivity.

To avoid the presence of "crazy photomultiplier signals" an additional cut on the imbalance is set: any cell with an imbalance between the two photomultipliers satisfying

$$\left|\frac{E_l - E_r}{E_l + E_r}\right| > 0.9$$

is not considered. This kind of cells are usually related to a *spark* in one of the photomultipliers, which makes the signal very large, although it is deposited only in one of the photomultipliers, since it is very unlikely to have sparks on both photomultipliers belonging to the same cell at the same time.

This scaling of the UCAL energy and the selection of cells above given threholds is done as the first step of the offline analysis. This means that all the quantities which are calculated in the following assume that this selection has been done. The only exception is the calculation of variables in the trigger, since it is done before the reconstruction and the offline analysis take place.

Trigger requirements

Since the real data is selected using a trigger system (see section 2.2.6), it is necessary to impose some requirements on which triggers have to be fired for a "good" event to be taken. In this way, it is possible to correct for trigger inefficiency by applying the same trigger chain to data and Monte Carlo samples.

The trigger chain has to be completely defined to avoid biases coming from additional slots. In the present analysis, the trigger chain required was the following one:

- At the FLT, at least one of the slots numbered 40, 41, 42 and 43 has to be fired for the event. These slots are set if some vetoes were not fired for the event and only in presence of large energy deposits in the calorimeter.

  In a simplified form, this FLT requirement is similar to the following sets of cuts

  $$E_{EMC}^{FLT} > 10 \text{ GeV} \text{ or}$$
  $$E^{FLT} > 15 \text{ GeV} \text{ or}$$
  $$E_{BEMC}^{FLT} > 3.4 \text{ GeV or}$$
  $$E_{REMC}^{FLT} > 2 \text{ GeV} \text{ or}$$
  $${E_T}^{FLT} > 11.6 \text{ GeV},$$

  where the superscript FLT indicates that the energy is the one available from the calorimeter (of the indicated section) at the FLT level, where no correction is applied. That means that the exact value after reconstruction can be slightly different.

  Furthermore, for the determination of calorimeter energy sums at the FLT, the three inner ring around the FCAL beampipe and the inner ring around the RCAL beampipe are excluded.

  These effects are taken into account in the offline analysis by imposing cuts which override these trigger cuts, and in this way the effects of the trigger requirements are minimized.

  Regarding the vetoes, the C5 Counter, the vetowall and the SRTD timing information are used to veto events having a beam gas timing or non-ep timing. That is, events related to processes produced outside the detector.

- As an additional cut, it is required that the FLT slot 49 was fired for the event. This slot requires that the event contains at least one "good track", where "good tracks" are defined as tracks for which the z position in the first superlayer of the CTD is between -50 cm and 80 cm.

- At the SLT, the analysis requires the events to pass through the SLT branch called SLT-HPP01. That is, the events are required to satisfy all the following conditions:

  - A reconstructed vertex with $-60 < z_{vtx} < 60$ cm.
  - $E - p_z > 8.0$ GeV, where E and p_z are the energy and the longitudinal momentum of the event, determined from the energy deposits in the calorimeter. For an ideal 4π detector this quantity is expected to be $2E_e \simeq 55$, as it was described in section 2.3.3. In a NC DIS event, this value will be very close to the expected one, so events with values very low are usually associated to proton beam gas or photoproduction events.
  - $E_T^{cone} > 8.0$ GeV, where E_T^{cone} is the sum of transverse energy in all calorimeter cells outside a cone of 10^o around the FCAL beampipe.
  - $E - p_z > 12$ GeV or $p_z/E < 0.95$ [3].
  - This SLT branch is only fired if one of the FLT slots 40, 41, 42 and 43 is fired and if the FLT slot 49 is fired. These are the FLT slots described above.

  Again, the quantities are computed with the information available at the SLT and not with the values after the event reconstruction.

  Additionally, some vetoes are applied at the SLT to reduce the number of non-ep events. These vetoes are listed below:

  - Only one photomultiplier tube in the calorimeter gives a signal.
  - $E-p_z > 75$ GeV, which is not possible if energy and momemtum conservation holds. A fluctuation of the energy measured by a sampling calorimeter can fake a violation of the energy-momentum conservation, but to obtain the value mentioned above is very unlikely.
  - The timing information of the calorimeter is used to veto beam gas and other non-ep events. The cuts are applied as mentioned in the description of the detector and the values are
    - $t_{down} - t_{up} > 10$ ns
    - $t_{FCAL} - t_{RCAL} > 8$ ns
    - $|t_{FCAL}| > 8$ ns
    - $|t_{RCAL}| > 8$ ns

  The event is vetoed if one or more of these conditions is satisfied.

- At the TLT more time is available to perform operations on the event information and, in particular, the software trigger allows to apply jet algorithms. In the analysis, the events are required to pass through one of the four TLT branches called TLT-HPP01, TLT-HPP02, TLT-HPP14 and TLT-HPP15.

  All these have the following common cuts:

  - A reconstructed vertex is found with $-60 < z_{vtx} < 60$ cm.

[3] Since proton beam gas are characterized by $p_z \simeq E$, this requirement allows to make a softer cut on $E - p_z$ without losing ability in discarding background events, which is very important if the rate has to be drastically reduced.

- The number of "bad tracks" has to be smaller than 6. "Bad track" is defined as a track which is long enough for a good reconstruction (i.e. it has more than 5 hits in axial superlayers and more than 5 hits in stereo superlayers, and more than 20 hits in total) and is pointing to a very backward vertex ($z_{vtx} < -75$ cm). As an additional requirement for a "bad track", a cut on transverse momentum ($p_T \geq 200$ MeV) and in pseudorapidity ($0.35 < \eta < 3.13$) is done.

  This cut on "bad tracks" reduces the number of proton beam gas background events, which usually contain forward-going tracks the backward region. It should be noted that good events in coincidence with a proton beam gas event are rejected.
- They required a bit logic over several SLT branches. This requirement is always fulfilled if the SLT-HPP01 branch is on, which was the required SLT logic for the analysis.

Apart for the common requirements, each one of the mentioned branches has its particular set of cuts and are listed below. Cuts on jet variables refer to jets reconstructed using a cone-type algorithm (EUCELL) applied to the cells in the laboratory frame.

- TLT-HPP01 requires the same cuts on $E - p_z$ and p_z/E than those described in the SLT-HPP01. In addition, it requires $E_T^{cone} > 25$ GeV.
- TLT-HPP02 requires that the event contains at least one jet with $E_{T,jet} > 10$ GeV and $\eta_{jet} < 2.5$.
- TLT-HPP14 requires $p_z/E < 1.0$ and that the event contains two or more jets with $E_{T,jet} > 4$ GeV and $\eta_{jet} < 2.5$.
- TLT-HPP15 requires $p_z/E < 1.0$ and that the event contains either two (or more) jets with $E_{T,jet} > 6$ GeV and $\eta_{jet} < 2.5$ or two (or more) jets with $E_{T,jet} > 4$ GeV and $\eta_{jet} < 1.5$.

It should be noted that the jet algorithm is applied to the whole calorimeter, including the cells belonging to the positron candidate. This procedure identifies the positron's energy deposit as a jet. For high Q^2, this procedure makes this trigger very efficient for neutral current DIS events containing jets, in spite of the fact that a positron candidate is not requested.

- As mentioned in section 2.2.7, during the offline reconstruction of the events some additional requirements are imposed and stored in a bit structure (DST bits). This works formally as a fourth trigger although the quantities calculated are based on the offline reconstruction information, i.e. closer to the final values.

  For the analysis presented here, at least one of the following DST bits has to be set: DST-b64, DST-b65, DST-b77 and DST-b78. Each of them is set on if the associated TLT branch is set on. The corresponging TLT branchs are TLT-HPP01, TLT-HPP02, TLT-HPP14 and TLT-HPP15, respectively. This has the advantage that the DST bits are fixed and completely independent of the offline reconstruction[4] and the DST bit selection is not introducing any additional cut since the TLT requirement is always applied in the selection of the events.

[4]This is due to the fact that in the reconstruction only the component information can change, and not the trigger information, including the information from the components which was used in each step of the online selection.

Cuts to remove non-ep background events

After the trigger selection some non-ep events are still present in the sample. These events are removed by applying further cuts. In the analysis these cuts are applied using the information of the final reconstruction and they are listed below:

- $\not{p}_T/\sqrt{E_T} < 2.5\ \mathrm{GeV}^{1/2}$ where $\not{p}_T$ is the total transverse momentum and E_T the total transverse energy of the calorimeter. These quantities are computed as

  $$\not{p}_T = \sqrt{\left[\sum_{cells} E_i \sin\theta_i \cos\phi_i\right]^2 + \left[\sum_{cells} E_i \sin\theta_i \sin\phi_i\right]^2} \quad \text{and} \quad E_T = \sum_{cells} E_i \sin\theta_i\ .$$

  It is expected that $\not{p}_T \sim 0$ for a NC DIS event, while E_T can reach large values. In a sampling calorimeter, the energy measurement displays fluctuations, but considering that these are of the order of $\sqrt{E_T}$ for the measurement of E_T , then the quantity $\not{p}_T/\sqrt{E_T}$ will be usually smaller than 1.

  On the other hand, for events originated by cosmic or halo muons, this quantity can be very large, since the muon can leave more energy in one side of the detector and produce an imbalance.

- The number of "bad tracks" is required to be smaller than 5. This requirement is needed to override the trigger cut on "bad tracks'.

  In the offline selection, the requirements for a track to be considered as "bad track" are the same as those described in the trigger.

Additional cuts

Some final cuts are applied to override the trigger conditions and at the same time to reject events whose reconstruction is not reliable. These requirements are listed below:

- $38.0 \leqslant E - p_z \leqslant 65$ GeV, which removes some background coming from photoproduction processes, and non-ep events. Furthermore, it removes NC DIS events with very low $E - p_z$. The kinematics of these events is not well reconstructed because of the initial QED radiation which escapes through the RCAL beam-pipe hole[5].

- The event must contain a vertex reconstructed using tracks of the CTD. At least two tracks have to be associated to the vertex and the χ^2/ndf related to the fitted vertex must be smaller than 10. If such a vertex exists, it has to fulfill the conditions of being centered in the detector, i.e. $-38 \leqslant z_{vtx} \leqslant 32$ cm. This cut on z_{vtx} overrides the trigger requirements on the vertex. It is better not to include events far from the nominal interaction point in the final selection since the detector looks very different in comparison with the "centered" events, mainly for the angular position of the components.

- The identification of a second positron candidate makes possible to reject background coming from elastic Compton processes ($ep \to ep\gamma$) in which two electromagnetic deposits

[5]If the incoming positron radiates a photon, the value of $E - p_z$ for the ep process is smaller. The Σ method (described in section 2.3.3) is insensitive to this radiation and allows the estimation of the correct $E - p_z$ value.

are the only signals in the calorimeter[6]. To apply this cut, the second candidate is selected using SINISTRA (see section 4.1.1) as the one with next-to-largest probability. In case that two or more candidates display the same probability, the one with highest energy is considered.

To improve the efficiency of selecting this second candidate, we apply the same cuts in the energy, probability, size and isolation as for the first candidate. It should be noted that no cut on y_{ele} is applied because this quantity is not adequate for a candidate which is not necessarily the scattered positron.

If the second candidate is accepted as a good candidate, the cut to remove elastic Compton events is applied by requiring that the energy in the whole calorimeter not belonging to the two electron candidates is less than 4 GeV. In this case, the event is discarded. If the second candidate is not accepted this cut is not applied.

All the above cuts are applied to the simulated events in the same way as for the data, although for Monte Carlo the trigger requirements are applied over ideal (or average) conditions in the detector. The only selection requirements which are different for data and simulated events is the cut on events for which the detector was in a "bad condition" and the scaling of the energies of the cells, as we mentioned above.

Reconstruction and selection of the kinematic region

After all the above cuts and selection criteria have been applied, the kinematic reconstruction is performed. In the analysis, the Double Angle method is used for the kinematic reconstruction from the calorimeter information, since it is the most convenient for the kinematic region we are interested in (see section 4.2.1).

In the analysis the following requirement is imposed to the virtuality of the exchanged boson:

$$Q^2{}_{DA} \geqslant 125 \text{ GeV}^2.$$

Figure 4.1 shows the distribution of the events in the kinematic plane (the $x - Q^2$ plane). The events were selected with the complete set of cuts applied in the analysis except that on $Q^2{}_{DA}$, which is shown as an horizontal line in the plot.

4.1.3 Jet selection

Jets are reconstructed from the energies measured in the calorimeter cells. Only those cells passing through the cuts described above are considered. Furthermore, all the cells belonging to the island identified as the positron candidate are not considered.

For each selected cell, its energy and the angular position with respect to the reconstructed vertex allow the reconstruction the four-momentum, being by definition massless and having as energy that of the cell. The position is reconstructed using the geometrical center of the cell.

These four-momenta are used to reconstruct the jets by applying the longitudinally-invariant k_T-cluster algorithm in the way described in section 1.4.1. Upon the application this algorithm, the set of variables $\{E_{T,jet}, \eta_{jet}, \phi_{jet}\}$ contains all the information for the selected jets.

[6]Since this cut only uses the calorimeter information, events consisting of e^+e^- pairs are also removed.

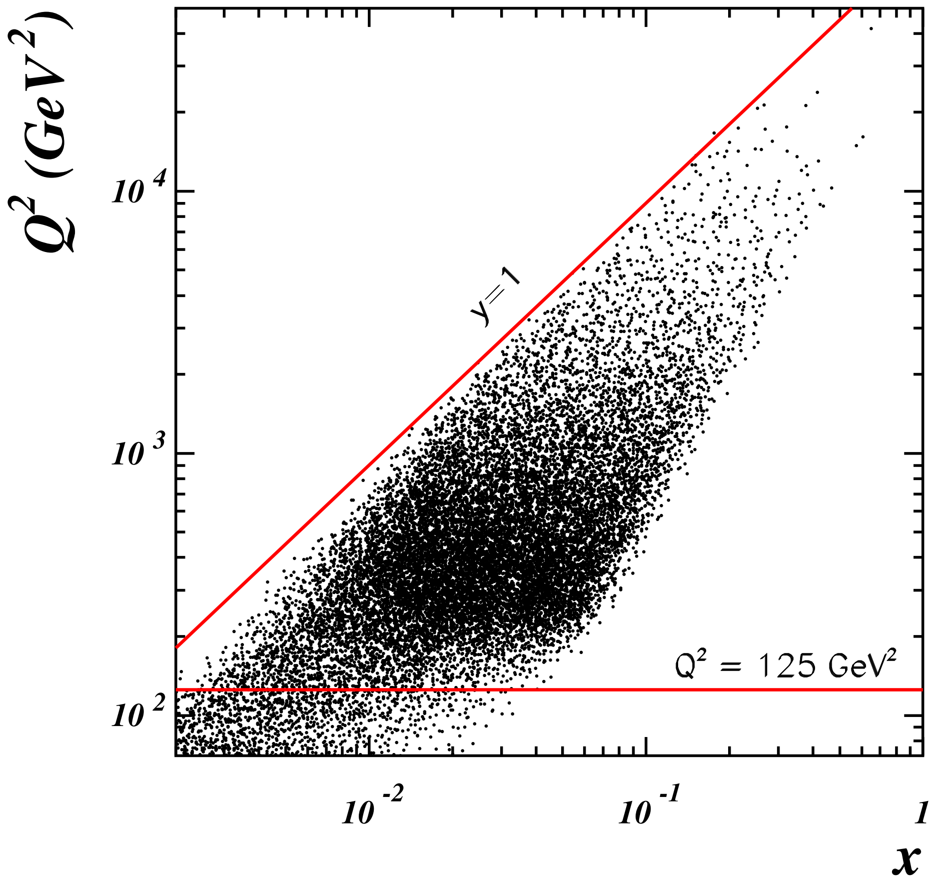


Figure 4.1: Distribution of the selected events in the $x - Q^2$ plane. The lines shows the applied cut on this plane ($Q^2 = 125\ GeV^2$) and the kinematic limit for the 96 and 97 running periods ($y = 1$).

The jets reconstructed from the calorimeter cells are preselected with a cut of $E_{t,jet} \geqslant 10$ GeV. Later on the jet energies are corrected for energy losses in the detector using the factors given by the Monte Carlo, as described in section 4.1.5. After this energy correction has been applied, the final cuts are applied to the jets and only those verifying the conditions

$$E_{T,jet} \geqslant 15 \text{ GeV} \quad \text{and} \quad -1 \leqslant \eta_{jet} \leqslant 2$$

are kept in the final sample.

As a final cut to improve the purity of the sample, the whole event is discarded if one of the final jets is closer by less than 1 unit to the positron candidate in the $\eta - \phi$ plane. This cut removes events in which the positron is probably a fake candidate. Even being a real positron, the closeness to a jet indicates that probably it is the product of a semileptonic decay of a heavy quark, for example $b \to c\,e^-\bar{\nu}$.

Further analysis is done over the jets found in the remaining events. After all these cuts, 37590 events constitute the final sample. Out of them, 36769 (97.8%), 797 (2.1%) and 24 (0.1%) events containing one, two and three jets are found. The total number of jets is therefore 38435.

One of these events is shown in figure 4.2 as seen in the ZEUS detector. The distribution of transverse energies in the $\eta - \phi$ plane is also shown. The clear scattered positron candidate allows to classify the event as a neutral current DIS event candidate. In addition to the positron, the event contains one jet which, when looked in more detail, contains two small subjets. The subjets are resolved for some values of the resolution scale.

4.1.4 Event selection at the hadron level in the MC samples

In simulated events, the same cuts are applied to select the events and the jets at the detector level. At the hadron level, a completely independent analysis is performed in order to obtain the information about the event before the detector modifies it.

To select the events at the hadron level, cuts are applied which are described in the following paragraphs.

Reconstruction of the kinematic variables

At hadron level, all the information is exactly known. To obtain the kinematics, the exact theoretical expressions are used. For example, the value of Q^2 is obtained by squaring the four-momentum of the exchanged boson in the event.

To select the kinematic region we are interested in, the following cut is applied:

$$Q^2 = -q_\mu q^\mu \geqslant 125 \text{ GeV}^2 .$$

Jet search and selection at hadron level

To obtain the jets at hadron level the same algorithm applied to the calorimeter cells is applied to the final-state hadrons. The only difference is that the hadrons are not massless and their three-momenta are scaled so as to make the mass equal to zero before the k_T-cluster algorithm is applied. This procedure sets the definition of transverse energy that is used in the analysis.

After applying the algorithm to the "massless" hadrons, the jets are obtained by selecting every jet in the event verifying the following conditions in the laboratory frame:

$$E_{T,jet} \geqslant 15 \text{ GeV} \qquad \text{and} \qquad -1 \leqslant \eta_{jet} \leqslant 2$$

One of the major issues is the definition of the hadron level. Since in the Monte Carlo event all the information is available, from the list of final-state particles the scattered positron and the possible radiated photons off the positron are removed. These radiated photons are assigned to the leptonic part of the events and are not included in the hadronic final state. Then we are left with the hadrons obtained after fragmentation. In principle, these hadrons can be considered as the hadronic final state, but a definition based on the lifetimes of the different particles is used here.

Definition of the hadronic final state

The idea is as follows: although the fragmentation process is considered to take place very close to the interaction, the generated hadrons show a time evolution dictated by the different particle decays. This suggests that a good way to define the final state, which is more independent of the fragmentation process, is the one based on the lifetime of the particles.

The selection of the hadrons which are considered as *stable* in the final state and used for the jet algorithm is done in terms of a *critical lifetime* τ_c: every particle belonging to the hadronic final state whose lifetime is longer than τ_c is considered as *stable*. If a particle belonging to the hadronic final state has a lifetime smaller than τ_c, then its decay products are considered. This leads to an iterative procedure which is stopped when all the particles generated by the fragmentation mechanism have been considered and they or their decay products form the list of particles (four-momenta) to be used in the jet search algorithm.

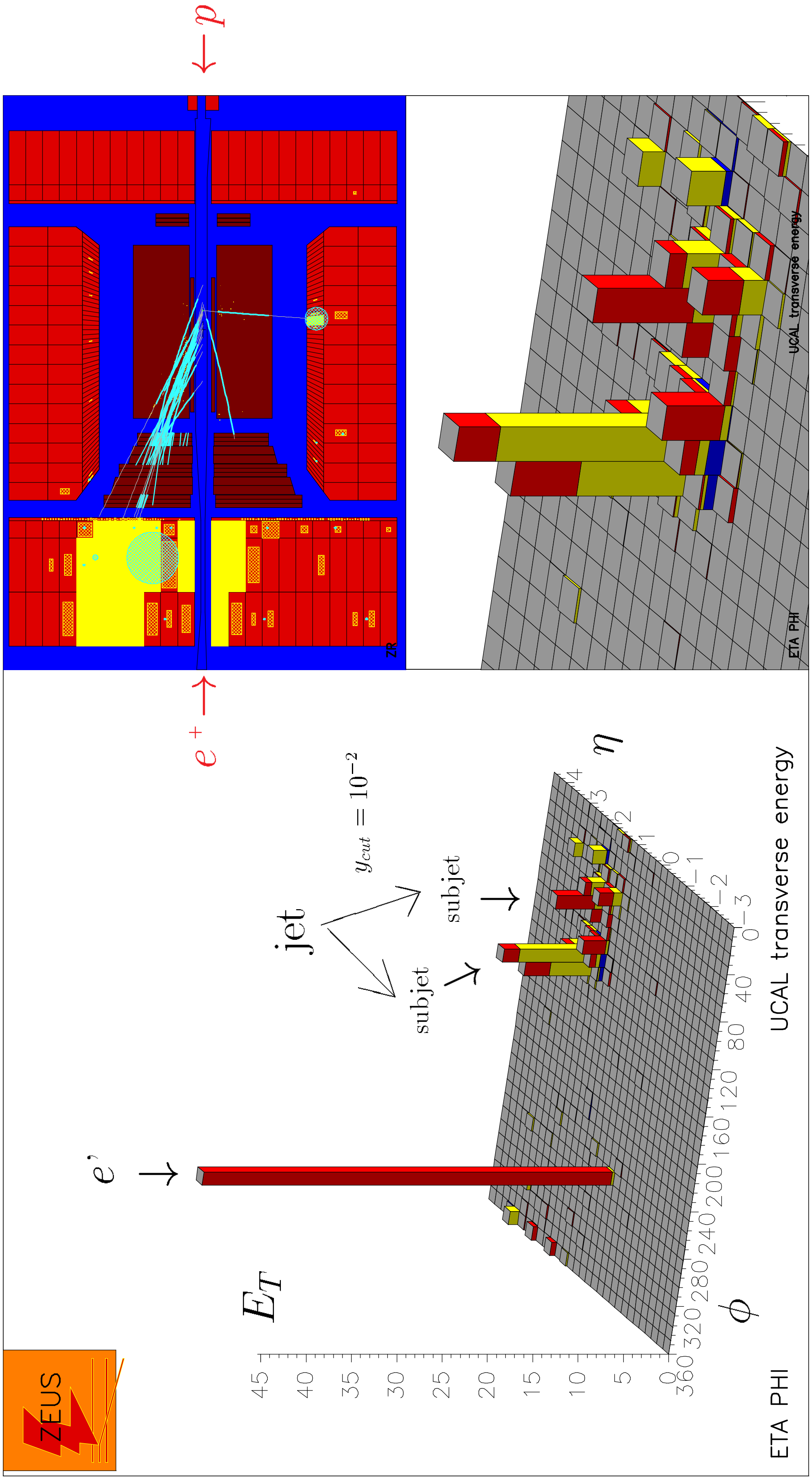


Figure 4.2: One event taken with the ZEUS detector and selected in the analysis. The small-transverse-size energy deposit is identified as the positron candidate. The additional transverse energy deposit is identified as a jet. It should be noted the presence of large energy deposits close to the direction of the incoming proton. Although this energy deposit, associated to the remnant, is very large, it does not show up in the transverse energy distribution, in which the positron and the jet are clearly observed. Looking at the jet in detail, as shown in the lower figure on the left, the internal structure is resolved, and two subjets can be distinguished.

The lifetimes of the particles are distributed in such a way that two good choices for the critical lifetime can be made. These two values are the following

$$\tau_c = 10^{-10}\ \mathrm{s} \qquad \text{or} \qquad \tau_c = 10^{-11}\ \mathrm{s}$$

The selection of the first one has an obvious disavantadge from the experimental point of view. If a particle with a lifetime between the two values above (for example, K_S, with $\tau = 0.89{\cdot}10^{-10}s$ has a sufficiently large energy, the mean distance can be large enough for the particle to reach the detector. For this reason, the second value is preferred and used in the present analysis: all the particles which are considered unstable have so small lifetimes that the detector effects are negligible.

It has to be noted that among the particles taken for the hadronic final state not only hadrons are present. Leptons and photons may appear in the final state as decay products of some hadrons. These particles are included in the hadronic final state since they were generated as decay products of the partonic final state resulting from the hard process.

4.1.5 Jet energy corrections

Jet energies need to be corrected since the final-state particles may interact with the detector before they and, even with an ideal calorimeter, the reconstructed energy is smaller than the original one.

Sometimes the effect of energy loss does not affect significantly the results. For example, the loss in the energy of the scattered positron is important in the reconstruction of the kinematic variables when using the electron method. However, by using the Double Angle method, the energy losses only introduce differences of second order.

In the analysis presented here, the kinematic variables are well reconstructed without a significant effect from the energy losses. On the other hand, the jet energy has to be corrected because the effects are very large.

There exist several methods to correct the energy of the jets. In this analysis, the standard procedure has been used. It is briefly described by the following steps:

- The standard cuts are applied to the hadron level to select the jets in a Monte Carlo sample.

