Diffractive Dijet Production in Deep Inelastic Scattering at ZEUS

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

Alessio Bonato

aus Torino

Hamburg
2008

Abstract

This thesis presents a measurement of dijet production in diffractive deep inelastic scattering ep collisions. This type of process is specially relevant for the experimental validity of the perturbative QCD approach to diffractive physics. The measurement was based on an integrated luminosity of 61 pb^{-1} collected at the HERA collider with the ZEUS experiment. The events were selected for virtualities of the photon, γ^*, $5 < Q^2 < 100$ GeV^2, and energies of the $\gamma^* p$ centre-of-mass, $100 < W < 250$ GeV. The jets were reconstructed from energy flow objects using the inclusive longitudinally-invariant k_T algorithm in the $\gamma^* p$ frame. The jets were required to have a transverse energy in the $\gamma^* p$ frame $E^*_{\mathrm{T,jet}} > 4$ GeV. The jet with the highest transverse energy was required to have $E^*_{\mathrm{T,jet}} > 5$ GeV. All jets were required to be in the pseudorapidity range $-3.5 < \eta^*_{\mathrm{jet}} < 0$ as measured in the $\gamma^* p$ frame. The selection of diffractive events was carried out by requiring a large rapidity gap in the direction of the scattered proton. The value of the fraction of initial proton momentum entering in the hard process, $x_{I\!\!P}$, was required to be $x_{I\!\!P} < 0.03$. The total cross section for the process was measured to be

$$\sigma^D_{\mathrm{TOT}}(ep \rightarrow ep\, \mathrm{jet}_1\, \mathrm{jet}_2\, X') = 91.5 \pm 1.2\ (stat.)\ ^{+3.3}_{-5.4}\ (syst.)\ ^{+6.4}_{-5.3}\ (corr.)\ \mathrm{pb}$$

Single and double differential cross sections were extracted and compared to leading-order predictions and next-to-leading-order QCD calculations. The latter used several diffractive parton densities extracted from inclusive diffractive deep inelastic scattering data. The agreement with the leading and next-to-leading order predictions is good and no hints of factorisation breaking are observed. The double differential measurement can be a precious input for the extraction of more accurate diffractive parton densities.

Kurzfassung

In dieser Arbeit wird eine Messung von Zwei-Jet Produktion in diffraktiver tief-inelastische Streuung vorgestellt. Diese Art von Prozess ist insbesondere für die experimentelle Überprufung des perturbativen QCD-Ansatzes für diffraktive Physik von Bedeutung. Die Messung basiert auf vom ZEUS-Detektor bei HERA aufgezeichneten Daten mit einer integrierten Luminosität von 61 pb^{-1}. Es wurden solche Ereignisse selektiert, in denen das Photon, γ^*, eine Virtualität, Q^2, von $5 < Q^2 < 100$ GeV^2 sowie eine Schwerpunktsenergie von $100 < W < 250$ GeV aufweist. Die Jets wurden aus Energieflussobjekten mit Hilfe des logitudinal-invarianten k_T-Algorithmuses im $\gamma^* p$-Bezugssystem rekonstruiert. Weiter wurde verlangt, dass die Jets eine im $\gamma^* p$-System gemessene Transversal-Energie von $E^*_{\mathrm{T,jet}} > 4$ GeV haben und sich im Pseudorapiditäts-Bereich von $-3.5 < \eta^*_{\mathrm{jet}} < 0$ befinden. Der Jet mit der höchsten Transversal-Energie musste zudem $E^*_{\mathrm{T,jet}} > 5$ GeV erfüllen. Es wurden solche Ereignisse als diffraktive Ereignisse angenommen, die eine grosse Rapiditätslücke in Richtung des gestreuten Protons aufwiesen. Für den Impulsbruchteil des Protons, der in die harte Streuung eingeht, $x_{I\!\!P}$, musste $x_{I\!\!P} < 0.03$ gelten. Der gemessene totale Wirkungsquerschnitt für den Prozess ist

$$\sigma^D_{\mathrm{TOT}}(ep \rightarrow ep\, \mathrm{jet}_1\, \mathrm{jet}_2\, X') = 91.5 \pm 1.2\ (stat.)\ ^{+3.3}_{-5.4}\ (syst.)\ ^{+6.4}_{-5.3}\ (corr.)\ \mathrm{pb}$$

Einzel- und doppeltdifferenzielle Wirkungsquerschnitte wurden bestimmt und mit Vorhersagen von QCD-Störungsrechnungen führender und nächst-führender Ordnung verglichen. In die Rechnungen nächst-führender Ordnung flossen diffraktive Parton-Verteilungsdichten ein, die aus Daten von tief inelastischer Streuung extrahiert worden sind. Die Vorhersagen führender und nächst-führender Ordnung stimmen gut überein und zeigen keinerlei Anzeichen vom Zusammenbruch der Faktorisierung. Die doppelt-differenziell gemessenen Wirkungsquerschnitte können benutzt werden, um diffraktive Parton-Verteilungsdichten mit höherer Genauigkeit zu bestimmen.

Contents

Chapter 1

Introduction

The Standard Model (SM) of particle physics is to date the most successful model able to describe the phenomena in the subatomic regime. Quantum ChromoDynamics (QCD) is the part of the SM that describes the strong interaction, the force responsible for the existence of hadrons and nuclei. QCD is a gauge theory with the two prominent features of asymptotic freedom (the strength of the interaction vanishes at short distances) and colour confinement (only particles neutral to the strong interaction can propagate freely at large distances). In the last decades, huge improvements in both theoretical and experimental understanding of QCD have been achieved. Many of them were obtained in the context of Deep Inelastic Scattering (DIS) experiments where a lepton collides against a hadron (typically a proton).

A particular subsample of reactions driven by the strong interaction is diffraction, when no quantum numbers are exchanged between the colliding particles. As a consequence either one or both the incoming particles might emerge intact from the interaction. Other striking experimental signatures are observed, like the presence of Large Rapidity Gaps (LRG), angular regions of the detector without hadronic activity. The study of diffraction was for a long time relegated outside the QCD context because of its intrinsic non-perturbative, large distance (i.e. small scale) nature. The observation of diffractive events in presence of a hard scale (hard diffraction) changed this picture, giving the idea to approach the study of diffraction in the framework of QCD. Diffractive DIS (DDIS) is an example of a process where in a soft, diffractive phenomenon, a hard scale is present, namely the virtuality of the exchanged boson, Q^2. By studying hard diffraction one has the opportunity to extend the understanding of the strong interaction, using the hard scale as a window on the soft regime which otherwise would not be accessible within perturbative QCD.

A significant progress in the task of providing a QCD motivated description of hard diffractive processes was the theoretical proof of the factorisation theorem for DDIS. This theorem states that diffractive process can be factorised into a short- and a long-distance part. The short-range part is the one containing the hard scale and is calculable by means of pQCD. The long-distance part consists of the diffractive parton densities (PDFs) that cannot be calculated a priori by perturbative QCD but once extracted can be used in calculations for other hard diffractive processes (universality of the dPDFs). The basic relevance of this theorem is the possibility of a QCD inter-

pretation of diffraction. This is the basis (alas not the solution) for inserting diffraction in the QCD framework. Activities are ongoing in proving experimentally the validity of this theorem. In the perspective of a QCD interpretation of diffraction, data from ep collisions at the HERA accelerator have a twofold importance. First, they can be used for extracting the dPDFs in a similar fashion to the proton PDFs. Second, HERA data can be used to check the factorisation theorem. It is important to note that the proof of factorisation only holds for DDIS but not for hadron-hadron scattering. Indeed, data from $p\bar{p}$ collisions showed a breaking of the factorisation explained by means of secondary soft interactions.

The study of dijets in diffractive DIS is an analysis well-suited for these goals. In fact the production of dijets simultaneously to the DIS process guarantees the presence of two hard scales, Q^2 and the transverse energy of the jets, E_T. Moreover, the production mechanism of dijets is very sensitive to the gluon content of the dPDFs. It was shown experimentally that gluons contribute to a large extent to the latter, thus the dijets data give a direct handle on dPDFs. Compared to other diffractive final states with similar characteristics, like the diffractive production of heavy quarks, diffractive jets benefit from higher statistics and harder scales. The study of dijets in DDIS can be useful for the same two reasons mentioned above: it is a stringent benchmark for the factorisation theorem and can be used to constrain the dPDFs in kinematic regions where the inclusive DDIS data have no sensitivity.

Experimentally, the main challenge of this analysis is the selection of diffractive events which is decreased in statistics compared to the standard dijet production. The diffractive selection is carried out by requiring the presence of a LRG in the direction of the scattered proton. This requires a very good knowledge of the detector and a careful simulation of the hadronisation process. Since the latter is not well known because of the non-perturbative nature of this part of the process, experimental techniques not too sensitive to it have to be considered. In order to increase the statistical significance of the sample, the thresholds of the jet selection were lowered. The reliable working of the jet clustering algorithm at low transverse energies has to be checked as well as the resolution of the detector. The uncertainties related to the calorimeter tend to increase as the energy of the jets decreases, thus experimental methods able to care for that are needed.

The measured data are compared to Next-to-Leading Order (NLO) QCD calculations using the dPDFs extracted from the inclusive DDIS data. The NLO calculation should describe the data both in shape and normalisation if the factorisation theorem is correct. At ZEUS, such NLO calculation was never performed for dijets in DDIS and no dedicated programs exist. Thus, the existent programs for QCD calculation in the standard, non-diffractive case were adapted to the needs of this analysis. Several dPDFs were employed in the calculation. The dijet data can be used as a benchmark for discriminating between the different dPDFs sets and selecting the ones better describing the data. The same calculation tools developed in the analysis can then be used to include the diffractive dijets data in a combined fit to the dPDFs together with the inclusive ones. Recent analyses demonstrated the large positive impact on the dPDFs accuracy on the inclusion of dijet data.

The thesis is organised as follows: Chapter 2 gives an overview on the theoretical

framework of the analysis, introducing first basic QCD concepts and then diffractive physics in particular in the framework of *ep* collisions. Chapter 3 describes the experimental device used, the ZEUS detector at the HERA collider. Chapter 4 explains how the physical quantities needed for the analysis were reconstructed from the detector output. Chapter 5 describes the detector simulation and the Monte Carlo samples used to correct the data for detector acceptance and resolution in addition to the NLO calculations compared to the measurement. Chapter 6 explains how the final sample of dijets in DDIS was extracted from the initial sample. Chapter 7 presents the results of the analysis: the extracted cross sections are compared to the LO Monte Carlo and to the NLO predictions. Double differential data useful for future fits to the dPDFs are also presented and discussed. Finally, the summary of the analysis and an outlook on possible future developments are given in Chapter 8.

Chapter 2

Theoretical framework

This chapter provides the basic theoretical knowledge needed for the rest of the thesis. The first part is a general introduction to Quantum Chromo Dynamics (QCD), the theory of strong interactions, and other important concepts like the factorisation theorem, the parton distribution function (PDFs) and the scale evolution of the PDFs. The theory of Deep Inelastic Scattering (DIS) and jet physics is also described: this has been one of the most important testing ground of QCD. The second part of the chapter is focused on diffractive physics. After the description of the typical features of the diffractive phenomena, the theoretical motivation for a perturbative QCD description of these processes is presented. Key issues of the latter are the factorisation theorem for diffraction and the diffractive PDFs (dPDFs). The understanding of these two subjects can be significantly improved by studying the production of dijets in diffraction, as explained in the last part of the chapter.

2.1   Quantum Chromo Dynamics

The strong force is one of the fundamental interactions acting in nature. It is responsible for the existence of all hadrons and, as consequence, of the nuclei that compose the universe. In the Standard Model [1, 2, 3] it is described by QCD. The main aspects of the theory are presented, together with the experimental tests of major interest for this thesis, i.e. DIS and jet physics.

2.1.1   The Standard Model

Four fundamental interactions are observed in Nature: the strong, the electromagnetic, the weak and the gravitational. The Standard Model (SM) is a field theory that provides at present the best understanding that we have of the former three fundamental interaction. A mechanism to include gravitation in the SM is still missing. The SM lagrangian is invariant under $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge rotations. The interactions between the particles are mediated by the exchange of vector bosons. There are four different types of vector bosons: gluons (carrier the strong force), photons (carrier the electromagnetic force), W and Z (both mediators of the weak force).
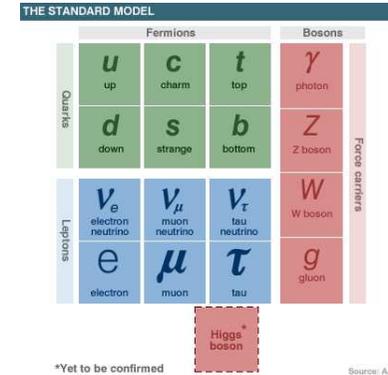
Figure 2.1: The elementary particles in the Standard Model.

There are 12 fundamental fermions, 6 leptons and 6 quarks. They are grouped in families, 3 for the leptons and 3 for the quarks (see Fig. 2.1). Only the quarks carry the strong interaction charge, called *colour*, and therefore they are the only ones subject to this force [4].

An additional scalar boson, the Higgs boson [5], is required by the SM in order to give mass to the vector bosons and the fermions. It is the last undetected particle of the SM and the search for it is one of the main goals of the experiments at the Large Hadron Collider (LHC) at CERN [6].

The actual status of the SM is debated: on one hand every experimental test supported its predictions at an excellent degree of accuracy. On the other hand it is known from theoretical arguments that the SM is an incomplete theory with some clear flaws, like the hierarchy problem. It is expected that new physics beyond the SM will manifest at the high and completely new energies probed at the LHC. Many different hypotheses on extensions of the SM have been proposed for solving the theoretically unpleasing aspects of the latter, among them supersymmetry, technicolor and large extra-dimensions.

2.1.2   The Quark-Parton Model

The concept of more fundamental constituents of the proton and the other hadrons arose in the 1960's with the development of hadron spectroscopy. The Quark-Parton Model (QPM) is an attempt to reduce the complexity of the observed large number of hadrons by introducing more fundamental constituents called *partons* [7]. The electrically charged partons are called *quarks* while the neutral ones are the *gluons*. The latter are the carriers of the strong force and will be introduced in Sect. 2.1.4. Quarks are point-like fermions that come in different types (*flavours*). At the time of their introduction, only three different types of quarks were considered: the up quark, u,

the down quark, d and the strange quark, s. Nowadays six different quarks have been experimentally found. Each flavour has its own mass and electric charge. The only possible values of the latter being either $\frac{2}{3}$ or $-\frac{1}{3}$. Hadrons are grouped into *baryons* and *mesons* containing three quarks or a quark-antiquark pair, respectively. The quarks need to have also an additional quantum number, otherwise hadrons containing quarks of the same type would violate the Pauli exclusion principle. This quantum number is called *colour charge*. It can have three different values, defined as red, green and blue. In the QPM, the proton is made only of three quarks, two up and one down. They are called *valence* quarks and they define the proton quantum numbers and share its total momentum.

2.1.3   Deep Inelastic Scattering

The QPM was first experimentally validated with experiments where a lepton, l, collides against a hadron or a nucleus, N, producing in the final state a scattered lepton (not necessarily the incoming one), l', and a hadronic final state, X [8]. DIS experiments are experiments of the kind $lN \to l'$X in the kinematic range of large four-momentum transferred between lepton and hadron and large invariant mass of the hadronic system produced in the collision. In a reference frame where the proton travels with very high momentum, the typical time of the interaction between the quarks and an external probe is much shorter than the typical time of the interactions between the partons in the proton (*impulse approximation*). This means that during the scattering process the internal structure of the proton is "frozen", and the interaction between the quarks can be ignored. We start by defining the important kinematical quantities used to describe this process.

2.1.3.1   DIS Kinematics

The inelastic scattering of a lepton off a proton[1] (see Fig. 2.2)

$$lP \to l'\text{X}$$

can be expressed (for two unpolarised beams at fixed centre-of-mass energy) as a function of two independent variables [2]. In order to have an experiment-independent kinematic description of the process, it is favourable to use Lorentz invariants for describing the kinematics. We define the following quantities

- k, the four-momentum of the incoming lepton. $k^\mu = (E_k, \vec{k})$, where E_k and $\vec{k}$ are the energy and the momentum of the incoming lepton, respectively;

- p, the four-momentum of the incoming proton. $p^\mu = (E_p, \vec{p})$, where E_p and $\vec{p}$ are the energy and the momentum of the incoming proton, respectively;

---

[1]For consistency with the rest of the work presented in the thesis, in the following the word proton is used also for other possible targets.

[2]The number of degrees of freedom of a process in which two incoming particles collide and n final states are measured is $3n - 4$.
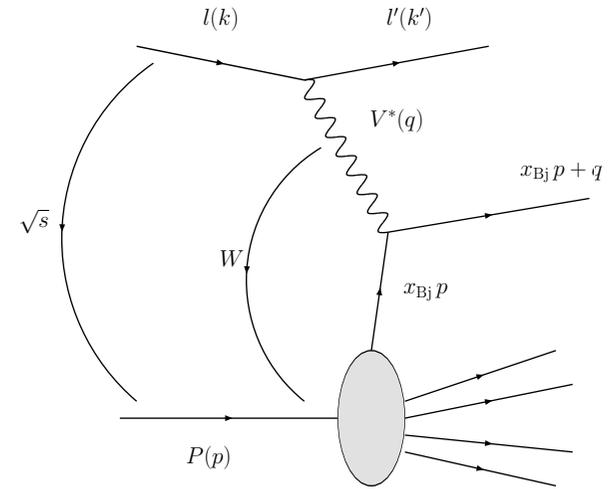
Figure 2.2:

- k', the four-momentum of the scattered lepton. $k'^\mu = (E_{k'}, \vec{k'})$, where $E_{k'}$ and $\vec{k'}$ are the energy and the momentum of the scattered lepton, respectively;

The centre-of-mass energy squared of the lP collision[3] is denoted as s

$$s = (p + k)^2. \tag{2.1}$$

The interaction between the lepton and the proton is mediated by either a photon, γ, a W or a Z boson[4]. In general we indicate this vector boson with V. The four-momentum of the exchanged boson is indicated with q and its value is given by

$$q = k - k'. \tag{2.2}$$

Under the approximation of a massless lepton, kinematics yields the four-momentum squared of the exchanged boson

$$q^2 \simeq -2\,E_k E_{k'}\,(1 - cos\theta_k) \tag{2.3}$$

where θ_k is the scattering angle between the incoming and the outgoing lepton. Thus q^2 is always $q^2 < 0$; this means that the exchanged boson is virtual, i.e. off its mass shell. Therefore an asterisc is added to the notation of the exchanged boson, V^*.

---

[3]In the following, the natural units will be adopted if not explicitly stated. The use of natural units fixes the value of the Plank constant and the speed of light in the vacuum to unity, $\hbar = c = 1$.

[4]If the lepton is a neutrino, the γ cannot be exchanged because of charge conservation.

$P(p)$

Proton remnant

$x_{\text{Bj}}\, p$

$V^*(q)$

$-x_{\text{Bj}}\, p$
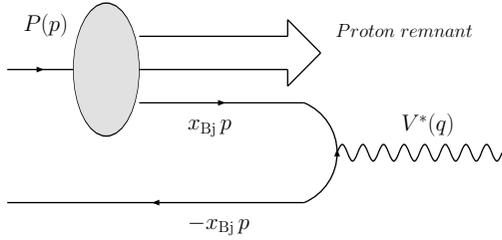
Figure 2.3: A schematic picture of a DIS interaction as seen in the Breit frame.

In order to work only with positive quantities, we define the virtuality of the exchanged boson, Q^2, as

$$Q^2 = -q^2. \qquad (2.4)$$

The interaction where the hadronic final state X is effectively produced is actually taking place between the proton and the boson. The boson probes the proton with a resolution given by the inverse of the square root of its virtuality. The centre of mass of the $\gamma^* p$ system is denoted by W:

$$W^2 = (p+q)^2 = M_{\text{X}}^2. \qquad (2.5)$$

Of course, the invariant mass of the final hadronic system, M_{X}, is equal to W because of conservation of the four momentum. An inelastic scattering is characterised by the condition $M_{\text{X}} \gg m_p$, where m_p is the mass of the proton. To resolve any internal structure of the proton, the resolution of the probe must be smaller than the proton size ($\sim 1\ fm$). Thus, the DIS kinematical regime is defined by the requirements

$$Q^2 \gg 1\ \text{GeV}^2, \qquad W \gg m_p. \qquad (2.6)$$

Another Lorentz invariant that can be defined is called inelasticity, y. It is defined as

$$y = \frac{p \cdot q}{p \cdot k} \qquad (2.7)$$

It has an intuitive physical interpretation in the reference frame where the proton is at rest. In this case Eq. (2.7) becomes

$$y = \frac{\nu}{E_k} = \frac{E_k - E_{k'}}{E_k} \qquad (2.8)$$

that is the fraction of the energy of the incoming lepton taken by the exchanged boson.

The boson in the DIS scattering interacts with one of the partons contained in the proton. In the QPM approximation, the partons move longitudinally along the proton direction and carry a fraction x_{Bj} of its total momentum. An easy calculation of x_{Bj} can

be obtained if we move to a particular reference frame where the quark and the vector boson collide head on and V^* transfers no energy but carries twice the momentum of the quark. Such a reference frame is called *Breit frame*. The DIS scattering in the Breit frame is represented in Fig. 2.3. The incoming quark four-momentum, f, is given by

$$f^\mu = (x_{\text{Bj}}\, p,\ x_{\text{Bj}} \vec{p})$$

having neglected the quark mass. The conservation of the four-momentum forces the scattered quark to have a final four-momentum, f' equal to

$$f'^\mu = (x_{\text{Bj}}\, p,\ -x_{\text{Bj}} \vec{p})$$

Thus the four-momentum of V^* in the Breit frame is

$$q^\mu = (0,\ -2x_{\text{Bj}} \vec{p}) \qquad (2.9)$$

giving a virtuality

$$Q^2 = -4x_{\text{Bj}}^2\, |\vec{p}|^2. \qquad (2.10)$$

On the other hand, the scalar product of p^μ and q^μ is

$$p \cdot q = -2x_{\text{Bj}}\, |\vec{p}|^2. \qquad (2.11)$$

By directly comparing Eq. (2.10) and (2.11), one obtains an expression for x_{Bj}

$$x_{\text{Bj}} = \frac{Q^2}{2\, p \cdot q} \qquad (2.12)$$

which is a Lorentz invariant and does not depend on the reference frame chosen.

The five Lorentz invariants presented here are the ones most commonly used. They are related to each other since only two of them are independent. Some of the equations relating these kinematical quantities are the following

$$W^2 = Q^2\, \frac{1 - x_{\text{Bj}}}{x_{\text{Bj}}} \qquad (2.13)$$

$$W^2 = sy - Q^2 \qquad (2.14)$$

$$Q^2 = s\, x_{\text{Bj}}\, y \qquad (2.15)$$

2.1.3.2 Derivation of the DIS cross section in the QPM

Once the kinematics of lepton-hadron scattering is specified, the cross section for the process can be calculated. In the following we will consider the case of unpolarised colliding beams and virtualities of the exchanged boson much lower than the mass of the W and Z bosons, such that the interaction can be approximated by photon exchange only. Another contribution to the cross section that is neglected in the following is

the exchange of more than one photon. The latter approximation is supported by the results of the comparison of the measured DIS cross sections using electron and positron beams which are very similar [3].