  As described in section 4.1.3, the jets at the detector level are selected using a cut on 10 GeV.

- For each event, the jets at hadron and detector levels are matched, that is, for each pair of jets, one at detector and one at hadron level, the distance in the $\eta - \phi$ plane defined as

  $$\Delta[hd] = \sqrt{(\eta_d - \eta_h)^2 + (\phi_d - \phi_h)^2}\ ,$$

  is calculated and the pair with the minimum distance is defined as a "matched pair" if this distance is smaller than 1. The procedure is repeated until no jet is left or the distances between the remaining pairs is larger than 1.

- For each pair, one takes the energy of the jet at hadron level as the original one and that of the jet at detector level as the reconstructed one. The correction of the jet energy is constructed from the correlation between the reconstructed energy and the original one. This correlation is parametrized with a liner function by fitting the distribution of the reconstructed energy as a function of the original energy The fitted function has the form

$$E_{T,jet}^{DET} = m \cdot E_{T,jet}^{HAD} + b \, .$$

  Then, for a given jet at detector level the corrected energy is obtained by inverting this function:

$$E_{T,jet}(corr) = \frac{E_{T,jet}(rec) - b}{m}$$

- Since the energy loss depends on the pseudorapidity, the actual parametrization is done separately for each region of η_{jet} by introducing a dependence on η_{jet} of the parameters m and b.

A complete description of the method and detailed plots are presented in appendix B.

After the correction of the transverse energy of the jet has been applied, the final selection cut on $E_{T,jet}$ is imposed to obtain the final sample of jets, as it was described in section 4.1.3.

4.2 Preliminary studies

Some preliminary studies using Monte Carlo event samples are presented to show how the detector effects and other contributions affect the measurements.

4.2.1 Reconstruction of the kinematic variables

The reconstruction of the kinematics for the selected events can be performed in any of the methods described in section 2.3. Figures 4.3 to 4.9 show the results of the different methods.

This study was made by using the Monte Carlo sample generated with ARIADNE and comparing the generated values of the kinematic variables (Q^2, x and y) with those reconstructed using a given method . For selecting the events, the cuts at hadron level are applied and the following requirements are made:

- A "good" positron candidate. The complete set of cuts on the positron candidate are imposed, included that to remove elastic Compton event.

- All the cuts to remove background events and assure a good reconstruction of the kinematics. This includes the cuts on $E - p_z$, vertex, bad tracks, ... as it was described in section 4.1.2.

- The trigger selection is not applied, and no cuts are imposed on the jets.

For each method the comparison between the generated and reconstructed values is shown as a "scatter" plot to show the level of correlation obtained. Below each plot, the mean value of the relative difference is presented. This plot is very useful to observe trends or biases in the reconstruction of one variable as a function of the same variable. The distribution in the relative difference for the complete sample is shown. Although this distribution mixes regions of good and bad reconstruction, it helps to analyze the global quality of the method.

The following list present the general results provided by each method. For the values of resolution, see table 4.1:

- Figure 4.3 shows the results obtained using the Double Angle method. This method is expected to give very good results in the kinematic region considered for the analysis. As shown in the plots the correlation is good except at low y.

- Figure 4.4 shows the results using the electron method. In this method, the resolution for Q^2 is slightly worse than those of the Double Angle method, and the reconstructed Q^2 shows a clear shift due to the energy loss of the positron.

- Figure 4.5 shows the results using the Jacquet-Blondel method. As expected, the results are worse than those from the above methods.

- Figure 4.6 shows the results using the Σ method. Again, the energy loss in the detector for the positron introduces a shift in the kinematic variables. The resolution in y is very good. Although this method seems to be worse than the Double Angle method, it should be observed that the correlation for Q^2 is less sensitive to QED radiation: the scatter plot does not present the "cloud" of points around the central value as in the Double Angle method.

  This property would make this method better than the Double Angle method, since the tails of the distributions are suppressed. However, it requires correcting the positron energy. This correction would worsen the resolution, so the Double Angle Method is preferred.

- To illustrate that the correction of the energy losses can solve the problem of the observed shifts, the Σ method is applied but now using for the energy of the positron that given by the Double Angle method, i.e.

  $$E'_{DA} = (1 - y_{DA})\, E_e + x_{DA} y_{DA}\, Ep = \frac{2E_e}{1 - \cos\theta + 2\left(\frac{y_{JB} E_e}{p_T^{had}}\right)\sin\theta}$$

  The result is shown in figure 4.7 and it is clearly seen that the shifts on Q^2 and x disappear. The improvement in the resolution for these variables is due to the better resolution for E'_{DA} than for the measured positron energy in the kinematic region considered in the analysis.

  The results of this method are very good, although still worse than that of the Double Angle method. However, the advantage of using the Σ method is lost since the dependence on the QED radiation is again present. Yet the effects of the QED radiation seem to be less important than for the Double Angle method.

- During the last years, the ZEUS Collaboration has developed a new kind of objets, called ZUFOS, to reconstruct the information of the hadronic system. The procedure combines the information from the CTD and the UCAL to reconstruct a list of four-momenta for the ZUFOS. The ZUFOS are then used in the same way as cells in the reconstruction of kinematic variables or the jet search.

Figures 4.8 and 4.9 show the results of the Double Angle and Σ methods using zufos. In the first case, the results are clearly worse than the same method based on the UCAL cells. In the second case, an improvement is observed using ZUFOS, but the resolutions are not as good as for the Double Angle method using the UCAL cells. This method has the disadvantage that again the energy of the positron has to be corrected before computing the expression for Q^2. In the plots shown, E'_{DA} was used for the reconstruction

Regarding the use of the Double Angle method using ZUFOS, it is possible to improve the results by using the transverse momentum of the positron (given by $E'_{DA} \sin\theta$, using cells) to estimate $\cos\gamma_h$. In this case the resolutions are very similar to the Double Angle method using UCAL cells[7].

The final conclusion from this comparison is that the Double Angle method is clearly the best method to reconstruct the kinematics in the kinematic region of the analysis. For this reason, this method has been used.

[7]This is obvious since most of the information is taken from θ and γ_h using UCAL cells. However, if the positron energy were corrected from energy losses with an external method, this method would be completely independent of the Double Angle using cells.

Method	Res. Q^2		Res. x		Res. y	
	mean[%]	rms [%]	mean[%]	rms [%]	mean[%]	rms [%]
DA method	0.7 (-0.9)	6.7 (2.5)	-3.3 (2.6)	14.7 (7.7)	5.1 (1.2)	12.3 (5.0)
electron	-6.1 (-6.0)	8.0 (6.2)	-27.7 (-)	24.9 (-)	39.1 (-)	46.2 (-)
Jacquet-Blondel	-30.1 (-30.3)	21.3 (18.4)	23.4 (21.2)	21.0 (14.3)	-11.9 (-12.1)	13.5 (10.1)
Σ method	-14.4 (14.5)	11.6 (10.8)	-11.7 (-1.7)	18.6 (16.4)	-2.0 (-4.0)	13.3 (8.7)
DA-e + Σ	-1.7 (-2.2)	9.3 (3.5)	1.5 (1.3)	15.5 (9.7)	-2.0 (-4.0)	13.3 (8.7)
DA with ZUFOS	4.7 (0.3)	9.0 (2.5)	11.5 (4.7)	24.1 (11.3)	-3.5 (-6.7)	14.5 (10.2)
DA-e + Σ with ZUFOS	4.7 (0.3)	9.0 (2.5)	11.5 (4.7)	24.1 (11.3)	0.9 (1.5)	12.8 (8.7)

Table 4.1: Values of the resolution in Q^2, x and y using different methods to reconstruct the kinematics. The values are the mean and the variance of the distributions for the relative difference between the reconstructed and the generated values, and shown as percentages. The numbers in parentheses are the corresponding values using a gaussian fit to the peak. The difference between the global values and those from the gaussian fit usually is related with asymmetries in the distribution (mean value) or with the presence of tails (rms values), and for this reason is useful to know both values. It has not been calculated when the distribution has a clear non-gaussian shape.

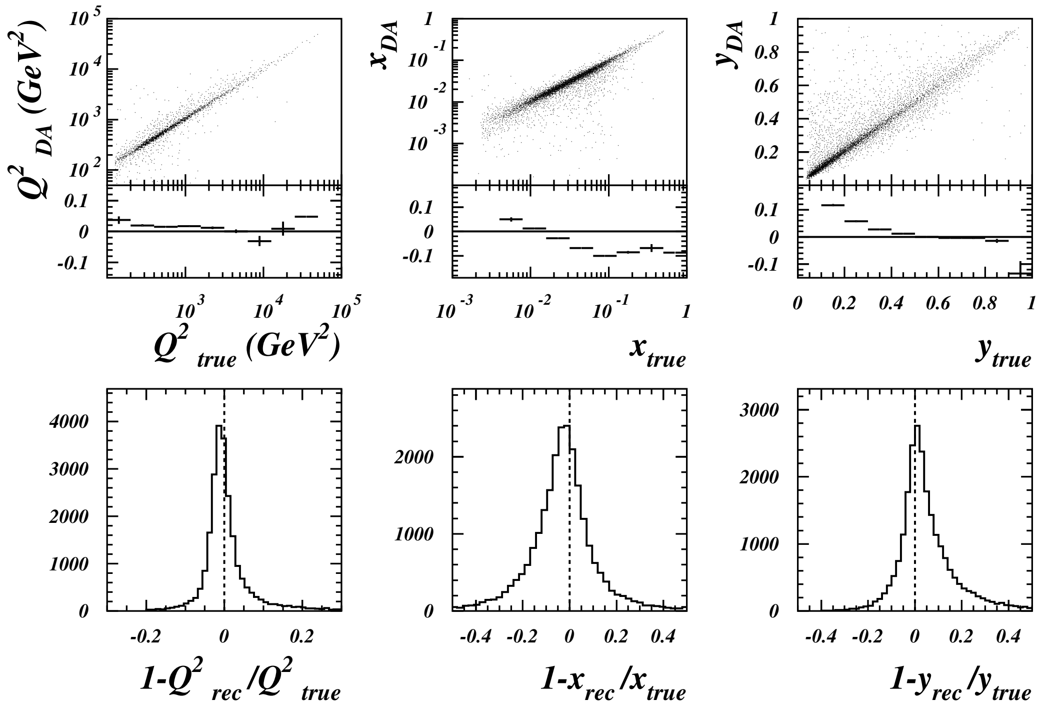


Figure 4.3: Resolution in the kinematic variables Q^2, x and y by using the Double Angle method in the ARIADNE sample. The plots above show the correlation between the reconstructed and generated value. The plots just below show $(1-Q^2_{DA}/Q^2_{true})$ as a function Q^2_{true} (and equivalently for x and y). The plots at the bottom show the distributions of the relative difference for each variable. The mean values and variances of these distributions are listed in table 4.1.

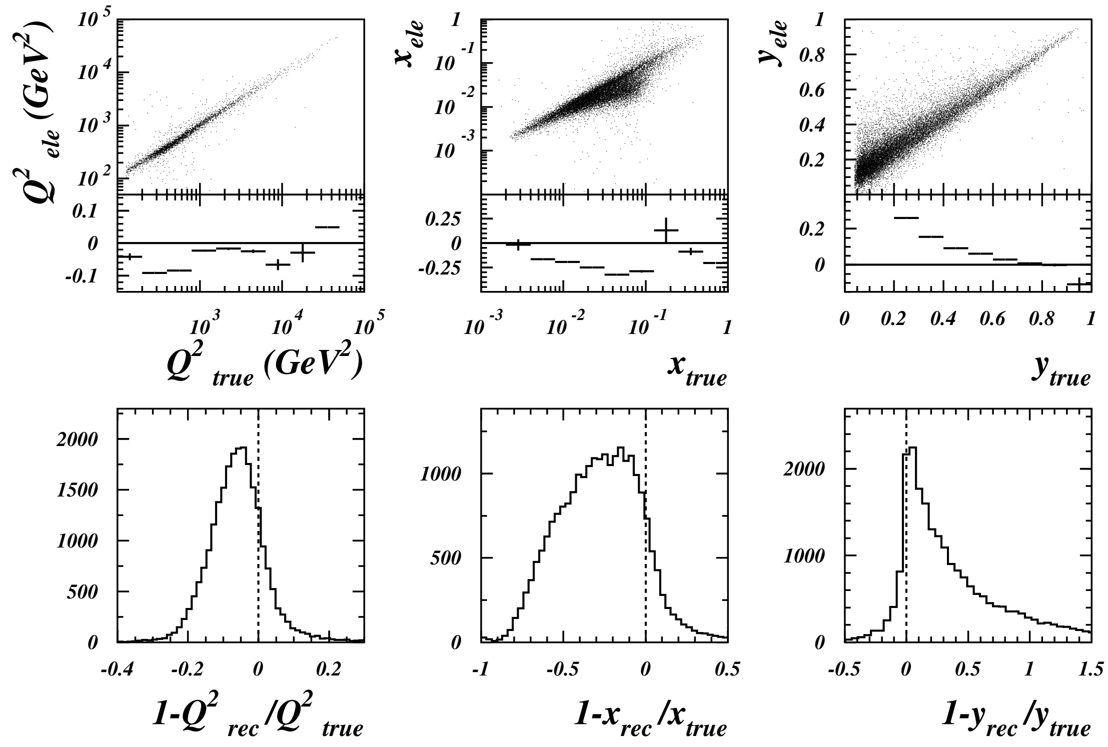


Figure 4.4: Resolution in the kinematic variables Q^2, x and y by using the electron method. Other details are as described in the caption to Fig. 4.3.

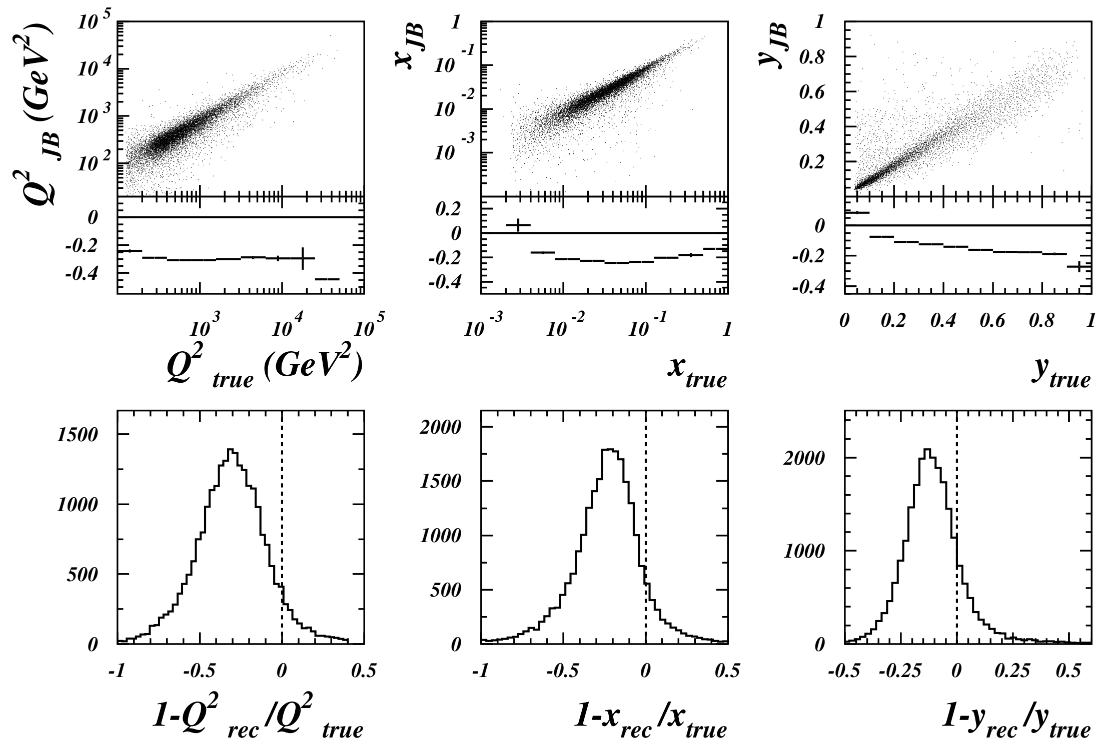


Figure 4.5: Resolution in the kinematic variables Q^2, x and y by using the Jacquet-Blondel method. Other details are as described in the caption to Fig. 4.3.

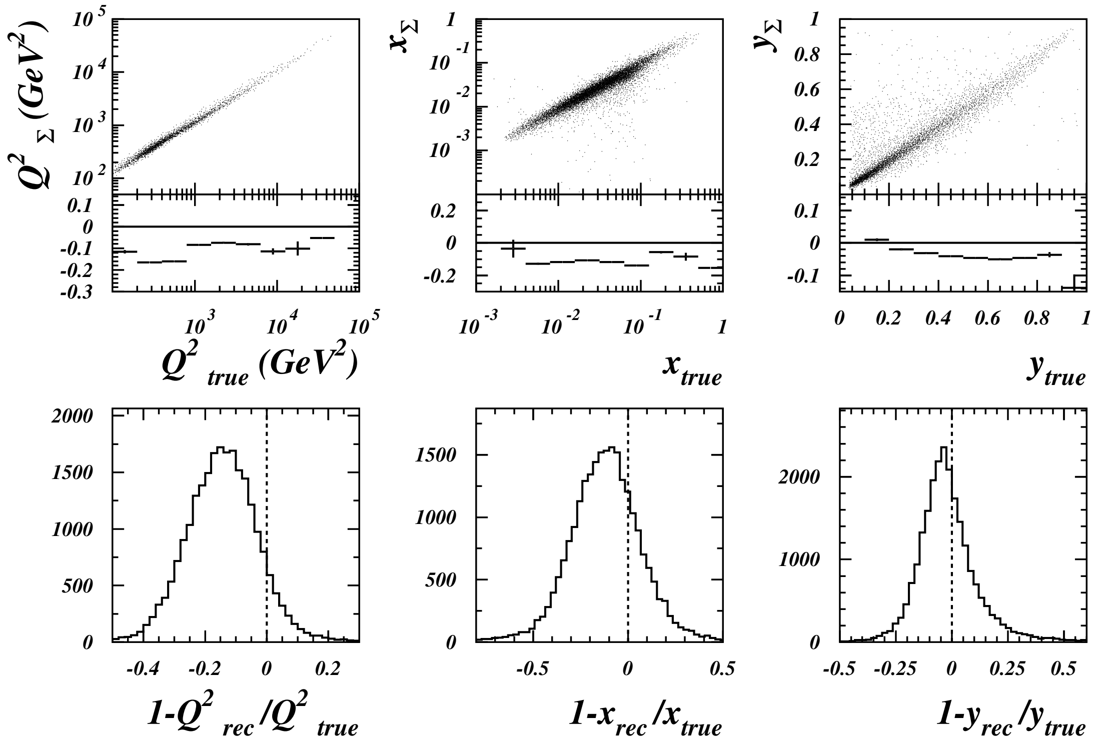


Figure 4.6: Resolution in the kinematic variables Q^2, x and y by using the Σ method. Other details are as described in the caption to Fig. 4.3.

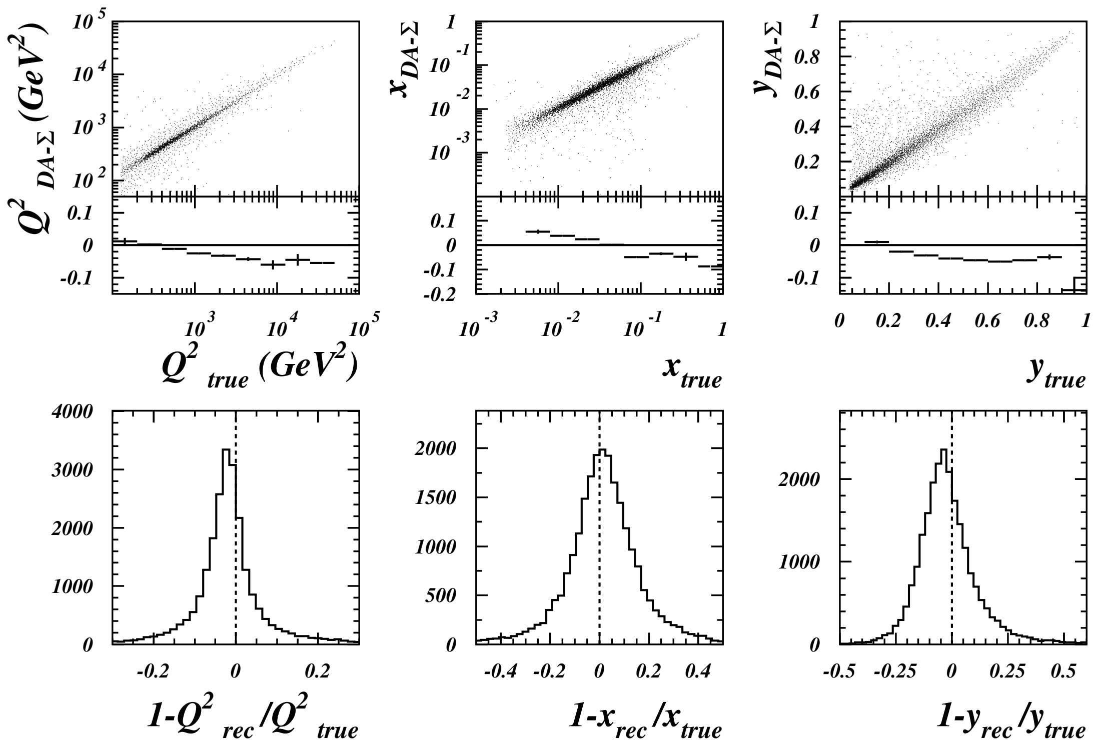


Figure 4.7: Resolution in the kinematic variables Q^2, x and y by using the Σ method; but in this case the positron energy was set to the value given by the Double Angle method (see text for description). Other details are as described in the caption to Fig. 4.3.

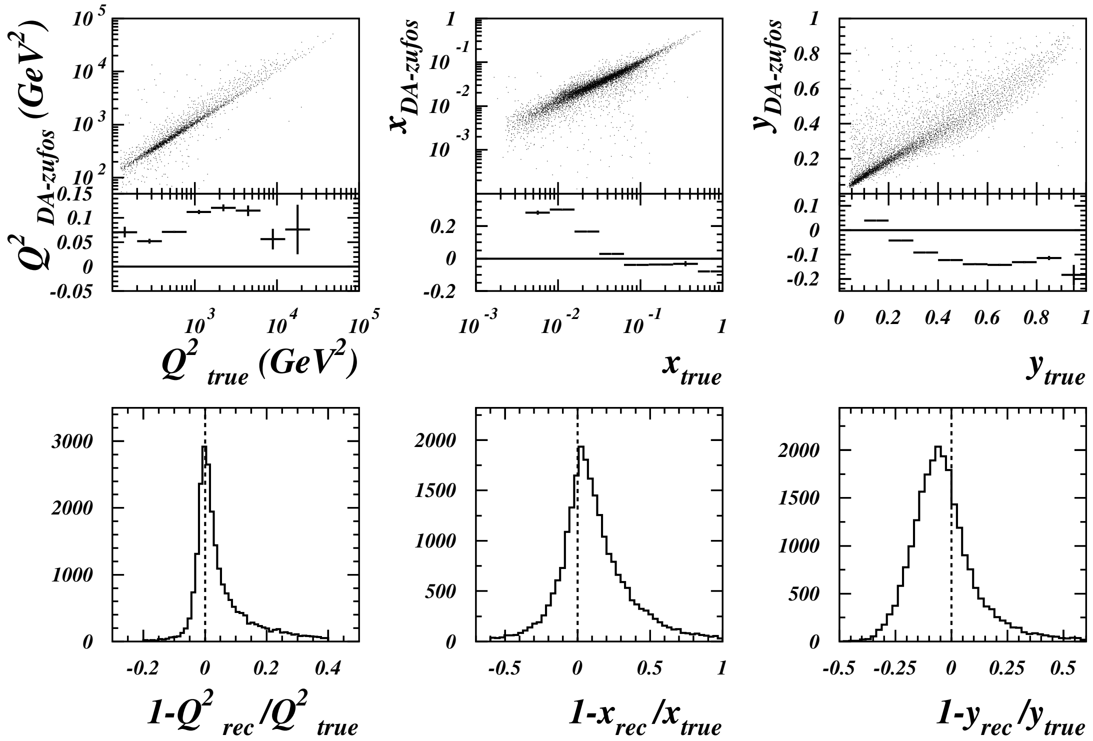


Figure 4.8: Resolution in the kinematic variables Q^2, x and y by using the Double Angle method. ZUFOS were used to reconstruct the hadronic system variables (see text). Other details are as described in the caption to Fig. 4.3.