The differential cross section for $lP \rightarrow l'X$ as a function of θ_k and $E_{k'}$ can be expressed as contraction of two tensors

$$\frac{d^2\sigma}{d\theta_k dE_{k'}} = \frac{\pi \alpha^2}{m_p q^4} \frac{E_{k'}}{E_k} L_{\mu\nu} W^{\mu\nu} \qquad (2.16)$$

where this expression was calculated in the laboratory (LAB) frame. Here $L_{\mu\nu}$ represents the leptonic tensor that expresses the EM transition from the initial to the final lepton. It is calculable in Quantum ElectroDynamics (QED) and is written as

$$L_{\mu\nu} = 2[k'_\mu k_\nu + k'_\nu k_\mu - k \cdot k' g_{\mu\nu}] \qquad (2.17)$$

In analogy, the hadronic tensor corresponds to the EM transition of the target hadron to all possible final states, X. Evaluating its expression is more difficult than for the leptonic tensor and we will limit ourselves to parametrise it in the most general form compatible with Lorentz invariance

$$W_{\mu\nu} = Ag_{\mu\nu} + Bq_\mu q_\nu + C(q_\mu p_\nu + q_\nu p_\mu) + Dp_\mu p_\nu \qquad (2.18)$$

with A, B, C and D being arbitrary coefficients. This expression can be simplified by considering the constraints imposed by the conservation of the EM current. The conventional way to write down the hadronic tensor is

$$W_{\mu\nu} = W_1(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}) + \frac{W_2}{m^2}(p_\mu - \frac{p \cdot q}{q^2}q_\mu)(p_\nu - \frac{p \cdot q}{q^2}q_\nu) \qquad (2.19)$$

where the coefficients have been redefined as W_1 and W_2, the *structure functions*. In analogy to nuclear experiments, they contain the information about the distribution of the electric charge in the hadron. Using the expressions for $L_{\mu\nu}$ and $W^{\mu\nu}$ in Eq. (2.16), one obtains the cross section in the LAB frame as a function of $E_{k'}$ and θ_k

$$\frac{d^2\sigma}{d\theta dE_{k'}} = \frac{\pi \alpha_{EM}^2}{2E_k^2 sin^4\frac{\theta_k}{2}}(2W_1 sin^2\frac{\theta_k}{2} + W_2 cos^2\frac{\theta_k}{2}) \qquad (2.20)$$

It is more convenient to express Eq. (2.20) in terms of two of the Lorentz invariants introduced in Sect. 2.1.3.1. We are going to choose x_{Bj} and Q^2 for their natural physical interpretation in a DIS experiment at a collider. It is also common to replace W_1 and W_2 with by the structure functions $F_1 = m_p W_1$ and $F_2 = \nu W_2$ with $\nu = p \cdot q$. After some algebra one obtains

$$\frac{d^2\sigma}{dx_{Bj}dQ^2} = \frac{2\pi \alpha_{EM}^2}{x_{Bj}Q^2}(Y_+ F_2(x_{Bj}, Q^2) - y^2 F_L(x_{Bj}, Q^2)) \qquad (2.21)$$

where we used the definitions $F_L = F_2 - 2xF_1$ and $Y_+ = 1 + (1 - y)^2$. In the QPM, F_L is neglected since it is suppressed by helicity conservation. This is known as *Callan-Gross relation*. In QCD, higher order terms contribute to give $F_L \neq 0$ (see Sect. 2.1.4).

2.1.3.3   Remarks on DIS results

In the QPM F_2 has a very intuitive statistical interpretation expressed by the following relation

$$F_2(x_{Bj}, Q^2) = \sum_{i=1}^{N_f} e_i^2 x_{Bj} f_i(x_{Bj}, Q^2) \qquad (2.22)$$

where the sum runs over all the quark flavours, e_i is the electrical charge (in units of the electron charge) of the quark of flavour i and $f_i(x_{Bj}, Q^2)$ are functions called *parton distribution functions* (PDFs) or also parton densities. The PDFs express the classical probability to find in the proton a parton carrying a fraction x_{Bj} of the total proton momentum.

It was noticed since the first DIS experimets at SLAC at the end of the 1960's that the dependence of F_2 on Q^2 was very weak in the phase space region probed (see Fig. 2.4) [9]. This can be understood considering that the quarks which are effectively interacting with the γ^* are point-like. As Q^2 increases, the distances probed by the γ^* decrease. In the DIS regime one can study the internal structure of the proton. But since there is no internal structure of the quark, there is no change in increasing the resolution of the probe. The quark will appear always the same to the photon. This independence from Q^2 is known as *scale invariance*. In the next chapter it will be shown that violations of the scale invariance are expected, the structure functions having a weak logarithmic dependence on Q^2.
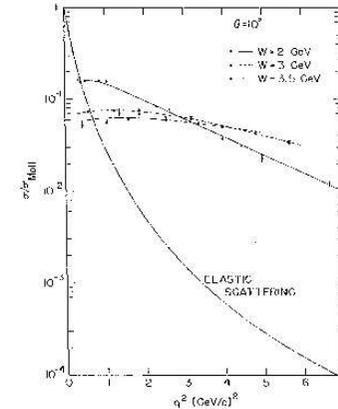

Figure 2.4: The weak dependence on Q^2 of the inealsitic cross section measured by the SLAC experiment. The same measurement taken at different values of W is presented with different markers (from [9]).

Although the QPM was able to motivate the results of the first DIS experiments, many problems still lacked a solution. Under the hypothesis that only quarks were present in the proton, the conservation of the momentum implies that

$$\sum_{i=1}^{N_f} \int_0^1 x f_i(x) dx = 1 \tag{2.23}$$

Instead the above integral is measured to be ~ 0.5. About one half of the proton momentum can not be directly detected in DIS experiments. This problem, together with the absence of detection of free quarks outside of the hadrons, stimulated the development of an extension of the QPM able to justify these experimental facts. This is achieved by means of the QCD theory.

2.1.4   Quantum Chromo Dynamics

QCD is the field theory that describes the strong interaction in the SM. The QCD lagrangian is gauge invariant like the QED one, just the gauge symmetry group changes[5]. The QCD lagrangian is invariant under local $SU(3)$ gauge transformations. The non-Abelian structure of the group is the peculiar aspect of QCD. It implies that there are three possible charge-states (*colours*) for the particles subject to the strong interaction (i.e. the quarks), as many as the dimension of the adjoint representation of the gauge group. Moreover, the non-Abelian nature of $SU(3)$ implies that, different to the QED case, also the gluons, the vector bosons that carry the force, can be in different colour states. There can be as many as the dimension of the fundamental representation of $SU(3)$, i.e. they can assume 8 different colour state. The lagrangian determines the dynamical properties of the interaction and allows to fix the Feynman rules for it. The set of Feynman rules can be found in Refs. [2, 3, 10]. Like in QED, the strength of the coupling between coloured particles is given by a constant, α_S, the strong coupling constant.
QCD has two main peculiar properties

- Asymptotic freedom. The coupling constant decreases at short distances while it increases at large distances. This means that at small length scales (i.e. large momentum transfer) the particles interact very weakly, as assumed in the QPM model. This allows to use perturbation theory for the calculations in the proper kinematical range.

- Confinement. No free coloured particles exists in nature. This property was deduced from the non-observation of coloured objects. Only particles which are colour singlets can be observed while the coloured partons are confined in the hadrons. This can be justified qualitatively from asymptotic freedom. As a quark and an antiquark get further from each other, the colour field becomes stronger and stronger until it is more favourable from the energetic point of view to create a new quark-antiquark pair that will bound to the former two creating

---

[5]We refer to Refs. [2, 3, 10] for a detailed description of the QCD lagrangian.
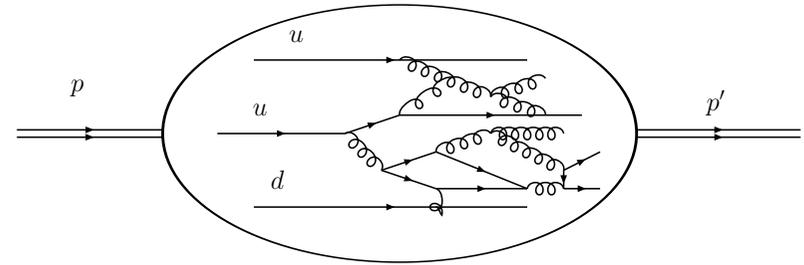
Figure 2.5: A schematic representation of the proton once that QCD is considered.

two hadrons. As it will be shown in the following, there are quantitative arguments that motivate asymptotic freedom. However, no rigorous demonstration is nowadays known for confinement.

The big difference introduced by QCD in the QPM picture is that now the partons in the proton can interact between themselves. The proton has become a dynamical object whose internal structure has a continous development (see Fig. 2.5).

Predictions for the physical quantities can be calculated with the Feynman rules in a perturbation expansion in orders of α_S. As for the QED case, there are divergencies in these calculations that have to be fixed in order to have finite and meaningful predictions. These divergencies are of different types: infrared and ultraviolet. The former comes from terms of the perturbative expansion in which a parton radiates a massless parton at very low angles, (collinear singularity) or at very low energies (infrared singularity).

The ultraviolet divergencies show up in higher order terms of the perturbative series, where loop diagrams introduce in the expression of the amplitude logarithmically divergent terms like

$$\int_0^\infty \frac{dk^2}{k^2 - m^2 + i\varepsilon}$$

where m is the mass of the particle exchanged in the loop and k its four-momentum. Examples of ultraviolet divergent diagrams at Next-to-Leading Order (NLO) in QCD (i.e. terms of the series proportional to α_S^2) are displayed in Fig. 2.6b and c.

These divergencies can be kept under control through a *renormalisation* procedure. A more detailed and rigorous description of the procedure followed to renormalise the QCD theory can be found in Refs. [3, 8]. In the following a few remarks about the most important features of the subject are mentioned [6]. The basic concept behind renormalisation is to redefine the parameters of the theory in such a way that the new expression of the perturbative series do not contain anymore the divergent integrals.

---

[6]The brief introduction to renormalisation presented here follows closely the one in Ref. [11]
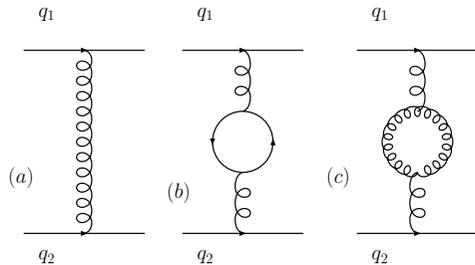
Figure 2.6: Feynman diagrams of a quark-quark QCD interaction at the (a) LO and (b and c) NL. The plot (b) has an internal quark loop while plot (c) has a internal gluon loop. (right) gluon loops.

Neglecting the quark masses, the only free parameter of QCD is $\alpha_0 = g^2/4\pi$, where g is the strenght of the coupling in the lagrangian. So we can write the value of a generic observable[7], $F(x)$, depending on a set of coordinates x, as a perturbative series in powers of α_0

$$F(x) = \alpha_0 + \alpha_0^2 F_1(x) + \alpha_0^3 F_2(x) + ... \qquad (2.24)$$

The problem is that the terms of the series, F_i, are divergent and make the calculation impossible. First the F_i need to be *regularised*. This means that we have to introduce a new set of functions $F_{i,\Lambda}$, related to the F_i, such that they are finite for any finite value of the parameter Λ, called *regulator*. The regularised functions tend to the former F_i in the limit for $\Lambda \to \infty$ and only at the very end of the calculation this limit is taken. This allows to perform the renormalisation procedure with well-defined functions. The practical relation between F_i and $F_{i,\Lambda}$ is called *regularisation scheme*. Several schems are possible, each with its own advantages and drawbacks. A possible choice can be to insert a cut-off Λ at the upper limit of the integral (2.1.4). A very famous regularisation scheme is the *dimensional regularisation* proposed by t'Hooft and Veltman in the 1970's. The relevant thing is that the final result of the calculation does not depend on the choice of the regularisation scheme, although the intermediate results will. The crucial observation is that since there is only one parameter for the theory, one measurement of $F(x)$ is enough to specify the theory itself. So we can define a new parameter $\alpha_S(\mu_R)$, function of a scale μ_R, called *renormalisation scale*. The new $\alpha_S(\mu_R)$ replaces α_0. We can rewrite the power series in terms of $\alpha(\mu_R)$ and $F_{i,\Lambda}$ and perform a measurement of the observable F at a given value μ_0 of μ_R, in order to fix the value of $\alpha_S(\mu_0)$. After these steps in the renormalisation of QCD, the terms in the perturbative series are well-behaved and we can take the limit for Λ such to restore the initial, Λ independent situation. Therefore, at the price of doing a measurement at a given point that we use as input, we are able to perform predictions at any other

---

[7]An example could be the cross section for a given QCD process.

values of the kinematic variables. It is important to stress that $\alpha_S(\mu_R)$ is the physical quantity relevant in the process under study. While α_0 is the bare coupling appearing in the QCD lagrangian, $\alpha_S(\mu_R)$ is the effective strength of the coupling between the partons after all the contributions at any order have been taken in account. Of course, since what we did was at the end just a reparametrisation, the divergences did not really cancel out, they are still somewhere. Indeed, they have been just moved in the relation between α_0 and $\alpha_S(\mu_R)$ and now $\alpha_0 \to \infty$ when the limit for Λ is taken, but this is not thought to be worrying since now α_0 is not a physical and observable quantity. The renormalisation procedure sketched here is used not only in QCD but also in QED, basically in the same fashion. In both cases, an important consequence is that the coupling constant now depends on the value of μ_R chosen for the process. The choice of μ_R is arbitrary and the physical quantity should not depend on it. This is imposed with the *renormalisation group equation* (RGE)

$$\mu_R^2 \frac{d}{d\mu_R^2} F = 0 \qquad (2.25)$$

Since F can be expressed in power series of $\alpha_S(\mu_R)$, the above equation sets a constraint on the μ_R-dependence of α_S too. We can define the variable $\beta(\alpha_S)$ as

$$\beta(\alpha_S) = \mu_R^2 \frac{\partial \alpha_S}{\partial \mu_R^2}. \qquad (2.26)$$

If the scale changes from a value μ_1 to a value μ_2, α_S will change in such a way that the physical quantity F will stay constant (*running coupling constant*). Integrating over the range (μ_1, μ_2) the definition (2.26) gives the relation between the change of μ_R and α_S

$$ln\left(\frac{\mu_R^2}{\mu_0^2}\right) = \int_{\alpha_S(\mu_0^2)}^{\alpha_S(\mu_R^2)} \frac{d\kappa}{\beta(\kappa)}. \qquad (2.27)$$

The exact nature of the dependence of α on μ_R is calculable in perturbative series

$$\beta(\alpha_S) = -b_0\, \alpha_S^2 - b_1\, \alpha_S^3 - b_2\, \alpha_S^4 + ...$$

and depends, at a fixed order, on all the possible contributions to the loop diagrams like those in Fig. 2.6. The peculiar thing that distinguishes QCD from QED, is the presence of gluon loops. This is ultimately related to the non-Abelian gauge structure of QCD that causes the vector bosons to have colour charge, while in QED the photon is electrically neutral. Because the gluons are spin-1 particles, the loop diagrams contribute to the scattering amplitude with opposite sign compared to the quark loops. While the fermionic loops strenghten the coupling constant with increasing scale, the gluon ones weaken it. The net effect depends by the number of possible flavours of quarks that can contribute to the total. In the specific case of QCD, at lowest order

$$\alpha_S(\mu_R) = \frac{\alpha_S(\mu_0^2)}{1 + \alpha_S(\mu_0^2)\frac{33-2N_f}{12\pi} ln(\mu_R^2\,/\,\mu_0^2)}$$
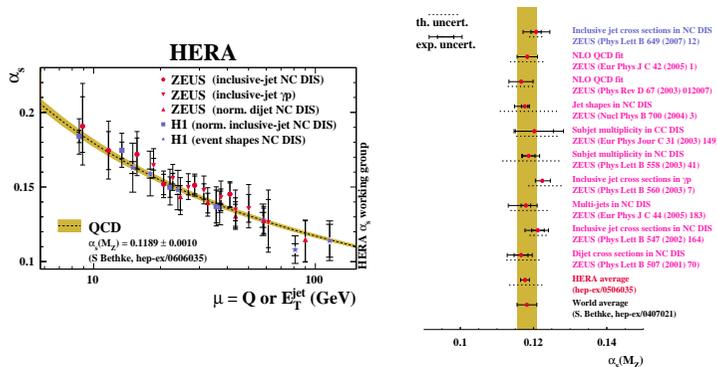
Figure 2.7: A collection of results of α_S measurements: (left) the running of α_S measured by the H1 and ZEUS collaborations, (right) a compilation of $\alpha_S(M_Z)$ measurements by the ZEUS Collaboration using different experimental methods together with the world average (from [13]).

2.1.5 Parton Distribution Functions and DGLAP evolution

As mentioned in Sect. 2.1.3.3, in the QPM parton densities (PDFs), $f_i(x)$, express the probability to find a parton of flavour i carrying a fraction x of the proton momentum. In the QPM, only quarks are considered in the PDFs and there is no dependence of the PDFs on the energy scale of the $\gamma^* p$ interaction. QCD changes dramatically this
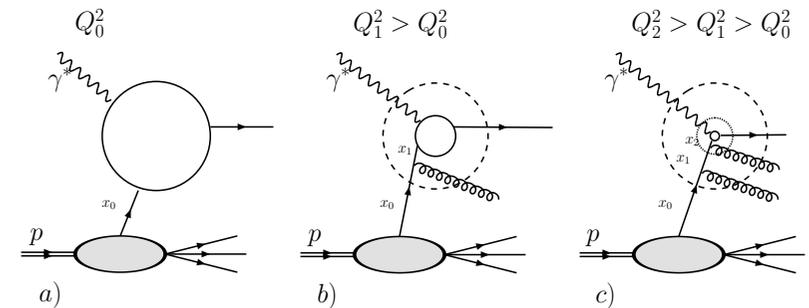

Figure 2.8: The virtual photon emitted by the electron has a spatial resolution of $\sim 1/\sqrt{Q^2}$. As Q^2 increases from (a) to (c), the photon is able to resolve more and more partons generated dynamically by the QCD field. This leads to a different proton internal structure for different Q^2 values.

scenario. Now the quarks interact one with each other exchanging gluons. Moreover, the gluons in the proton can create a quark-antiquark virtual pair. The quarks created dynamically in this way are called *sea quarks*. Sea quarks can create more gluons and more sea quarks. The gluons and sea quarks carry a fraction of the proton momentum too, so they have to be considered in the PDFs. Having neglected them is the reason why the experimental results seemed to violate the momentum sum rule (2.23).

This continous creation and annihilation of gluons and sea quarks introduces a scale dependence of the PDFs and the structure functions. The qualitative explanation for that can be understood with the help of Fig. 2.8. The spatial resolution of the γ^* is of the order $1/\sqrt{Q^2}$. For an initial low value of Q^2, Q_0^2, the photon is able to resolve only a certain PDF $f(x, Q_0^2)$. As the virtuality increases, the photon "sees" more and more gluons and partons emitted from the former quark. Therefore the momentum of the former quark seen at Q_0^2 now is shared between many other partons and the photon will interact with one of them. The net effect will be that one will observe more low-x partons and less high-x partons as Q^2 increases. This Q^2 dependence is known as *scaling violation* and is a peculiar QCD effect.

The analytical treatment of the scaling violations can be done in the pQCD framework but needs some care. When calculating the cross sections for QCD processes one encounters infrared divergences caused by the emission of soft and collinear gluons and quarks. The collinear emissions introduce in the cross section calculation logarithmically divergent terms

$$\int_0^{p_{T,max}^2} \frac{dp_T^2}{p_T^2}$$

where $p_{T,max}^2 = \frac{Q^2(1-z)}{4z}$ and $z = x_0/x_1$ is the fraction of the initial parton momentum taken by the outgoing quark. The solution for that is to introduce a lower cut-off in

---

where N_f is the number of flavours. If $N_f = 6$, as presently believed, the strong coupling constant decreases as the scale μ_R increases. This can be seen as a quantitative proof of the asymptotic freedom.

As said above, the final result must be independent of the regularisation scheme adopted. It should also be independent of the choice of the renormalisation scale, which is not a physical parameter. But these considerations are valid only if one considers the whole perturbation series. The parts of the calculation depending on the regularisation scheme and the renormalisation scale cancel in a complex way between different terms of the series. If one does not consider all the terms, some residual contribution do not have a counterpart that cancel them. Practically, the QCD calculations are done up to the NLO level, in a few cases to the Next-to-NLO level. This means that a scale dependence is always left and one must specify both in which scheme the calculation was carried out and at which scale the renormalisation was done.

The experimental tests have given in the last decades a strong support to the QCD theory [12]. The running of α_S is nicely proven by the experimental measurements (see Fig. 2.7) [12, 13].

the integration, κ, and a term $ln(\frac{Q^2(1-z)}{\kappa z})$ appears in the cross section formula. In order to have results independent of this arbitrary cut-off, a procedure similar to the one used for renormalising the theory is carried out. This procedure is called *collinear factorisation* (see Ref. [10] for a review). A (collinear) factorisation scale, $\mu_F >> \kappa$, is introduced to separate the process into a hard, short-distance part and a soft, long-distance and non-perturbative one. The latter is described with the parton densities. The infrared divergences are moved inside the PDFs which also acquire a Q^2 dependence, as discussed qualitatively above. As for the renormalisation procedure, the way this reparametrisation is done is arbitrary and the final predictions of the physical quantities should not depend on it. Two very common prescriptions for factorising the infrared divergences are the $\overline{MS}$ and the DIS schemes. The former tries to move inside the PDFs as few terms as possible. Conversely, the latter moves inside the PDFs as many as possible.

The statement that it is possible to factorise the QCD process into a short- and a distant-range part with the procedure just sketched is known as *QCD factorisation theorem*. The demonstration of the QCD factorisation theorem for diffractive DIS [14] is a complex issue beyond the scope of this thesis and we comment only on its results. The consequences of the factorisation theorem make the predictiveness of QCD extremely powerful. The short-range part of the process is the one containing the hard-scale (i.e. high-energy) subprocess between the particles that actually interact. For instance, in the DIS case this is the partonic cross section for the photon-quark collision. This hard scattering cross section is calculable in QCD by means of perturbative expansion in powers of α_S which is small because of the property of asymptotic freedom of QCD (see Sect. 2.1.4). The long-range (hence non-perturbative) component is identified with the parton densities. A crucial aspect of the factorisation theorem is that it states that the PDFs are *universal*, i.e. they depend only on the hadron type and not on the kind of process considered in the calculation. One can use the same PDFs at different colliders and for different final states. This universality is one of the most appealing aspects of the PDFs. Notice that although strictly related to the PDFs (as in Eq. (2.22)), the structure functions are not equivalent to the PDFs for this reason, i.e. they are not universal but depend upon the process studied.

Thanks to the factorisation theorem, the differential cross section can be expressed with the following formula

$$\Delta\sigma(x,Q^2) = \sum_{i=q,\bar{q},g} \int_{\Delta Q^2} dQ^2 \int_{\Delta x} d\xi \,\, \hat{\sigma}\left(x,Q^2;\frac{\mu_F^2}{\mu_R^2},\alpha_S(\mu_R^2)\right) \, f_i(\xi,Q^2;\mu_F^2,\mu_R^2) \quad (2.28)$$

and is visually represented in Fig. 2.9. Notice that in order to factorise the cross section we need to specify the separation between what is soft and what is hard. This is done setting the factorisation scale, μ_F. It is important to stress that the validity of the theorem relies on approximations. The main one is to ignore higher-twist QCD terms in the process amplitude. Higher-twist terms are higher-order contributions to
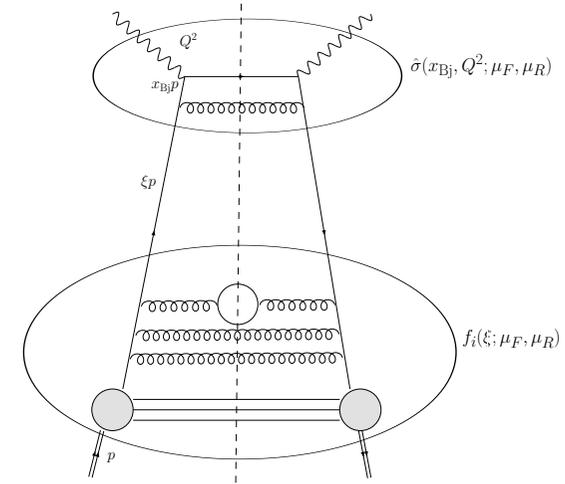
Figure 2.9: A schematic representation of the QCD factorisation in DIS processes. The interaction between γ^* and p is factorised in a soft part (the PDFs, f_i) containing the infrared divergences, typical of the p but process-independent and a hard part, process-dependent, calculable in pQCD ($\hat{\sigma}$). The separation between soft and hard is set by the factorisation scale, μ_F (adapted from [8]).

the structure functions. The structure functions can be expanded in series of $\frac{1}{Q^2}$

$$F_2(x_{\text{Bj}},Q^2) = \sum_n B_n(x_{\text{Bj}},Q^2)\left(\frac{1}{Q^2}\right)^n$$

where for large Q^2 considering the leading term $n=0$ (leading-twist) is a good approximation. Nonetheless higher-twist terms exist and they are not covered by the factorisation theorem.