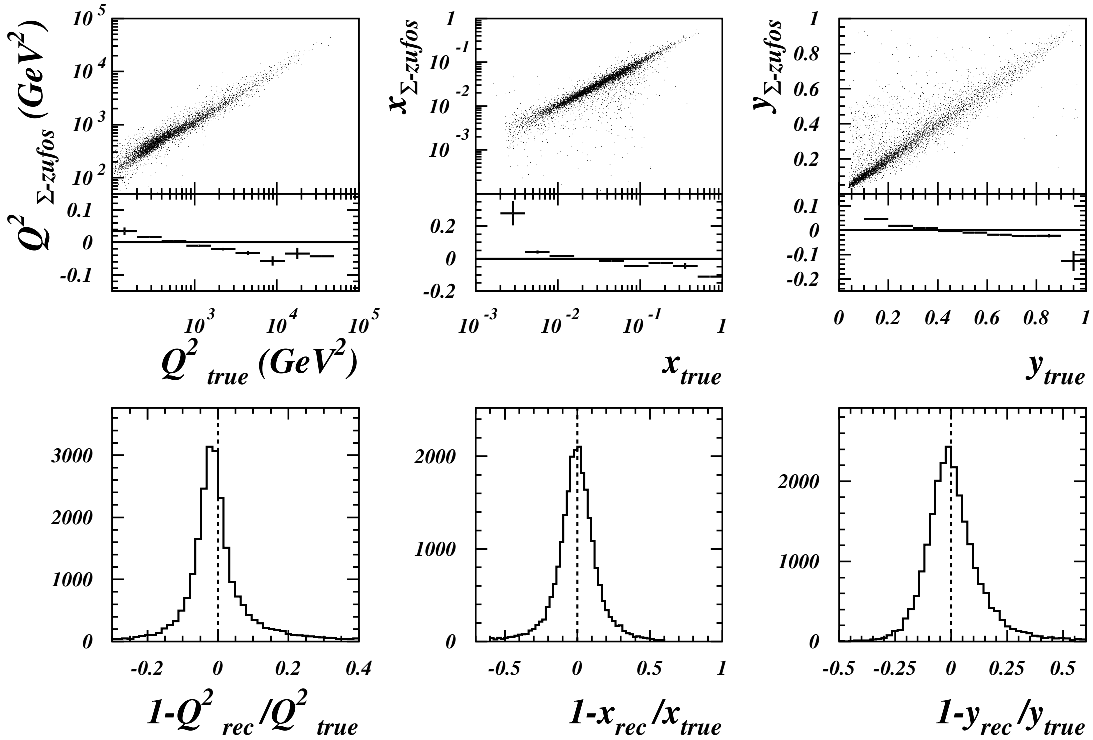


Figure 4.9: Resolution in the kinematic variables Q^2, x and y by using the Σ method. ZUFOS were used to reconstruct the hadronic system variables (see text). Other details are as described in the caption to Fig. 4.3.

4.2.2 Reconstruction of the jet variables

After applying the same cuts as for studying the resolution of the kinematic variables (see section 4.2.1 for details about the cuts), a matching between the selected jets at hadron level and detector levels is performed.

This is the same selection as the one described in section 4.1.5 when the matching of jets at hadron and detector level was used to correct the energy of the jets. Now the aim is to study the correlation and resolution in the reconstruction of the jet variables.

Figures 4.10 to 4.13 show the results of this study. The plots on the left-side of each figure show the correlation between the reconstructed and generated values and the dependence of the difference with respect to the variable. It should be noted that the resolution of the angular variables is not studied by means of the relative difference but the difference itself.

The values of the mean and variance for each variable are presented in table 4.2.

The figures show a very good resolution in the jet angular variables. The small shift in ϕ_{jet} is due to the magnetic field in the inner detector. Since its magnitude is negliglible as compared with the resolution, no correction is applied for this effect. The resolution in η_{jet} is slightly worse in the region $\eta_{jet} > 1.5$.

For the transverse energy, the resolution plot in figure 4.10 shows the shift which we have to correct for. After applying the corrections the resolution is slightly worse, but the mean value is very close to zero, as shown in figure 4.11. It is observed that the resolution improves as $E_{T,jet}$ increases.

	Res. $E_{T,jet}$		Diff. η_{jet}		Diff. ϕ_{jet}	
	mean[%]	rms [%]	mean	rms	mean	rms
reconstructed	-14.3 (-12.4)	10.2 (8.4)	-0.009 (-0.006)	0.05 (0.03)	0.31 (0.26)	3.0 (2.0)
corrected	0.5 (1.4)	10.4 (9.1)				

Table 4.2: Values of the resolution in the jet variables. For the transverse energy, the distribution of the relative difference is used to obtain the values, here shown as percentages. For the angular variables the absolute differences between the reconstructed and the generated value are considered. The numbers inside parentheses are the corresponding values using a gaussian fit to the peak. The difference between the global values and those from the gaussian fit usually is related with asymmetries in the distribution (mean value) or with the presence of tails (rms values), and for this reason is useful to know both values.

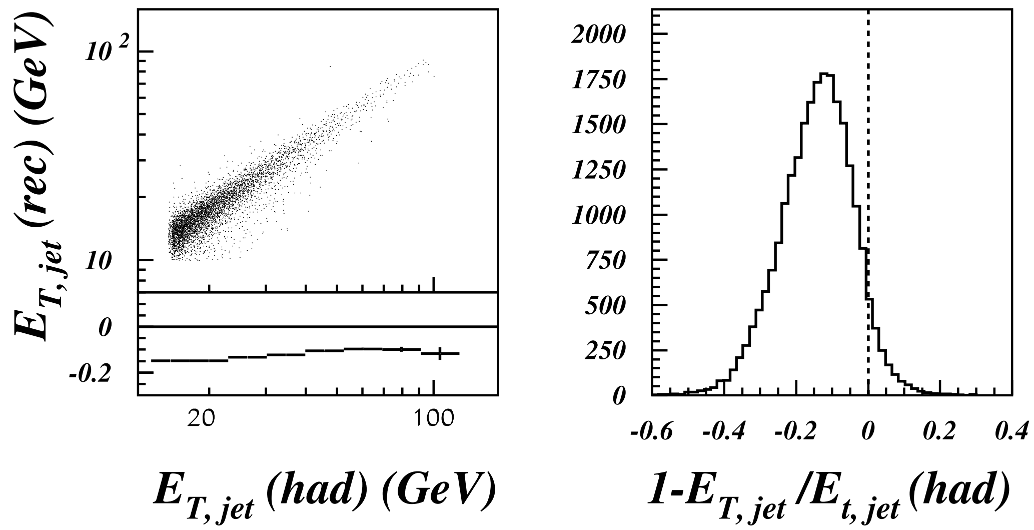


Figure 4.10: Resolution in the the transverse energy of the jets using the ARIADNE sample. The value at detector level is the reconstructed one, i.e. before applying any correction.

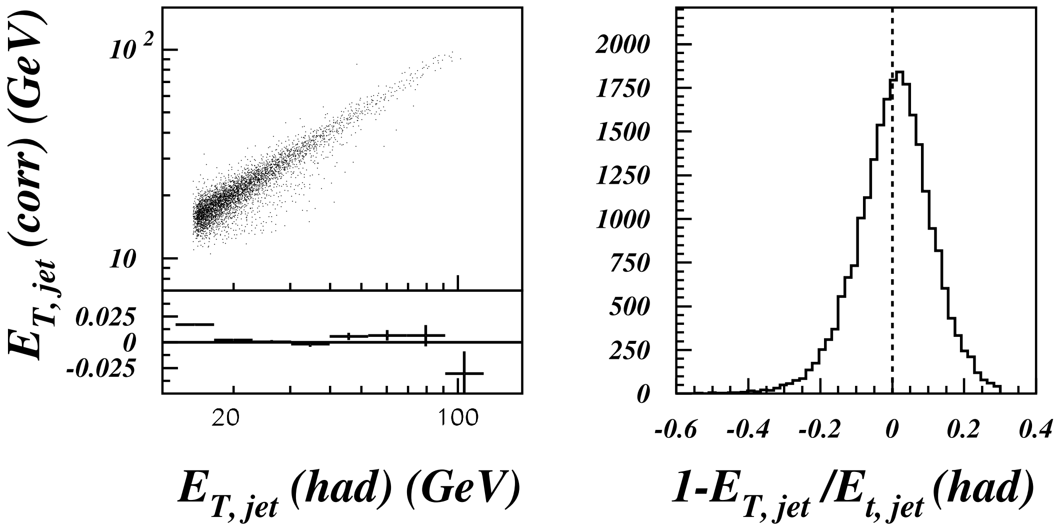


Figure 4.11: Resolution in the transverse energy of the jets after correcting the energy of the jets from energy loss in the detector. The ARIADNE sample was used.

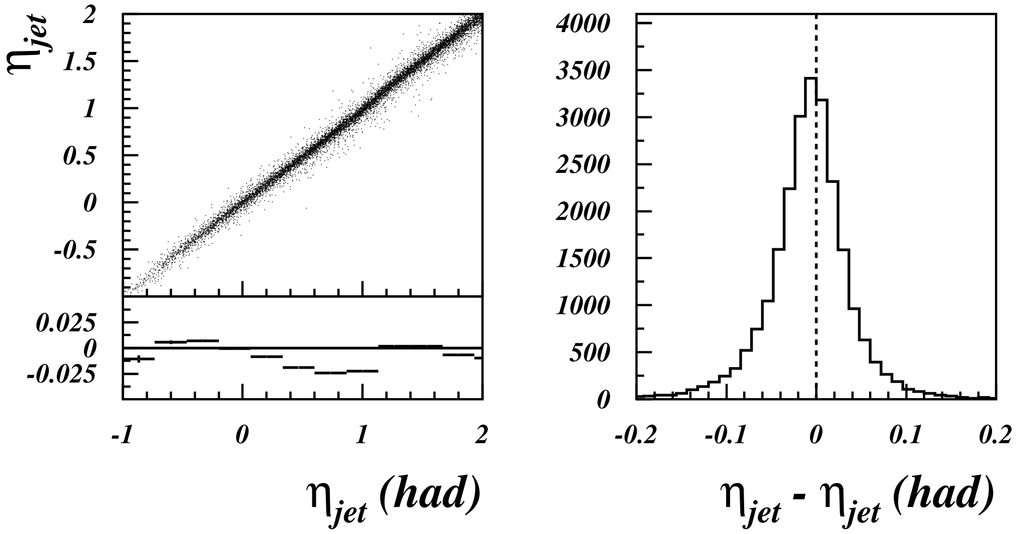


Figure 4.12: Resolution in the jet pseudorapidity using the ARIADNE sample.

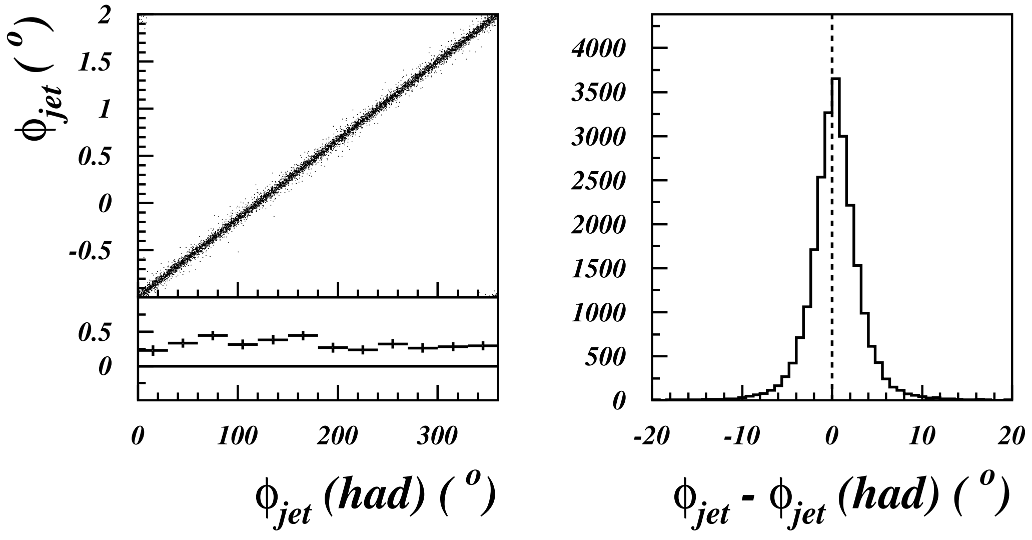


Figure 4.13: Resolution in the azimuthal angle of jets using the ARIADNE sample.

4.2.3 Study of the detector effects on jets

The experimental selection of events introduce modifications in the observed final state coming from the detector resolution or the selection cuts. Although these modifications are qualitatively understood, they have to be quantified using the detector simulation and the Monte Carlo samples.

The detector effects can be studied in many ways, being the most important the value of the *correction factors*, as it is described in section 4.2.4. This gives information on how similar or different are the measured quantities with respect to the generated values at hadron level.

It should be noted that this is usually not enoguh since the result could be very similar to the generated value but only by chance. That is why it is very common to complete the study of the detector effects with the calculation of the purity and efficiency, defined as follows:

- The purity is the percentage of jets which are identified at detector level in the same bin than the corresponding jet at hadron level. That is, it is the fraction

$$\mathcal{P} = \frac{\#generated \text{ AND } reconstructed}{\#reconstructed}$$

- The efficiency is the percentage of jets which are found at detector level in the same bin than the corresponding event (jet) at hadron level. That is, it is the fraction

$$\mathcal{E} = \frac{\#generated \text{ AND } reconstructed}{\#generated}$$

To study all these quantities, the standard analysis is performed at hadron and detector levels and, for each event, a matching procedure is applied to decide which jet at hadron level corresponds to which jet at detector level. This matching is done in the same way as described in section 4.1.5. After matching the jets, a jet is added to a distribution only if the jet at hadron and detector levels is in the same bin. This allows the computation of the numerators in the definitions given above. The denominators are the results of the standard analysis, counting the number of jets in each bin.

In the analysis presented here, the purities and efficiencies have been computed for the entire samples. Using the sample generated with ARIADNE (LEPTO-MEPS) the efficiency is 72% (74%) and the purity are around 81% (78%). The study has been done for the different bins in $E_{T,jet}$ and η_{jet}, which are the variables with respect to which the subjet multiplicity has been studied. Figures 4.14 and 4.15 show the result and the comparison between ARIADNE and LEPTO-MEPS. It should be emphasized the very good agreement between the two samples. Regarding the values themselves, it should be noted that the purity is always above 50% except in the lowest $E_{T,jet}$ bin. The efficiency is always above 40% except in that same bin. The results for that bin are understood since the selection cut on Q^2 affects mainly the events in that bin. The drop in the first bin in η_{jet}, although it is not so important since the efficiency and the purity are still very large, probably is due to similar reasons and not to problems in the jet reconstruction.

As a last remark, the high efficiency is mostly due to the correction of the transverse energy of the jets. Without that correction, the resulting values would be very small, specially for the bins in $E_{T,jet}$.

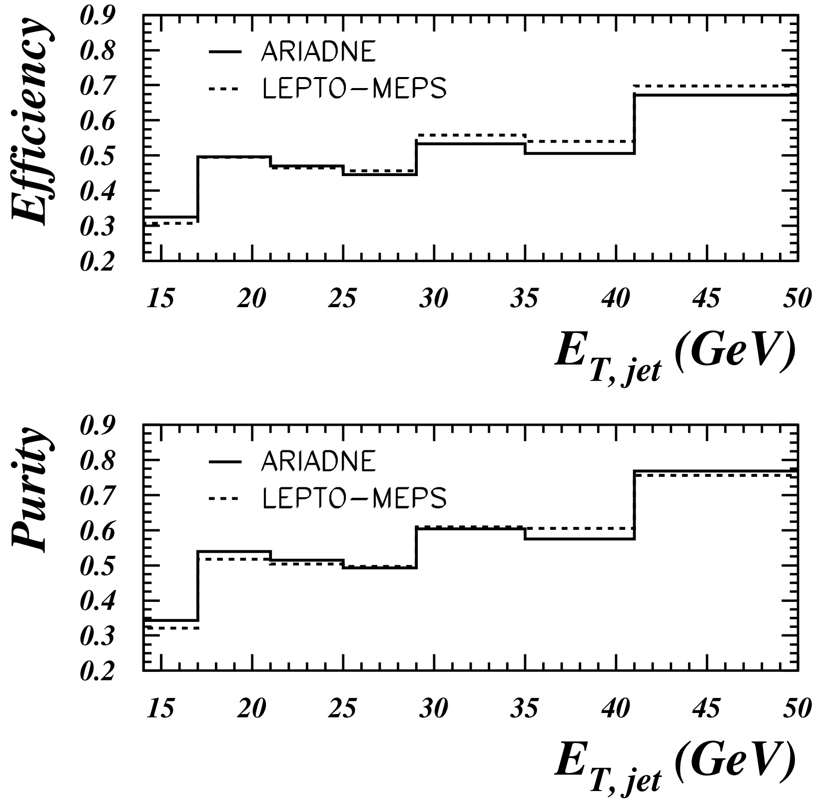


Figure 4.14: Efficiency (top) and purity (bottom) as a function of $E_{T,jet}$ computed using the ARIADNE and LEPTO-MEPS event samples.

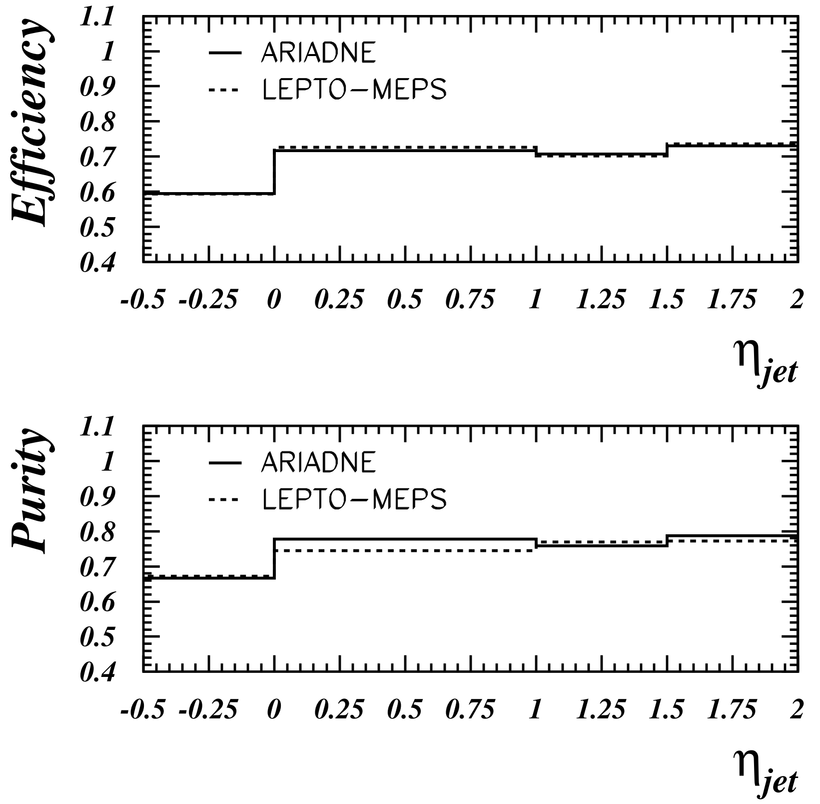


Figure 4.15: Efficiency (top) and purity (bottom) as a function of η_{jet} computed using the ARIADNE and LEPTO-MEPS event samples.

n_{sbj} hadrons → detector ↓	1	2	3	4	5	6	7	8
1	22.9	4.13	0.554	0	0	0	0	0
2	56.1	45.0	19.2	4.95	0.590	0	0	0
3	19.1	41.9	52.2	35.7	16.5	4.35	0	0
4	1.92	8.49	24.4	44.7	47.2	34.8	0	0
5	0	0.485	3.51	13.5	30.7	43.5	100	0
6	0	0.022	0.0792	1.19	5.01	13.0	0	0
7	0	0	0.016	0	0	4.35	0	0
8	0	0	0	0	0	0	0	0

n_{sbj} hadrons → detector ↓	1	2	3	4	5	6
1	53.9	14.8	2.28	0	0	0
2	43.1	68.0	40.3	12.9	0	0
3	2.98	16.6	52.7	65.6	0	0
4	0.065	0.599	4.69	20.4	100	0
5	0	0	0	1.08	0	0
6	0	0	0	0	0	0

n_{sbj} hadrons → detector ↓	1	2	3	4
1	94.14	37.4	0	0
2	5.86	62.1	75.0	0
3	0.005	0.490	25.0	0
4	0	0	0	0

Table 4.3: Fraction of jets (in per cent) that contains a given number of subjets at detector (row) and hadron (column) levels. It should be noted that the numbers are scaled in such a way that those in the same column add up to 100%. The three tables correspond to the complete sample of jets and for $y_{cut} = 5 \cdot 10^{-3}$ (top), $y_{cut} = 10^{-2}$ (bottom left) and $y_{cut} = 3 \cdot 10^{-1}$ (bottom right). The correlation increases as y_{cut} approaches 1. In general, the jets with low multiplicity contain more subjets at detector than at hadron level, whereas those with high multiplicity contain less subjets at detector level.

Since the main variable under study is the subjet multiplicity, an additional check has been performed to study the correlation between detector and hadron levels. In this case we cannot expect a perfect correlation, since detector effects introduce a smearing on this number. We have studied this correlation by looking at the distribution of the number of subjets for the jets at the detector level with a matched jet at hadron level which contains a given number of subjets.

Table 4.3 shows the correlation for different values of the resolution scale y_{cut}. The values show that for low multiplicity at hadron level, the jet have more subjets at detector level. On the other hand, for high multiplicity at hadron level, the jets have less subjets at the detector level. This effect is less pronounced as y_{cut} approaches 1.

A similar behaviour is observed at a fixed y_{cut} with increasing transverse energy of the jets; the results are shown in table 4.4. It is observed that the correlation increases when y_{cut} or $E_{T,jet}$ become larger. This behaviour leads to small correction factors in those regions, as discussed in the next section.

The use of simulated events makes it possible to recover in a consistent way the information of the hadronic final state by removing the detector effects. The way the simulated events are used to correct the values at detector level is presented in the next section.

n_{sbj} hadrons → / detector ↓	1	2	3	4	5
1	45.1	13.2	1.88	0	0
2	50.0	67.9	36.1	14.3	0
3	4.64	18.2	58.6	66.7	0
4	0.189	0.659	3.45	19.0	0
5	0	0	0	0	0

n_{sbj} hadrons → / detector ↓	1	2	3	4	5
1	52.9	14.0	0.926	0	0
2	44.3	70.4	34.3	28.6	0
3	2.71	15.2	58.3	42.9	0
4	0.082	0.363	6.48	28.6	0
5	0	0	0	0	0

n_{sbj} hadrons → / detector ↓	1	2	3	4	5
1	62.0	18.9	0	0	0
2	36.7	66.1	36.8	0	0
3	1.33	13.7	63.2	0	0
4	0	1.29	0	100	0
5	0	0	0	0	0

Table 4.4: Fraction of jets (in per cent) that contains a given the number of subjets at detector (row) and hadron (column) levels. The three tables correspond to jets with transverse energy in the ranges $15-17$ GeV (top left), $21-25$ GeV (to right) and $29-35$ GeV (bottom). In the three cases the resolution scale was taken as $y_{cut} = 10^{-2}$. Other details are as described in the caption to table 4.3.

4.2.4 Correction of detector effects using Monte Carlo techniques

The removal of detector effects is done by using a *correction factor* which is given by the Monte Carlo samples and the detector simulation.

In the present analysis, the correction has to be applied to the mean subjet multiplicity as reconstructed at detector level. The correction is given by the ratio of the result at hadron level over that at detector level. That is, we compute the bin-by-bin correction factor with the expression

$$\mathcal{F}_{\mathcal{C}} = \frac{\langle n_{sbj}\rangle|^{had}_{MC}}{\langle n_{sbj}\rangle|^{det}_{MC}} \tag{4.1}$$

This correction factor has been calculated for each value of y_{cut} using either the Monte Carlo sample of ARIADNE and LEPTO-MEPS. Figures 4.16 to 4.18 show the correction factor $\mathcal{F}_{\mathcal{C}}$ as a function of y_{cut}, $E_{T,jet}$ and η_{jet} for ARIADNE and LEPTO-MEPS.

It should be noted that the two MC samples give similar results, especially for the shape of the distribution, although the normalization is slightly different. The relative difference between the two MC is always less than 4%.

From the figures, it is observed that the correction factor gets smaller when y_{cut} increases. The same behaviour is observed for a given y_{cut} and increasing $E_{T,jet}$.

The correction factor has a small dependence on the pseudorapidity, and is larger in the forward region.

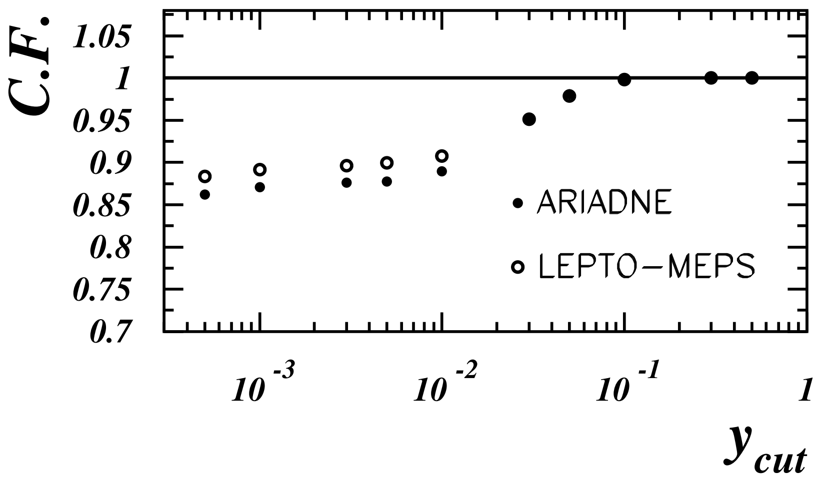


Figure 4.16: Correction factor for the mean subjet multiplicity as a function of y_{cut} computed using the ARIADNE and LEPTO-MEPS event samples.