In the factorisation procedure, only the largest terms are considered and resummed over all orders, namely the ones proportional to $\alpha_S^n \, ln(Q^2)^n$. This approximation is called *leading logarithm approximation* (LLA). It is interesting to remark that this approximation implies a strong ordering in the transverse momentum, p_T, of the partons emitted before the hard interaction.

In Fig. 2.10 a complex diagram with many emissions is shown. The LLA imposes that the transverse momentum of the partons increases after each emission up to the value of the hard scale of the process.

$$\kappa^2 \leq p_{T,1}^2 \leq p_{T,2}^2 \leq ... \leq p_{T,n-1}^2 \leq p_T^2 < Q^2$$

while the fraction of x carried by the subsequent emissions decreases
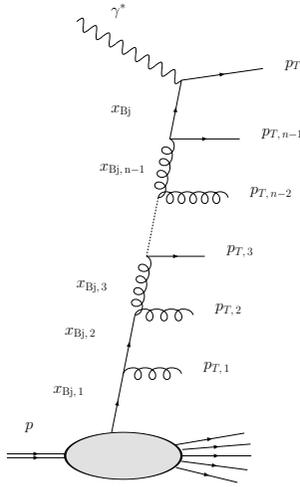
Figure 2.10: Higher order contributions to $\gamma^*q \to qX$.

$$x_1 < x_2 < ...x_{n-1} < x.$$

The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution [15, 16] uses the LLA for describing analytically the dynamical internal structure of the hadrons. The Q^2 evolution of the PDF is formally described by the "Altarelli-Parisi" equations [16]

$$\frac{dq_i(x,Q^2)}{d\,lnQ^2} = \frac{\alpha_S(Q^2)}{2\,\pi} \int_x^1 \frac{dy}{y} \left[q_i(y,Q^2)P_{qq}(x/y) + g(y,Q^2)P_{qg}(x/y)\right]$$

$$\frac{dg_i(x,Q^2)}{d\,lnQ^2} = \frac{\alpha_S(Q^2)}{2\,\pi} \int_x^1 \frac{dy}{y} \left[\sum_i q_i(y,Q^2)P_{gq}(x/y) + g(y,Q^2)P_{gg}(x/y)\right]$$

(2.29)

where the *splitting functions*, $P_{ij}(z)$, represent the probability that a parton of type j (either a quark or a gluon), carrying an initial momentum xP, emits a parton i carrying a fraction z of its momentum (see Fig. 2.11).

The analytic expressions for the LO splitting functions shown in Fig. 2.11 are the following

Figure 2.11: The splitting function diagrams at the lowest order in α_S.

$$P_{qq}(z) = \frac{4}{3}\frac{1+z^2}{1-z}$$

$$P_{qg}(z) = \frac{1}{2}\left[z^2 + (1-z)^2\right]$$

$$P_{gq}(z) = P_{qq}(1-z) = \frac{4}{3}\frac{1+(1-z)^2}{z}$$

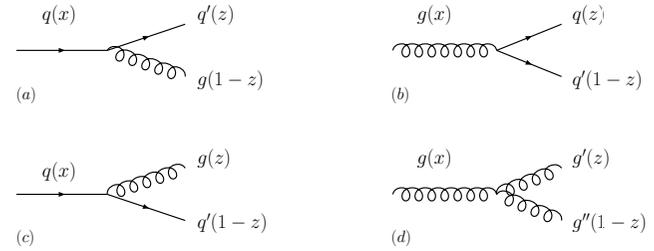$$P_{gg}(z) = 6\left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right]$$

Although perturbative QCD is not able to calculate the PDFs a priori, thanks to the DGLAP evolution one is able to predict the value of the PDFs over a large kinematic region once their value at a point in Q^2 is given. But one has to be careful in their use because they rely on the LLA approximation which is not everywhere valid. In fact at low x, $ln(1/x)$ terms become important and should not be excluded anymore as done in the LLA. In a moderately low x region the *double leading logarithm approximation* (DLLA) prescribes how to sum leading terms in $ln(1/x_{Bj})$ when they appear in the calculation accompanied by leading terms in $ln(1/Q^2)$. The steep rise of F_2 at low x_{Bj} suggests that even the DLLA is not enough. This has led to another model for the evolution of the PDFs proposed initially by Balitzky, Fadin, Kuraev and Lipatov (BFKL) [18]. The BFKL evolution equations sum leading terms in $ln(1/x)$, independently of their lnQ^2 dependence. The strong ordering in p_T of the DGLAP evolution is lost ("random p_T walk") in the BFKL evolution.

Because of the QCD confinement, a coloured particle like a quark or a gluon cannot propagate freely after the interaction. Rather it will tend to radiate other partons and, finally, to merge with other partons in order to form a colourless hadron. This process is called *hadronisation* and is a predominantly non-perturbative process which is described exploiting once again collinear factorisation in a fashion similar to the PDFs. The hard subprocess is factorised from the hadronisation process at a scale μ that is usually chosen for convenience equal to the factorisation scale μ_F. The non-perturbative, infrared divergent part is described by *fragmentation functions*, $D_{h/i}(z;\mu)$, the probability that a parton of type i hadronises into a hadron of type h,
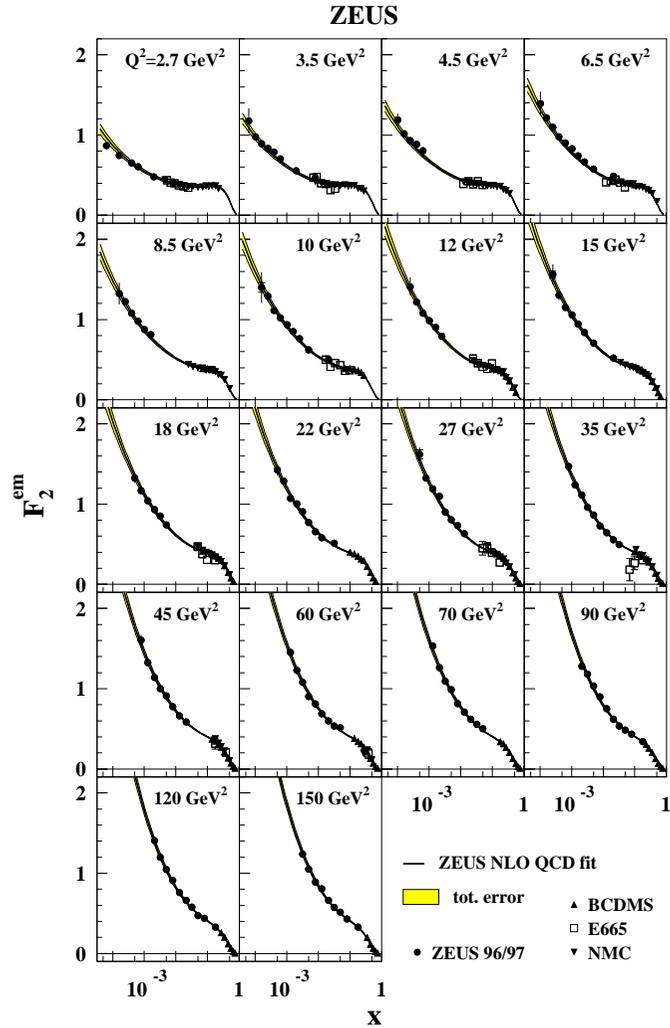
Figure 2.12: A collection of measurements of the F_2 structure function from several experiments (markers). F_2 is presented as a function of x_{Bj} in different bins of Q^2 and is compared to the NLO prediction using the DGLAP evolution. The scaling violations are clearly visible as a change of the steepness at low and high x_{Bj} as Q^2 increases (from [17]).
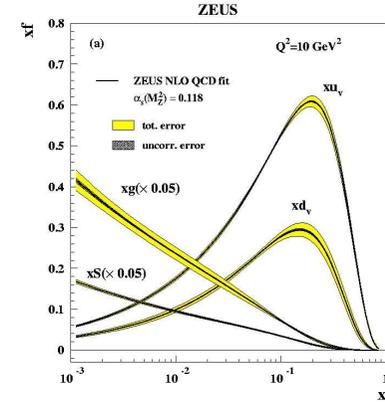
Figure 2.13: The curves represent the PDFs, f, multiplied by x for different kind of partons as extracted by the ZEUS collaboration. The PDFs are presented at a fixed value of $Q^2 = 10 \;\; GeV^2$ as a function of x_{Bj} (from [17]).

the latter carrying a fraction z of the initial parton momentum. Hence the physical meaning of the fragmentation functions is strongly related to the parton densities.

2.1.6   Saturation model

In the previous description of the DIS processes, we worked in a reference frame where the proton moves very fast. The same process can be seen from a different and complementary point of view, the *colour dipole model*[8, 19]. In the colour dipole model, the interaction is studied in a reference frame where the proton is at rest. In this case the photon splits into a $q\bar{q}$ pair (*dipole*) far upstream the target. The transverse separation of the $q\bar{q}$ pair, r, is proportional to $1/\sqrt{Q^2}$. The life time of the dipole is proportional to $1/x_{\mathrm{Bj}}$ and much longer than the interaction time. Thus, the transverse size of the dipole does not change during the process. The interaction itself is between the proton and the dipole, as depicted in Fig.2.14. A theoretical model of particular success in describing the measured inclusive DIS data is the *saturation model*, originally proposed by Golec-Biernat and Wüsthoff [20]. The $\gamma^* p$ interaction can hence be factorised into two parts: first the photon splits into a dipole of radius r where the (anti-)quark carries a fraction z $(1-z)$ of the initial photon momentum. Then the dipole interacts with the proton. The effective dipole-proton cross section is indicated by $\hat{\sigma}$ and depends on x_{Bj} and r. Thus the $\gamma^* p$ cross section can be written as

$$\sigma_{T,L}(x_{\mathrm{Bj}}, Q^2) = \int d^2 \mathbf{r} \int_0^1 dz |\Psi_{T,L}(z, \mathbf{r})|^2 \hat{\sigma}(x_{\mathrm{Bj}}, r^2) \qquad (2.30)$$

Figure 2.14: Schematic picture of the γ^*p DIS scattering in the saturation model.

where the indices T and L refer to transversely and longitudinally polarised photons, respectively, and $\Psi_{T,L}(z, \mathbf{r})$ is the squared photon wave function. The dynamics of the process is defined by the effective dipole cross section. Different choices for $\hat{\sigma}$ can be made. The saturation model chooses the following simple expression

$$\hat{\sigma}(x, r^2) = \sigma_0 \left[1 - exp \left(-\frac{r^2}{4R_0^2(x)} \right) \right] \qquad (2.31)$$
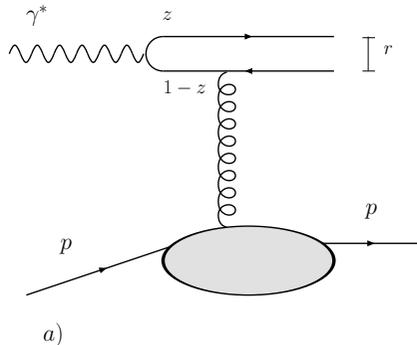
where σ_0 is a normalisation factor and the quantity R_0 is an x_{Bj}−dependent saturation scale. The parametrisation for R_0 is

$$R_0(x) = \frac{1}{\mathrm{GeV}} \left(\frac{x}{x_0} \right)^{\lambda/2}. \qquad (2.32)$$

The parameters σ_0, x_0 and λ are not given by the theory and need to be extracted from the experimental data. With the definitions (2.31) and (2.32), one ensures that the final cross section is proportional to r^2 for small r (colour transparency) and approaches asymptotically a constant value for large r (saturation). The saturation regime sets up when $r \sim 2R_0$. The density of partons in the proton increases with x_{Bj} as seen in Sect. 2.1.5. This increases the total cross section since the number of scattering targets increases. However, if this density becomes higher than the photon resolution (i.e. the dipole radius), further increases of it do not cause more interactions and the cross section saturates. Thus, the process is determined by the ratio between r and R_0, the former being inversely proportional to the square root of the photon virtuality. The dipole cross section depends only on the quantity $\tau = Q^2 R_0^2(x)$. The invariance of the cross section at fixed values of τ is known as geometric scaling [21, 22] and is experimentally nicely demonstrated at HERA (see Fig. 2.15).
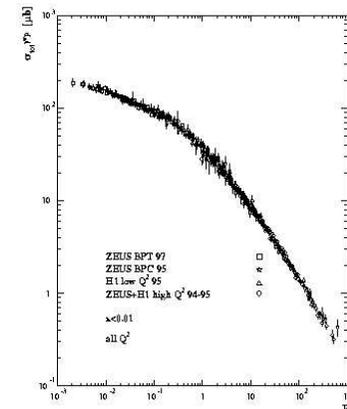
Figure 2.15: The cross section for γ^*p scattering measured at HERA as a function of the scaling variable $\tau = Q^2 R_0^2(x_{\mathrm{Bj}})$. The measurement is presented for $x_{\mathrm{Bj}} < 0.01$ and $0.045 < Q^2 < 450\,GeV^2$ (from [22]).

2.1.7 Jet physics

Because of the confinement property of QCD , free partons can not be observed in Nature. This makes a direct study of the final state products of the strong interaction process impossible. Within the typical time scale of the strong interaction ($\approx 10^{-24}\,s$) the partons create hadrons in a complex and non-perturbative process of merging (the hadronisation described in Sect. 2.1.5). These final hadrons are the particles actually detected by the experiment.

The key feature is that if the hadrons origin from a high p_T parton they will concentrate in a collimated angular region around the direction of the former parton. These "sprays" of hadrons are labeled as jets and are one of the main tools for studying QCD dynamics [23]. The kinematics of the hadronic system in the jet are the same as of the initial parton, thus they are the connection between the interaction and the observation. In order to make this connection in a reliable way, the jet kinematics must be measured with as little bias as possible. This means not only to measure accurately the kinematical properties of the single hadrons emerging from the interaction but also merging them in the right way and reconstructing the resulting jet kinematics in an unbiased way with respect to the initial parton. Moreover, the jet must be infrared safe, i.e. the outcome must be insensitive to the emission of collinear or soft partons produced in higher-order QCD processes.

There are several jet algorithms suitable for QCD studies. They all exploit the notion that the particles belonging to the same jet should be close to each other in phase space. Cone algorithms [24, 25, 26] merge particles which are geometrically close

to each other. The basic strategy of the algorithm is to define *seeds*, particles with a transverse energy E_T higher than a certain threshold $E_{T,cut}$ and to merge in the same jet all the particles inside a cone of radius R_{cut} around the seed. This algorithm is inclusive, i.e. not all the particles are assigned to a jet. In pp collisions this allows the treatment of the proton remnant. Cone algorithms are still used a lot thanks to their ease of implementation although theoretical issues affect it. Modifications to the cone algorithm have been proposed in order to improve this [27].

Other algorithms merge the particles according to their momenta rather than their directions. The *JADE algorithm* [28] defines for each pair of particles, i and j, a distance, m_{ij}

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

where E_i and E_j are the energies of the $i-$th and the $j-$th particle respectively and θ_{ij} is the polar angle between them. In other words, the distance m_{ij} is the invariant mass squared of the system made by the two particles. The particles are merged in a jet only if this invariant mass is lower than a certain threshold, $M_{cut}^2 = y_{cut} M^2$, where y_{cut} is a resolution parameter and M^2 a reference mass. Differently from the cone algorithm, with the JADE algorithm all the particles in the event are assigned to a jet, making problematic the treatment of the proton remnant.

The k_T-algorithm is similar to the JADE algorithm to some extent. In this case the distance between two particles, $k_{T,ij}$, is the transverse momentum of one relative to the other

$$k_{T,ij}^2 = 2(1 - cos\theta_{ij})\,min\{E_i^2, E_j^2\} \qquad (2.33)$$

The treatment of the beam remnant is implemented in the k_T-algorithm by defining another parameter, $k_{T,iP}$, which defines the transverse energy of the $i-$th particle relative to the beam axis

$$k_{T,iP}^2 = 2(1 - cos\theta_{iP})\,E_i^2 \qquad (2.34)$$

where θ_{iP} is the polar angle between the particle i and the incoming proton beam direction. For each iteration of the algorithm, the minimum of all the $\{k_{T,ij}, k_{T,iP}\}$ is taken and tested against a threshold E_T^{min}. If the smallest value is lower than the threshold, the particle is merged either to its closest neighbour (if $k_{T,ij}$ was the smallest value) or to the beam remnant (if $k_{T,iP}$ was the smallest). The prescription for combining two particles forming a new "pseudo-particle" is given by a *recombination scheme*. Several schemes are valid for the jet reconstruction. The procedure is iterated until only objects with distances above the threshold are left.

The *longitudinally invariant k_T algorithm* [29, 30] combines the advantages of the cone and the k_T algortihm. The definition of the distance between two particles is different from Eq. (2.33) as well the beam distance definition is different from Eq. (2.34). The longitudinally invariant k_T algorithm is the one used in this analysis and is described in more detail in Sect. 4.2.5 together with the choice of the recombination scheme.

The study of jets can give a significant contribution to the test and understanding of QCD. A milestone was the experimental discovery of the gluons at the e^+e^- PETRA
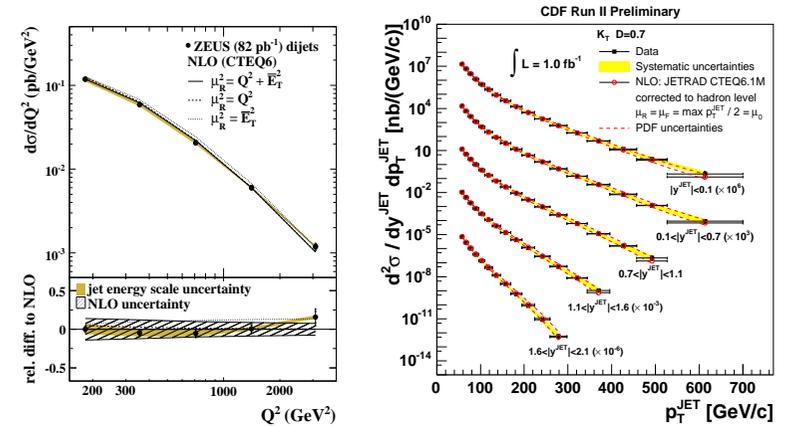
Figure 2.16: Two examples of jet measurements. The left-hand plot (a) shows the measurement of the differential cross section for dijet production in ep collisions as a function of the virtuality of the exchanged photon, Q^2. The measurement was performed by the ZEUS collaboration and is compared to the several NLO predictions differing for the choice of the renormalisation scale (from [37]). The right-hand plot shows the measurement of the differential cross section for dijet production in $p\bar{p}$ collisions as a function of the transverse momentum of the jet, p_T, in different bins of rapidity, Y. The measurement was performed by the CDF collaboration and is compared to the NLO prediction.

collider through the study of three-jets coplanarity [31]. More generally, jets give the opportunity to test predictions in a kinematic regime where QCD is expected to work. For example, the NLO QCD predictions of the jet differential cross section are presented in Fig. 2.16a, showing the high degree of precision of the theory over a wide phase space. Another possible test of QCD with jets is the study of the azimuthal asymmetries in events with two or more jets (*dijet production*) [32]. Jets are also used as inputs for fits to the parameters of the theory. There are several methods for estimating the value of α_S from jets. It can be done by measuring the ratio of the yields for dijets and trijets [33, 34] or from the study of the internal substructure of jets [35]. These quantities are determined essentially by parton radiation which depends on the value of the strong coupling constant. Moreover, jets can be combined with inclusive DIS data in order to better constrain the PDFs [36]. The reduction of the parton densities uncertainties, especially for the gluon, is sizeable and indicates once again the power of this experimental tool for QCD studies.

2.2   Diffraction in strong interactions

Since the 1960's, the existence of a subset of strong interaction processes with very peculiar characteristics called *diffraction* is known [39]. Diffraction was observed in pp collisions before the birth of QCD and models like Regge phenomenology [40, 41] were developed in order to include the effects of diffraction in the observed measurements. Although QCD is used for describing the strong interactions since more than twenty years, diffraction was hard to fit in this framework because of its intrinsical soft and non-perturbative nature. A possible way to study diffraction in the QCD framework is to require the presence of a hard scale in the diffractive process, given by e.g. the exchange of a high virtuality boson, jets or heavy quarks. The QCD theory can be remarkably boosted by the study of diffractive processes: the simultaneous presence of a soft and a hard scale allows to test the border line between perturbative and non-perturbative processes in a more sophisticated fashion. Diffraction concentrates in the low-x region and the complex dynamics of this still not well-known kinematic region can be studied in detail.

2.2.1   Soft diffraction

In the early experiments of pp collisions in the 1970's, a surprising feature of the data observed was the slow rise of the total cross section, σ_{tot}, as a function of s (σ_{tot} was expected to saturate at a certain energy). It was also noticed that the elastic cross section exhibited a particular behaviour as a function of t, the four-momentum transferred squared at the proton vertex defined as

$$t = (p' - p)^2 \tag{2.35}$$

where it should be noted that t assumes negative values. In Fig. 2.17 this trend is shown: a peak around $|t| \sim 0$ (the *diffractive peak*) followed by an exponential decrease,

$$\frac{d\sigma_{el}}{dt} = a\, e^{b\, t}$$

where a and b are two parameters. A subsequent minimum followed by a secondary maximum is observed. A particular feature of the $t-$distribution of the pp data also shown in Fig. 2.17 is the shrinkage of the width of the diffractive peak (i.e. the increase of b) with the increase of the energy.

In times predating the birth of the QCD, models like Regge phenomenology were used to interpret these results[8]. In a Yukawa-model approach, an interaction is mediated by the exchange in the $t-$channel of an object with mass squared $m^2 = t$ between the two colliding particles. For a given hadronic reaction, only specific particles can be exchanged in order to conserve all the relevant quantum numbers. For example, in the reaction $\pi^- p \rightarrow \pi^0 n$, only the ρ, a_2 and f_6 mesons can be exchanged. Remarkably, the relation between mass squared and spin connects these particles, as shown in

---

[8]For an extensive review on the subject of Regge phenomenology the reader is referred to Refs. [41, 43]
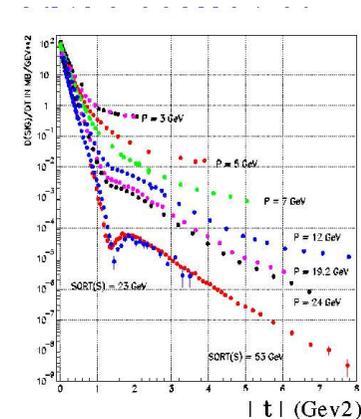
Figure 2.17: The differential cross section as a function of t in pp elastic collisions for different values of the centre-of-mass energy squared, s. In case of fixed target experiments measurements, the momentum of the incoming proton beam, P, is given (from [42]).

Fig. 2.18. The straight line on which the mesons lie is called a *Regge trajectory*. The same property is valid for the other mesons and other Regge trajectories are shown in Fig. 2.18.