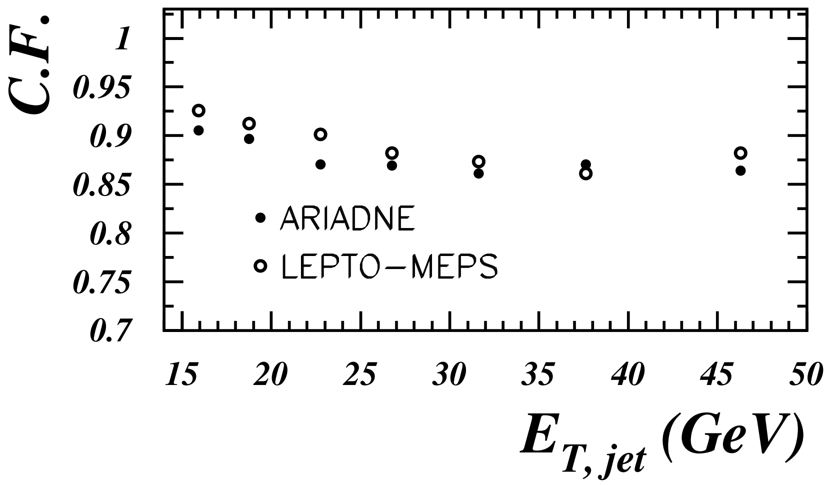


Figure 4.17: Correction factor for the mean subjet multiplicity as a function of the transverse energy of the jets for $y_{cut} = 10^{-2}$ computed using the ARIADNE and LEPTO-MEPS event samples.

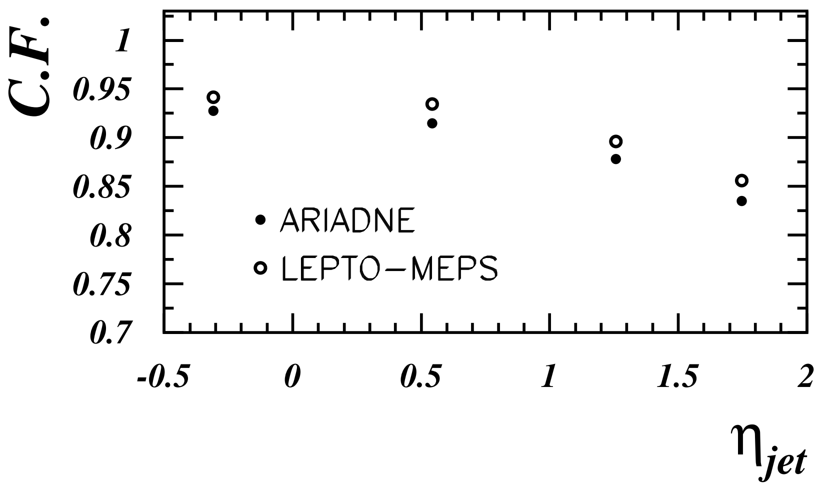


Figure 4.18: Correction factor for the mean subjet multiplicity as a function of η_{jet} for $y_{cut} = 10^{-2}$ computed using the ARIADNE and LEPTO-MEPS event samples.

4.3 Measurement of the subjet multiplicity

In this section, the measurements are presented, while the comparison with the NLO QCD predictions is described in the next chapter.

4.3.1 The mean subjet multiplicity

After the final selection of jets, we start by studying the subjet multiplicity as resulting from the algorithm as described in section 1.4.2. For each jet the number of subjets at each value of y_{cut} is counted and the average value of this number for the sample of jets is computed.

The measurements are done for the complete sample and for samples of jets in different regions of $E_{T,jet}$ and η_{jet}. This allows the study of the dependence of the mean subjet multiplicity on these variables. The regions for these variables are given by the following boundaries:

$$E_{T,jet} = 15,\ 17,\ 21,\ 25,\ 29,\ 35,\ 41,\ 55 \text{ GeV}$$

$$\eta_{jet} = -1,\ 0,\ 1,\ 1.5,\ 2.$$

It should be mentioned that in the plots the measured points are plotted at the corresponding average value of the variable in the given bin.

4.3.2 Comparison between data and Monte Carlo predictions

The first step was the comparison of the results at detector level between data and Monte Carlo simulations. Figures 4.19 and 4.20 show the comparisons for Q^2 and the jet variables. Both ARIADNE and LEPTO-MEPS give a reasonable description of the data. Thus the global properties of the jets in the data are well reproduced by the Monte Carlo models.

The measured mean subjet multiplicity is shown as a function of the resolution scale in figure 4.21. The values shown are those for the sample of jets reconstructed in the detector. The values reconstructed using the Monte Carlo samples are also shown. The mean subjet multiplicity in the data is well described by the Colour Dipole Model (ARIADNE) while the parton-shower(LEPTO-MEPS) approach seems to overestimate the data.

Figures 4.22 and 4.23 show the comparisons of the mean subjet multiplicity in data and simulations as a function of the jet variables. ARIADNE gives a good description of the data. LEPTO-MEPS predictions are similar to those of ARIADNE although the former are above the data.

The number of subjets within a jet can be considered as an additional jet variable, and this means that we can study the distribution of the number of jets containing a given number of subjets at a given y_{cut}. This is shown in figures 4.24 and 4.25 for the complete sample of jets and different resolution scales, and in figures 4.26 and 4.27 for the different regions in $E_{T,jet}$ at $y_{cut} = 10^{-2}$. It should be observed that the distributions in the data are well described in shape by the Monte Carlo predictions, although a small shift is observed in LEPTO-MEPS. This shift explains the difference observed in the mean subjet multiplicity between the data and the predictions of LEPTO-MEPS.

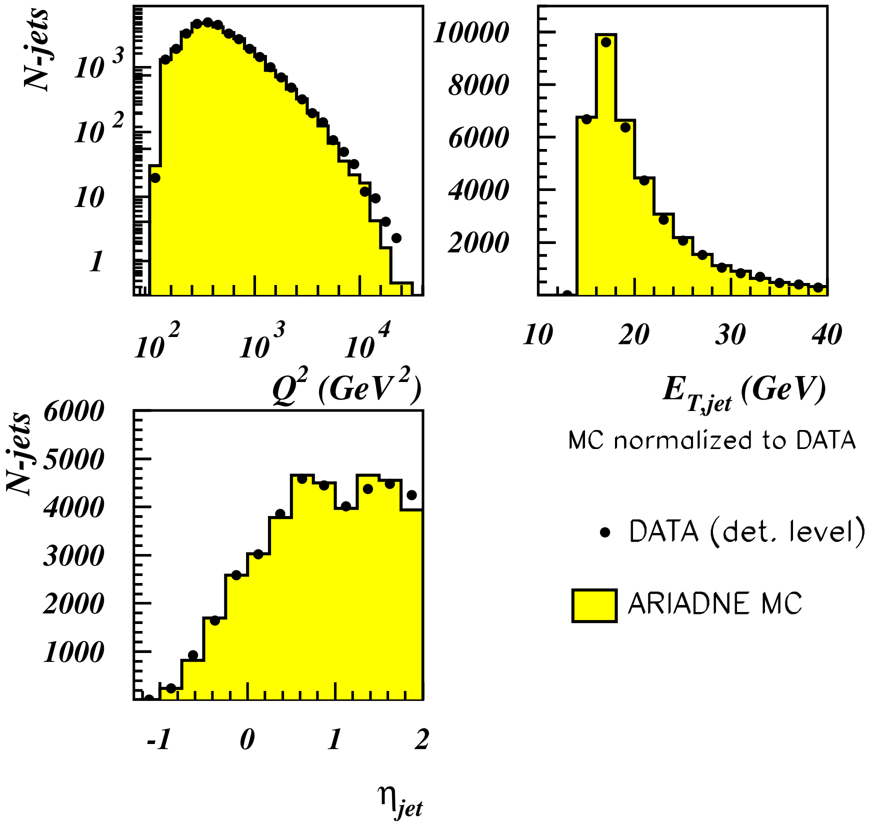


Figure 4.19: Distribution of Q^2 and jet variables at detector level in the data (black dots). The predictions from the ARIADNE sample (histograms) are also shown. The Monte Carlo prediction has been normalized to the data in order to compare the shape of the distributions.

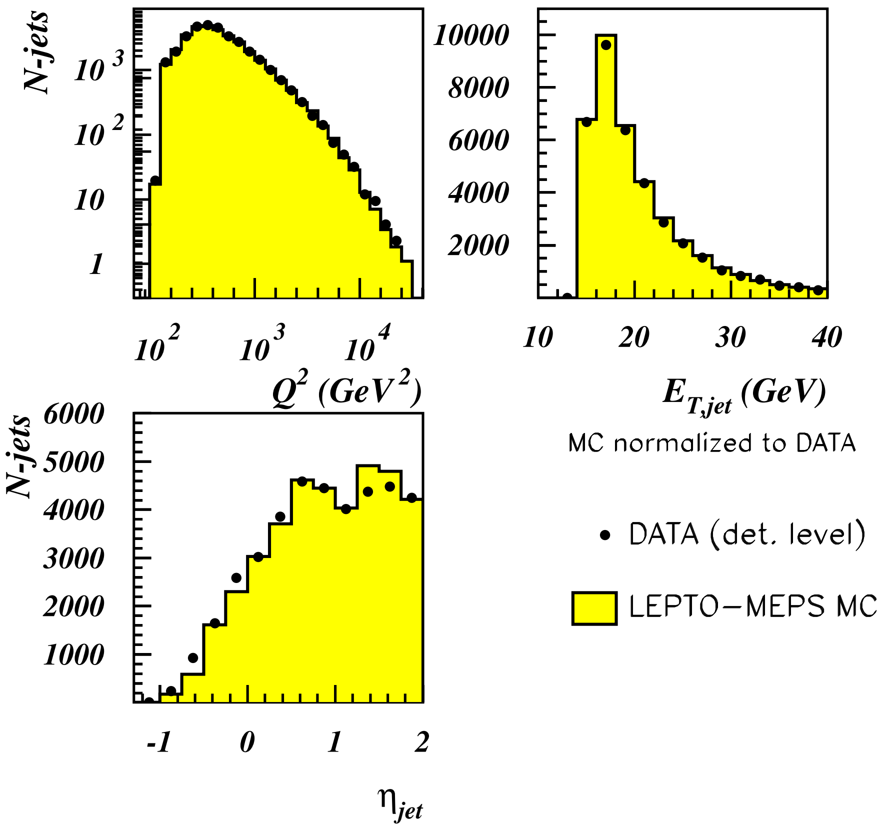


Figure 4.20: Distribution of Q^2 and jet variables at detector level in the data (black dots). The predictions from the LEPTO-MEPS sample (histograms) are also shown. The Monte Carlo prediction has been normalized to the data in order to compare the shape of the distributions.

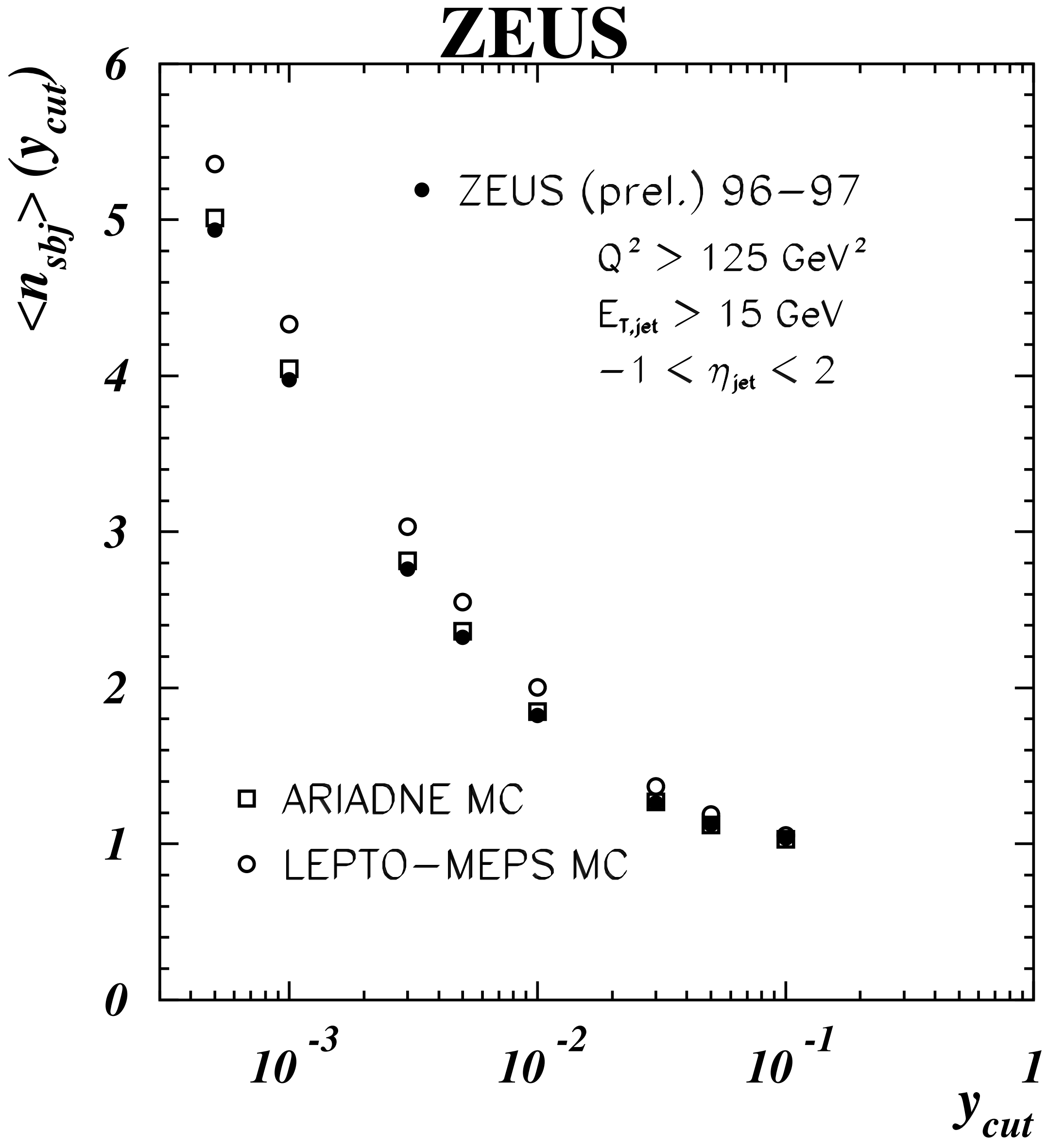


Figure 4.21: Mean subjet multiplicity in the data as reconstructed with the calorimeter as a function of y_{cut} (black dots). The measurements are compared with the predictions of the Monte Carlo samples: ARIADNE (open squares) and LEPTO-MEPS (open circles).

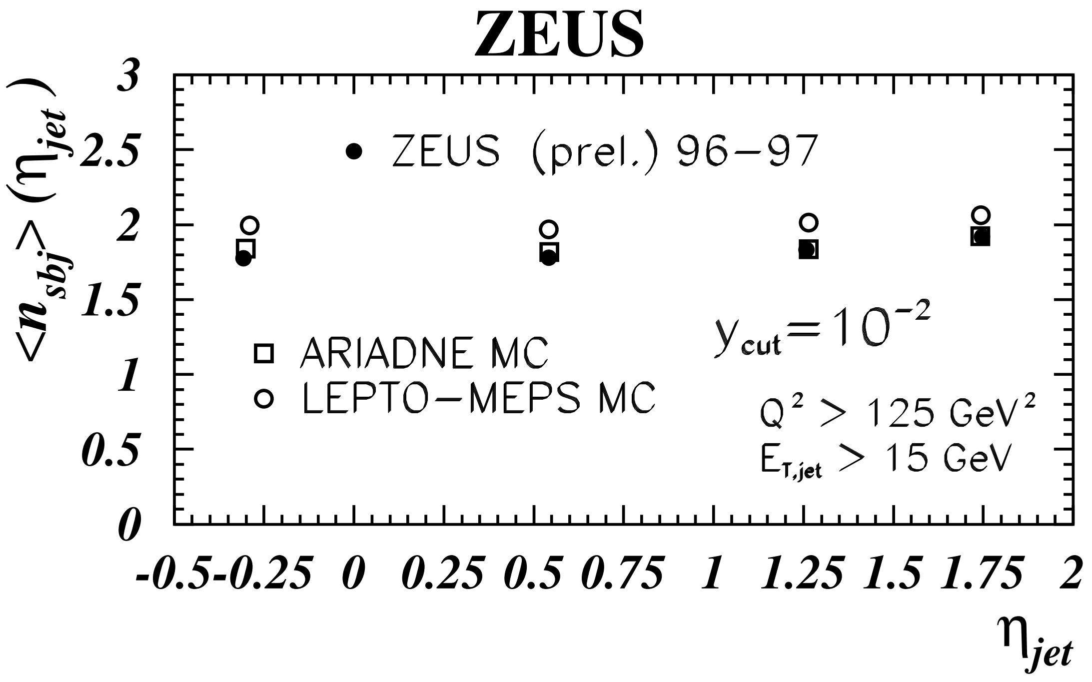


Figure 4.22: Mean subjet multiplicity in the data as reconstructed with the calorimeter as a function of η_{jet} (black dots). Other details are as described in the caption to figure 4.21.

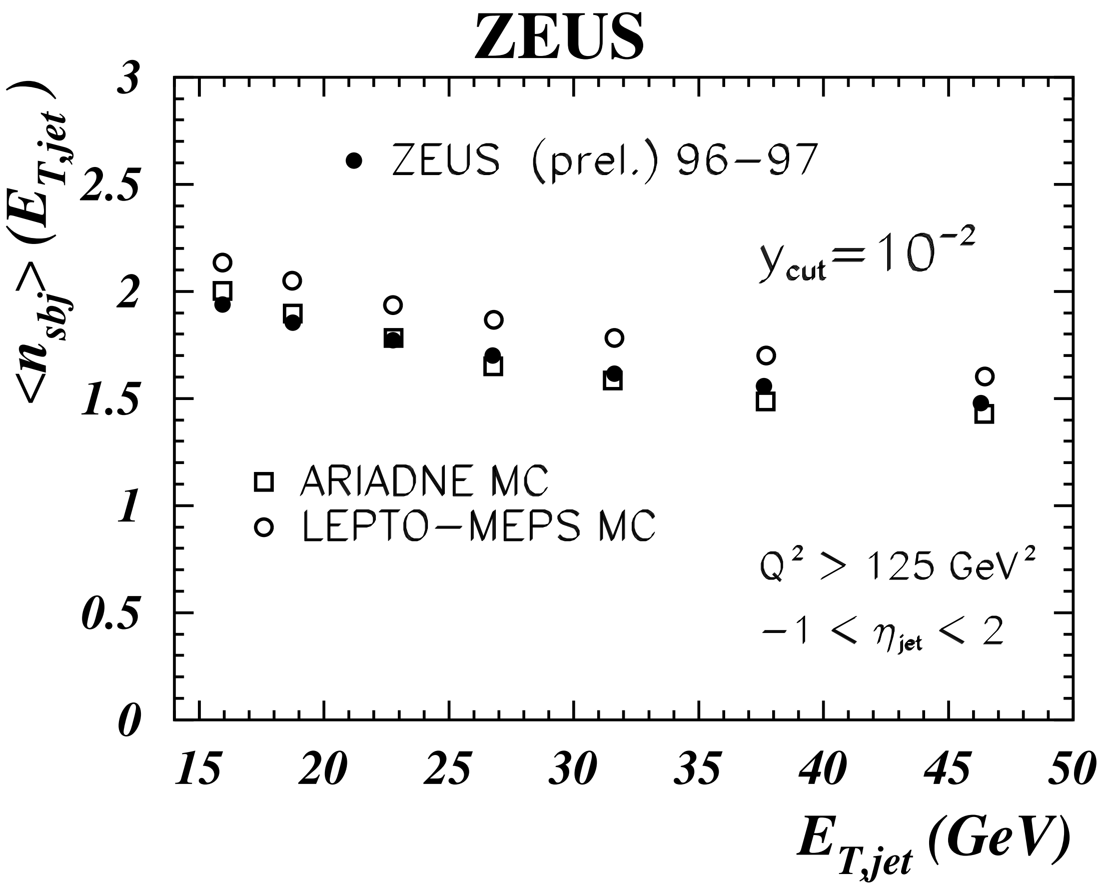


Figure 4.23: Mean subjet multiplicity in the data as reconstructed with the calorimeter as a function of $E_{T,jet}$ (black dots). Other details are as described in the caption to figure 4.21.

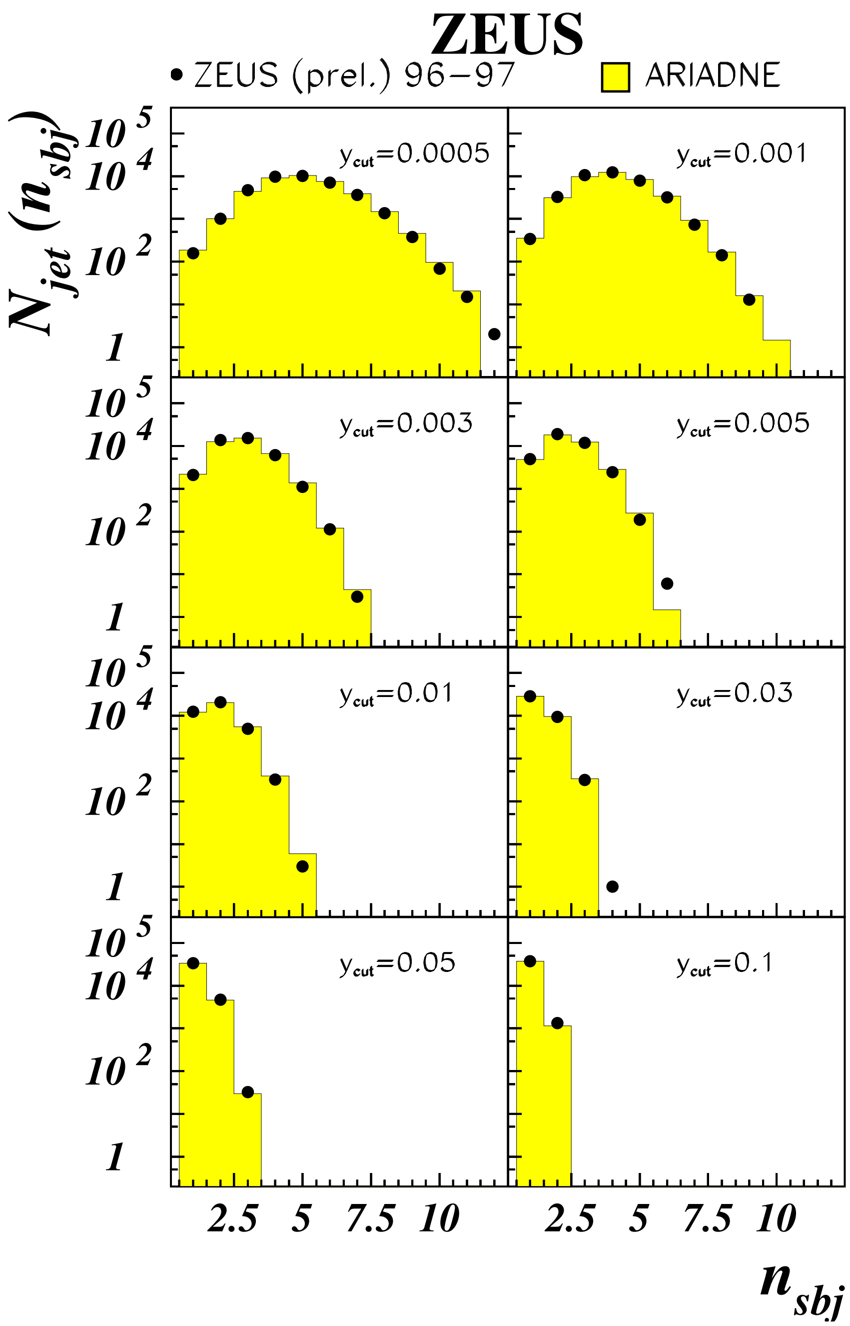


Figure 4.24: Distribution of the number of subjets in the data for different values of y_{cut} (black dots). The predictions of ARIADNE (histograms) are normalized to the data and are also shown.

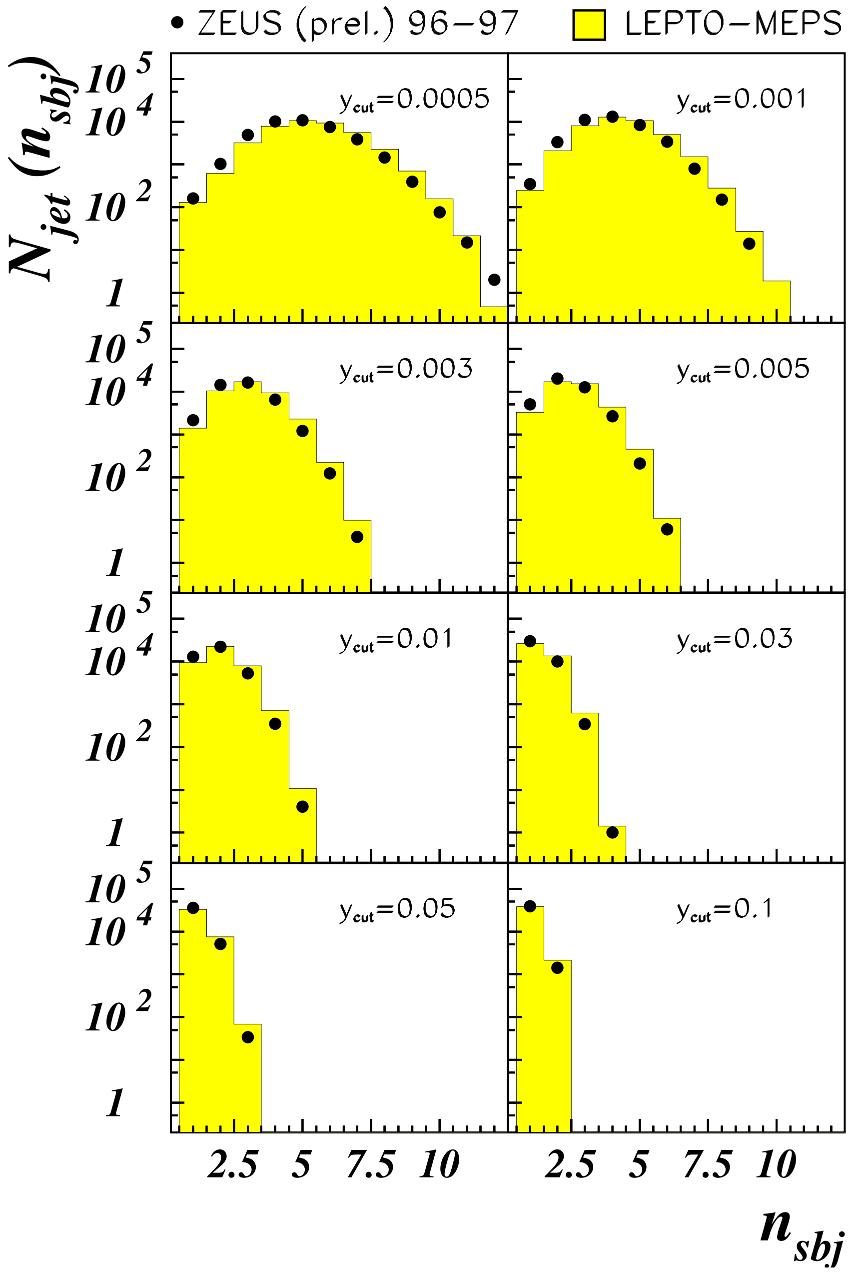


Figure 4.25: Distribution of the number of subjets in the data for different values of y_{cut} (black dots). The predictions of LEPTO-MEPS (histograms) are normalized to the data and are also shown.