Regge phenomenology connects the asymptotic high-energy behaviour of the cross section to the singularities in the complex angular momentum of the partial wave amplitudes in the crossed channel. By analytically continuing the partial wave expansion to complex values of the angular momentum, one can realize that, because of very generic properties like unitarity, analyticity and crossing, the asymptotic high-energy behaviour of the cross section is connected to the singularities in the crossed t-channel that arise in the calculation. Rather than the exchange of a particle like in the Yukawa model, Regge phenomenology considers the collective effect of the exchange of all the mesons belonging to the same Regge trajectory. The trajectory can be parametrised as

$$J = \alpha(t) = \alpha(0) + \alpha' t \tag{2.36}$$

where $\alpha(0)$ and α' are the intercept and the slope of the trajectory[9] (see Fig. 2.18). The Regge trajectories for mesons are called *reggeons* and the typical values for their parameters are $\alpha(0) \sim 0.5$ and $\alpha' \sim 1\,\text{GeV}^{-2}$. Only trajectories with the proper quantum numbers can contribute to a given reaction. For example the elastic $\pi^- p$ process receives contributions only from the ρ, the f_2 and the $I\!P$ trajectories (the

---

[9]Although this parametrisation is expected to be valid only for small values of $|t|$, it is experimentally seen to hold up to $|t| \approx 5 - 6\,\text{GeV}^2$.
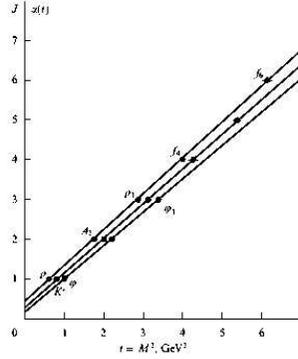
Figure 2.18: The correlation between the squared mass and the spin of several mesons. Mesons belonging to the same Regge trajectory lie on the same straight line. The Regge trajectories of the ρ, ω, f_2 and a_2 are degenerate and superimpose.

latter will be introduced later in this section). The cross section for elastic interaction of hadrons A and B at high centre of mass energy, s, and fixed t is then predicted to be

$$\frac{d\sigma_{\rm el}^{AB}}{dt} = \sum_k \frac{\beta_{Ak}^2(t)\beta_{Bk}^2(t)}{16\pi} s^{2\alpha_k(t)-2} \tag{2.37}$$

where the sum runs over all the allowed trajectories and the functions β are called *residue functions* and express the coupling between the $k-$th trajectory and the hadron A (B). Considering Eq. (2.36), one can define the *slope* parameter $b = b_0 + 2\alpha' \, ln(s)$ and rewrite Eq. (2.37)

$$\frac{d\sigma_{\rm el}^{AB}}{dt} = \sum_k \frac{\beta_{Ak}^2(t)\beta_{Bk}^2(t)}{16\pi} s^{2\alpha(0)-2} \, exp(bt) \tag{2.38}$$

showing that Regge phenomenology predicts the shrinkage of the forward peak as seen in the data[10]. The total cross section is related to the elastic via the optical theorem and can be written as

$$\sigma_{rmtot}^{AB} = \sum_k \beta_{Ak}(0)\beta_{Bk}(0)s^{\alpha_k(0)-1} \tag{2.39}$$

---

[10]This shrinkage is at the end a consequence of the limit for high s and fixed t at the base of Regge theory. Considering the two colliding hadrons, A and B, as massless and using the definition in Eq. (2.35), one can write t as $t \sim -2E_A E_B(1 - cos\theta)$ where E_A and E_B are the energies of the hadrons and θ the scattering angle. As s increases, the only way for t to stay constant is that θ must become smaller and smaller, i.e. closer and closer to the forward direction.
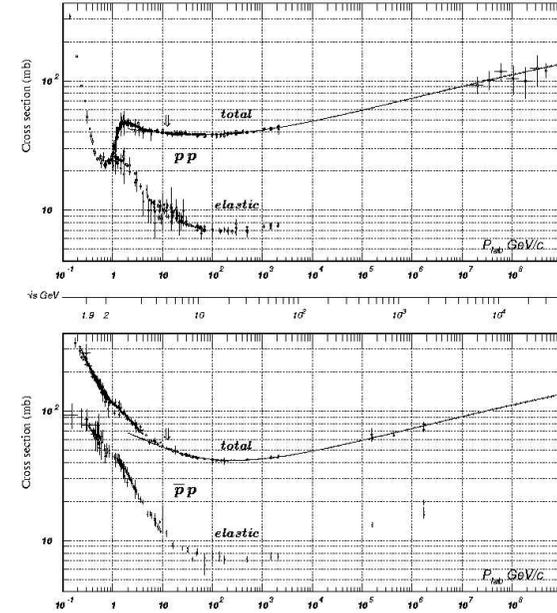
Figure 2.19: Total and elastic cross section in (upper plot) pp and (lower plot) $\bar{p}p$ interactions. From Ref. [12].
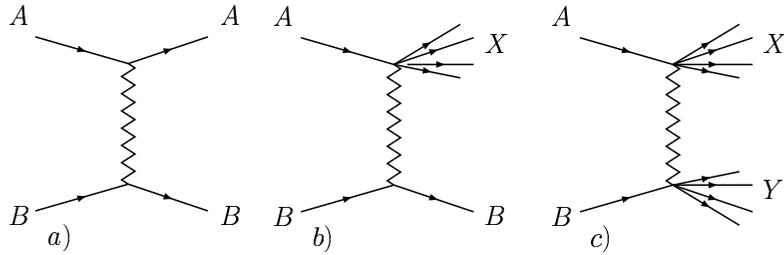
Figure 2.20: Schematic diagrams for (a) elastic, (b) single diffractive and (c) double diffractive pp collisions.

It was noticed that the slow rise of the total cross section could not be described with the known trajectories using the known mesons (see Fig. 2.19). Since the intercept $\alpha(0)$ has approximately the same value of 0.5 for all the mesonic trajectories, this would cause a decrease of the total cross secetion. This induced the introduction of a new trajectory, the *pomeron* ($I\!P$), that dominates at high energies [44]. According to a generally accepted fit to the total cross sections for different processes [45], the intercept of the pomeron is higher than in the case of the mesonic trajectories ($\alpha_{I\!P}(0) = 1.08$) and the slope is $\alpha'_{I\!P}0.25\,\mathrm{GeV}^{-2}$. The peculiar property of the $I\!P$ is to carry no quantum numbers but for the spin and the parity (equivalently said, the $I\!P$ carries the vacuum quantum numbers). Within the Regge theory, *diffraction* is that class of processes in which a $I\!P$ is exchanged between the interacting particles although it is important to stress that the $I\!P$ must not be misunderstood as a real particle. More in general, the most valid definition from the theoretical point of view is that diffraction is the dominant high-energy process when no quantum numbers are exchanged between the interacting particles. Using this definition we can include among the diffractive processes also the ones having one or both protons dissociating in a low-mass resonant state (*single* and *double dissociation* processes, see Fig. 2.20), typically of the order of $2-3$ GeV.

The two features of diffractive events mentioned above, the exchange of the vacuum quantum numbers and the slight perturbation that the incoming hadron undergo, imply very specific experimental signatures that are exploited experimentally for tagging an event as diffractive (see Sect. 2.2.3). The incoming hadron in a diffractive process can either stay intact or dissociate in a low-mass state: in the former case and in most events of the latter, the outgoing hadron will escape in the beam pipe hole leaving no signal in the central detector. A typical experimental signature of the diffractive signal is the presence of one or more *large rapidity gaps* (LRG). The rapidity of a particle, Y is defined as

$$Y = \frac{1}{2} ln \frac{E + p_Z}{E - p_Z} \qquad (2.40)$$

where E and p_Z are the energy and the $Z-$component of the momentum of the particle.
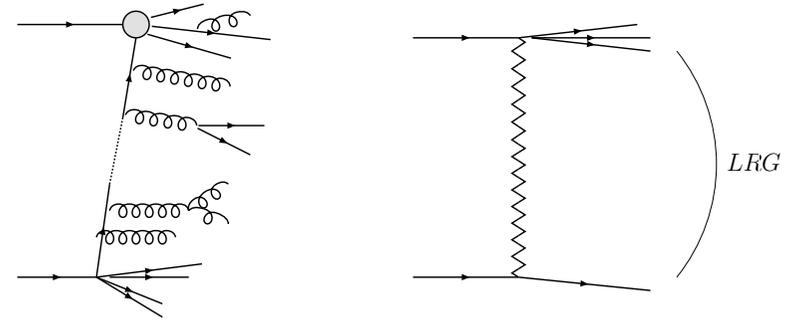
Figure 2.21: A qualitative justification of the presence of LRG in diffractive events is shown. In standard proton-proton strong interactions (left) a parton is exchanged between the hadrons and on his path emits, as predicted in QCD, other partons that will subsequently hadronise. In diffractive interactions, the exchange has no colour charge and therefore does not emit any parton.

A quantity that estimates Y for massless particles is the pseudorapidity, η, defined as

$$\eta = -ln \left[tan \left(\frac{\theta}{2} \right) \right] \qquad (2.41)$$

where θ is the polar angle of the momentum of the particle. Since only the vacuum quantum numbers are exchanged between the hadrons, no colour charge is exchanged. As depicted in Fig. 2.21, this absence of colour flow causes a region in rapidity in the direction of the scattered proton without hadrons that otherwise would be generated by the QCD radiation. Thus, using the LRG one can tag an event as diffractive by requiring a rapidity region around the beampipe with no particle flow. The size of this rapidity region between the scattered proton and the most forward particle produced in the central system, ΔY, depends by the kinematics of the process (see Sect. 2.2.3). It should be stressed that the background from non-diffractive events that accidentally have a LRG is strongly suppressed. In fact, if one assumes that the average number of partons radiated by the strong field in an interval of rapidity dY, is constant

$$< \frac{dP_{\mathrm{QCD}}}{dY} > \approx k$$

then the probability to emit accidentally no particles in ΔY is

$$P_{no-rad} \propto e^{-k\,\Delta Y} \qquad (2.42)$$

showing that the size of the rapidity gap in non-diffractive events is exponentially suppressed.

The diffractive physics described here is known as *soft* diffraction because elastic and total cross section are processes with no hard (i.e. high-energy) scale reactions.

This makes the calculations with perturbative QCD described in Sect. 2.1 impossible. However, the question of how to interprete the pomeron in QCD can still be addressed. The basic ideas for studying the partonic structure of the diffractive exchange (that means in turn a QCD interpretation of diffractive processes) are presented in the following subsections.

2.2.2   Hard diffraction

The first proposal to study diffraction in presence of a hard scale was suggested in 1985 by Ingelman and Schlein [46]. The basic idea in what is known as *resolved pomeron model* is that as the reggeons are ensembles of virtual mesons, the pomeron can be treated as a particle that is exchanged between the two hadrons. The nature and the properties of this quasi-particle are a priori unknown. The diffractive hadron-hadron reaction can be thus modelled as a two steps process. First the diffractive exchange is emitted from a proton with momentum transfer t. Then the diffractive exchange interacts with the other proton. The interesting thing is to require that in the second step of the process a hard scale is involved, like jets or heavy quark production. This would allow to test in pQCD the $I\!\!P$ internal hadronic structure like any other hadron[11]. The factorisation adopted in the resolved pomeron model is called *Regge factorisation* or also *proton vertex factorisation*. Thus, the cross section for a diffractive interaction between two hadrons, A and B, is approximated by

$$\sigma_{AB}^{\mathrm{D}} = f_{I\!\!P/A} \cdot \sigma_{BI\!\!P} \qquad (2.43)$$

where $f_{I\!\!P/A}$ represents the *pomeron flux factor*, the probability of emitting a $I\!\!P$ with given kinematic characteristics, and $\sigma_{BI\!\!P}$ represents the total cross section for the $BI\!\!P$ interactions. Regge phenomenology arguments say that the expression of the pomeron flux depends only on two variables, t and $x_{I\!\!P}$, where the latter is defined as the fraction of the momentum of the hadron A taken away by the pomeron. Its definition is

$$x_{I\!\!P} = \frac{(p_A - p'_A) \cdot p_B}{p_A \cdot p_B} \qquad (2.44)$$

where p_A is the four-momentum of the incoming hadron A, p_B is the four-momentum of the incoming hadron B, and p'_A is the four-momentum of the diffractively scattered hadronic system A'. There are many different definitions for the pomeron flux, the ones mostly used being proposed by Ingelman and Schlein [46] and by Donnachie and Landshoff [47]. It is important to remind that QCD is not playing a role at this stage and there is no motivation for the resolved pomeron model in QCD.

The first experimental proof of the presence of hard scales in diffractive processes was given by the UA8 collaboration in pp collisions at the SPS collider (Fig. 2.22) [48]. Diffractive events were selected by requiring the presence of a proton emerging intact from the interaction and detected far away from the interaction point [49]. In a subsample of diffractive events, jets with a high transverse energy were found, proving

---

[11]Of course this could sound weird at first, since it is known that the $I\!\!P$ is not a standard hadron at all. It cannot be observed outside the proton and no momentum sum rules hold for it.
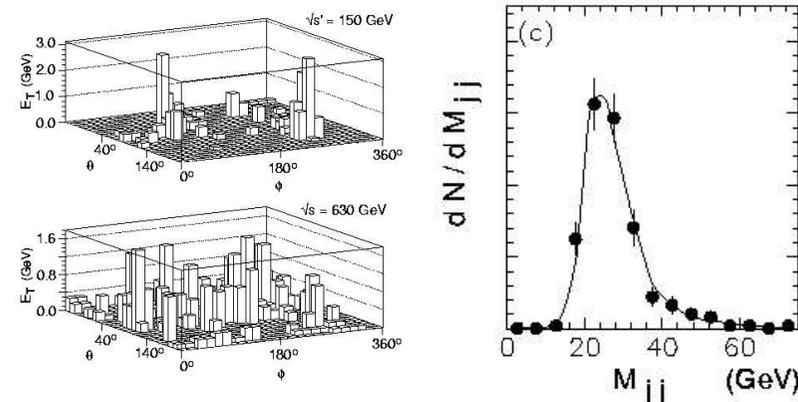
Figure 2.22: Results from the UA8 experiment. (Left) The energy distribution in the calorimeter for an event with a detected dijet system and a leading proton (upper plot) compared to the one for an event without a leading proton (lower plot). Notice that in the diffractive case there are two LRG in the low and high polar angle (θ) region of the detector. (Right) The spectrum of dijet invariant mass for events with a leading proton (from [48]).

that it was possible to find a hard scale in a diffractive process and hence to use pQCD in this kind of physics. How to do it in a practical way was nonetheless mysterious and only with subsequent data from ep collisions at HERA and $p\bar{p}$ collisions at Tevatron a clearer understanding in terms of QCD was achieved.

2.2.3   Diffraction in ep collisions

As already mentioned in Sect. 2.1, the analysis of ep collisions resulted in a great improvement of the knowledge of QCD and the internal structure of the proton. Since the study of ep interactions is a powerful tool for studying the strong interaction, it is not surprising that our understanding of diffraction was significantly improved by ep measurements. In the following, the specific case of the HERA ep collider will be considered, even though the same arguments are valid for any experiment of this kind. The motivations for studying diffraction at HERA are basically the same for other QCD studies:

- The hard scale needed by pQCD is given by the virtuality of the exchanged boson, Q^2. When producing jets and heavy quarks, additional hard scales are available and one can use pQCD also for the description of *photoproduction* reactions, i.e. processes in which the exchanged photon is almost real with $Q^2 \approx 0$.

Figure 2.23: A schematic picture of a diffractive DIS interaction as seen in a reference frame where the proton moves very fast.

- Having a lepton in the initial state reduces the hadronic multiplicity of the final state, making the experimental environment "cleaner" and easier to analyse.

- The kinematic range of HERA is very wide, ranging from a maximum Q^2 of approximately $2 \cdot 10^4 \, \text{GeV}^2$ to a minimum x_{Bj} of approximately 10^{-5}. In the following it will be shown that the latter remarkably low value is especially important for diffraction.

In ep diffractive interaction at HERA, the actual interaction takes place between the virtual photon and the proton. The γ^* dissociates into a hadronic (diffractive) final state measured in the central detector. The p can either stay intact and escape in the beam pipe (single dissociation) or dissociate in a low-mass resonant state (double dissociation). Considering processes in the kinematic region defined in Eq. (2.6), one can study diffraction in the hard DIS regime. This process is called *diffractive DIS* (DDIS). The goal of such a study is to give a quantitative QCD description of this process. The DDIS reaction can be expressed as $ep \to e$XY, where X represents the diffractive system produced centrally and Y the hadronic system at the proton vertex (in the case of single dissociation Y $\equiv$ p). The number of independent kinematic variables is five, higher than in the standard case. The usual choice for them is Q^2, t, $x_{I\!P}$, β, M_{Y}.

The variables Q^2 and t were already introduced and we refer to Eq. (2.4) and Eq. (2.35) for their definitions. The definition of $x_{I\!P}$ introduced in Eq. (2.44) can be rewritten in the DIS case as

$$x_{I\!P} = \frac{(p - p') \cdot q}{p \cdot q} \tag{2.45}$$

Although its physical interpretation stays the same as the one introduced in Sect. 2.2.2, it has to be stressed that we did not yet introduced any Regge factorisation in the definition. The variable β plays in diffractive DIS a similar role as x_{Bj} in standard DIS.

It is defined as

$$\beta = \frac{Q^2}{2\,(p - p') \cdot q} \simeq \frac{Q^2}{Q^2 + M_{\text{X}}} \tag{2.46}$$

and it can be verified immediately by comparing Eq. (2.46) with Eq. (2.45) and Eq. (2.12) that

$$x_{\text{Bj}} = x_{I\!P} \cdot \beta. \tag{2.47}$$

The variable M_{Y} is the invariant mass of the hadronic system Y. In case the proton stays intact $M_{\text{Y}} = m_p$.

In order to pursue our goal to describe quantitatively the diffractive DIS process, the first idea is simply to repeat the treatment adopted for the standard non-diffractive DIS just adding the requirement that the proton undergoes a diffractive scattering. The suffix "D" will be added to stress the latter additional requirement. A first simplification that can be done is to consider only single dissociative events, i.e. diffractive events where the proton stays intact. This choice has practical justifications. The detection of the system Y is often impossible experimentally (the hermeticity of a collider experiment cannot be total and the system Y escapes in the beam pipe). Moreover, even if one has managed to detect it, one has no more experimental signatures but the fact that only vacuum quantum numbers have been exchanged in the interaction. In fact also the exchange of reggeons produces similar experimental signatures except changing the Y quantum numbers. Nonetheless, even if we ignore this kind of reactions in the theoretical treatment, they are still produced in the collisions. The experimental way to deal with them will be explained later in Sect. 6.5. The differential DDIS cross section then depends only on four variables and can be written as

$$\frac{d^4\sigma^D}{dQ^2 d\beta dt dx_{I\!P}} = \frac{2\pi\alpha^2}{x_{\text{Bj}}Q^2} \left[(1 + (1-y)^2) F_2^{D(4)}(\beta, Q^2, t, x_{I\!P}) - y^2 F_L^{D(4)}(\beta, Q^2, t, x_{I\!P}) \right] \tag{2.48}$$

If the scattered proton is not detected in the experiment, Eq. (2.48) is integrated over t and becomes

$$\frac{d^3\sigma^D}{dQ^2 d\beta dx_{I\!P}} = \frac{2\pi\alpha^2}{x_{\text{Bj}}Q^2} \left[\left(1 + (1-y)^2\right) F_2^{D(3)}(\beta, Q^2, x_{I\!P}) - y^2 F_L^{D(3)}(\beta, Q^2, x_{I\!P}) \right] \tag{2.49}$$

Often Eq. (2.48) and Eq. (2.49) are expressed in terms of a *reduced cross section*, $\widetilde{\sigma}^{D(4)}$, defined as

$$\widetilde{\sigma}^{D(4)} = F_2^{D(4)}(\beta, Q^2, t, x_{I\!P}) - \frac{y^2}{(1 + (1-y)^2)} F_L^{D(4)}(\beta, Q^2, t, x_{I\!P})$$

As for standard DIS, the longitudinal diffractive structure function, F_L^D, is usually neglected. In this case $\widetilde{\sigma}^D$ and F_2^D match each other. The treatment of the DDIS is noticeably simplified by the use of the resolved pomeron model introduced in Sect. 2.2.2.

Then, the expression of F_2^D, which in general depends on four variables, is factorised in two parts, each one depending on two separate variables

$$F_2^{D(4)}(\beta, Q^2, t, x_{I\!\!P}) = f_{I\!\!P/p}(t, x_{I\!\!P}) \cdot F_2^{I\!\!P}(\beta, Q^2) \qquad (2.50)$$

where $f_{I\!\!P/p}$ is the pomeron flux factor and $F_2^{I\!\!P}$ the pomeron structure function. The former can be modeled by Regge phenomenology, the latter is the quantity extracted from the measurement. Again, in case one does not measure t, the expression of Eq. (2.50) is integrated over this variable. In the resolved pomeron model, the physical meaning of β arises naturally. As x_{Bj} is the fraction of the proton longitudinal momentum taken by the parton which is struck by the photon, β is the fraction of the pomeron longitudinal momentum taken by the parton struck by the photon.

Experimentally, at HERA diffraction is tagged using three different methods. They exploit different experimental signatures of diffraction and can be considered as complementary to each other.

- The single dissociation events can be directly detected by means of a forward instrumentation called *proton spectrometer* [50, 51, 52, 53, 54, 55]. Because of the low t, the outgoing p is scattered at very low angles with respect to the initial direction and one needs to place the proton spectrometer very far from the interaction point and very close to the beam axis. An experimental technique adopted at HERA is the use of *roman pots*. These devices are able to move the detectors (e.g. silicon microstrips detectors) in the direction of the beam. When the beam is not well focused yet, the detectors are kept far from it; when the beam is stable the roman pots move them close to the beam (the typical distance at HERA is $\approx 1\ cm$). The dipoles of the collider are used to bend the scattered proton such to separate them from the beam line (making their detection possible) and to allow a measurement of their momentum. If the interaction was diffractive, the scattered (*leading*) proton is perturbed only slightly. This means that it must have lost only a small fraction of its momentum. If one measures with a proton spectrometer the quantity x_L, defined as

$$x_L = \frac{|\vec{p'}|}{|\vec{p}|} \qquad (2.51)$$

one finds a pronounced peak at $x_L \approx 1$ (see Fig. 2.24). That is the diffractive peak and diffractive events concentrate at values $x_L \gtrsim 0.95$. The detection of a leading proton is a very clear indication that a diffractive process happened in the event considered. It is also rich of information because it measures directly t. However, it has the drawback of low acceptance and, consequently, low statistics and it does not reject the background from reggeon exchange.

- The large rapidity gap (LRG) method tags the diffractive events by requiring an angular region in the direction of the scattered proton without particle flow [56, 57, 58, 59, 60]. Using the definitions of rapidity, Y, and pseudorapidity, η, in Eq. (2.40) and Eq. (2.41), one can write the dependence on the kinematic
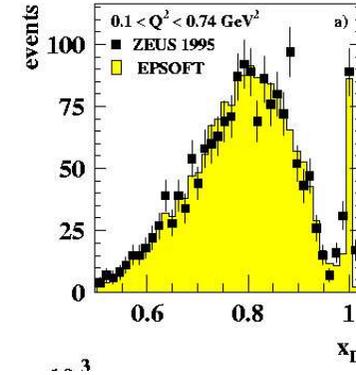
Figure 2.24: The x_L spectrum measured by the ZEUS collaboration. The points with the error bars show the data while the histogram the prediction from the Monte Carlo. The diffractive peak is clearly visible at $x_L \approx 1$ (from [52]).

variables of DIS of the size of the LRG between the diffractive system X and the scattered proton

$$Y_p - Y_{\mathrm{X}} = \Delta Y \approx ln\frac{W^2}{m_p\,M_{\mathrm{X}}} \qquad (2.52)$$

Typical values at HERA are $W = 200\,\mathrm{GeV}$ and $M_{\mathrm{X}} = 20\,\mathrm{GeV}$. Thus one expects a LRG of the size of approximately 7.7 units of rapidity. But the hadronisation of the diffractive system causes a spray of particles that decreases the detected size of the LRG, so typical requirements on the LRG size are of the order 2.5 - 3 units of rapidity. Experimentally one measures the pseudorapidity of the most forward particle in the detector (η_{MAX}) and requires that the η range between it and the edge of the forward detector instrumentation is large enough. The diffractive events concentrate therefore at low values of η_{MAX}, corresponding to large values of $\Delta\eta$ (see Fig. 2.25). The LRG method has the advantage of a much higher statistics compared to the proton spectrometer method, but since the scattered proton is not detected the source of information on t is lost. Also, the contribution from reggeons is still present. Although reggeons do not contribute to diffraction, they are a colour singlet as well and cause a LRG. It is proven that the reggeon contribution vanishes at sufficiently low values of $x_{I\!\!P}$, but a precise limit is not defined yet (at $x_{I\!\!P} < 0.01$ the reggeon contribution is absolutely negligible while at $x_{I\!\!P} < 0.03$ a small but significant contribution is present [61]). Another drawback of the LRG method is the sensitivity on the model of hadronisation which is still not well-known theoretically.

- The M_{X} method is closely connected to the LRG method. If the size of the LRG is large, Eq. (2.52) states that the ratio M_{X}/W must stay small. The
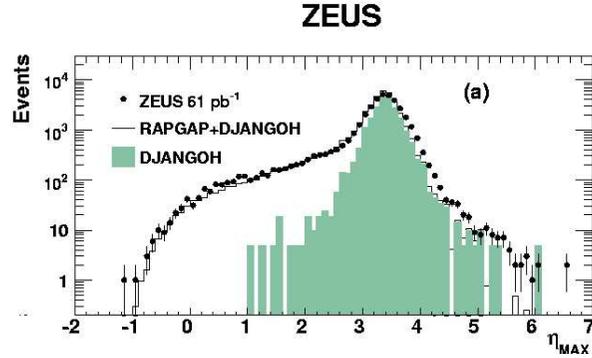
Figure 2.25: The η_{MAX} distribution as measured by the ZEUS collaboration. The points with the error bars show the data. The Monte Carlo predictions for the diffractive and non-diffractive contributions are shown as histograms. The diffractive signal gives the tail at low η_{MAX} (from [78]).

diffractive sample is found as an excess at low values of M_{X} compared to the expectations from standard DIS (see Fig. 2.26) [62, 63]. The M_{X} distribution in different bins of W is plotted and the non-diffractive contribution estimated from MC is statistically subtracted from it, leaving a diffractive sample. Like for the LRG method, also the M_{X}-method has high statistics. Moreover, the selection over the M_{X} distribution rejects also the background from reggeon exchange. The sensitivity to the hadronisation models is present also here, affecting in particular the subtraction of double dissociation events.

Once a clean diffractive sample is available, one can extract the diffractive differential cross section and the diffractive structure functions using Eq. (2.48) and Eq. (2.49). The diffractive structure functions, F_2^D, as a function of Q^2 are shown in Fig. 2.27 together with the standard DIS structure functions, F_2. While F_2 is presented in bins of x_{Bj}, F_2^D is presented in bins of β, which has a similar physical meaning.

From Fig. 2.27 an important conclusion can be made. From QCD we know that the presence of gluons causes scaling violations in the structure functions. Indeed, such violations can be seen in the F_2 plot as the Q^2 dependence of the structure function becomes steeper as x_{Bj} decreases. The same happens in F_2^D but in a much more pronounced way. The rise as a function of Q^2 starts at values of β lower than the ones of x_{Bj} in the corresponding non-diffractive plot. From this one can realise that the internal structure of the diffractive exchange is not like the one of the proton but is richer in gluons. This fact is an important reason for using jets as a probe for studying diffraction, as will be explained in Sect. 2.2.8.
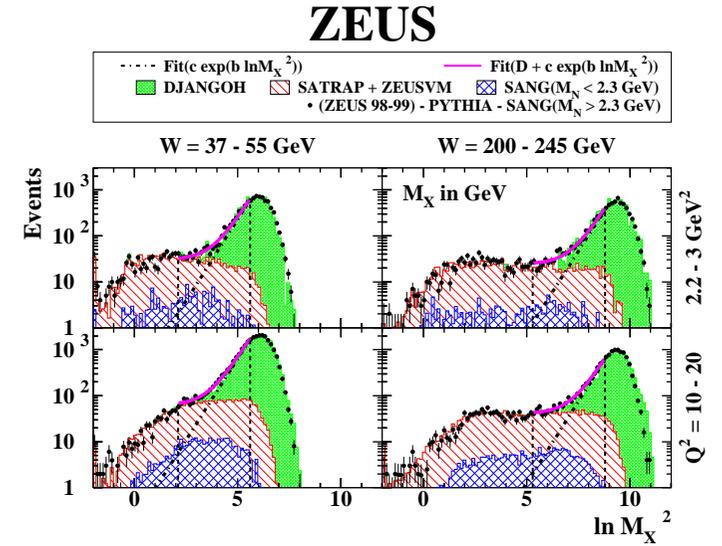
Figure 2.26: Distributions of $\ln M_{\mathrm{X}}^2$ in different (W, Q^2) bins measured by the ZEUS collaboration. The points with the error bars show the data. The hatched areas represent the MC predictions for the signal and the backgrounds. The histograms represent the different contributions as estimated with the simulations. The diffractive contribution is larger in the tail at low M_{X} (from [63]).
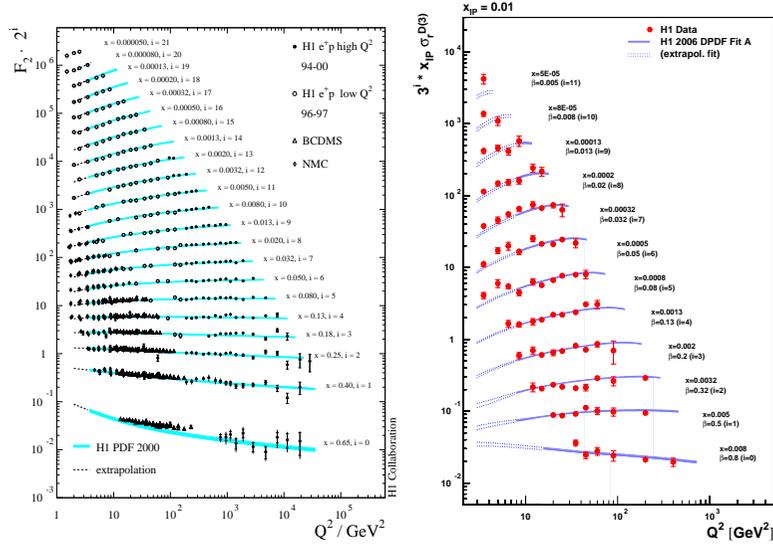
Figure 2.27: The left hand plot shows the proton structure function F_2 as a function of Q^2 in bins of x_{Bj} measured by the H1 collaboration (from [64]). The right hand plot shows the diffractive reduced cross section, $\tilde{\sigma}^D$, measured by the H1 collaboration as a function of Q^2 in bins of β at a fixed value of $x_{I\!\!P} = 0.01$ (from [60]). For visibility, the inclusive (diffractive measurements was scaled by a factor 2^i (3^i) with i as indicated. The $\tilde{\sigma}^D$ was scaled as well for different $x_{I\!\!P}$ values. The scaling violations in the right-hand plot are steeper than the left-hand plot.
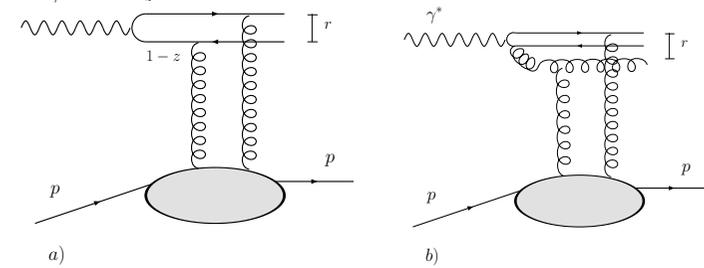

Figure 2.28: Schematic picture of the γ^*p diffractive DIS scattering in the saturation model for (a) the $q\bar{q}$ contribution and for (b) the higher-order $q\bar{q}g$ contribution.
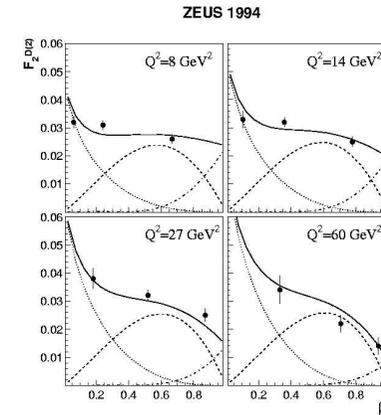
2.2.4   Saturation model in diffraction

The saturation model turns out to be a powerful tool to describe also diffraction [20, 66]. Since diffraction is concentrated at low-x_{Bj}, this kind of events are strongly influenced by saturation effects. The parameters of the model in diffractive ep scattering are the same as extracted from inclusive DIS data [20]. The same power-behaviour in x_{Bj} of inclusive and the diffractive DIS cross sections can be explained by means of saturation [20]. A schematic picture of a diffractive interaction is shown in Fig. 2.28a. The $q\bar{q}$ process is dominant at low M_X. At high M_X the contribution from an additional emission of a gluon from the dipole becomes important (Fig. 2.28b). This can be seen in Fig. 2.29 where F_2^D measured by the ZEUS collaboration is compared to the prediciton based on the saturation model. The $q\bar{q}g$ contribution is dominant at low β hence because of Eq. (2.46) at high M_X.

Figure 2.29: The diffractive structure function, $x_{I\!\!P}F_2^D(x_{I\!\!P}, \beta, Q^2)$ for $x_{I\!\!P} = 0.0042$ as a function of β measured by the ZEUS collaboration. The prediction from the saturation model is shown as a solid line. Also shown are the individual contribution (dashed lines) from the $q\bar{q}$ for transverse photons, (dot-dashed lines) from the $q\bar{q}$ for longitudinal photons and (dotted lines) from the $q\bar{q}g$ component (from [65]).

2.2.5   QCD factorisation in diffraction

The QCD-based study of diffractive DIS data relies on an important theorem only recently proven by J. Collins in 1998 [67]. The *QCD factorisation theorem for diffractive DIS* states, in analogy with the QCD factorisation theorem for DIS (see Sect. 2.1.5), that in a process with a sufficiently hard scale, the diffractive DIS reaction factorises in two parts. The short-distance part is characterised by the presence of a hard scale that includes the hard scattering of a parton off the virtual boson. This part is calculable in pQCD and depends on the process under study (inclusive DIS, jet production, heavy quark production, etc.). Since it is a pure pQCD result, the prediction for the hard subprocess does not depend on whether the reaction is diffractive or not. The second part includes the long-range soft processes and the infrared divergences are included in it. It is the equivalent of the proton PDFs and in the diffractive case is called *diffractive parton distribution functions* (dPDFs). The dPDFs can be defined as the standard PDFs with the additional requirement that the proton underwent a diffractive interaction. They depend only on the type of hadron considered and can be used in many different tuypes of processes. The differential diffractive cross section for $\gamma^* p$ interaction can then be expressed as the convolution of the short- and long-distance terms

$$\Delta\sigma^D = \sum_{i=q,\bar{q},g} \int_{\Delta t} dt \int_{\Delta x_{I\!P}} dx_{I\!P} \int_{\Delta Q^2} dQ^2 \int_{\Delta\beta} d\xi \, \hat{\sigma}^{\gamma^* i}(Q^2,\xi) f_i^D(x_{I\!P}, t, \xi, Q^2) \qquad (2.53)$$

where $\hat{\sigma}^{\gamma^* i}(Q^2,\xi)$ represents the partonic cross section for the hard interaction between the γ^* and the parton i, and $f_i^D(x_{I\!P}, t, \xi, Q^2)$ are the dPDFs for a given type of parton i.

It is wortwhile to remind that the validity of the factorisation theorem relies on neglecting higher-twist terms (see Sect. 2.1.5). This statement is valid for every available demonstration of the factorisation theorem but is particularly relevant for the diffractive case where higher-twist terms can be larger than in the inclusive case. Thus, in order to provide a solid basis to the QCD interpretation of diffraction, it is fundamental to prove experimentally the validity of this very important theorem. Once it is found to be compatible with data, one can use for diffractive processes all the QCD machinery described previously and thus include diffraction in the QCD framework.

At HERA, the dPDFs have been determined within the QCD DGLAP formalism by means of fits to inclusive diffractive DIS measurements with a procedure similar to that used to extract the standard proton PDFs from inclusive DIS data. An example of the outcome of one of these fits is shown in Fig. 2.30.

In most of the dPDF parameterisations, the Regge factorisation introduced in Sect. 2.2.2 is assumed in order to factorise the $(x_{I\!P}, t)$ from the (β, Q^2) dependence of the cross section. In the Regge approach, diffractive scattering proceeds via the exchange of the Pomeron trajectory. The dPDFs are then written as the product of $f_{I\!P/p}$, the Pomeron flux (dependent on $x_{I\!P}$ and t) and $f_{I\!P}^D$, the parton distributions in
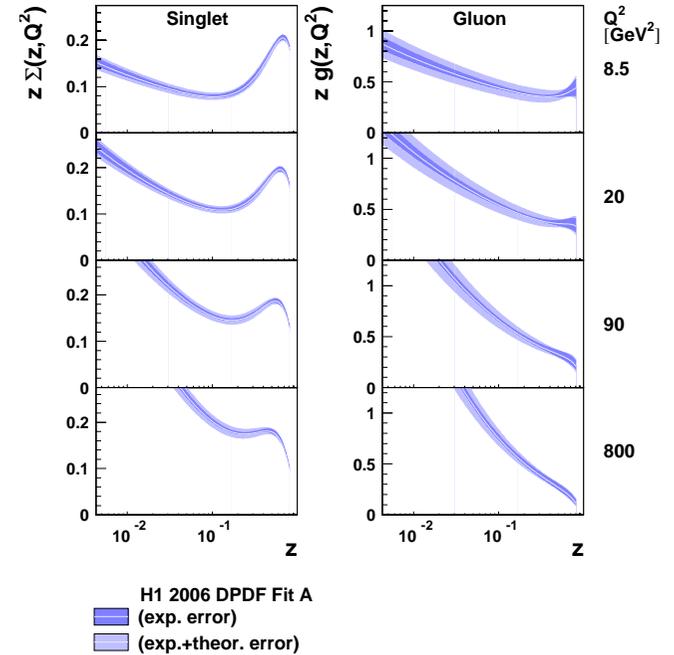

Figure 2.30:   The diffractive parton densities (dPDFs) estimated with the H1 2006 − FitA *extracted from inclusive diffractive DIS data by the H1 collaboration (from [60]). The dPDFs are presented as a function of β for different bins in Q^2. The left-hand plots show the contribution from the quark colour singlet while the right-hand plot the contribution from the gluon colour singlet. Notice the different scales on the $y-$axis.*

the Pomeron (dependent on β and Q^2)

$$f_i^D(x_{I\!\!P}, t, \xi, Q^2) = f_{I\!\!P/p}(x_{I\!\!P}, t) \cdot f_i^{I\!\!P}(\beta, Q^2). \tag{2.54}$$

The diffractive structure function can be expressed as in Eq. (2.50)

It has to be kept in mind that from the experimental point of view the selected signal can still contain a contribution from reggeons that mimic a diffractive process (this depends on the experimental technique adopted). In this case Eq. (2.54) becomes

$$f_i^D(x_{I\!\!P}, t, \xi, Q^2) = f_{I\!\!P/p}(x_{I\!\!P}, t) \cdot f_i^{I\!\!P}(\beta, Q^2) + f_{I\!\!R/p}(x_{I\!\!P}, t) \cdot f_i^{I\!\!R}(\beta, Q^2) \tag{2.55}$$

where in analogy with the diffractive case, $f_{I\!\!R/p}$ parametrises the flux of reggeons from the proton and $f_i^{I\!\!R}$ are the reggeon parton densities. Also Eq. (2.50) changes in

$$F_2^{D(4)}(x_{I\!\!P}, t, \xi, Q^2) = f_{I\!\!P/p}(x_{I\!\!P}, t) \cdot F_2^{I\!\!P}(\beta, Q^2) + f_{I\!\!R/p}(x_{I\!\!P}, t) \cdot F_2^{I\!\!R}(\beta, Q^2) \tag{2.56}$$

where $F_2^{I\!\!R}$ is the reggeon structure function.

2.2.6   Factorisation breaking in $p\bar{p}$ collisions

One of the most relevant assertions of the factorisation theorem is that the dPDFs are universal, i.e. they can be employed in different kind of interactions and for different kind of final states. Thus one should be able to use the dPDFs extracted in ep collisions at HERA also in $p\bar{p}$ collisions at Tevatron. A diffractive process with a hard scale that can be studied at Tevatron is for example the diffractive production of dijets. This measurement was performed and the structure function for diffractive dijet production, F_2^{JJ}, extracted. The measurement is then compared to the NLO QCD predictions using the HERA dPDFs in Fig. 2.31 [68].

In the comparison an evident discrepancy is observed between data and NLO predictions. The fact that the dPDFs extracted in ep collisions cannot be used directly at a $p\bar{p}$ collider means that the collinear factorisation is broken. The widely accepted explanation for such a breakdown focuses on the secondary interactions between the spectator partons in the protons. After the diffractive reaction it is still possible that two partons - that did not enter in the former - interact. The second interaction is typically soft and spreads particles towards the direction of the incoming hadrons. This makes the experimental detection of the diffractive scattering impossible. The rapidity gap is filled with the products of the second soft interaction and the proton breaks up and can not be detected in the proton spectrometer. Therefore, there are fewer diffractive events than expected without secondary rescattering. One can define the probability that a rapidity gap produced in hadron-hadron collisions is not filled by secondary processes in the same collision. This probability is called rapidity gap *survival probability*, ω. The survival probability can be expressed as the product of two terms [69]
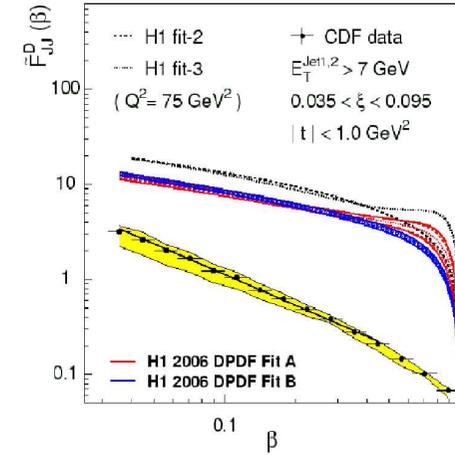
$$\omega = S^2 T^2$$

Figure 2.31: The diffractive dijet structure function measured by the CDF collaboration compared to a NLO QCD calculation using the HERA dPDFs. The same calculation using two different dPDFs (the H1 2006 - Fit A *and the* H1 2006 - Fit B*) is shown. A significant overestimation of the data is observed in the theoretical predictions, implying a breaking of the collinear factorisation in diffractive $p\bar{p}$ interactions.*
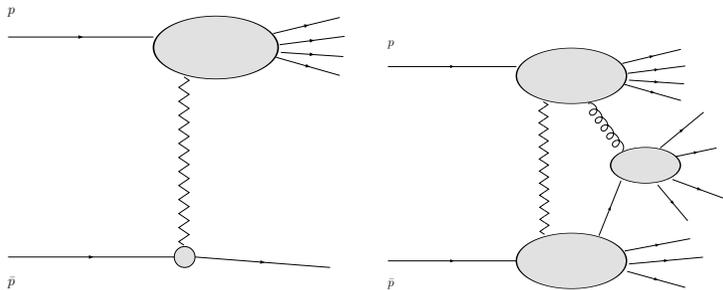
Figure 2.32: A schematic representation of the mechanism of factorisation breaking in $p\bar{p}$ collisions. In the left plot a diffractive event is depicted. No secondary interactions happen and the final state in this example has one rapidity gap and one hadron stays intact and is detectable with a proton spectrometer. In the right-hand plot two partons that did not participate to the hard diffractive process interact softly. The latter interaction spoils the rapidity gap(s) and perturbates enough the leading hadron(s) such to make impossible the diffractive tagging of the first hard process.

First there is the term S, where $(1 - S^2)$ is the probability that the rapidity gap may be filled by secondaries produced (via parton rescattering) in the underlying soft interaction. Second, there is also the probability $(1 - T^2)$ that the gap may be populated by extra-gluons emitted in the hard diffractive subprocess. The final value of ω depends on the type of interaction, the energy of it and the final state considered. The actual calculation of ω was performed by many theoretical groups [70, 71] and is able to reproduce the factor $\omega \approx 0.05 - 0.2$ needed to describe the Tevatron results. It is important nonetheless to confirm the general validity of these calculations in order to apply them in other experimental environments like the Large Hadron Collider [71]. In this task, HERA can provide precious informations.

2.2.7   Test of QCD factorisation in ep collisions

HERA diffractive data can be used not only to extract the dPDFs but also to verify their universality. The proposed explanation of the factorisation breaking at Tevatron described in Sect. 2.2.6 can be tested also in ep collisions. A photon can fluctuate in a quark-antiquark pair. In the Vector Meson Dominance model (VMD) [72] this $q\bar{q}$ fluctuation behaves approximately like a vector meson (ρ, ω, ϕ...). The lifetime of this fluctuation is proportional to $1/\sqrt{Q^2}$ and for low virtualities it can be sufficiently long that the photon exhibits hadronic structure during the interaction with the proton. This means that in photoproduction events (γp), i.e. when the photon emitted by the lepton is quasi-real ($Q^2 \approx 0$), the interaction may actually occur between two hadronic objects. In this way one is able to test whether soft rescattering effects occur also in photoproduction and if the models used for $p\bar{p}$ collisions also describe the γp scattering.
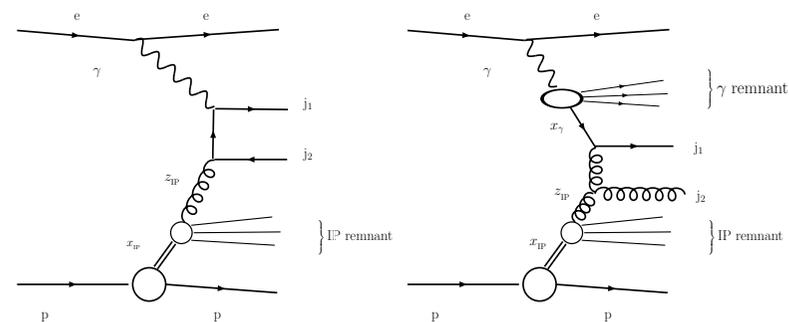
Figure 2.33: Direct and resolved photon LO diagrams contributing to the production of diffractive dijets. Regge factorisation is assumed in these plots.

In LO, the γp events are grouped in *direct* and *resolved* γp (see Fig. 2.33). In the former case, the photon, although real, does not fluctuate and couples directly to the quark in the proton. The resolved photon processes are those where the photon exhibits a hadronic structure. As for the proton, a collinear factorisation is applied for the γ, defining a hard subprocess between the parton coming from the p and the parton coming from the γ, and photon PDFs extracted experimentally from e^+e^- data [73]. The fraction of the photon longitudinal momentum entering the hard subprocess is called x_γ. In the photon, the variable x_γ plays the same role as x_{Bj} in the proton and can be defined as

$$x_\gamma = \frac{p \cdot u}{p \cdot q} \tag{2.57}$$

where u is the four-momentum of the parton originating from the γ entering the hard subprocess.