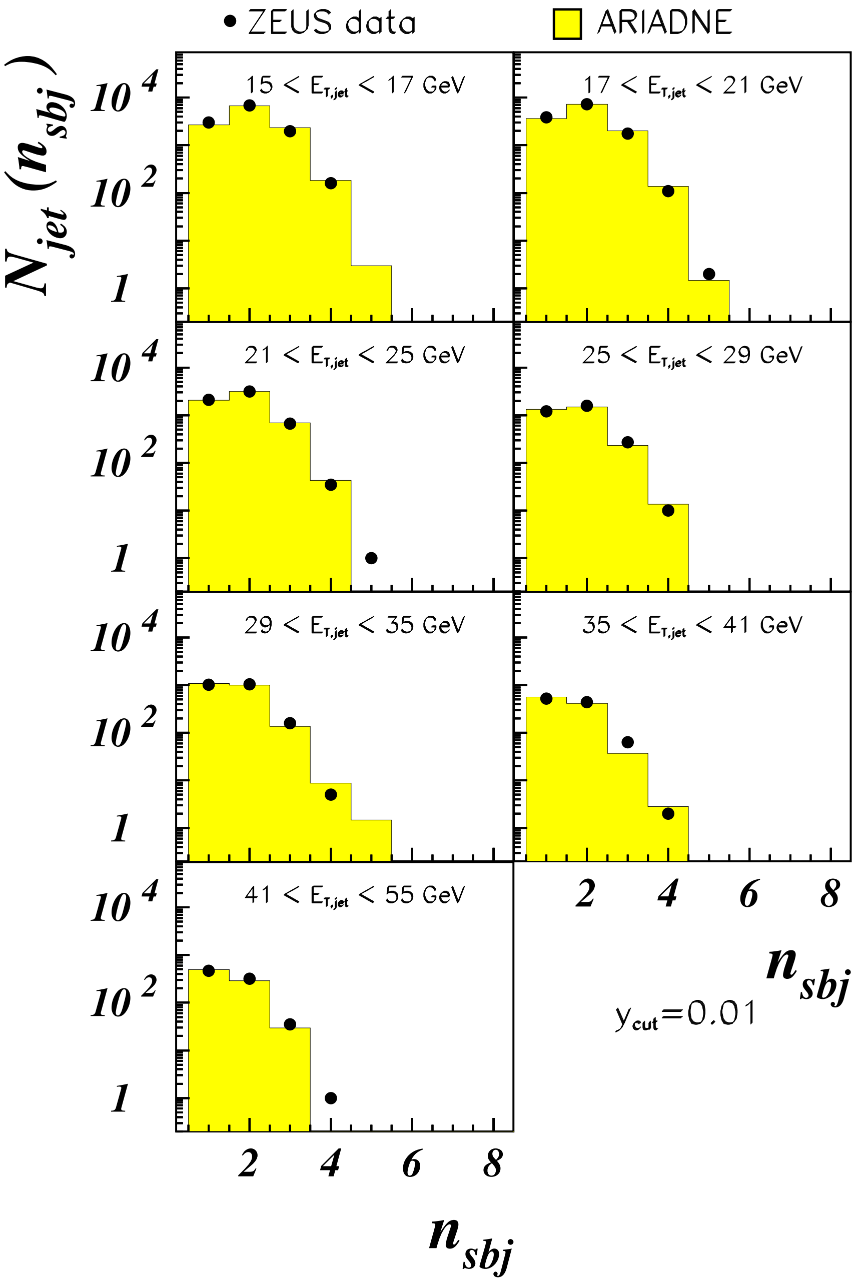


Figure 4.26: Distributions of the number of subjets in the data for $y_{cut} = 10^{-2}$ and different regions in $E_{T,jet}$ (black dots). The predictions of ARIADNE (histogram) are normalized to the data and are also shown.

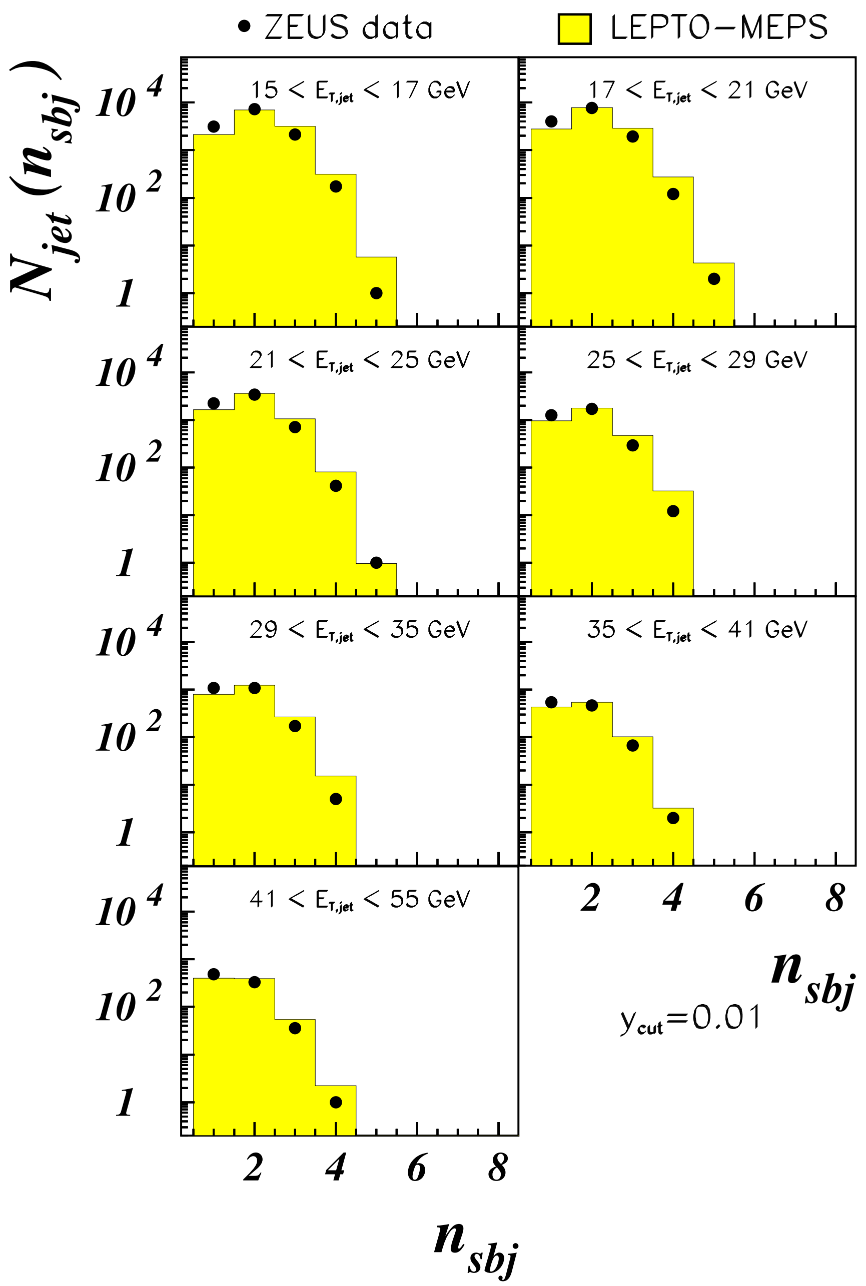


Figure 4.27: Distributions of the number of subjets in the data for $y_{cut} = 10^{-2}$ and different regions in $E_{T,jet}$ (black dots). The predictions of LEPTO-MEPS (histogram) are normalized to the data and are also shown.

The good description of these distributions in the data given by the Monte Carlo of these distributions supports that the simulations are describing the detector effects in a correct way. For this reason we are confident that correcting for the detector effects using these Monte Carlo samples, at least with ARIADNE, is sensible.

The correction using LEPTO-MEPS is more problematic since it does not give a good description of the observed values as ARIADNE. In any case, the correction using LEPTO-MEPS is used as a systematic check; it should be noted that this estimation of the systematic uncertainty might be an overestimation due to the poorer description given by LEPTO-MEPS in comparison with ARIADNE.

4.3.3 Measurements corrected for detector effects

Since ARIADNE gives a good description of the data, we correct the mean subjet multiplicity of the data by applying the correction factor as described in section 4.2.4. The corrected value is the measurement at hadron level.

Figure 4.28 shows the corrected mean subjet multiplicity as a function of y_{cut}. The measured values in the data are compared with the predictions given by the ARIADNE and LEPTO-MEPS. The comparison is similar to that at detector level described before. This shows that although LEPTO-MEPS displays a relatively large shift in the predicted values from the data, the correction factor is very similar to the one given by ARIADNE. The reason is that the shift appears both at the hadron and detector levels and, thus, the correction is similar.

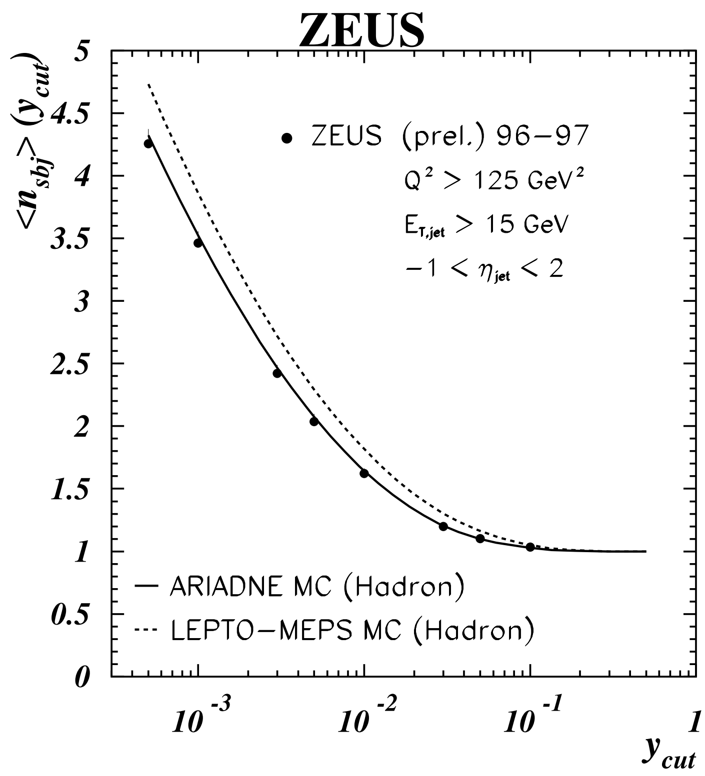


Figure 4.28: Corrected mean subjet multiplicity as a function of y_{cut} (black dots). The predictions of ARIADNE (solid line) and LEPTO-MEPS (dashed line) at hadron level are also shown.

Figures 4.29 and 4.30 show the measurements at $y_{cut} = 10^{-2}$ as a function of the jet variables.

The mean subjet multiplicity shows no dependence on the jet pseudorapidity. This is consistent with the QCD prediction that the sample of jets is dominated by quark-initiated jets in the entire η_{jet} range. This result and the interpretation are discussed in more detail in section 5.1.

Regarding the $E_{T,jet}$ dependence, the mean subjet multiplicity decreases as $E_{T,jet}$ increases. This result has been observed in other studies of the jet internal structure [92, 93, 94].

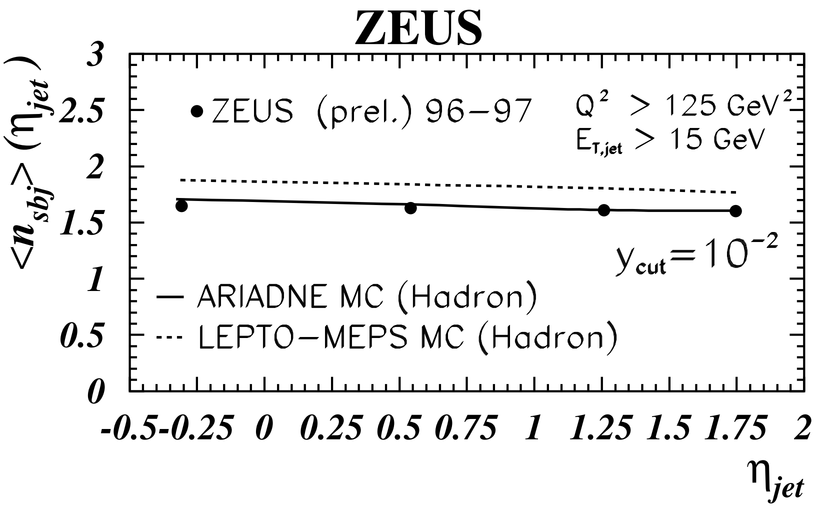


Figure 4.29: Corrected mean subjet multiplicity as a function of η_{jet} for $y_{cut} = 10^{-2}$ (black dots). The predictions of ARIADNE (solid line) and LEPTO-MEPS (dashed line) at hadron level are also shown.

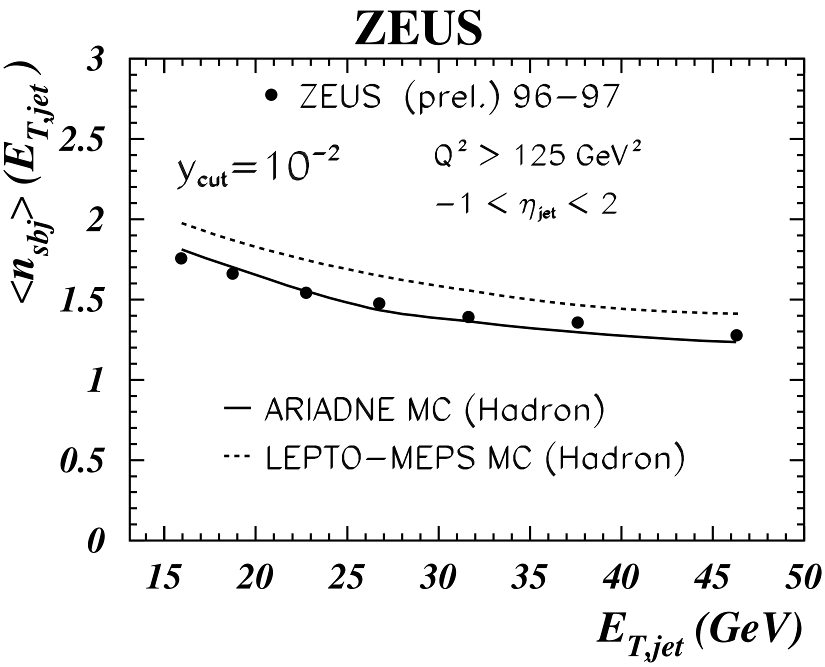


Figure 4.30: Corrected mean subjet multiplicity as a function of $E_{T,jet}$ for $y_{cut} = 10^{-2}$ (black dots). The predictions of ARIADNE (solid line) and LEPTO-MEPS (dashed line) at hadron level are also shown.

4.3.4 Study of systematic uncertainties

The experimental procedure described in the previous sections is affected by systematic uncertainties. Most of them come from the incomplete knowledge of detector effects and some from the definitions and methods used to analyze the data.

In order to take these effects into account, the analysis has been repeated with some modifications. The observed difference is then considered as the systematic uncertainty coming from the corresponding aspect of the analysis. For example, in the previous sections we have already mentioned the difference in the correction factors between ARIADNE and LEPTO-MEPS; this difference has to be taken into account, since there is no argument to assure that the values given by ARIADNE are more reliable than those of LEPTO-MEPS. Of course, they are preferred since they give a better description of the data, and that is the reason why we use that model for the central values.

A preliminary study of the systematic uncertainties has been done. In addition to the comparison between the two Monte Carlo models, the following sources of uncertainty have been considered:

- The uncertainty coming from that on the energy scale of the calorimeter is estimated by changing the energies in the detector level of the simulated events by $\pm 3\%$. The contribution is separated into the following ones; it should be noted that each contribution is computed independently.
  - The $\pm 3\%$ variation of the transverse energy of the jets in simulated events introduces a relatively large effect, around 1-2%. This effect comes from the dependence of the mean subjet multiplicity on the transverse energy.
  - The $\pm 3\%$ variation of the energy of the positron candidate induces a very small effect in comparison with other contributions.
  - The $\pm 3\%$ in the value of $E - p_z$ gives rise to a negligible effect..
- The cut on y_{ele} was set to 0.9 instead of the default 0.95. This cut was changed in data as well as in simulated events. The effect is very small.
- The cut on the minimum energy for the positron candidate to be considered a valid candidate was raised to 11 GeV. This cut is changed both in data and simulated events. The effect is very small.
- The cut on $\not{p}_T/\sqrt{E_T}$ was changed to 3 GeV$^{1/2}$ and 2 GeV$^{1/2}$ from the standard one at 2.5 GeV$^{1/2}$. The effect is very small.

Figures 4.31 to 4.33 show the size of the systematic uncertainties for the measurements as functions of y_{cut},$E_{T,jet}$ and η_{jet}. Since most of them are negligible in comparison with the statistical error, only the largest ones are shown.

The conclusion is that the most important source of uncertainty is that of using LEPTO-MEPS to correct the data. The uncertainty due to that of the energy scale of the calorimeter is important for the measurements as a function of $E_{T,jet}$. These two sources of systematic uncertainty are the only ones which are of relative importance, as can be seen in the comparison with the statistical errors.

To obtain the total systematic uncertainty, all the systematic uncertanties are added in quadrature. This procedure assumes that these uncertainties are independent. Although this is not exactly true, since the main contributions satisfy this condition, it is a good approximation.

To obtain the complete experimental uncertainty, the systematic and the statistical uncertainties are added in quadrature. This experimental uncertainty is shown for every point in all the figures showing the results of the analysis. The statistical uncertainty is shown as inner error bars while the complete experimental uncertainty is shown as the outer error bars.

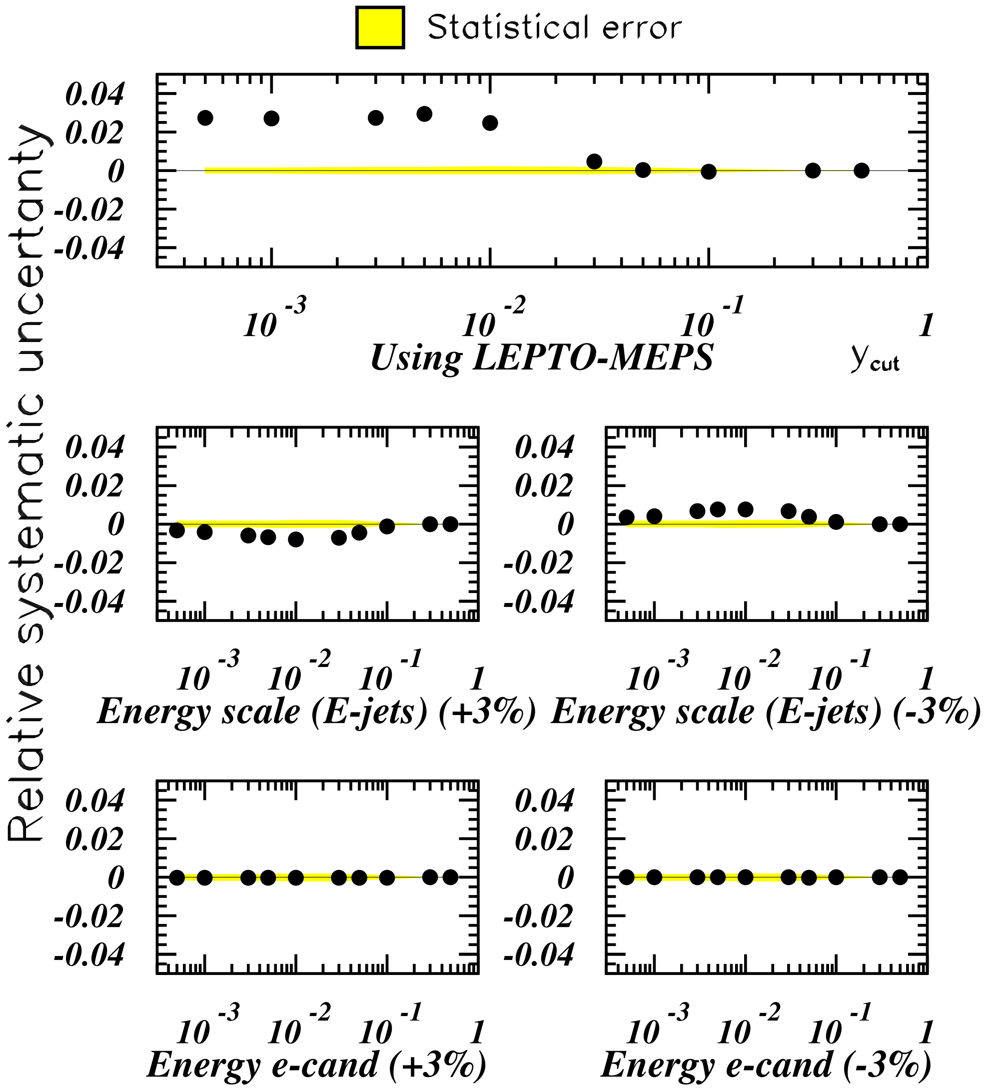


Figure 4.31: Systematic uncertainties of the measurement of the mean subjet multiplicity as a function of y_{cut} (black dots). For comparison, the statistical error is shown as a shaded band (too small to be seen).

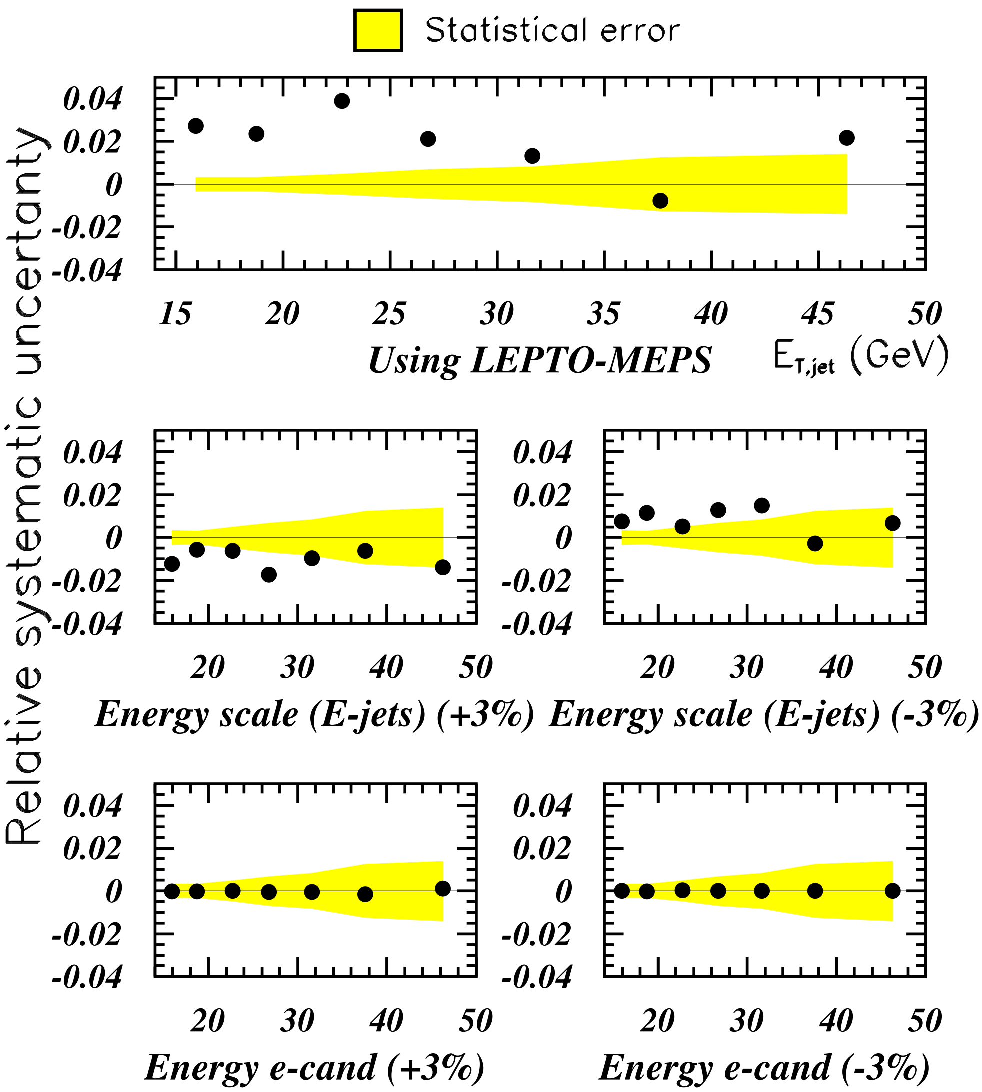


Figure 4.32: Systematic uncertainties of the measurement of the mean subjet multiplicity as a function of $E_{T,jet}$ for $y_{cut} = 10^{-2}$ (black dots). For comparison, the statistical error is shown as a shaded band.