Secondary rescattering is not expected to happen in DIS interactions and direct photon processes simply because the proton does not have a hadronic counterpart to rescatter on. Conversely, the resolved part should be suppressed. The experimental separation between direct and resolved is based on the value of x_γ. For direct processes, $x_\gamma = 1$ (neglecting detector resolution effects) while for resolved processes the value of this variable can be significantly lower.

The experimental strategy for testing the QCD factorisation at HERA is to compare the data from a diffractive final state to the NLO prediction using the dPDFs previously extracted from inclusive DIS. Of course, the data sample under test must not have been used in the dPDFs extraction. If the factorisation holds, the NLO should describe the data, both in shape and normalisation. Thus, according to the models used to describe the $p\bar{p}$ diffractive reactions, one expects to see in the low x_γ region of the phase space a suppression compared to the NLO prediction using the dPDFs. Such a suppression should not be visible in DIS and direct γp events.
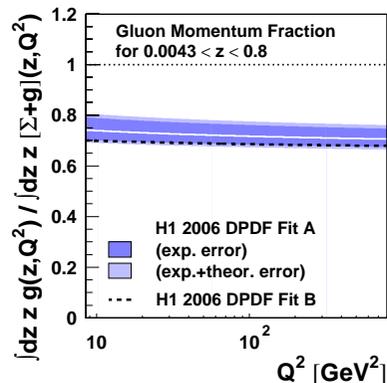
Figure 2.34: The ratio of gluons in the dPDFs as a functions of Q^2 estimated from the fits H1 2006 − FitA *and* H1 2006 − FitB*. In this plot the variable called β in the text is denoted by z (from [60]).*

The diffractive reactions suitable for the test of QCD factorisation must satisfy the following requirements:

- it must have a hard scale in the final state such as to allow the use of pQCD, like the factorisation theorem and the DGLAP evolution equations;

- the process must have a sensitivity to the gluon content of the diffractive exchange. This requirement is due to the enhancement of gluons in the diffractive exchange. Fig. 2.34 shows that the gluon content in the diffractive exchange is approximately $70 - 80\%$. Thus it is important to choose processes that are sensitive to the gluons in order to have a direct handle on the main contribution to the dPDFs.

Two processes satisfying these requirements and widely used at HERA for QCD diffractive studies are the production of open charm and of dijets [75, 76, 77, 78, 79, 80, 81]. Charm is produced dynamically by the gluons in the diffractive exchange (see Fig. 2.35) and hadronises in a fraction of the events in D^* mesons whose detection is then used as signature of charm production.

The consistency observed between the cross sections for the semi-inclusive DIS processes considered and the respective calculations based on Eq. (2.53) using the available dPDFs represents an experimental support of the validity of the QCD factorization hypothesis in diffractive DIS [77, 80]. Nonetheless the theoretical uncertainty on the NLO calculations is large and there is still the need for a better experimental verification of the theorem. The expected suppression in γp is observed only in dijet production and not in open charm [76, 77]. The latter result can be understood considering that the charm production is suppressed when the photon has a hadronic structure [76]. In
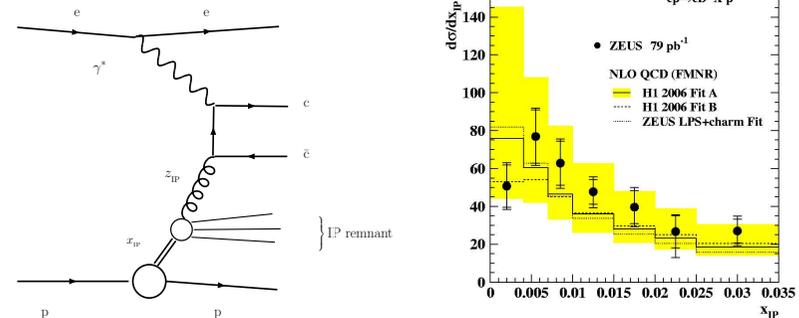
Figure 2.35: The left plot shows a schematic representation of diffractive charm production in the Regge factorisation framework. The right plot is a recent result from the ZEUS collaboration on diffractive production of D^* mesons in γp: the differential cross section as a function of $x_{I\!P}$ is shown (points and error bars) and compared to the NLO calculation using different dPDFs as indicated in the legend (from [79]).

the case of dijets in diffractive photoproduction, experimental results show an overestimation of the cross section by the NLO calculation [79, 80], although in different amount among the experiments. In both cases the suppression is observed in the entire x_γ region, contradicting the model which works at the Tevatron. The subject is still under study at the moment of writing this thesis.

2.2.8 Diffractive dijets in DIS

The study of dijet production in DDIS can provide very relevant informations about diffraction. This process is a perfect candidate to fulfill the requirements specified in Sect. 2.2.7. There are two hard scales in the process, the virtuality of the photon and the transverse energy of the jets. They guarantee to select events in the perturbative regime needed by the QCD calculation. The sensitivity to the gluon content in the diffractive exchange is given by the production mechanism of the dijets. At the LO, the dijets can be produced via the QCD Compton (QCDC) and the Boson-Gluon Fusion (BGF) processes (see Fig. 2.36). At HERA, the BGF process is dominant in the kinematic region $Q^2 \lesssim 500$ GeV^2 while for higher Q^2 the QCDC process becomes more important [82]. On the other hand, diffraction is enhanced at low values of Q^2. Thus, the diffractive dijet production is a BGF dominated process and has a direct sensitivity to the gluon dPDFs. Compared to diffractive open charm production, dijets have typically harder scales and higher statistics. The latter asset is given both by the higher cross section and by the better experimental detection efficiency.

Studying dijets an additional variable is required, $z_{I\!P}$, the fraction of the momentum
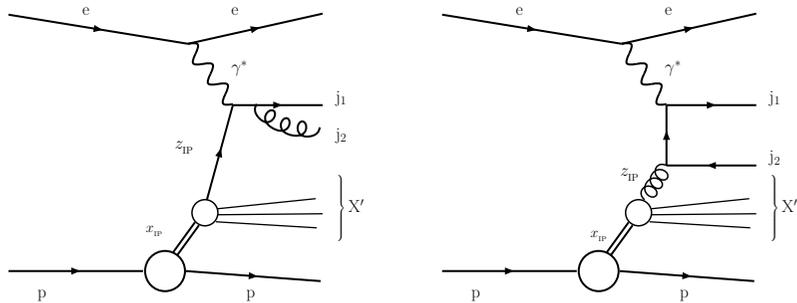
Figure 2.36: The leading-order QCD diagrams for dijet production in single dissociation diffractive DIS. The Regge factorisation is assumed. The left-hand plot represents the QCD Compton (QCDC) production while the right-hand plot the Boson-Gluon Fusion (BGF) diagram.

of the diffractive exchange carried by the parton participating in the hard process and defined as

$$z_{\mathbb{P}} = \frac{q \cdot v}{q \cdot (p - p')} \tag{2.58}$$

where v is the four-momentum of the parton originating from the colourless exchange entering the hard subprocess with the γ^*. In diffractive production of dijets, the variable $z_{\mathbb{P}}$ replaces β as the kinematical variable on which the dPDFs depend . Eq. (2.53) is rewritten as:

$$\Delta\sigma_{jj}^D = \sum_{i=q,\bar{q},g} \int \Delta t dt \int_{\Delta x_{\mathbb{P}}} dx_{\mathbb{P}} \int_{\Delta Q^2} dQ^2 \int_{\Delta z_{\mathbb{P}}} d\xi \, \hat{\sigma}_{jj}^{\gamma^* i}(Q^2, \xi) f_i^D(x_{\mathbb{P}}, t, \xi, Q^2), \tag{2.59}$$

where now $z_{\mathbb{P}}$ is the variable sensitive to the dPDFs and the subprocess cross section $\sigma^{\gamma^* i}$ is replaced by the cross section, $\sigma_{jj}^{\gamma^* i}$, for the reaction $\gamma^* i \rightarrow$ jet1 jet2.

The dijets in DDIS can be used as a benchmark for the dPDFs and the factorisation theorem, as described in Sect. 2.2.7. On the other hand, if one assumes the validity of the factorisation theorem one can use the same data for improving the accuracy of the dPDFs. The dijets data can be used to constrain the dPDFs in a combined fit with the inclusive data. This technique was already used in the context of proton PDFs [83] and it has proven to significantly reduce the uncertainties in the PDFs, especially the gluon PDFs. The reason is that in a fit using only inclusive data the gluon parton densities are extracted only in an indirect way from the Q^2 dependence of F_2 (which is connected to the amount of gluons via the DGLAP equations) and the conservation of the momentum sum rules in the proton. By selecting dijets, the data measure directly the fraction of proton momentum taken by the gluon entering the hard
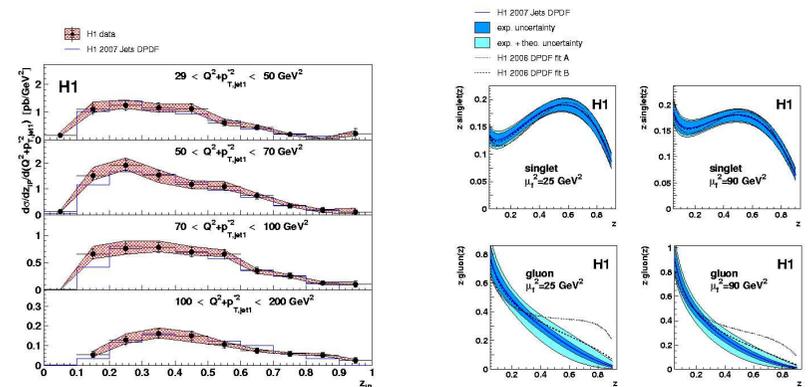
Figure 2.37: The left-hand plot shows the measurement performed by the H1 collaboration of the double differential cross section for production of dijets in DDIS as a function of $z_{\mathbb{P}}$ and $Q^2 + p_{T,jet1}^{*\,2}$ (points with error bars). The curves represent the NLO predictions using the fit including inclusive and dijets data (H1 2007 Jets dPDFs). The right-hand plot presents the "H1 2007 Jets" dPDFs compared to two older fits that used only inclusive data as a function of $z_{\mathbb{P}}$ in bins of the factorisation scale, μ_F, separately for the quark and the gluon contribution (from [81]).

subprocess[12]. This property makes the inclusion of dijets data in a fit of the dPDFs even more attractive. In fact the particularly richness of gluons in the dPDFs increases the impact of an improvement in the gluon parton density estimation. Moreover, the diffractive data have a larger experimental uncertainty compared to the standard DIS ones. This reflects in an even worse precision in measuring the Q^2 dependence of F_2^D and, consequently, the gluon dPDFs have much larger relative uncertainties compared to the proton PDFs. This uncertainty is so relevant that the inclusive data alone are able to constrain the gluon parton densities only up to $\beta \lesssim 0.8$ and, in addition, make specific assumptions on their initial parametrisation at the starting scale of the DGLAP evolution [60].

The inclusion of dijets in DDIS in the fit of the dPDFs was recently performed by the H1 collaboration [81]. In Fig. 2.37 two results from this analysis are shown in order to give a feeling of the status of the inclusion of dijets in the dPDFs fit. Fig. 2.37a shows the double differential cross section for production of dijets in DDIS as a function of $z_{\mathbb{P}}$ and the chosen estimator of the hard scale of the process, $Q^2 + p_{T,jet1}^{*\,2}$, where $p_{T,jet1}^*$ is the transverse energy of the jet with the highest transverse energy as measured in the reference frame where the $\gamma^* p$ centre-of-mass is at rest. The precision, kinematic coverage and robustness of the fit are greatly improved.

Thus, the measurement of dijets can dramatically improve our knowledge of the

---

[12]In non-diffractive analysis this quantity, analogous to $z_{\mathbb{P}}$, is called ξ.

diffractive dynamics in a QCD framework and have an impact on future analyses also at other colliders. The study of diffraction at the LHC, the new pp collider with a centre-of-mass energy $\sqrt{s} = 14\,TeV$, has a broad and interesting physics program [84]. Similarly to the non-diffractive case, a fundamental ingredient of any calculation of diffractive processes are the dPDFs. Reducing the dPDFs uncertainty can strongly reduce the uncertainty on at the LHC, e.g. diffractive Higgs production. The study of the production of dijets in diffractive DIS was never pursued with the ZEUS detector.

Chapter 3

Experimental setup

The first part of this chapter describes the *ep* collider HERA and its main features including a brief description of the upgrade carried out in the years 2000-2003; the second part is focused on the general purpose detector ZEUS and its components which played a relevant role in the analysis.

3.1 The HERA *ep* collider

The Hadron Elektron Ring Anlage (HERA) is a collider located at the DESY research center in Hamburg, Germany. The tunnel where the accelerator is placed is 6.3 Km long and at a depth of about 20 meters below the ground. At HERA either electrons or positrons were collided against protons. The leptons and the protons were circulating in two separate rings and brought to collision in four different interaction points. The leptons were accelerated using normal and superconducting cavities while for the protons normal cavities were used. The leptons were kept in orbit using 0.3 T room-temperature dipole magnets; superconducting 5 T dipoles were adopted instead for the heavier protons. The first period of operation of HERA (HERA I) lasted from 1992 to 2000 colliding both electrons and positrons at 27.5 GeV against protons. The starting proton beam energy was 820 GeV (centre-of-mass energy of 300 GeV) and switched to 920 GeV (centre-of-mass energy of 318 GeV) in 1998. In 2000 the operations were stopped for allowing a luminosity upgrade of the machine bringing to a planned increase of the specific luminosity[1] by a factor 5. In addition the lepton beam was longitudinally polarised. This second stage of operations was called HERA II. During the HERA II run the proton energy changed from the initial 920 GeV to 460 GeV and 575 GeV. The energy of the lepton beam was kept fixed to 27.5 GeV. A more detailed description of the machine parameters and the data taking periods can be found in Tables 3.1 and 3.2.

At the interaction points four experiments collected data. In the North Hall the multi-purpose experiment H1 worked. In the West Hall the HERA-B collided the beam-halo protons against a fixed target in order to produce B mesons for performing CP violation studies; HERA-B stopped the data taking in 2003. HERMES was the

---
[1]What the specific luminosity is

HERA parameters

Parameter	Value	
Circumference	6336 m	
Centre of mass energy	225 / 251 / 300 / 318 GeV	
Beam Cross Over	96 ns	
Luminosity	$1.6\times10^{31}\,cm^{-2}s^{-1}$ / $7.0\times10^{31}\,cm^{-2}s^{-1}$	
	Proton beam	Electron beam
Nominal energy	460 / 575 / 820 / 920 GeV	27.6 GeV

Table 3.1: The main parameters of the HERA collider.

experiment in the East Hall: it used the lepton beam with a hydrogen fixed target to study the spin structure of the proton. In the South Hall the second multi-purpose detector at HERA worked, ZEUS. A more detailed description of ZEUS is present in the next section.

3.2 The ZEUS detector

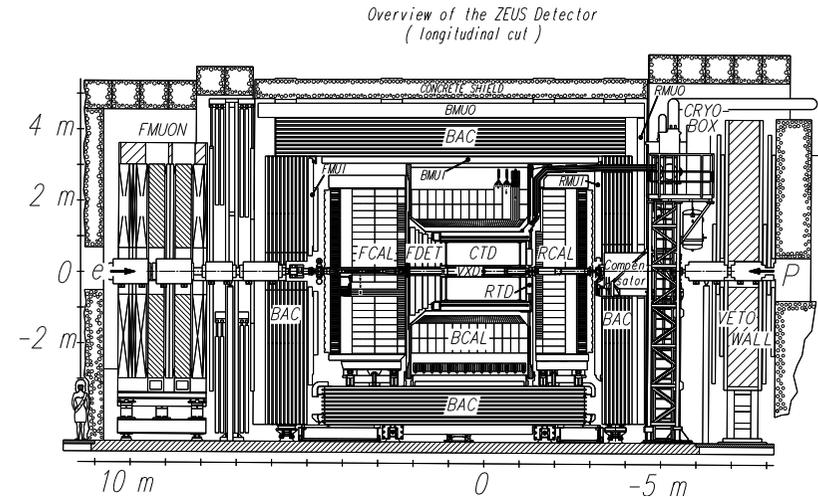### 3.2.1 Overview of the ZEUS Detector

Figure 3.1: View of the ZEUS detector along the beam direction.

The ZEUS detector [85] was a general purpose detector designed to study various aspects of electron–proton scattering. It has been in operation from 1992 to 2007.

Various components were installed to measure final state hadrons and leptons and to characterise observed final state in terms of particle energy, direction, and type.

The experiment consisted of the main detector, surrounding the nominal interaction point and several small components positioned along the beam line in both directions (positive and negative Z)[2]. The schematic view of the main detector is shown in Figs. 3.1 and 3.2. The design of the detector was not symmetric with respect to the nominal interaction point ($Z = 0$). The difference in the energy of the electron (positron) beam (27.5 GeV) and proton beam (820/920 GeV) resulted in a large boost of the centre-of-mass system in the direction of the proton beam and in the large forward–backward asymmetry of the particle production. Therefore the forward part of the detector was more instrumented than the rear one.
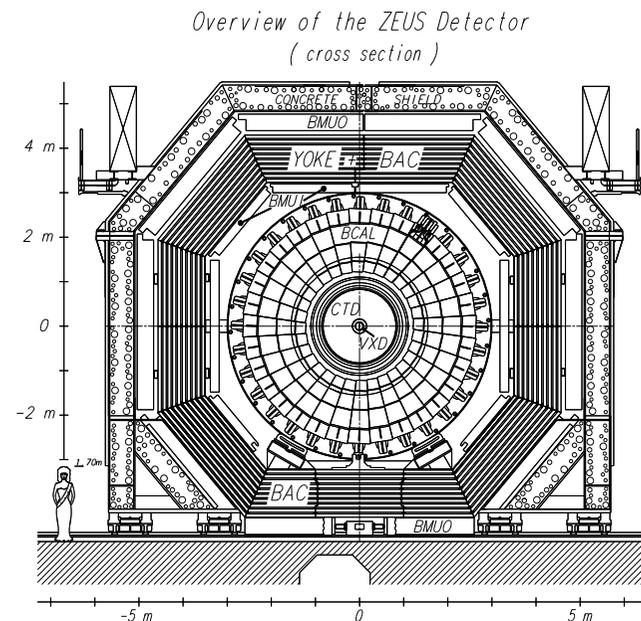
HERA luminosity

Period	Colliding particles	Proton beam energy (GeV)	Luminosity delivered (pb^{-1})
1993	$e^- p$	820	0.54
1994	$e^+ p$	820	5.2
1994	$e^- p$	820	1.0
1995	$e^+ p$	820	12.3
1996	$e^+ p$	820	17.1
1997	$e^+ p$	820	36.4
1998	$e^- p$	920	8.0
1999	$e^- p$	920	17.1
1999	$e^+ p$	920	28.5
2000	$e^+ p$	920	66.4
2003	$e^+ p$	920	6.5
2004	$e^+ p$	920	77.9
2004/05	$e^- p$	920	204.8
2006	$e^- p$	920	86.1
2006	$e^+ p$	920	118.4
2007	$e^+ p$	920	62.2
2007	$e^+ p$	460	17.7
2007	$e^+ p$	575	9.4

Table 3.2: The luminosity delivered by HERA.

Figure 3.2: View of the ZEUS detector perpendicular to the beam direction. See text for a description of the components.

The tracking system enclosed by a superconducting solenoid producing an axial magnetic field of 1.43T formed the inner part of the main detector. The main component of the tracking system was the *Central Tracking Detector* (CTD), a cylindrical drift chamber, surrounding the beam pipe at the interaction point. The CTD measured

[2]The ZEUS coordinate system is a right-handed Cartesian system, with the Z axis pointing in the proton beam direction, referred to as the "forward direction", and the X axis pointing left towards the center of HERA. The coordinate origin is at the nominal interaction point.See also Appendix A.

charged tracks in a polar angle range of $15° < \theta < 164°$. The CTD will be described in more details in Sect. 3.2.2. In order to extend the angular coverage of track reconstruction in the forward and backward directions, the CTD was supplemented by the *Forward Tracking Detector* (FTD) and the *Rear Tracking Detector* (RTD). Both FTD and RTD were composed by three sets of planar drift chambers. In the forward direction the FTD planes were interleaved with the *Transition Radiation Detector* (TRD) planes[3]. The *Small angle Rear Tracking Detector* (SRTD) was placed behind the RTD and covered the surface of the RCAL from the beam–pipe hole to a radius of about 34 cm. Its purpose was to measure electrons scattered at very small angles ($\theta > 170°$) outside the RTD acceptance. At the start of the operations, the ZEUS detector was also equipped with the vertex detector (VXD) designed to enhance reconstruction of the event vertex and possible secondary vertices, and to improve the momentum and angular resolutions of charged tracks measured with the CTD and other tracking detectors. It was however removed during the shutdown of 1995/96 as it could be not operated continously due to the high beam background levels. During the HERA upgrade in 2001 a silicon microstrip detector was installed at the same place providing high-precision tracking measurements.

The high resolution uranium calorimeter (UCAL) surrounded the central part of the detector, i.e. the tracking detectors and the superconducting solenoid. The UCAL was subdivided into the forward (FCAL), barrel (BCAL), and rear (RCAL) parts. It was used to measure energies of produced particles as well as to reconstruct their position and to discriminate between electrons and hadrons. The UCAL will be described in detail in Sect. 3.2.3. The *Forward Plug Calorimeter* (FPC) was a calorimeter installed during the data taking period 1998–2000 in the forward direction in order to extend the angular coverage of the UCAL. The *Hadron Electron Separator* (HES) was a plane of silicon diodes designed to improve the identification of electromagnetic objects, including non-isolated ones. The HES was installed inside the the RCAL (*Rear Hadron Electron Separator*, RHES) and inside the FCAL (*Forward Hadron Electron Separator*, FHES) at a depth of 3.3 radiation lenghts, approximately the position of the electromagnetic shower maximum.

The UCAL was surrounded by an iron yoke made of 7.3 cm thick iron plates. The yoke provided a return path for the solenoid magnetic field. In the addition to the return field of the solenoid, the yoke is magnetised to 1.6 T by copper coils producing toroidal field. At the same time it was instrumented with the *the backing calorimeter* (BAC). The BAC consisted of proportional chambers making it possible to measure energy leakages out of the UCAL and to reconstruct high energy muons. The limited streamer tubes (*muon chambers*) were mounted inside and outside of the iron yoke in the barrel (BMUI, BMUO) and the rear (RMUI, RMUO) regions to enhance muon identification and to measure their momentum. As the average particle density and the muon momentum in the forward direction was higher than in the barrel and rear directions, the muon system in the forward direction was more complex. It consisted of limited streamer tubes mounted inside of the iron yoke (FMUI) as well as drift chambers and limited streamer tubes mounted in front of the iron yoke (FMUO). Two

---

[3]After the HERA 2000 upgrade the TRD was replaced by the *Straw Tube Tracker* (STT). The STT improved tracking efficiency in events with high multiplicities.

iron toroids provide a toroidal magnetic field of 1.7 T for measurement of forward muon momenta.

In the backward direction, at Z = −7.3 m, the so called VETO WALL detector was placed. It consisted of concrete wall and large scintillator planes, and was used to reject background events coming from proton–beam-gas reactions in the HERA tunnel. Downstream of the electron beam, luminosity was measured by the luminosity monitor (LUMI). The LUMI detector could also be used for physics analysis, to tag the so called photoproduction events.