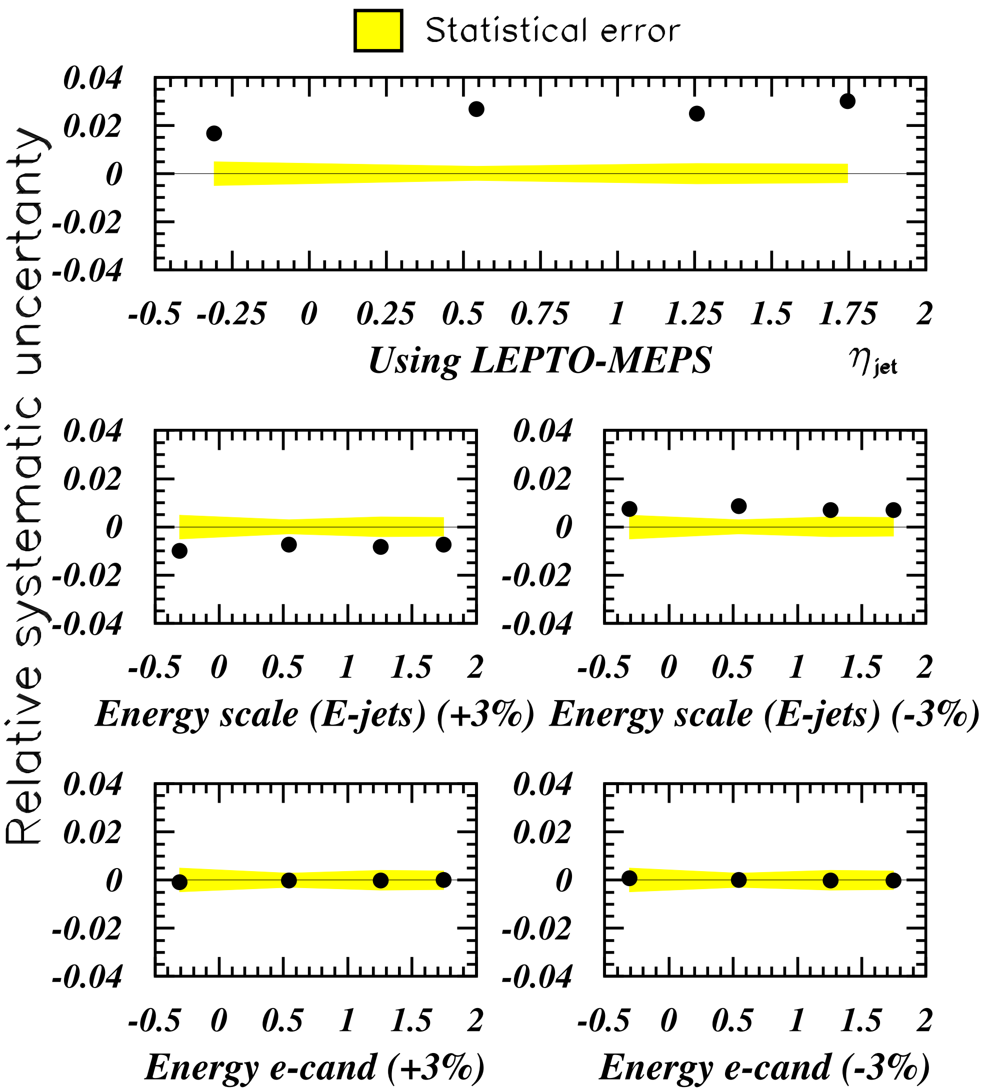


Figure 4.33: Systematic uncertainties of the measurement of the mean subjet multiplicity as a function of η_{jet} for $y_{cut} = 10^{-2}$ (black dots). For comparison, the statistical error is shown as a shaded band.

Chapter 5

Discussion of the results

The results obtained in the previous chapter are discussed in the following sections to draw the final conclusions. In this chapter, the next-to-leading order pQCD predictions are obtained, corrected to hadron level and compared with the data. The comparison allows the extraction of the value of α_s.

5.1 Quark and gluon-initiated jets

As mentioned in section 1.4.3, QCD predicts differences between the internal structure of the quark and gluon jets. A complete analysis of these differences can only be done by tagging the quark and the gluon jets. At present quark tagging in NC DIS cannot be made with sufficient statistical accuracy and the differences between quark and gluon initiated-jets have been analyzed by using the Monte Carlo samples. The measurements are then compared to the MC predictions for pure samples of quark and gluon jets. From this comparison, conclusions have been obtained concerning the quark and gluon content of the data sample.

It is possible to obtain predictions for the mean subjet multiplicity in jets initiated by quarks or by gluons using the sample of events generated with ARIADNE. The first step is to separate the jets in two subsamples by assigning the jets a given type of parton. This is done using the method described in appendix C.

Having done this separation, the analysis is performed for each subsample, and the same observables considered before are measured for quark and gluon-initiated jets. The predicted values are then compared to the data, as shown in figures 5.1 and 5.2. The measured values in the data are consistent with the expectation that the sample is dominated by quark-initited jets over the entire η_{jet} range; this dominance is reflected in the independence of $\langle n_{sbj} \rangle$ on η_{jet}.

The results of a similar analysis in photoproduction [92] are shown in figure 5.2 for comparison. A clear increase of the mean subjet multiplicity is observed as η_{jet} increases in the photoproduction regime.

In the photoproduction case the fraction of gluon-initiated jets is expected to increase with η_{jet}, and the observed dependence of the mean subjet multiplicity on η_{jet} is consistent with expectations. The constant behaviour observed in the DIS regime confirms the interpretation given to the absence of a variation of the mean subjet multiplicity on η_{jet} in terms of the quark dominance for the entire η_{jet} range.

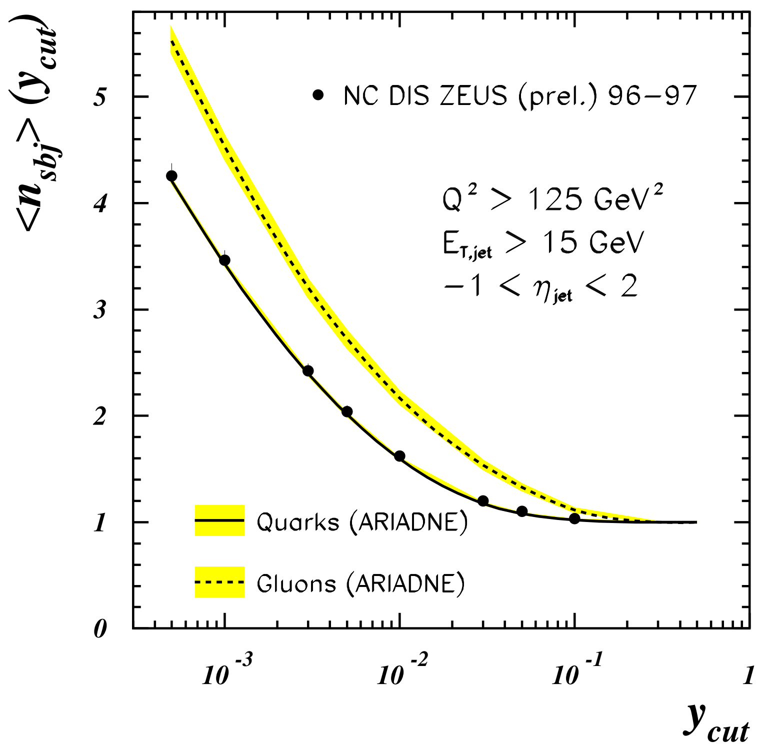


Figure 5.1: Measurements of the mean subjet multiplicity as a function of y_{cut} (black dots). For comparison, the predicted values for quark (solid line) and gluon-initiated (dashed line) jets are also shown. The predictions and the uncertainties (displayed as shaded bands) were extracted by using the ARIADNE model as described in detail in appendix C. It should be observed that the measured values are in agreement with the sample being dominated by quark-initiated jets.

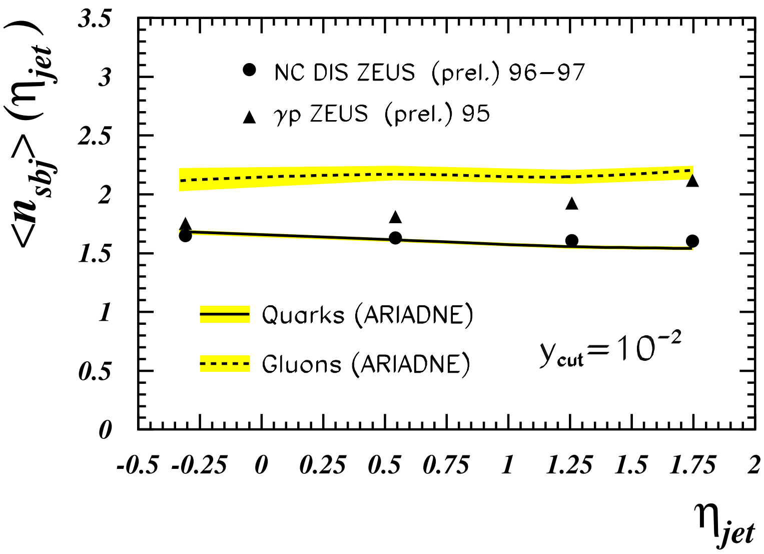


Figure 5.2: Measurements of the mean subjet multiplicity as a function of η_{jet} for y_{cut} = 0.01 (black dots). For comparison, the predictions for quark (solid line) and gluon-initiated (dashed line) jets are also shown. In addition, results for the photoproduction of jets are shown. The measured values for jets in DIS are consistent with a quark-dominated sample of jets in the entire η_{jet} range, whereas the measurements in photoproduction display an increase of the number of subjets towards the forward region, which can be interpreted as due to the increase in the fraction of gluon-initiated jets.

5.2 NLO QCD predictions

To make a precise comparison with the theory, pQCD predictions for the mean subjet multiplicity has been made to the second order in α_s. The Monte Carlo models earlier considered only contain the first-order contribution and introduce the effects of higher orders in an approximate way. On the other hand, for jets defined in the laboratory frame, it is at present possible to make predictions of the jet internal structure for the first two orders.

The way these predictions are obtained and corrected for hadronization is explained in the following sections. Finally, in section 5.3, the comparison to the data is presented.

5.2.1 Description of the NLO QCD calculation

The mean subjet multiplicity at a fixed order in the perturbative QCD expansion depends upon the number of partons inside a jet. At lowest order the jet is only formed by one parton and the mean subjet multiplicity is exactly 1, independently of the resolution scale. This order is referred to as the *trivial order*.

If two partons are inside the jet, the maximum number of subjets is 2, but depending on how these partons are distributed inside the jet, it will be possible to find 1 or 2 subjets. Therefore there is a dependence on the resolution scale. The contributions to the configuration in which a jet contains 2 subjets are corrections to the trivial order and constitute the first non-trivial contribution to the mean subjet multiplicity. This order is called the leading order.

To calculate the leading order for the mean subjet multiplicity, it is needed to divide the number of jets containig two subjets at the given resolution scale over the number of jets,

$$\langle n_{sbj} \rangle = 1 + \frac{N_{jets}\,(sbj = 2)}{N_{jets}} + \cdots = 1 + A_1 \alpha_s + \ldots \;, \tag{5.1}$$

where both numerator and denominator are calculated at the first non-trivial order.

In the analysis presented here, since jets are reconstructed in the laboratory frame, the first order contribution to the denominator is the Quark-Parton model (non QCD) diagram.

To obtain the next-to-leading order predictions, it is needed to generalize the expression above to the second order. As the order is increased, the number of partons which are able to be inside a jet is increasing. To compute $\langle n_{sbj} \rangle$ we use the following expression

$$\langle n_{sbj} \rangle - 1 = \frac{\sum_{jets} ({n_{sbj}}^{jet} - 1)}{\sum_{jets} 1} = \frac{a_1 \alpha_s + a_2 {\alpha_s}^2 + \ldots}{b_0 + b_1 \alpha_s + \ldots} \;, \tag{5.2}$$

where the last term is the expression at next-to-leading order. This expression allows the computation to be reduced to the calculation of the number of jets and subjets in a consistent way.

At present the calculation of the mean subjet multiplicity to second order in pQCD is only possible in DIS for jets defined in the laboratory frame; in other frames, terms which have not been calculated are required.

5.2.2 The program DISENT

The calculations of the mean subjet multiplicity for the analysis presented here are obtained using DISENT [95], which contains all the matrix elements needed for the calculation at NLO in perturbative QCD. The analytical calculation is too complicated and the predictions are obtained by solving the integrations numerically, by using Monte Carlo methods.

The program DISENT generates different configurations, which correspond to the terms in the perturbative expansion, for the partons in the initial and final states, each of which has an associated weight. After convoluting each configuration with the parton distribution functions, the total contribution to the cross section is obtained.

The k_T-cluster algorithm is then applied to the partons in the final state in the same way it was applied to the hadron and detector levels. The selection cuts are applied in the same way as for the hadron level, and finally, if the configuration is accepted, the weight is counted so much times as accepted jets in the configuration to obtain: a) the cross section to produce jets and b) the cross section to produce jets with a given number of subjets. The integration procedure consists on generating a large number of configurations; the final result for $\langle n_{sbj} \rangle$ is then obtained by dividing both cross sections and adding unity, the trivial order.

The calculation of the matrix elements for each configuration is made internally in DISENT. The infrared divergences are removed by using the so-called *Dipole Formalism* [96] based on the *substraction method* [97].

In addition to the matrix elements, the following ingredients must be introduced to complete the calculation:

- **The PDF set**: the convolution of the weights given by DISENT with the parton densities is essential to obtain the final result. In this analysis, the parametrizations of the parton densities obtained by the CTEQ4 collaboration are used for the central predictions.
- **The factorization scale**: as described in section 1.3.3 the presence of a hadron in the initial state leads to the introduction of a factorization scale to separate the processes inside and outside the hadron. In the present analysis, this scale was set equal to $\sqrt{Q^2}$. It is expected that the results do not depend significantly on this scale.
- **The renormalization scale**: as described in section 1.3.2 the mechanisms used to remove the UV divergences in the predictions of perturbative QCD introduce a dependence on the renormalization scale which, in turn, can be used to estimate the contribution of higher orders to the observable under study. This is done in the analysis presented here, as described in the following section.

5.2.3 Results obtained with DISENT

The predictions for the mean subjet multiplicity have been obtained with DISENT by using the procedure described in section 5.2.1. Figure 5.3 shows the NLO QCD results for $\langle n_{sbj} \rangle$ as a function of y_{cut}, η_{jet} and $E_{T,jet}$.

The predictions have been obtained for different values of α_s. This is performed by using in each case a set of parton densities which was determined assuming a given value of α_s. In this way, the correlation between the α_s value in the matrix elements and that assumed in the

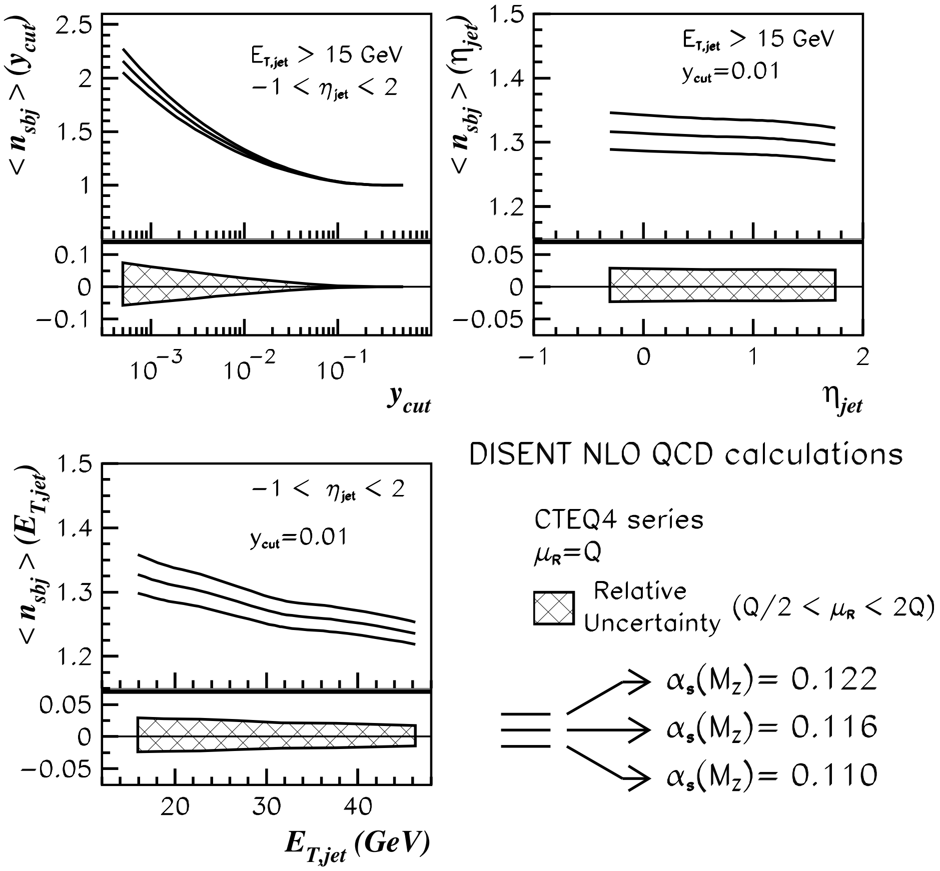


Figure 5.3: NLO QCD predictions for $\langle n_{sbj} \rangle$ as a function of y_{cut} (top left), η_{jet} (top right) and $E_{T,jet}$ (bottom) obtained with DISENT. The three curves in each plot show the predictions for different values of $\alpha_s(M_Z)$. The hatched band below the plot shows the relative uncertainty coming from the residual dependence on the renormalization scale, which is estimated by varying it by a factor of 2.

parametrization of the parton densities is taken into account. The results shown in the figure were obtained with three of the five sets of the CTEQ4A series [32]. A clear dependence with the value of α_s is observed.

In addition, calculations have been done using as renormalization scale $\frac{\sqrt{Q^2}}{2}$ and $2\sqrt{Q^2}$. This is the standard way to quantify the dependence on the renormalization scale. This dependence allows the estimation of the uncertainty due to the higher-order terms not included in the calculations, and it is shown as a hatched band below each plot. It is observed that the dependence on the renormalization scale gets clearly smaller when the resolution scale is larger. However, this uncertainty displays only a slightly dependence on the transverse energy of the jets or the jet pseudorapidity.

For a more complete description of the dependence on the renormalization scale, figure 5.4 shows this dependence over 3 orders of magnitude in μ_R for the 7 bins in $E_{T,jet}$ at $y_{cut} = 10^{-2}$. It shows that at NLO the μ_R dependence is still relatively large This means that higher order contributions are still important for the predictions of this observable. This importance is translated in the analysis into a relatively large theoretical uncertainty.

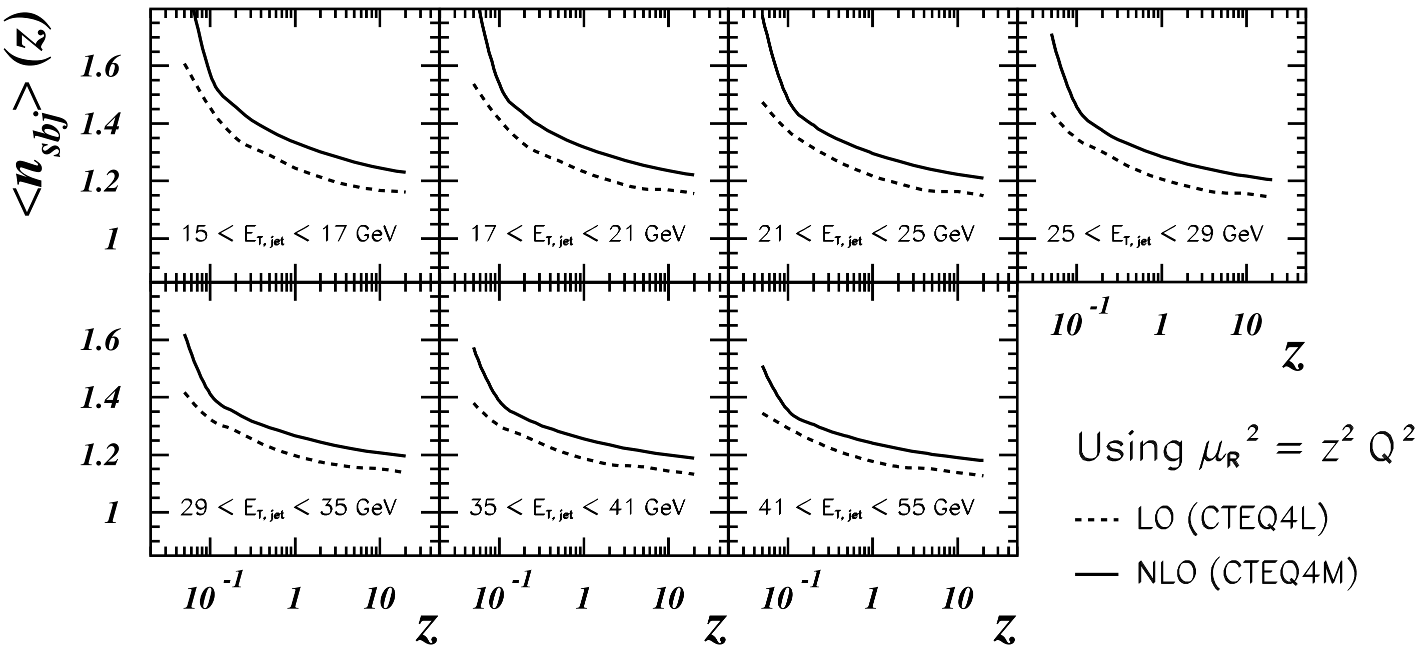


Figure 5.4: NLO QCD calculations of the mean subjet multiplicity as a function of z ($\mu_R{}^2 = z^2Q^2$) at $y_{cut} = 10^{-2}$ in different $E_{T,jet}$ regions. The leading order (dashed lines) and the next-to-leading order (solid lines) predictions display a similar dependence, but the normalization is different.

5.2.4 Parton-to-hadron correction

To compare the theoretical predictions with the data, it is necessary to correct the predictions to the hadron level. This correction is not calculable in perturbative QCD and has to be estimated by using phenomenological approaches. In the analysis presented here the parton-to-hadron correction has been estimated by using the Monte Carlo models ARIADNE and LEPTO-MEPS.

The correction is estimated by using the ratio of the value at hadron level over the corresponding one at parton level. The values at hadron level are the same described in section 4.1.4.

For the parton level, all the partons which are produced after the parton radiation takes place are considered. These are the partons that were used as the starting point for the hadronization procedure. These partons are handled in the same way as the hadrons. First all, the three-momenta are scaled to obtain massless four-momenta for the partons. In a second step, the k_T-cluster algorithm is applied to these four-momenta to obtain the jets and the number of subjets for each jet at each resolution scale. Finally, the same selection for kinematic variables and jets as at hadron level is performed. It should be noted that the kinematic variables for a given event are exactly the same at hadron and parton levels.

After obtaining the final sample of jets at the parton level, the observables are computed. The mean subjet multiplicity is calculated in the same bins as for the data. The hadronization correction is defined as

$$\mathcal{C}_{\mathcal{H}} = \frac{\langle n_{sbj}\rangle|^{had}_{MC}}{\langle n_{sbj}\rangle|^{par}_{MC}} . \tag{5.3}$$

Figure 5.5 shows the hadronization corrections applied to the NLO QCD predictions as estimated by using either ARIADNE or LEPTO-MEPS. The correction is smaller as y_{cut} or $E_{T,jet}$ increases as expected since in those cases harder processes are resolved and the soft ones are less and less important.

It should be noted that the uncertainty in this correction, estimated by comparing the two phenomenological models, becomes smaller when the correction itself is smaller. In the higher-$E_{T,jet}$ region, the two models predict very similar results, so it is expected that hadronization corrections in that region are under control.

5.2.5 Theoretical uncertainties

A detailed study of the uncertainties in the theoretical predictions has been done and is described here.

The considered sources of uncertainty are listed below:

- Higher-order effects: they have been estimated by varying the renormalization scale between $\frac{\sqrt{Q^2}}{2}$ and $2\sqrt{Q^2}$ by a factor of 2. This represents the largest uncertainty.
- The hadronization correction: it was estimated by comparing the corrections obtained with ARIADNE and LEPTO-MEPS.
- The renormalization scale was set equal to the transverse energy of the jet with highest $E_{T,jet}$ The results are very similar to those using $\sqrt{Q^2}$ and always within the uncertainty of varying the renormalization scale by a factor of 2. For this reason, this variation is not considered as an additional source of uncertainty.
- The calculations were repeated using the MRST98 parametrizations [30] of the parton densities. No significant difference was found with respect to the predictions obtained using CTEQ4.