3.2.2 The Central Tracking Detector

The Central Tracking Detector (CTD) [86] was a cylindrical gas-filled wire chamber. It provided measurements of the charge and momentum of charged particles. The fiducial volume of the CTD extended from $Z = −1.01 \ m$ to $z = 1.06 \ m$. It covered a polar angle of $15° < \theta < 164°$ and the full range of the azimuthal angle ϕ. The outer and inner diameter were 159 cm and 36 cm respectively. The chamber is filled with a mixture of argon, CO_2, and ethane. The CTD was subdivided into eight sections (octants) and nine superlayers. One octant is shown in Fig. 3.3. Each octant consisted of 72 drift cells equipped with eight sense wires each. For each sense wire, the position in the (X, Y) plane of the charged particle track in the cell (so called hit) was reconstructed from the drift time measurement. The total number of cells in the CTD increased from 32 in the innermost superlayer to 96 cells for the outermost superlayer. Every other superlayer had its sense wires rotated by a certain angle with respect to the beam axis. The angles for each superlayer are given in Fig. 3.3. With this configuration the Z position of a track could be reconstructed with an accuracy of approximately 2 mm.
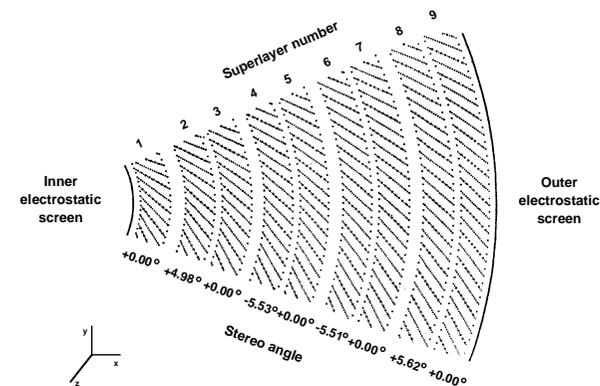

Figure 3.3: Layout of a CTD octant. Each octant has nine superlayers with the even numbered ones declined with respect to the beam axis ('Stereo angle').

The position resolution in $r - \phi$ was approximately 230 μm and the resolution of the transverse momentum, p_t, measurement for charged particles crossing all layers was:

$$\frac{\sigma(p_t)}{p_t} = 0.0058 \cdot p_t(\text{GeV}) \oplus 0.0065 \oplus \frac{0.0014}{p_t}, \tag{3.1}$$

where the first term was due to the resolution in the hit position determination, the second term to smearing from multiple scattering within the CTD and the last term to multiple scattering before the CTD. The best transverse momentum resolution of about 0.8% was obtained for $p_t \approx 0.5$ GeV and the measurement better then 10% was possible for
$p_t \leq 17$ GeV.

From the fit to all reconstructed tracks, the position of the interaction point could be reconstructed with a typical resolution of 0.1 cm in X and Y and of about 0.4 cm in Z. Charged particle identification in the CTD was possible in the limited kinematic range by measurement of the mean energy loss, dE/dx, along the tracks.

3.2.3   The Uranium-Scintillator Calorimeter

The ZEUS calorimeter (UCAL) was a sampling calorimeter, built of depleted uranium absorber plates interleaved with scintillator layers. The scintillating light coming from the latter was collected with an optical readout. The goal was to obtain an almost hermetic calorimeter with a nearly full solid–angle coverage and a very good hadronic energy resolution. The latter was achieved by the compensation, i.e. equal response to electromagnetic and hadronic cascades.

The layout of the UCAL is shown in Fig. 3.4. It was divided into three parts, which covered different polar angles [87, 88, 89].

- the *forward calorimeter* (FCAL) covering $\theta = 2.2°$ to 39.9°, with a total depth of 7.1 interaction length; the front face of the FCAL was placed at $z = 234.4$ cm, i.e. 234.4 cm away from the nominal IP,

- the *barrel calorimeter* (BCAL) covering $\theta = 36.7°$ to 129.1°, with a total depth of 5.3 interaction length; the inner radius of the BCAL was $R = 134.5$ cm,

- the *rear calorimeter* (RCAL) covering $\theta = 128.1°$ to 176.5°, with a total depth of 4.0 interaction length; the RCAL face was placed at $z = 162.2$ cm.

Uranium was found to be a very suited absorber for hadron calorimetry, since it provides a high yield of spallation neutrons, which in turn can transfer their energy to the atoms in the scintillator. Together with an additional contribution of photons from neutron capture in the uranium, this helps compensating the energy losses in hadronic cascades arising from the binding energy, nuclear fission and from undetected decay products. Electrons and photons do not suffer such losses. Hence, in typical non–compensating sampling calorimeters, average response to electrons or photons is about $20 - -40\%$ higher than to hadronic cascades with the same energy. Due to fluctuations in the contribution of the electromagnetic component (originating mainly from the π^0 decays) to the hadronic cascades this leads to poor hadronic resolution.
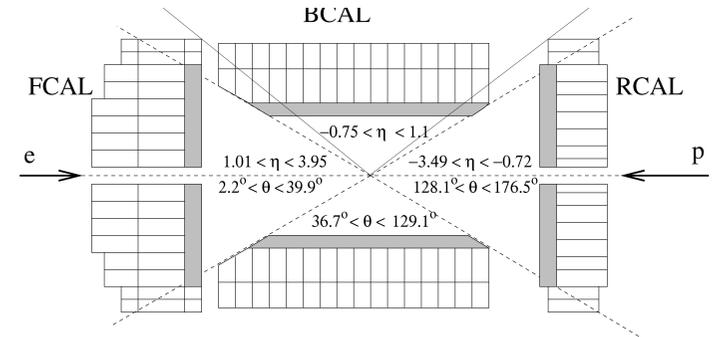
Figure 3.4: Layout of the ZEUS Uranium Calorimeter.

All parts of the ZEUS calorimeter, FCAL, BCAL and RCAL were built of alternating layers of 3.3 mm thick depleted uranium plates with 0.2 or 0.4 mm stainless steel, and plastic scintillator plates (SCSN38), 2.6 mm thick. With this proportion between the thickness of the inactive and active material, the ZEUS UCAL had a ratio between the calorimeter response to electrons and hadrons $e/h = 1.00 \pm 0.03$. Thanks to compensation mechanism the UCAL offered a very good hadronic energy resolution. The performance of the calorimeter, i.e. linearity of response and the energy resolution for electrons and hadrons had been studied in details with test beams at CERN SPS. The result can be summarised as follows [85] (energy E in GeV):

- non–linearity of response to hadrons below 1%;

- non–linearity of response to electrons below 1%;

- energy resolution for hadrons

$$\sigma_E/E = 0.35/\sqrt{E} \oplus 2\%; \tag{3.2}$$

- energy resolution for electrons

$$\sigma_E/E = 0.18/\sqrt{E} \oplus 2\%; \tag{3.3}$$

- module-to-module energy scale uncertainty $1 - 2\%$;

- time resolution ≤ 0.7 ns for electrons above 15 GeV.

However, these results were obtained with no inactive material in front of the UCAL. In the ZEUS detector, inactive material between IP and UCAL surface (beam pipe, CTD, solenoid) could significantly influence energy measurement for both electrons and hadrons. Corrections applied to reduce this effect are describe in Sect. 4.2.3.

The three calorimeter parts were subdivided into modules. The modules were transversally separated into towers and the towers were in turn longitudinally divided
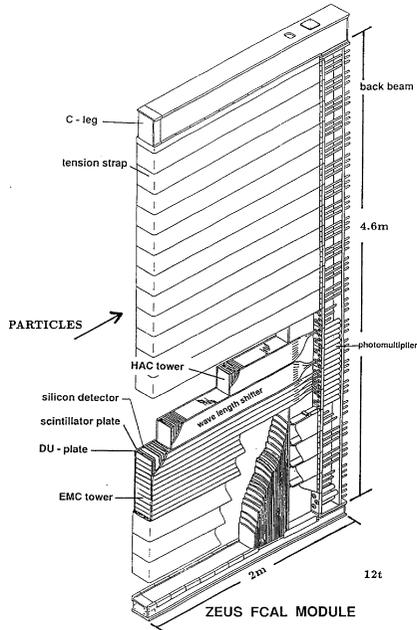
Figure 3.5: Layout of the FCAL module. The module is subdivided into electromagnetic (EMC) and hadronic (HAC) towers, which in turn are divided into cells. A cell is read out on both sides by two wave–length shifter bars.

into electromagnetic (EMC) and hadronic sections (HAC). The design of a FCAL module is shown in Fig. 3.5.

The FCAL and RCAL modules were planar and their face was perpendicular to the beam axis (see Fig. 3.1), while the BCAL modules were wedge-shaped and the EMC section is projective in the polar angle. The EMC and the HAC sections were further segmented into cells. Each EMC section was segmented transversally into four (in FCAL and BCAL) or two cells (in RCAL) for better electron identification and position measurement. The HAC towers in the FCAL and the BCAL were longitudinally subdivided into two hadronic cells (HAC1, HAC2). Scintillator plates of each cell were read out by two wave–length shifter plates (WLS) attached on both sides of the module. Light from WLS was transferred via light guides to photomultipliers placed behind the module. Information from two photomultiplier tubes per cell provided reconstruction of the particle position in the cell (from light attenuation in the scintillator) and an additional check of the uniformity of the readout. The natural radioactivity of ^{238}U was used as a reference signal to calibrate the readout channels to a precision of $< 0.2\%$.

3.2.4   Forward Plug Calorimeter

The Forward Plug Calorimeter (FPC) [90] was a lead-scintillator sandwich calorimeter placed in the forward direction of the ZEUS experiment. It was installed in ZEUS in 1998 and took data until the HERA I data taking in 2000. The FPC was installed in the 20×20 cm^2 forward beam-hole of ZEUS. It had an internal aperture in order to host the HERA beam-pipe. The main purpose of the FPC was to increase the angular coverage of the UCAL. The pseudorapidity coverage of the calorimeters in the forward direction increased from $\eta \approx 4$ to $\eta \approx 5$ with the introduction of the FPC. This improved the ability of the detector to collect the proton remnants, having a big impact on diffractive analyses based on the large rapidity gap or the M_X-method (see Sect. 2.2.3).
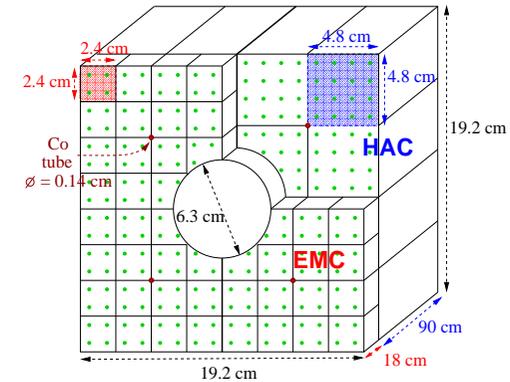

Figure 3.6: Front view of the FPC. The readout cells and the position of the wavelenght shifters are shown.

A schematic view of the FPC is hown in Fig. 3.6. The active part of the FPC had outer dimensions of 192×1080 mm^3. The 6.3 cm diameter central hole accomodated the HERA beampipe. The characteristics of the FPC were chosen to be similar to the ones of the UCAL since a large fraction of the hadronic shower was expected to be shared with the main calorimeter. The radiation length, $X_0 = 0.68$ and the nuclear absorption length, $\lambda = 20$ cm, are similar to those of the FCAL. The FPC is subdivided longitudinally into an EMC and a HAC section which are readout separately. Each section is further divided into cells. The size of the cells was 24×24 mm^2 in the EMC part. This size was taken in order to match the Moliere radius of the FPC, $R_M = 2$ cm. In the HAC part the cells were bigger, 48×48 mm^2.

The energy resolution, σ_E/E, for pions from test-beam was measured to be [91]

$$\frac{\sigma_E}{E} = \frac{0.53 \text{ GeV}^{-0.5}}{\sqrt{E}} \oplus 0.11 \oplus 0.03 \, log \left(\frac{E}{1 \text{ GeV}} \right)$$

where the energy of the pion, E, was measured in GeV.

3.2.5 Hadron Electron Separator

The Hadron Electron Separator (HES) [92] consisted of a layer of silicon pad detectors. Its task was to improve the identification of the EMC clusters and the reconstruction of their position. The rear part of the HES (RHES) was located in the RCAL at a depth of 3.3 radiation lengths, i.e. the approximate position of the EMC shower maximum. Each silicon pad had an area of 28.9×30.5 mm^2, providing a spatial resolution of about 9 mm for a single hit pad. If more than one adjacent pad was hit by a shower, a cluster consisting of at most 3×3 pads around the most energetic pad was considered, allowing to reconstruct the position of the incident particle with the resolution of 5 mm. The forward part (FHES) was located in a similar position in the FCAL and had the same general properties.

3.2.6 Small Rear Tracking Detector

The Small Rear Tracking Detector (SRTD) [93] was made of two planes of scintillator strips positioned in front of the RCAL close to the beampipe. It was installed in 1994. Its main task was the detection and position measurement of low-angles electrons that otherwise would have escaped outside the RCAL geometric acceptance. The SRTD covered the area 68×68 cm^2 centered around the beampipe. A hole of size 8×20 cm^2 accomodated the beampipe. The strips of the two planes had a perpendicular orientation, one along the $X-$axis and the other along the $Y-$axis. The strip width was 1 cm and its thickness was 0.5 cm. Its position resolution was ~ 3.5 mm and its time resolution was better than 2 ns.

3.2.7 The Luminosity Measurement

The integrated luminosity, $\mathcal{L}$, relates the expected number of events, N, with the cross section for given process, σ: $N = \mathcal{L}\sigma$. Thus, a precise determination of the luminosity is essential for any cross section measurement in a high energy physics experiment. The luminosity of ep–collisions at HERA was determined from the measurement of the rate for the bremsstrahlung process $ep \rightarrow e\gamma p$ [94]. As the theoretical cross section is known to an accuracy of 0.5% from QED calculations, a measurement of the photon rate permits a precise determination of the ep–luminosity.

Fig. 3.7 shows the layout of the HERA magnet system in the backward direction ($Z < 0$), where the ZEUS luminosity detectors were placed. Luminosity measurement was performed in the two lead/scintillator electromagnetic calorimeters placed at Z $= -34$ m (LUMIE) and Z $= -107$ m (LUMIG).

Photons with a polar angle. θ_γ, $|\pi - \theta_\gamma| < 0.5$ $mrad$ originating from the bremsstrahlung process $ep \rightarrow e\gamma p$ were detected by the LUMIG detector [95, 96]. The energy resolution of LUMIG was measured, under test-beam conditions, to be $18\%/\sqrt{E}$ with the energy E measured in GeV. However, the carbon lead filter placed in front of the detector to shield it against synchrotron radiation degraded the resolution to $23\%/\sqrt{E}$. The impact position of incoming photons could be determined with a resolution of 0.2 cm in X and Y, using 1cm wide scintillator strips installed at a depth of $7 X_0$
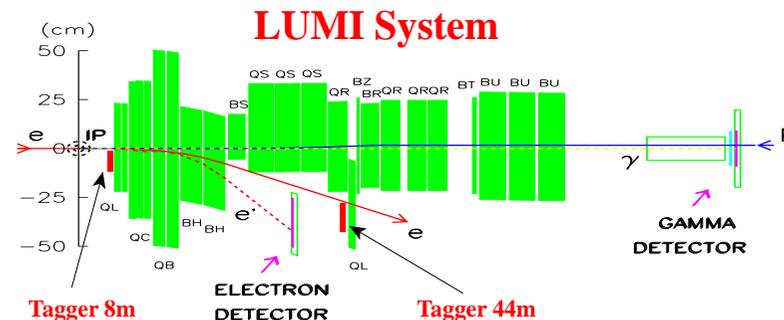
Figure 3.7: Location of ZEUS detectors in negative Z-direction. Shown are the gamma (LUMIG) and electron detectors (LUMIE) used for the luminosity measurement.

within the detector. LUMIG was also used to determine the electron beam tilt and to measure photons from initial–state radiation.

The LUMIE calorimeter [95, 96] was designed to measure the electrons from the process $ep \rightarrow e\gamma p$ in coincidence with photons measured in the LUMIG detector. It was placed at Z $= -35$ m and detected electrons with polar angles of less than 5 mrad with respect to the electron beam direction. The measurement was restricted to the limited energy range, from 7 to 20 GeV. For these energies, electrons deflected by the HERA magnet system left the beam pipe at Z $= -27$ m through an exit window in front of LUMIE (similar to the one in front of the LUMIG detector). The LUMIE detector had an energy resolution of $18\%/\sqrt{E}$ (under test-beam conditions). It turned out that electron tagging in LUMIE is not necessary for a precise measurement of the luminosity. On the other hand, the LUMIE detector proved to be very useful for physics analysis, to tag the so called photoproduction events.

3.2.8 Trigger and Data Acquisition Systems

The short time between bunch crossings at HERA, 96 ns, (equivalent to a rate of about 10 MHz), was a technical challenge for detector construction and put stringent requirements on the ZEUS trigger and data acquisition systems. The rate of interesting ep physics events which should be measured in the ZEUS detector and stored for the analysis was of the order of a few Hz [97, 98, 99, 103], while the total interaction rate (any signal in the detector), which was dominated by background from upstream interactions of the proton beam with residual gas in the beam pipe, was of the order 10–100 kHz. Other important background sources were electron beam gas collisions, beam halo and cosmic events. ZEUS detector readout was based on a sophisticated three-level trigger system developed to efficiently select ep physics events while reducing the rate to a few Hz. A schematic diagram of the ZEUS trigger system is shown in Fig. 3.8.
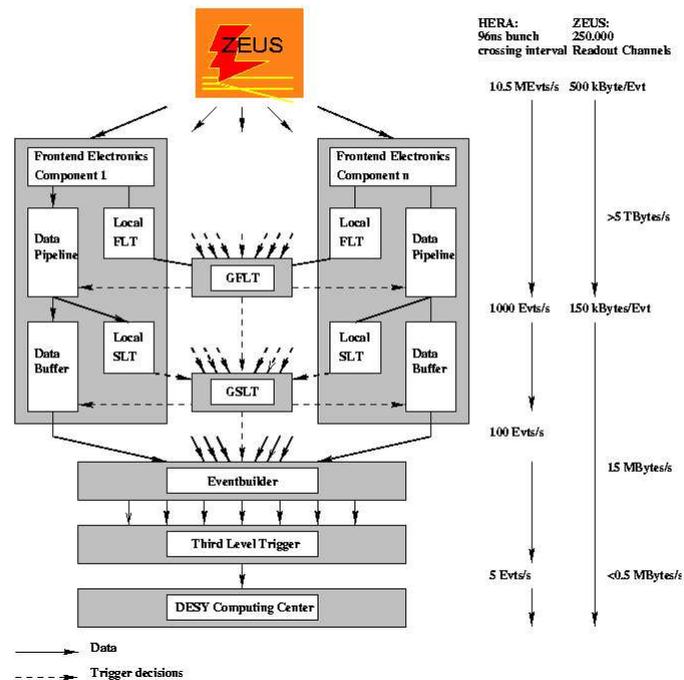
HERA:
96ns bunch
crossing interval

ZEUS:
250.000
Readout Channels

10.5 MEvts/s   500 kByte/Evt

>5 TBytes/s

1000 Evts/s   150 kBytes/Evt

100 Evts/s

15 MBytes/s

5 Evts/s   <0.5 MBytes/s

Figure 3.8: Schematic diagram of the ZEUS trigger and data acquisition systems.

The First Level trigger (FLT) [98] was a hardware trigger, designed to reduce the input rate below 1 kHz. Each detector component had an own FLT-dedicated part that developed a trigger decision within 2 μs after the bunch crossing. The decisions from the local FLTs were passed to the Global First Level Trigger (GFLT), which decided whether to accept or reject the event, and returned this decision to the readout components within 4.4 μs. The relevant information used at the GFLT was based on UCAL energies (e.g. total transverse energy, missing transverse momentum), CTD tracks (e.g. number of tracks, vertex position) and hits in the muon chambers. For the time needed to develop final GFLT decision the component data were stored in analog or digital pipelines.

If the event was accepted, the data was read out and transferred to the component Second Level Trigger (SLT). The event variables reconstructed at the SLT level had a better resolution than those at the FLT. Moreover, some new information was avail-

able at the SLT. The most important one was the UCAL timing, which was very useful in rejecting non–ep background events. The SLT was designed to reduce the event rate to about 50-100Hz. Each detector component had its own SLT, which passed calculated trigger quantities to the Global Second Level Trigger (GSLT) [100].

If the event is accepted by the GSLT, all detector components sent their complete data to the Event Builder (EVB) [101], which combined all the data for an event into a single record of a database table (ADAMO tables [102]). This was the data structure used for the ZEUS data storage. Combined data were sent to the Third Level Trigger (TLT) [103] which was based on the offline reconstruction code and selected events according to specific physics requirements. It was designed to reduce the rate to a few Hz.

Chapter 4

Data sample and event reconstruction

In this chapter the data sample analysed is described, including all the Data Quality Monitoring (DQM) controls required. The data sample collected is processed by a reconstruction chain which extracts the basic physical quantities needed for the analysis. The scattered electron reconstruction, the extraction of the most important DIS quantities, the measurement of the four-momenta of the particles produced in the collision, the kinematical properties of the jets and the calculation of diffraction related variables are described. The treatment of the inactive regions of the ZEUS detector that bias the reconstruction is also explained.

4.1 Data sample used

The data analysed were collected during the years 1999 and 2000. Data coming from both e^+p and e^-p collisions were used. The centre-of-mass energy was 318 GeV.

In order to obtain a reliable dataset for the following analysis all the main components of ZEUS needed to be active and well-working. This was checked by the ZEUS "take" routines which set a run-based veto in case a particular component of the detector malfunctioned or did not function at all. The routines themselves consists of DQM jobs performed by the component crews. The main requirement for the analysis (EVTAKE routine) was the simultaneous correct working of the tracking chamber (CTD), the calorimeter (CAL), the trigger and the luminosity detectors. Moreover the overall number of FPC channels with hardware problems (most of the times related to the high-voltage power supply) was recorded on a run-by-run basis (FPCTAKE routine). Runs in which the FPC was not taking data or had any hardware problems were rejected. During the year 2000, data with the vertex position shifted with respect to the nominal z-position were also taken: the corresponding integrated luminosity of $0.815 \, \mathrm{pb}^{-1}$ was also excluded from the analysis.

The total integrated luminosity analysed is 61.3 pb^{-1}, as summarised in Tab. 4.1.

Year	Interaction	Luminosity (pb^{-1}) after EVTAKE	Luminosity (pb^{-1}) after EVTAKE+FPCTAKE
1999	e^-p	3.2	3.2
1999	e^+p	16.2	12.7
2000	e^+p	45.4	45.4
Total		64.8	61.3

Table 4.1: ZEUS data samples used for the analysis divided into year and interacting beams. The luminosity is given after the EVTAKE flag selection in the third column and with the additional requirement of no veto from FPCTAKE in the fourth column.

4.2 Event Reconstruction

The basic informations needed to perform the analysis are the four-momenta of the charged particles obtained from the tracks detected by the tracking system and the energy clusters in the calorimeters. From them, one can extract the kinematic properties of the event and reconstruct all the quantities needed for the signal selection. In the following all the reconstruction steps relevant in this analysis are described.

4.2.1 Track and energy clusters reconstruction

4.2.1.1 Tracking

Charged particles passing through the CTD cause signals on the CTD wires (*hit*) which are digitised in terms of time and amplitude. This raw information is transformed offline into the spatial positions of the hits which are then passed as input to the track reconstruction package VCTRACK [104] for pattern recognition, track fit and vertex fit.

The pattern recognition associates the CTD hits compatible with the same track. First a "seed" is set by looking for three hits in three different outer axial superlayers lying in a common area in the XY plane. Only the outer CTD superlayers are used because there the track density is lower and the pattern recognition is therefore easier. Then a two-step procedure is performed starting from the seed: in the first step, a circle in the XY plane is fitted to the space hits. Then a first rough projection in the z-direction is provided by a z-by-Timing algorithm which exploites the signal propagation time along the CTD wires in order to restrict the regions where to look for stereo hits, improving further the 3D estimation of the trajectory. Once all the seeds in the outer superlayers have been associated to a track, the track seed is extrapolated inward. The trajectory parameters are updated with increasing precision as additional hits are gathered. Track segments can be associated to only one track candidate.