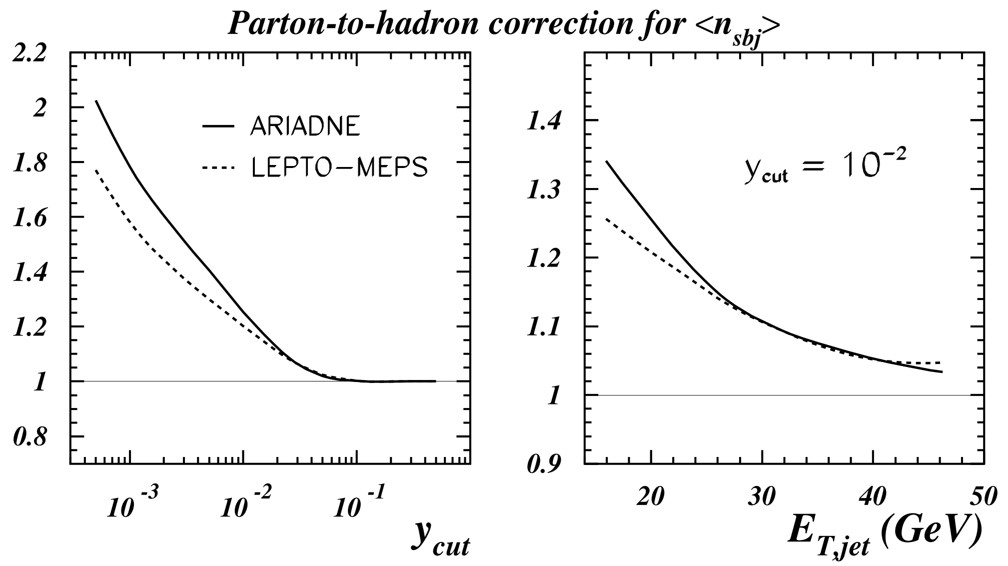


Figure 5.5: Parton-to-hadron correction factor as estimated by using ARIADNE (solid lines) and LEPTO-MEPS (dashed lines) as a function of y_{cut} (left) and $E_{T,jet}$ (right).

- The uncertainty due to that on the gluon content of the proton was estimated by repeating the calculations with sets of parton densities which make different assumptions on the gluon content of the proton. Three sets of the MRST98 series permit to make this check. Differences smaller than 0.1% are observed .

It should be emphasized the small uncertainty in $\langle n_{sbj} \rangle$ coming from that of the parton distribution functions. This represents an advantage of this analysis for studying QCD.

5.3 Comparison between the NLO QCD predictions and the measurements

Once the NLO QCD predictions have been corrected to the hadron level, it is possible to compare them directly to the data.

Figure 5.6 shows the comparison of the predictions and the measurements of the mean subjet multiplicity for the entire sample as a function of y_{cut}. It should be noted that the NLO predictions give a very good description of the data even in the region where parton-to-hadron corrections are very large, which is indicated in the plot as the shaded region. On the other hand, leading order predictions are clearly below the measurements.

The plot also shows the predictions for different values of $\alpha_s(M_Z)$, which have been obtained in the same way we described in section 5.2.3.

Figure 5.7 shows the comparison of the NLO QCD predictions and the measurements for the mean subjet multiplicity at a given resolution scale ($y_{cut} = 10^{-2}$) as a function of the transverse energy of the jets. The agreement between data and NLO QCD is very good. The predictions using different values of $\alpha_s(M_Z)$ show that the measurements are sensitive to this value.

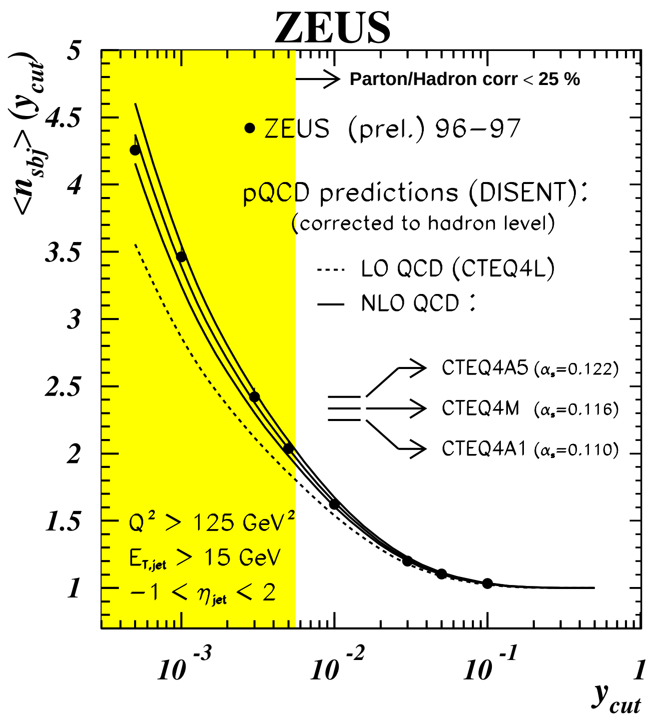


Figure 5.6: The measured mean subjet multiplicity corrected to the hadron level $\langle n_{sbj} \rangle$ as a function of y_{cut} for inclusive jet production in NC DIS (black dots). For comparison, NLO QCD calculations corrected for hadronization effects, obtained using three different CTEQ4 sets of the proton parton densities are shown (solid lines). LO QCD calculations corrected for hadronization effects, obtained with the CTEQ4L set of proton parton densities are also shown (dashed line).

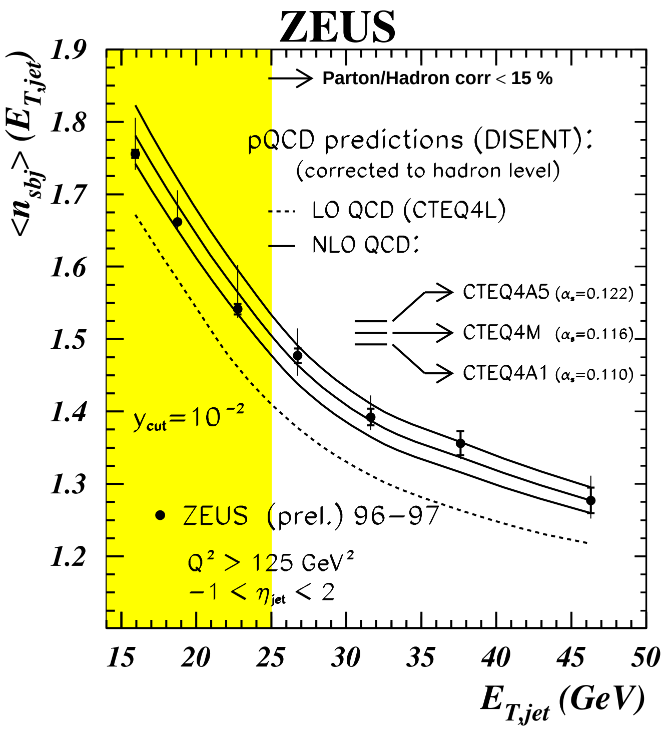


Figure 5.7: The measured mean subjet multiplicity corrected to the hadron level $\langle n_{sbj} \rangle$ as a function of $E_{T,jet}$ at $y_{cut} = 10^{-2}$ for inclusive jet production in NC DIS (black dots). The inner error bars represent the statistical uncertainties of the data. The outer error bars show the statistical and systematic uncertainties added in quadrature. Other details are as described in the caption to Fig. 5.6.

5.4 Extraction of $\alpha_s(M_Z)$

Given the good agreement between data and theory, and the sensitivity of the measurements to the value of α_s, a precise determination of α_s has been performed. In this section, the method is described as well as the result of the extraction.

5.4.1 Description of the method

As mentioned in section 5.2.3, the values of $\langle n_{snj} \rangle$ obtained using different values of α_s can be used to study the dependence of the predicted mean subjet multiplicity on the value of α_s. To make this analysis consistently, the sets of parton densities have to be comparable, which is the case if the same functional form and fitted procedure are assumed and the only change is the α_s value.

Using this method, the complete dependence of $\langle n_{snj} \rangle$ on the value of α_s is studied by taking into account the correlation between the value of α_s in the matrix element and that assumed in the determination of the parton densities. The dependence of $\langle n_{snj} \rangle$ on α_s has been parametrized by fitting the predicted values to the analytical expression

$$\langle n_{sbj} \rangle = \alpha_s(M_Z) \cdot a_1 + \alpha_s^2(M_Z) \cdot a_2 \; . \tag{5.4}$$

(a_1, a_2) are the parameters of the fit. This has been done for each value of y_{cut} in each $E_{T,jet}$ region. A value for α_s has been extracted by inverting Eq. (5.4) and using the measured values.

For the analysis, the high-$E_{T,jet}$ ($E_{T,jet} > 25$ GeV) and $y_{cut} = 10^{-2}$ has been considered as the most reliable to determine the value of α_S. The main criteria was the size of the hadronization correction, which suggests to avoid the region of low-$E_{T,jet}$ or very small y_{cut}. In addition, the final region was selected such that a large sensitivity on α_s was achieved.

To obtain the parametrization, the five sets of the CTEQ4A-series [32] with different values of α_s have been used. The parametrization mentioned above in Eq. (5.4) has been obtained for each measurement at $E_{T,jet} > 25$ GeV and for each data point a value of α_s has been determined.

Figure 5.8 shows the parametrizations and the way the measured values of $\langle n_{sbj} \rangle$ are translated into values of $\alpha_s(M_Z)$. At the same time, the statistical error of the measurement is propagated into the statistical contribution to the uncertainty of α_s. Figure 5.9 presents the four values of $\alpha_s(M_Z)$. It should be emphasized the good agreement among the four determinations, which are completely independent from the experimental point of view. The systematic uncertainties shown in the figure are discussed in the next section.

To obtain a combined value, an additional χ^2 fit is performed to the four values considering only the statistical uncertainties. The error obtained from the fit is taken as the statistical uncertainty of the final value of α_s. This value which is presented in figure 5.9, is

$$\alpha_s(M_Z) = 0.1185 \pm 0.0016 \,(\text{stat})^{+0.0067}_{-0.0048} \,(\text{syst.})^{+0.0089}_{-0.0071} \,(\text{th.}) \; .$$

where the procedure to obtain the systematic and theoretical uncertainties is discussed in the next section.

This value of $\alpha_s(M_Z)$ is in good agreement with the world average value, 0.1180 ± 0.0020 [98], and the uncertainty is competitive with other determinations of the strong coupling constant.

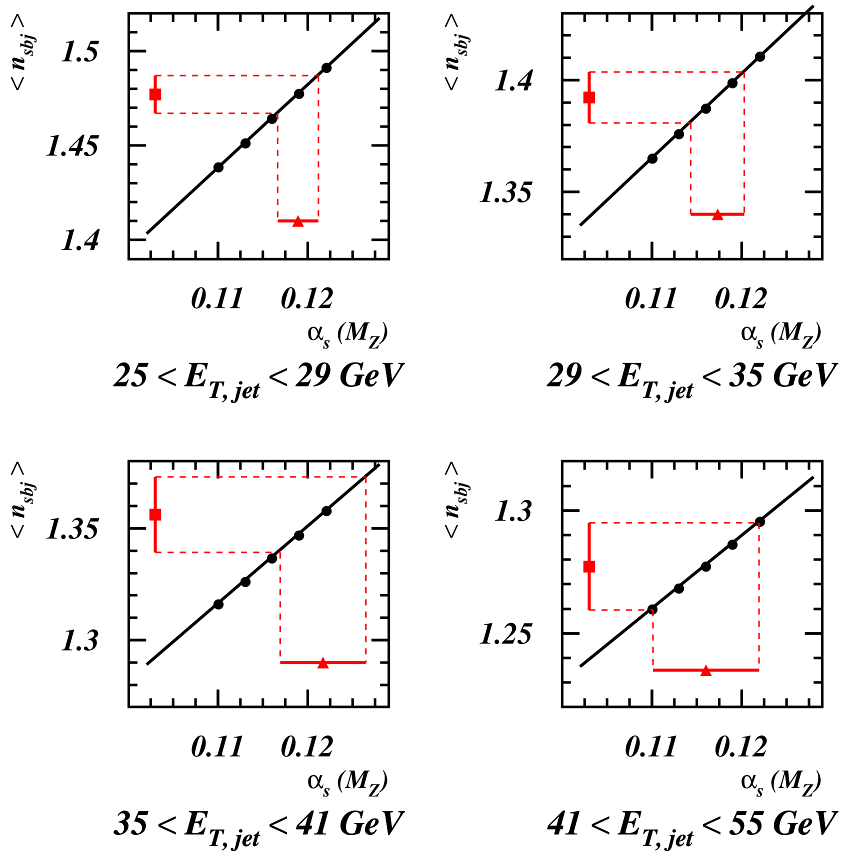


Figure 5.8: NLO QCD calculations of $\langle n_{sbj} \rangle$ for different values of $\alpha_s(M_Z)$ (black dots) in various $E_{T,jet}$ regions. The parametrization (solid lines) according to Eq. (5.4) are also shown. The measured $\langle n_{sbj} \rangle$ values (solid squares) are used to determine a value for α_s (solid triangles) in each plot.

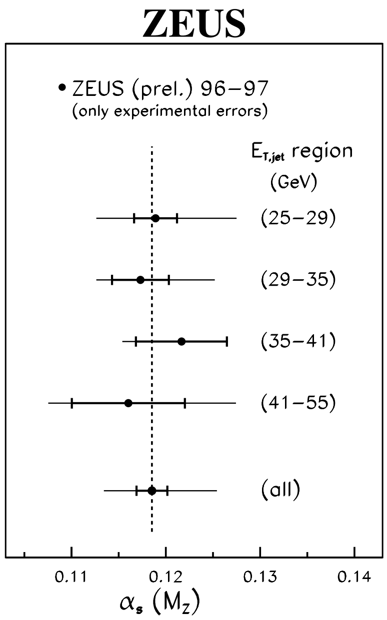


Figure 5.9: The $\alpha_s(M_Z)$ values determined from the QCD fit of the measured mean subjet multiplicity $\langle n_{sbj} \rangle$ at $y_{cut} = 10^{-2}$ in the different $E_{T,jet}$ regions (black dots). The combined value of $\alpha_s(M_Z)$ obtained by using all the $E_{T,jet}$ regions is shown (black dot at the bottom). The inner error bars represent the statistical error of the data. The outer error bars show the statistical and systematic uncertainties added in quadrature.

5.4.2 Systematic and theoretical uncertainties on α_s

The systematic and theoretical uncertainties are obtained by repeating the entire procedure, but by making the corresponding variation either in the measured value or the theoretical assumption.

All the sources of systematic uncertainty considered for the measurements are propagated in this way to the value of α_s. The most important contribution is the uncertainty on the energy scale of the jets and the use of LEPTO-MEPS instead of ARIADNE to correct the data for detector effects.

For the theoretical uncertainty, all the sources described in section 5.2.5 have been taken into account. However, only the contribution due to the hadronization correction and the dependence with the renormalization scale are considered in the final theoretical uncertainty. The reason is that other theoretical uncertainties are either negligible with respect to the two above or are considered as cross-checks and not as additional sources of uncertainty.

The theoretical uncertainties in the value of α_s are listed below

- Variation of the renormalization scale by a factor of 2:

$$\Delta\alpha_s(M_Z) = ^{+0.0088}_{-0.0070}$$

- The uncertainty on the modelling of the parton-to-hadron correction has been estimated by comparing the corrections with ARIADNE and LEPTO-MEPS and amounts to

$$\Delta\alpha_s(M_Z) = \pm 0.0015$$

- The whole procedure was repeated using those sets of parton densities of MRST99, central, $\alpha_s \uparrow\uparrow$ and $\alpha_s \downarrow\downarrow$. The difference in the resulting $\alpha_s(M_Z)$ value is less than 0.5%. This uncertainty was neglected.

- All the NLO QCD calculations were repeated using as renormalization scale the $E_{T,jet}$ of the highest-E_T jet in the event. The difference of the resulting $\alpha_s(M_Z)$ with respect to the central value is less than 0.1%. It was not considered as a contribution to the theoretical uncertainty because this variation is already included in the uncertainty coming from the variation of μ_R by a factor of 2.

It has become clear that terms beyond NLO are needed to reduce the theoretical uncertainty further. As shown in section 5.2.5, the theoretical uncertainty due to the missing higher orders in the perturbative series is by far the most important contribution to the uncertainty on α_s.

Chapter 6

Summary and conclusions

This document presents the description and discussion of an analysis of the jet internal structure using 38.6 pb^{-1} of data taken with the ZEUS detector at HERA. The analysis is performed by measuring the mean subjet multiplicity in an inclusive sample of jets produced in neutral current deep inelastic positron-proton scattering .

This observable has been studied as a function of the resolution scale y_{cut} as well as a function of the jet variables. The measurements have been compared with the next-to-leading order QCD predictions.

The main conclusions are:

- The measured mean subjet multiplicity has been observed to be consistent with that of a sample dominated by quark-initiated jets.

- No dependence on the mean subjet multiplicity with respect to the pseudorapidity of the jet is observed. This is consistent with the theoretical predictions, which predict that the sample is dominated by quark-initiated jets for the entire pseudorapidity range.

- The mean subjet multiplicity is observed to decrease as the transverse energy of the jet increases. This means that the jets become narrower as the transverse energy increases.

  Several sources contribute to this result. They include the running of α_s with the scale and the hadronization correction. This result is in good agreement with similar measurements done in other high-energy processes.

- NLO QCD predictions for $\langle n_{sbj} \rangle$ have been obtained. They give a very good description of the data. The leading order prediction is not sufficient to describe the data. It has been shown that hadronization corrections are under control if the scale to resolve the subjets and the transverse energy are large enough.

  This good description confirms the validity of QCD as the theory to describe the interaction between quarks and gluons.

From the comparison between the NLO QCD calculations and the data, a QCD analysis has been developed to extract the strong coupling constant. The value of $\alpha_s(M_Z)$ has been extracted in a region where the parton-to-hadron correction are small enough. The extracted value of $\alpha_s(M_Z)$ is

$$\alpha_s(M_Z) = 0.1185 \pm 0.0016\ (\mathrm{stat})^{+0.0067}_{-0.0048}\ (\mathrm{syst.})^{+0.0089}_{-0.0071}\ (\mathrm{th.})\ ,$$

which is in good aggrement with the world average value and the uncertainties are competitive with other determinations of this quantity.

Appendix A

The laboratory and Breit frames

The analysis presented in the previous chapters has been performed in the laboratory frame. In this reference system the proton and the positron collide head-on with energies $E_p = 820$ GeV and $E_e = 27.5$ GeV.

On the other hand, the final-state partons which are provided by the program DISENT (see section 5.2.2) have their four-momenta as seen in the Breit frame [14, 99]. In this frame, the proton collides head-on with the exchanged boson, which is purely space-like with 3-momentum $\mathbf{q} = (0, 0, -Q)$.

To obtain the NLO QCD predictions of $\langle n_{sbj} \rangle$ for jets in the laboratory frame, it is necesary to transform the four-momenta of the partons back to the laboratory frame. This is very simple if the exchanged boson is known in the laboratory frame. Since the kinematic variables (for example, x and Q^2) are Lorentz-invariant, the reconstruction of this four-momentum is straightforward, and the components are

$$\begin{aligned} q_0 &= y \cdot (E_e - x\, E_p) \\ q_1 &= \sqrt{Q^2 \cdot (1 - y)} \\ q_2 &= 0 \\ q_3 &= -y \cdot (E_e + x\, E_p) \ , \end{aligned}$$

where the second component is set to zero by convention. A rotation in the first two spatial components (here q_1 and q_2) does not modify the results.

With this four-momentum, the transformation from the Breit frame to the laboratory frame is performed using the following relation for any four-momentum p_B.

$$p_L = \mathcal{L}_{bl}\, p_B$$

where the matrix $\mathcal{L}_{bl}$ is the following 4x4 matrix,

$$\mathcal{L}_{bl} = \begin{bmatrix} \frac{q_0}{Q} + \frac{Q}{q_0 - q_3} & -\frac{q_T}{q_0 - q_3} & 0 & -\frac{q_0}{Q} \\ \frac{q_1}{Q} & -\frac{q_1}{q_T} & \frac{q_2}{q_T} & -\frac{q_1}{Q} \\ \frac{q_2}{Q} & -\frac{q_2}{q_T} & -\frac{q_1}{q_T} & -\frac{q_2}{Q} \\ \frac{q_3}{Q} + \frac{Q}{q_0 - q_3} & -\frac{q_T}{q_0 - q_3} & 0 & -\frac{q_3}{Q} \end{bmatrix}$$

where the first row and column are those related to the time component, $Q = \sqrt{Q^2}$ and $q_T = \sqrt{{q_2}^2 + {q_2}^2}$.
Thus, upon application of this matrix to the four-momenta of the final state partons provided by DISENT, the jets can be reconstructed in the laboratory frame.

Appendix B

Jet energy corrections

As discussed in section 4.1.5, it is necessary to correct the energy of the jets reconstructed from the UCAL for energy losses. A more complete description of the method is presented here.

The method consists of obtaining the corrected energy of the jet by using a correction factor given by the detector simulation and the Monte Carlo samples. This factor is defined in terms of the expression

$$E_{T,jet}(corr) = \frac{E_{T,jet}(rec) - b(\eta_{jet})}{m(\eta_{jet})}$$

where the parameters $b(\eta_{jet})$ and $m(\eta_{jet})$ contain the information given by the simulated events.

To obtain these parameters, the following steps are followed:

- The sample of jets is selected at the hadron level of the Monte Carlo with the complete set of cuts of the analysis. The jets are then matched with the jets found at detector level. The cuts and the method for matching the jets at both levels were described in section 4.1.5.

- The entire η_{jet} range is divided into bins bins. The matched jets are then separated according to the η_{jet} value for the jet at hadron level. For this analysis, 12 bins of 0.25 units each are considered. The parameters $b(\eta_{jet})$ and $m(\eta_{jet})$ are kept constant within each bin.

- Correlation plots are made between the jet transverse energy at detector and hadron levels. This correlation plot is done for each bin in η_{jet}. The correlation displays the mean value of the transverse energy of the jets at the detector level as a function of that of the jets at hadron level.

- The resulting $\langle E_{T,jet}(det)\rangle(E_{T,jet}(had))$ values are then fitted to obtain the values of the parameters $b(\eta_{jet})$ and $m(\eta_{jet})$ for each bin in η_{jet}

  $$E_{T,jet}(rec) = b(\eta_{jet}) + m(\eta_{jet}) \cdot E_{T,jet}(had) \ .$$

  The fitted parameters are used to correct the energy of the jets at detector level both in data and simulated events..

- Since the lowest $E_{T,jet}$ range is slightly biased by the cut $E_{T,jet}(det) > 10$ GeV used to preselect the jets at the detector level, the first bins in $E_{T,jet}$ are excluded in the fit.

- It should be noted that although the binning in η_{jet} is done using the jet at hadron level, no significant effect is expected if the η_{jet} reconstructed at the detector level is used instead since the good resolution in η_{jet} is very good.

Figure B.1 shows the correlation plots for the 12 bins in η_{jet} and the corresponding fit.

Figure B.2 shows the values of the two fitted parameters for the 12 bins in η_{jet}. It should be emphasized that the amount of dead material in front of the UCAL is clearly visible in this distribution: a larger correction (m differs more from unity) is displayed in the BCAL-FCAL boundary ($\eta_{jet} \sim 1.1$) and BCAL-RCAL boundary ($\eta_{jet} \sim -0.75$) The same figure shows the comparison (and good agreement) between the values of the fitted parameters obtained by using the two Monte Caro samples.

Finally, figure B.3 shows the values of the two parameters when the correlation plots are done after applying the energy corrections. The parameter b is close to zero while m is very close to one. This means a good correlation of the energies at hadron and at detector levels and shows that the correction method works.

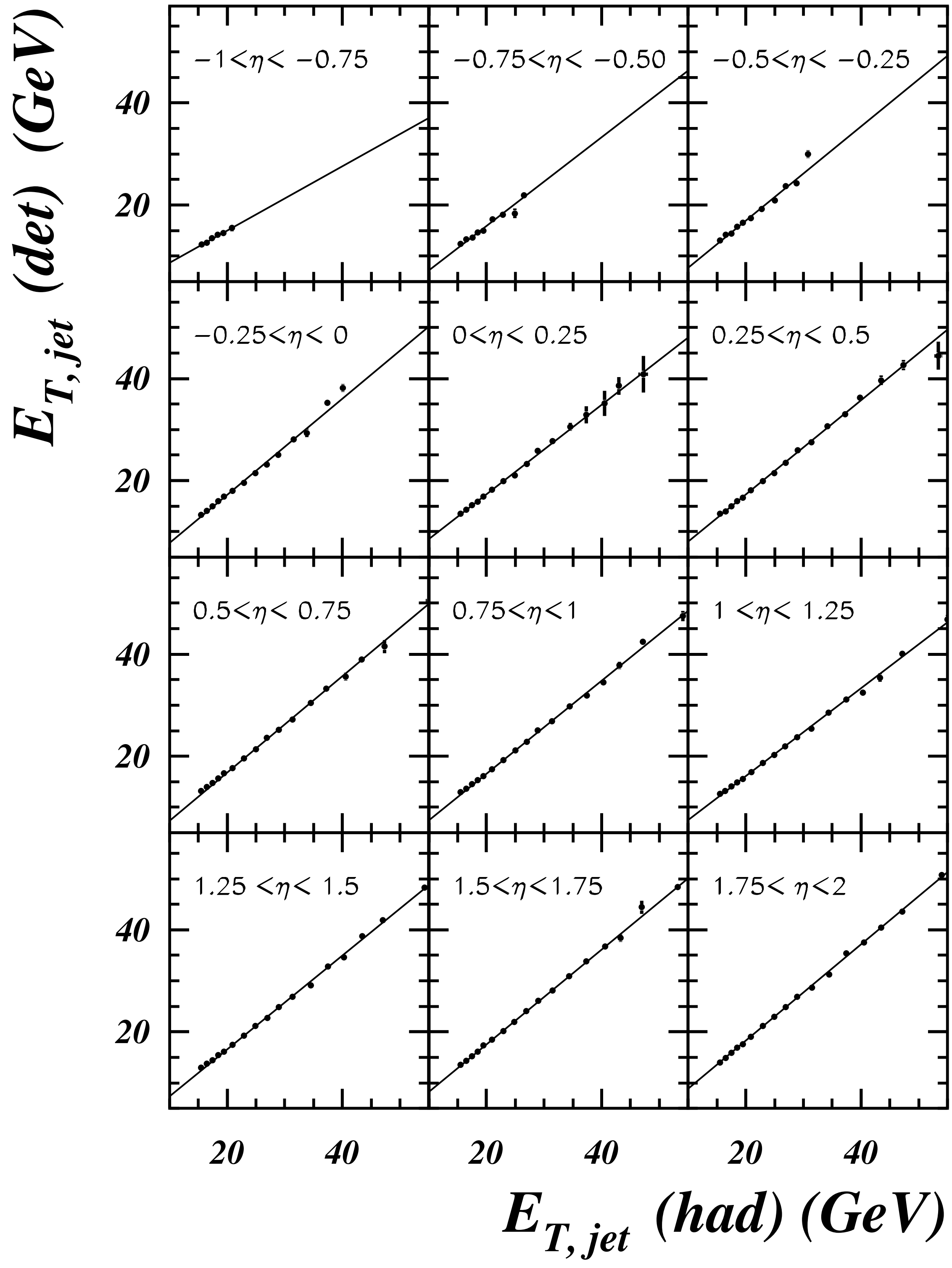


Figure B.1: $\langle E_{T,jet}(det)\rangle$ as a function of $E_{T,jet}(had)$ using the matched sample of jets in the ARIADNE sample. The lines show the results of the fits.

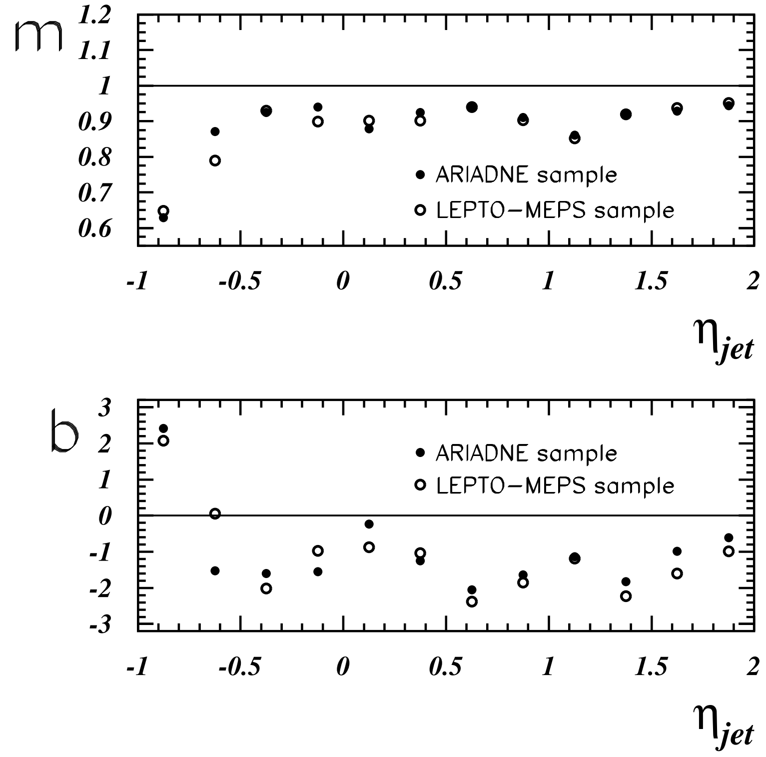


Figure B.2: The fitted parameters $m(\eta_{jet})$ (top) and $b(\eta_{jet})$ (bottom) as functions of η_{jet} as obtained from the ARIADNE (black dots) and LEPTO-MEPS (open circles) samples.

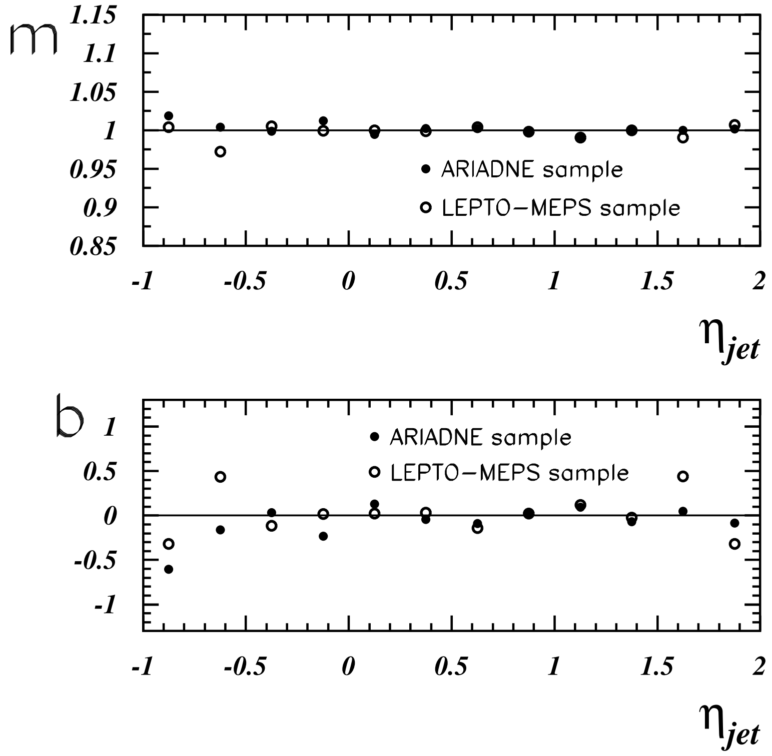


Figure B.3: The fitted parameters $m(\eta_{jet})$ (top) and $b(\eta_{jet})$ (bottom) as functions of η_{jet} as obtained from the ARIADNE (black dots) and LEPTO-MEPS (open circles) samples after the jet energy corrections have been applied.

Appendix C

Method to obtain samples of quark and gluon jets

To obtain the predictions for quark and gluon-initiated jets which were presented in section 5.1, separated samples of the two types are needed. This separation can be achieved using Monte Carlo samples by matching the jets at hadron level to the final-state partons in the event.

The following procedure has been developed as a criterion to classify a given jet as a quark or gluon-initiated jet. This procedure is applied to each jet in the event.

- The distance in the $\eta - \phi$ plane of each parton i to the jet is computed:

$$\Delta_i = \sqrt{(\eta_i - \eta_{jet})^2 + (\phi_i - \phi_{jet})^2}$$

- by using all the partons with $\Delta_i \leq R_0$, the following "flavoured" sum is computed:

$$F_{jet} = \sum_i f_i \ ,$$

  where f_i is $+1$ for quarks, -1 for antiquarks and 0 for any other type of partonic particle (gluons and diquarks). F_{jet} is the quantity used to extract the "flavour" of the jet.

- If F_{jet} is cero, the jet is classified as a gluon-initiated jet. If not, the jet is classified as a quark jet. It should be mentioned that depending on the sign, the jet could be classified as quark or antiquark-initiated jet, but this is not needed in the present analysis.

- If there is no parton satisfying the condition $\Delta_i \leq R_0$, then the jet is not classified.

This method provides very good results when compared with other procedures [100]. The dependence with the parameter R_0 allows to estimate the uncertainty in the predictions in a straightforward way. The value for the parameter in the present analysis was selected to be $R_0 = 0.7$ since the results are more stable for this choice. For the systematic uncertainty, radii of $R_0 = 0.5$ and 1.0 were used.

For a larger radius, a larger dependence of the result on gluon splitting into quark-antiquark pairs is observed[1]. On the other hand, the use of a smaller radius translates into a less efficient method for matching the jets.

For $R_0 = 0.7$, more than 99.8% of the jets were classified as quark or gluon-initiated jets. Out of them, 8.7% were identified as gluons and 91.3% as quarks. The uncertainties in these fractions are $\sim 3\%$ and have been estimated by varying the radius as mentioned before.

[1]The dependence on the presence of quark-antiquarks pairs from soft gluon splitting was estimated by applying a preselection of partons. From the complete list, quark-antiquark pairs of the same flavour were removed if they were close enough (distance in $\eta - \phi$ smaller than $R_0/2$). Since the difference using this preselection with respect to the standard one is smaller than the dependence with R_0, the preselection step was considered not to be necessary.

Bibliography

[1] O. González, *Determination of α_s using subjet multiplicities*, to be published in the Proceedings of the 9^{th} Iternational Workshop on Deep Inelastic Scattering (ed. G. Bruni, G. Iacobucci and R. Nania), Bologna, Italy, 27^{th} April - 1^{st} May, 2001.

[2] ZEUS Collab., *Measurements of jet substructure in NC DIS and determination of α_s at HERA*. Contributed paper to the International Europhysics Conference on High Energy Physics, Budapest, Hungary (2001). Abstract 641.

[3] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc. (1984).

[4] B.H. Bransden and C.J. Joachain, *Physics of atoms and molecules*. Addison Wesly Longman Limited (1983).

[5] Y. Ne'eman, Nucl. Phys. **26** (1961) 222;
M. Gell-Mann, Phys. Rev. **125** (1962) 1067.

[6] M. Gell-Mann, Phys. Lett. **8** (1964) 214;
G. Zweig, CERN-8192/TH 401 (1964); CERN-8419/TH 412 (1964).

[7] O.W. Greenberg, Phys. Rev. Lett. **13** (1964) 598;
M.Y. Han and Y. Nambu, Phys. Rev. **139** (1965) B1006.

[8] N. Cabibbo, Phys. Rev. Lett. **10** (1961) 531;
M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** (1973) 652.

[9] S. Glashow, Nucl. Phys. B**22** (1961) 579.

[10] S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264;
A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm, Almquist and Forl, Stockholm (1968).

[11] R.K. Ellis, W.J. Stirling and N.R. Webber, *QCD and Collider Physics*. Cambridge University Press (1996).

[12] P.W. Higgs, Phys. Lett. **12** (1964) 132; Phys. Rev. Lett. **13** (1964) 508; Phys. Rev. **145** (1966) 1156;
R. Brout and F. Englert, Phys. Rev. Lett. **13** (1964) 321.

[13] J.D. Bjorken, Phys. Rev. **179** (1969) 1547.

[14] R.P. Feynman, *Photon Hadron Interactions*, W.A. Benjamin, New York (1972).

[15] SLAC-MIT collaboration, E.D. Bloom et al., Phys. Rev. Lett. **23** (1969) 930;
SLAC-MIT collaboration, E.D. Bloom et al., Phys. Rev. Lett. **23** (1969) 935;
SLAC-MIT collaboration, G. Miller et al., Phys. Rev. D**5** (1972) 528.

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

[17] C.G. Callan and D. Gross, Phys. Rev. Lett. **22** (1969) 156.

[18] Gargamelle Neutrino Collab., H. Deden et al., Nucl. Phys. B**85** (1975) 269.

[19] TASSO Collab., R. Bradelick et al., Phys. Lett. B**86** (1979) 243.

[20] A. Quadt, *Measurement and QCD Analysis of the Proton Structure Function F_2 from the 1994 HERA Data using the ZEUS Detector.* Doctoral Thesis. Christ Church, University of Oxford (1996).

[21] ZEUS Collab., M. Derrick et al., Phys. Lett. B**316** (1993) 412;
H1 Collab., I. Abt et al., Nucl. Phys. B**407** (1993) 515.

[22] ZEUS Collab., S. Chekanov et al., DESY-01-064, accepted by Eur. Phys. J. C

[23] G. Parisi, *Proc.* 11^{th} *Rencontre de Moriond*, J. Tran Tranhm (ed.), ed. Fronties (1976) 83;
G. Altarelli and G. Parisi, Nucl. Phys. B**126** (1977) 298;
G. Altarelli, Nucl. Phys. B**81** (1981) 1.

[24] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438;
L.N. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 96;
Y.L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641.

[25] Y. Balitzki and L.N. Lipatov, Phys. Rev. Lett. **28** (1978) 822;
E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Phys. Rev. Lett. **44** (1976) 443; Phys. Rev. Lett. **45** (1977) 199.

[26] M. Ciafaloni, Nucl. Phys. B**296** (1988) 49;
S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. B**234** (1990) 339; Nucl. Phys. B**336** (1990) 18.

[27] J. Kwiecinski, A.D. Martin and P.J. Sutton, Phys. Rev. D**52** (1995) 1445.

[28] M.G. Ryskin and E.M. Levin, Phys. Rep. **189** (1990) 267;
A.H. Mueller, Nucl. Phys. B**335** (1990) 115.

[29] L.V. Gribov, E.M. Levin and M.G. Ryskin, Phys. Rep. **100** (1983) 1; Nucl. Phys. B**188** (1981) 155;
A.H. Mueller and J. Qiu, NPB**260** (1986) 427.

[30] A.D. Martin et al., Eur. Phys. J. C**4** (1998) 463.

[31] A.D. Martin et al., Eur. Phys. J. C**14** (2000) 133.

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

[33] M. Glück, E. Reya and A. Vogt, Z. Phys. C**49** (1991) 501;
M. Glück, E. Reya, Dortmund DO-TH 93/27 (1993);
M. Glück, E. Reya and A. Vogt, Phys. Lett. B**306** (1993) 391;
M. Glück, E. Reya and A. Vogt, Z. Phys. C**67** (1995) 433; Eur. Phys. J.**5** (1998) 461.

[34] M. Milite, *The Internal Structure of Charmed Jets in Photoproduction at HERA*. Doctoral Thesis, University of Hamburg (2001).

[35] G. Sterman and S. Weinberg, Phys. Rev. Lett. **39** (1977) 1436.

[36] JADE Collab., W. Bartel et al., Z. Phys. C**33** (1986) 23;
JADE Collab., S. Bethe et al., Phys. Lett. B**213** (1988) 235.

[37] S. Catani, Yu.L. Dokshitzer and B.R. Webber, Phys. Lett. B**285** (1992) 291;
S. Catani et al., Nucl. Phys. B**406** (1993) 187.

[38] UA1 Collab., G. Arnison et al., Phys. Lett. B**123** (1983) 115.

[39] S.D. Ellis and D.E. Soper, Phys. Rev. D**48** (1993) 3166.

[40] J. Huth et al. *Proc. of the 1990 DPF Summer Study on High Energy Physics*, Snowmass, Colorado. E.L. Berger (ed.), World Scientific, Singapore (1992) p. 134.

[41] OPAL Collab., G. Abbiendi et al., Eur. Phys. J. C**17** (2000) 373;
ALEPH Collab., D. Buskulic et al., Z. Phys. C**76** (1997) 1991.

[42] *HERA: A Proposal for a Large Electron Proton Colliding Beam Facility at DESY* (1981), DESY-HERA-81/10.

[43] ZEUS Collaboration, "The ZEUS Detector", Status Report 1993, U. Holm (ed.), DESY (1993) (unpublished).
Available on http://www-zeus.desy.de/bluebook/bluebook.html.

[44] C.B. Brooks et al. (ZEUS-UK Collab.), *Development of the ZEUS central tracking detector*. Nucl. Instr. and Meth. A**283** (1989) 477;
B. Foster et al., *The design and construction of the ZEUS central tracking detector*. Nucl. Instr. and Meth. A**338** (1994) 254.

[45] O. Deppe, *Measurement of $D^{*\pm}$ Electroproduction at HERA*. Doctoral Thesis. DESY-THESIS-2000-006 (2000).

[46] J.A. Blissett et al., *Development of a tapered cell drift chamber II: Operation of high magnetic fields and different gas pressures*. Nucl. Instr. and Meth. A**245** (1986) 291.

[47] R. Hall-Wilton et al., *The CTD tracking resolution*. ZEUS-Note 99-024 (1999)

[48] I. Redondo, *Charm electroproduction at HERA*. Doctoral Thesis, U. Autónoma de Madrid (2001).

[49] W.R. Leo, *Techniques for Nuclear and Particle Physics Experiments*. Springer-Verlag, New York (1997).

[50] D.H. Perkins, *Introduction to High Energy Physics*. Addison-Wesley (1982).

[51] E. Lohrmann, *Hochenergiephysik.* Teubner (1990).

[52] U. Amaldi, *Fluctuations in Calorimetry Measurements.* Physica Scripta **23** (1981) 409-424.

[53] Y. Iga and S. Schlendstedt, *Status and Improvement of the ZEUS Calorimeter Simulation.* ZEUS-Note 97-019 (1997).
C. Catterall and Y. Iga, *Improved Simulation of Hadronic Cascades in the Uranium Calorimeter.* ZEUS-Note 98-077 (1998).

[54] T. Woeniger, *Influence of Calorimeter Calibration Errors on the Measurement of Deep Inelastic Scattering.* Doctoral Thesis, University of Hamburg (1991).
I. Bohnet, *Investigations on Uniformity, Radiation Resistance and Calibration of ZEUS Calorimeter Components at HERA.* Doctoral Thesis. DESY-THESIS-1999-041 (1999).

[55] C.W. Fabjan, *Calorimetry in High Energy Physics.* CERN-EP/85-54.

[56] A. Fürtjes, *Kalibration des ZEUS Vorwärts- und Rückwärtskalorimeters.* Doctoral Thesis, University of Hamburg. DESY-THESIS-93-03 (1992).

[57] C. Fabjan et al., Nucl Instr. Meth. **141** (1977) 61-80.
H. Brückman, U. Behrens, B. Anders, *Hadron Sampling Calorimetry - A Puzzle of Physics.* DESY 86-155 (1986).

[58] E. Bernardi et al. (ZEUS Collab.), Nucl. Instr. and Meth. A**262** (1987) 229;
G. García López, *Two Different Cases of Calorimetry in High Energy Physics: the ATLAS Liquid Argon Electromagnetic End Cap and the ZEUS Forward Plug Calorimeter.* Doctoral Thesis, U. Autónoma de Madrid. DESY-THESIS-2000-10 (2000).

[59] J. Straver, *Design, Construction and Beam Test of the High Resolution Uranium Scintillator Calorimeter for ZEUS.* Doctoral Thesis (1991)

[60] J. Krüger, *The Uranium Scintillator Calorimeter for the ZEUS Detector at the Electron-Proton Collider HERA.* DESY-Report, DESY F35-92-02 (1992).

[61] M. Derrick et al., Nucl. Instr. and Meth. A309 (1991) 77;
A. Andresen et al., Nucl. Instr. and Meth. A309 (1991) 101;
A. Caldwell et al., Nucl. Instr. and Meth. A321 (1992) 356;
A. Bernstein et al., Nucl. Instr. and Meth. A336 (1993) 23.

[62] L. Hervás, *The Pipelined Readout for the ZEUS Calorimeter.* Doctoral Thesis, U. Autónoma de Madrid (1991). DESY F35D-91-01.

[63] H. Bethe and W. Heitler, Proc. Roy. Soc. A**146** (1934) 83.

[64] M. van der Horst, Doctoral Thesis, University of Amsterdam (1991).

[65] A. Bamberger et al., *The Small Angle Rear Track Detector for ZEUS.* ZEUS-Note 95-095 (1995).

[66] W.H. Smith et al., Nucl. Instr. and Meth. A**335** (1995) 278-294.

[67] L.A.T. Bauerdick et al. *The Physics Analysis Enviroment of the ZEUS Experiment.* DESY 95-236 (1995).

[68] R. Brun et al., CERN DD/EE/81-1 (1987).

[69] F. Jacquet and A. Blondel, *Proceedings of the study for an ep facility for Europe*. DESY 79-48 (1979) p.391.

[70] S. Bentvelsen, J. Engelen, P. Kooijman, *Proceedings of the 1991 Workshop on Physic at HERA*, Vol. 1, p.23. W. Buchmüller and G. Ingelman (ed.) (DESY, Hamburg, 1992).

[71] U. Bassler, G. Bernardi, Nucl. Instr. and Meth. A**361** (1995) 197.

[72] J.C. Collins, D.E. Soper and G. Sterman, Nucl. Phys. B**308** (1988) 833.

[73] M. Bengtsson and T. Sjöstrand, Phys. Lett. B**185** (1987) 435; Nucl. Phys. B**289** (1987) 810.

[74] Y. Azimov et al., Phys. Lett. B**165** (1985) 147;
G. Gustafson, Phys. Lett. B**175** (1986) 453;
G. Gustafson and U. Petersson, Nucl. Phys. B**306** (1988) 746;
B. Andersson et al., Z. Phys. C**43** (1989) 625.

[75] A. Petersen et al., Phys. Rev. D**37** (1988) 1;
OPAL Collab., M.Z. Akrawy et al., Phys. Lett. B**246** (1990) 285.

[76] V.V. Sudakov, Zh.E.F.T. **30** (1956) 87; Sov. Phys. JETP **30** (1956) 65.

[77] A.H. Mueller, Phys. Lett. B**104** (1981) 161;
B.I. Ermolaev and V.S. Fadin, JETP Lett.**33** (1981) 269;
A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Rep. **100** (1983) 201.

[78] R.D. Field and R.P. Feynman, Phys. Rev. D**15** (1977) 2590; Nucl. Phys. B**138** (1978) 1.

[79] B. Andersson, G. Gustafson and B. Söderberg, Z. Phys. C**20** (1983) 317; Nucl. Phys. B**264** (1986) 29;
B. Andersson et al., Phys. Rep. **97** (1983) 33.

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

[81] D. Amati and G. Veneziano, Phys. Lett. B**83** (1979) 87;
A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Lett. B**83** (1979) 207;
G. Marchesini, L. Trentadue and G. Veneziano, Nucl. Phys. B**181** (1980) 335.

[82] R.D. Field and S. Wolfram, Nucl. Phys. B**213** (1983) 65.

[83] G. Ingelman, A. Edin and J. Rathsman, Comp. Phys. Comm. **101** (1997) 108; version used: 6.5.

[84] H. Plothow-Besch, Comput. Phys. Commun. **75** (1993) 396; versions used: 7.08 and 7.09.

[85] K. Kwiatkowski, H. Spiesberger and H.-J. Möhring, Comp. Phys. Comm. **69** (1992) 155;
H. Spiesberger, *HERACLES - An Event Generator for ep Interactions at HERA Including Radiative Processes (Version 4.6)*, 1996, available on `http://www.desy.de/~hspiesb/heracles.html`

[86] K. Charchula, G.A. Schuler and H. Spiesberger, Comp. Phys. Comm. **81** (1994) 381;
H. Spiesberger, *DJANGO6 version 2.4 - A Monte Carlo Generator for Deep Inelastic Lepton Proton Scattering Including QED and QCD Radiative effects*, 1996, available on `http://www.desy.de/~hspiesb/django6.html`

[87] T. Sjöstrand, Comp. Phys. Comm. **39** (1986) 347;
T. Sjöstrand and M. Bengtsson, Comp. Phys. Comm. **43** (1987) 367; version used: 7.4.

[88] L. Lönnblad, Comp. Phys. Comm. **71** (1992) 15; Z. Phys. C**43** (1995) 285; version used: 4.08.

[89] G. Marchesini et al., Comp. Phys. Comm. **67** (1992) 465; hep-ph/9607939; version used: 5.9.

[90] H. Abramowicz, A. Caldwell and R. Sinkus, *Neural network Based Electron Identification in the ZEUS Calorimeter*, DESY 95-054;
R. Sinkus, Nucl. Instr. and Meth. A**361** (1995) 290.

[91] J. Edmonds, J. Grosse-Knetter and A. Quadt, *Neutral Current Cross Sections at High* Q^2, ZEUS-Note 99-016 (1999).

[92] ZEUS Collab. *Contributed paper to ICHEP'99 (Tampere, July 1999) N-530.*

[93] D0 Collab., B. Abbott et al., Nucl. Phys. Proc. Suppl.**79** (1999) 494.

[94] ALEPH Collab., D. Buskulic et al., Phys. Lett. B**346** (1995) 389.

[95] S. Catani and M.H. Seymour, Nucl. Phys. B**485** (1997) 291.

[96] S. Catani and M.H. Seymour Phys. Lett. B**378** (1996) 287.

[97] R.K. Ellis, D.A. Ross and A.E. Terrano, Nucl. Phys. B**178** (1981) 421.

[98] D.E. Groom et al., Eur. Phys. J. C**15** (2000) 1.

[99] K.H. Streng, T.F. Walsh and P.M. Zerwas, Z. Phys. C**2** (1979) 237.

[100] E. Rodrigues, ZEUS-Note in preparation.

Acknowledgments

The analysis presented here would not have been performed without the work of the HERA machine group and the people who constructed the ZEUS detector and make possible the data taking.

I would like to mention that the analysis was developed with the essential help of different people of the ZEUS collaboration, specially those belonging to the Hadronic Final States (now QCD) Working Group. I thank them all for their comments and useful discussions, and even for their own work, from which my analysis has benefited a lot, mainly in the use of programs and methods which were used and developed before I have need of them.

Finally, I would like to thank the HEP group of the Universidad Autónoma de Madrid to give me the opportunity of joining in the efforts of the ZEUS experiment and analyzing the data taken in a high-energy collider. Specially, I wish to emphasize the anormous support and guiding I receive from Juan Terrón, my thesis supervisor. Without his guidance and teaching, this work probably would never have reached the end or even started at all. I would like to mention also Ignacio Redondo for his comments and suggestions and his help to solve a lot of important tasks in DESY and Hamburg.