Once the pattern recognition phase is accomplished, a fit to the hits associated to the track candidates estimates the track parameters. The fitted trajectory of the particle passing through the CTD is a five-parameter axial helix since the magnetic solenoidal field is almost parallel to the z-axis. During this procedure, several corrections for bad hits, drift velocities inhomogeneities and Coulomb scattering are applied.

The last part of the track reconstruction is the reconstruction of the vertices, both the primary and the secondary ones. The vertex information is used then to constrain the tracks and another track fit like the one described above is performed.

4.2.1.2   Calorimeter clustering

The electric signals coming from the calorimeter-cell photomultipliers (PMTs) are converted into energies according to calibration constants determined from calorimeter-module test beam measurements. The signals are corrected for energy absorption by inactive material in front of the CAL and noisy cells - i.e. cells with not-properly working PMTs. A cell is tagged as noisy if no neighboring cell has a signal and the cell signal amplitude is less than 80 MeV for cells in the electromagnetic section (EMC) and 140 MeV for cells in the hadronic section (HAC). The cells are then grouped into *cell islands* for estimating the total energy deposit by a particle in the CAL. The cells are merged only if their signal is above the noise threshold and if they are contiguous. The clustering algorithm works on each CAL part (Front, Barrel, Rear) and each section separately. Cell islands in EMC and HAC are then merged into a *cone island* only if matched to a common angular cone pointing to the primary vertex.

4.2.2   Identification of the scattered electron

In any DIS analysis the correct and efficient identification of the scattered electron is fundamental. The reconstruction of its kinematical properties is also needed for the evaluation of many kinematical quantities used in the analysis. The scattered electron identification was carried out by means of the SINISTRA Neural Network (NN) algorithm [105]. The NN exploits the different transverse and longitudinal development of showers in the calorimeter coming from electrons or photons and those from hadrons. Typically, electromagnetic particles deposit their energy mostly in the EMC and the transverse size of the shower is smaller than that of hadrons. SINISTRA uses as input quantities related to the energies read by the PMTs of the cells belonging to a common island and gives as output the probability that the considered island was originated by an electron or by a photon. The NN was trained on MC events. In this analysis, the island is considered as an electron candidate if the output probability is $P_{\text{elcluster}} > 0.9$. The probability distributions for hadronic and electromagnetic clusters given by SINISTRA running on a MC sample is shown in Fig. 4.1.

The efficiency and purity [1] as a function of the energy of the scattered electron for a given MC sample is shown in Fig. 4.2. It may happen that more than one electron candidate is found. In this case the one with the highest energy is considered as DIS scattered electron.

The electron position was first reconstructed using the CAL information alone. The CAL-estimated position is obtained from the logarithm-weighted centre-of-mass

---

[1] In this context, purity is defined as the fraction of clusters with probability higher than 0.9 that are effectively electrons and not misidentified hadrons.
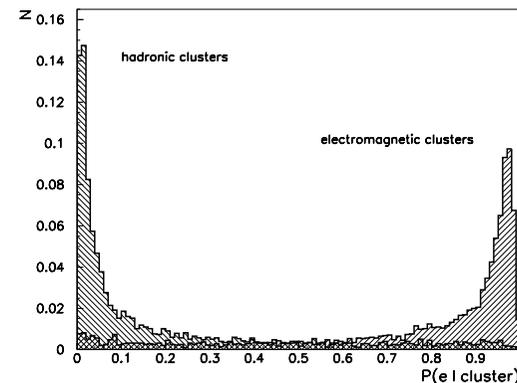
Figure 4.1: Probability distribution for a given cluster to be an electromagnetic cluster using the SINISTRA *electron finder (from [105]).*
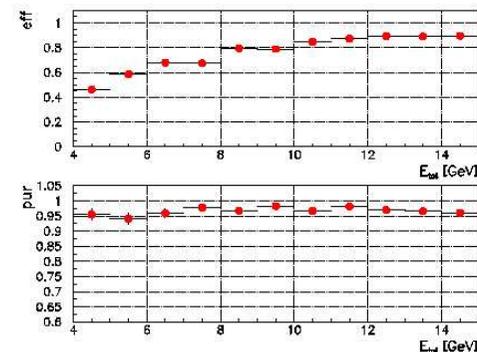

Figure 4.2: Efficiency and purity to identify the scattered lepton with SINISTRA *as a function of reconstructed energy for a DIS MC generated with* $Q^2 > 2.2\,GeV^2$ *and* $y > 0.4$ *(from [105]).*

energy[2] of the cells belonging to the cluster according to the following formula [106]:

$$\vec{r} = \frac{\sum_i w_i \vec{r_i}}{\sum_i w_i}, \tag{4.1}$$

with

$$w_i = max(0, [W_0 + ln(\frac{E_i}{\sum_i E_i})]),$$

where the sums run over the cells belonging to the cluster and $\vec{r_i} = (x_i, y_i, z_i)$ is the position of the center the i-th cell and the parameter W_0 is a cut off determined in order to reduce systematic biases. In case the CAL-estimated position are within the acceptance of the rear part of the hadron-electron separator (RHES) or of the small rear tracking detector (SRTD), the information of these two other detectors was used. The granularity in the RHES and SRTD is higher than in the CAL, therefore an improvement in the position reconstruction can be achieved if the information coming from the former two components is used. An algorithm decides which information to use according to the position of the electron, preferring first the SRTD, then the RHES and only at last the rear CAL (RCAL). The resolution obtained in this way on the electron scattering angle, for electrons with energy higher than 10 GeV detected in the RCAL, varies from the 7 mrad using the CAL to the 2 mrad using the SRTD[3], as shown in Fig. 4.3.

4.2.3   Energy Flow Objects reconstruction

The relative energy resolution of the CAL degrades with the decrease of the energy of the incoming particle (see Eq. (3.2)). It is therefore convenient to exploit the information of the CTD for low-momentum charged particles by means of Energy Flow Objects (EFOs)[107, 108]. The EFO algorithm determines the energy of a particle using either the track momentum or the island energy. EFOs are made by either a single track with no CAL cluster matched or a cone island without matched tracks or cone islands matched to tracks (see Fig. 4.4). In the latter case, the algorithm chooses which information to use according to the following criteria [108]:

- in case of a 1-to-1 track-island match, the following requirements have to be satisfied in order to prefer the CTD information to the CAL one:

  - $\frac{E_{CAL}}{p_T} < 1.0 + 1.2 \cdot \sigma\left(\frac{E_{CAL}}{p_T}\right)$; the transverse momentum of the track, p_T, must be higher than the cone island energy, E_{CAL}, considering also the uncertainty on this ratio, $\sigma\left(\frac{E_{CAL}}{p_T}\right)$.

---

[2]With this logarithmic weights, biases due to the varying cell projectivity as seen from the vertex are reduced. Furthermore, in this way the exponential falloff of the shower energy is considered.

[3]The resolution here is defined as the half width at the half maximum of the distribution of the residuals.
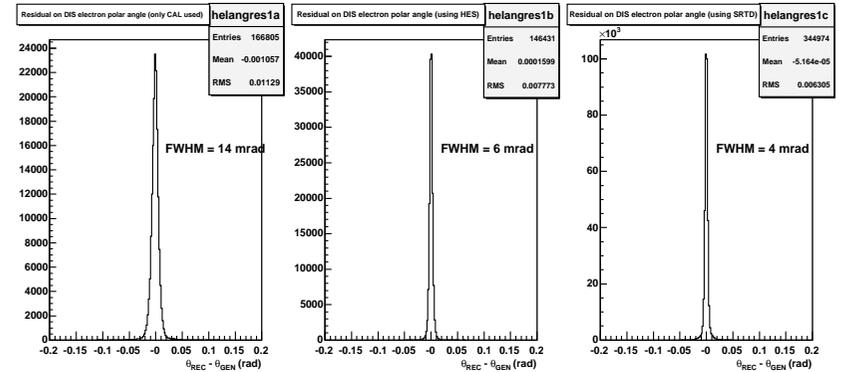
Figure 4.3: Resolution on the electron scattering angle for MC events using CAL, RHES or SRTD for reconstructing the electron position. Events with a reconstructed vertex, $45 < E - p_Z < 65$ GeV and a scattered electron candidate with an energy $E'_e > 10$ GeV were used.

- $\frac{\sigma(p_T)}{p_T} < \frac{\sigma(E_{CAL})}{E_{CAL}}$, the resolution on the track transverse momentum must be better than the one on the calorimeter cluster. The two resolutions are shown in Fig. 4.5.

- for cases like 1-to-2, 2-to-1, 1-to-3 the above requirement for the track-island match are still used with the prescription of replacing the quantities related to cluster energy and transverse momentum with the total sum over the different clusters or tracks.

In all cases the 4-momentum of the EFO is obtained by assuming that the particle is a pion.

The EFO algorithm has also the task of identifying *backsplashes*. Backsplashes are energy deposits at very large polar angles in the RCAL that are not caused by particles generated in the primary interaction. Possible sources are neutral particles generated in the front CAL (FCAL) which travel through the entire detector, particles showering in the inactive material in front of the RCAL or noisy cells. The algorithm that tags a cluster as backsplash uses the angle and energy of the deposit as inputs. Backsplashes can alter significantly the $E - p_Z$ measured in an event because of their large angle and therefore they are rejected in many analyses. In a diffractive analysis the typical values of x_{Bj} are very small; this boosts the system backwards, making the identification of backsplashes problematic and increasing the risks of excluding good clusters. For this reason no backsplash subtraction was applied in this analysis. A study on the impact of the backsplashes has been carried out. In Fig. 4.8 the distribution of the $E - p_Z$ for a sample of diffractive DIS events is shown with and without the backsplash subtraction.
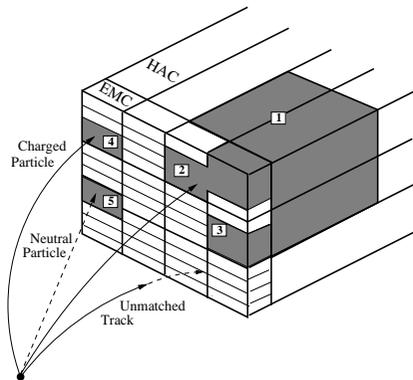
Figure 4.4: Schematic picture of the EFOs reconstruction. The EMC cell islands 2 and 3 are matched to the HAC cell island 1 to form a cone island. The cone island and the isolated EMC cell island 4 are matched to a track to form two separate EFOs. Other two EFOs are made by the isolated EMC cell island 5 and the unmatched track.

One can notice how the peak of the distribution is closer to the kinematical peak at 55 GeV.

The EFOs are corrected for inactive material using the DMCO corrections [109]. Such corrections take into account the energy loss of the EFOs before the calorimeter. They were estimated on a MC basis using GEANT 3.13 [110] for simulating the detector and are applied to both data and MC simulation. The amount of inactive material estimated is shown in Fig. 4.6. The energy losses due to the inactive material can be relevant for low-energies particles produced in the physics sample under study. Thus, a study on an alternative parametrisation of the effect of the inactive material was carried out in order to understand the impact of different choices in the detector description. The so-called "Vosselbeld-Ochs" (VO) corrections were estimated by requiring the balancing in E_T of the hadronic system and the DIS electron in high-Q^2 DIS events [111, 112]. The correction factors in bins of θ have the following analytical form

$$Corr = 1 + \frac{A_1}{E^{A_2}} \qquad (4.2)$$

where E is the energy of the EFO and A_1 and A_2 are two parameters depending on the polar angle and are different for data and MC. The advantage of such a method is to have two independent parametrisations, one for the data and one for the MC. On the other hand the particular functional form chosen is unfortunate for a diffractive analysis. As seen in Fig. 4.7, for very forward angles the relatively large size of the parameters and the functional form of Eq. (4.2), which diverges for low energies, can give very high corrections. This causes the rejection of good diffractive events, for which the presence of a large rapidity gap and little energy in the forward direction
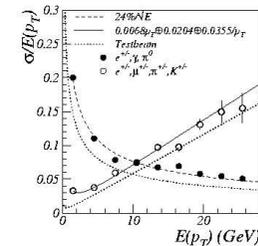
Figure 4.5: Resolution of the EMC energy for a single particle MC (full dots) and resolution of CTD on the transverse momentum (open circles) (from [108]).

are required. In Fig. 4.8 the distributions for the events selected from the data sample obtained with the two different sets of inactive material corrections are compared. It can be seen that the total number of events selected using the DMCO corrections is about 10 % higher than the one obtained using the VO corrections. The final impact on the cross sections is smaller, anyway, due to a change also in the acceptance estimated with the MC. It was chosen to keep the DMCO corrections because of the higher acceptances obtained using them and the large size of the VO corrections at forward polar angles.

The description of the inactive material and the energy scale of the calorimeter by the MC can be checked by comparing the balancing of the transverse momentum between the scattered electron and the hadronic system. The total momentum is obtained by summing the momentum vectors of the scattered electron and the hadronic system composed by all the reconstructed EFOs. The total transverse momentum, p_T^{TOT}, after the final selection (see Sect. 6.1) is presented in Fig. 4.9. As expected from momentum conservation, the distribution is strongly peaked at zero. The tail at higher values of p_T^{TOT} is due to detector effects. However, the most important aspect is the correct description of the data by the MC. This guarantees a good simulation of the calorimeter energy scale and the energy losses due to the inactive material.

4.2.4   Reconstruction of DIS kinematics

The quantities Q^2 and x_{Bj}, that define the kinematics of a DIS event, were reconstructed from the particles measured in the detector. Several methods have been developed exploiting the overconstrained kinematics of the DIS process. In this analysis the *Double Angle method* (DA) was used [113]. The measured quantities needed for estimating the kinematic variables with this method are:

- E_e, the initial energy of the incoming electron;
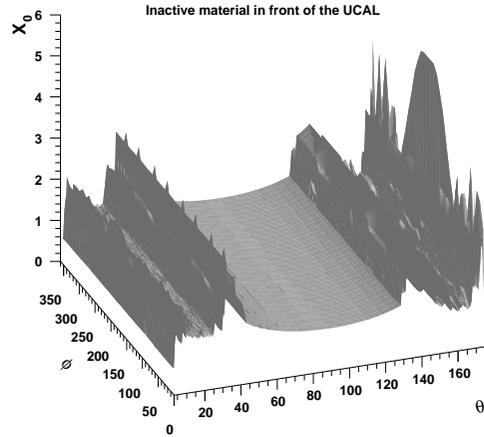
- E_p, the initial energy of the incoming proton;

Figure 4.6: Distribution of the inactive material in front of the CAL in units of radiation lengths, X_0, in the $\theta - \phi$ plane as implemented in the MC simulation of the detector (from [109]).

- $\theta_{e'}$, the polar angle of the scattered electron;

- $\delta_h = \sum_{i=1}^{N_{had}} E_i - p_{Z,i}$, the sum of $E - p_Z$ of the whole hadronic system, i.e. all the particles measured in the detector except the DIS electron;

- $p_{T,h} = \sqrt{(\sum_{i=1}^{N_{had}} p_{X,i})^2 + (\sum_{i=1}^{N_{had}} p_{Y,i})^2}$, the total transverse energy of the hadronic system.

- $\gamma_h = a cos \frac{p_{T,h}^2 - \delta_h^2}{p_{T,h}^2 + \delta_h^2}$, the polar angle towards which the whole hadronic system is directed.

From these variables one can reconstruct Q^2 and x_{Bj} with the following relations [113]:

$$Q_{DA}^2 = 4 E_e^2 \frac{\sin \gamma_h (1 + \cos \theta_{e'})}{\sin \theta_{e'} + \sin \gamma_h - \sin(\theta_{e'} + \gamma_h)} \tag{4.3}$$

$$x_{Bj,DA} = \frac{E_e}{E_p} \frac{\sin \theta_{e'} + \sin \gamma_h + \sin(\theta_{e'} + \gamma_h)}{\sin \theta_{e'} + \sin \gamma_h - \sin(\theta_{e'} + \gamma_h)} \tag{4.4}$$

In the QPM picture, γ_h estimates the polar angle of the scattered quark. The value of W is then calculated exploiting Eq. (2.13).
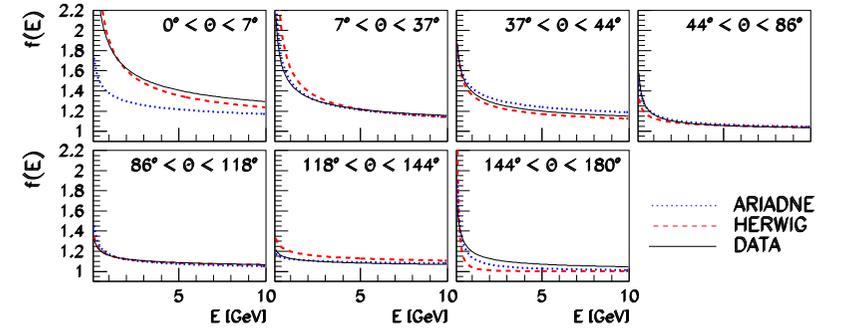
Figure 4.7: Energy corrections for different θ bins applied to EFOs for data and MC according to the VO parametrisation.

The basic feature of the DA method are that

- experimentally the directions of the particles are measured with better precision than their energies;

- the resulting kinematics are weakly sensitive to calorimeter miscalibrations, since angles and not energies are entering in the formula;

- the very good hermeticity of ZEUS allows a very good reconstruction of the hadronic system. The proton-remnant particles lost in the forward region not covered by the detector contribute only to a small fraction of the values of δ_h and $p_{T,h}$.

This method is the best performing over a large part of the phase space. The resolution achieved on the kinematic variables is shown in Appendix C.

4.2.5 Jet reconstruction

Due to the confinement property of coloured particles, no isolated quark can be observed in nature. The only observables are hadrons produced in the collision which can merge into jets as described in Sect. 2.1.7. Therefore an algorithm is needed to merge the hadrons originating from the same parton into a reconstructed jet whose kinematic properties resemble as much as possible the ones of the primordial parton. In this analysis the longitudinally-invariant k_T algorithm was used [29, 30]. The main features that this algorithm provides are [29]:

i) infrared (IR) and collinear safe; this allows the cross section to be calculated in perturbation theory absorbing IR and collinear divergencies. The first divergency is due to the emission of arbitrarly soft (i.e. small energy) particles from the partons originated in the dijet process; the characteristics of the jets reconstructed
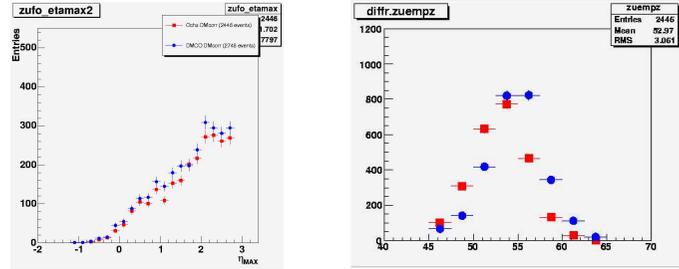
Figure 4.8: Distributions of the events selected from the data sample using the DMCO corrections without the backsplash subtraction (dots) and the VO corrections with backsplash subtraction (squares) as corrections for the inactive material. The distributions are shown as a function of η_{MAX}, the pseudorapidity of the most forward EFO reconstructed in the event (left) and as a function of the total reconstructed $E - p_Z$ (right). Only a subsample of the total available luminosity was used for these plots.

by the k_T algorithm remains unchanged by adding any number of these soft particles. The collinear divergency occurs when two particles are emitted in the same direction. The k_T algorithm merges the two collinear particles into one with their combined momenta;

ii) simple to use in experimental analyses;

iii) simple to use in theoretical calculations;

iv) subject to small hadronisation corrections;

v) able to factorise initial-state collinear singularities into universal parton densities; this means that the algorithm has to be able to conserve the QCD factorisation.

vi) not strongly affected by contamination from hadron remnants and the underlying soft event.

This algorithm was run on the EFOs present excluding the one associated to the DIS scattered electron. The merging procedure was the following:

1. for every pair of objects with four-momenta i and j, a *closeness*, d_{ij}, between them was defined as

$$d_{ij} = [(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2] \, min\{p_{T,i}^2, p_{T,j}^2\} \qquad (4.5)$$

2. a closeness to the beam, d_i, was calculated for each four-momenta:

$$d_i = p_{T,i}^2 \, R \qquad (4.6)$$

where R is the $\eta - \phi$ radius of the cone enveloping the proton remnant. In this analysis R was set to 1.
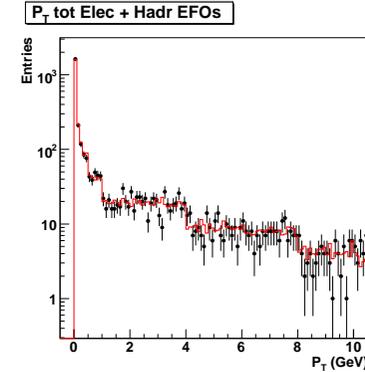
Figure 4.9: The distribution of the total transverse momentum, p_T^{TOT}, of the reconstructed final state for the sample passing the entire selection chain. The simulation produced with RAPGAP (solid line) is compared to the measured data (full dots) after having been rescaled to the area of the data distribution.

3. the next step depends on what is the smallest value among d_{ij}, d_i.

- if it is d_i, exclude the i-th object from the next iterations. Since now the i-th object is defined as a *macrojet*.

- if it is d_{ij}, the two four-momenta are merged into a *protojet* whose four-momentum is defined by a prescription called *recombination scheme*.

4. this procedure is repeated until only macrojets are remaining. Note that at the end all the initial four-momenta are associated to macrojets by the algorithm. Since at every iteration one four-momentum is excluded, the number of iterations needed is equal to the initial number of four-momenta.

There are several different recombination schemes suitable to an analysis in ep collisions. The one used here was the p_T-weighted scheme. Under this prescription the transverse momentum, $p_{T,(ij)}$, pseudorapidity, $\eta_{(ij)}$, and azimuthal angle, $\phi_{(ij)}$, of the protojet coming from the merging of the i-th and j-th particle are defined as[4]

$$p_{T,(ij)} = p_{T,i} + p_{T,j} , \qquad (4.7)$$

$$\eta_{(ij)} = \frac{p_{T,i}\eta_i + p_{T,j}\eta_j}{p_{T,(ij)}} , \qquad (4.8)$$

$$\phi_{(ij)} = \frac{p_{T,i}\phi_i + p_{T,j}\phi_j}{p_{T,(ij)}} . \qquad (4.9)$$

---

[4]The transverse momenta enter these formulae as scalar quantities.

The jets are reconstructed after having boosted the final state system in a reference frame where the photon and the proton collide head on. This reference frame can be the Breit frame (see Sect. 2.1.3.1) or the $\gamma * p$ frame. In the latter frame, the photon-proton centre of mass is at rest. These two different frames are equivalent except for a boost along the proton direction. Because of a property of the algorithm, the final outcome will be the same, i.e. the particles will be merged always in the same jets. The choice of the $\gamma * p$ frame is very suited for jet analyses in DIS. In fact, the quark struck by the γ^* inverts its direction while the proton remnant keeps travelling in the former direction. In this way one achieves a maximal separation between the two systems, the hard final state and the proton remnant.

4.2.6 Reconstruction of diffractive variables

As introduced in Eq. (2.45) and (2.58), the fraction of the proton longitudinal momentum carried by the diffractive exchange,$x_{I\!P}$, and the fraction of the diffractive exchange longitudinal momentum carried by the parton entering the hard subprocess, $z_{I\!P}$, require the four-momenta of partons that are not directly measured in the detector. Experimentally, one uses estimators for the two variables, $x_{I\!P}^{\text{obs}}$ and $z_{I\!P}^{\text{obs}}$. They are defined as follows

$$x_{I\!P} = \frac{(p - p') \cdot q}{p \cdot q} \sim x_{I\!P}^{\text{obs}} = \frac{Q^2 + M_{\text{X}}^2}{Q^2 + W^2} \qquad (4.10)$$

$$z_{I\!P} = \frac{q \cdot v}{q \cdot (p - p')} \sim z_{I\!P}^{\text{obs}} = \frac{Q^2 + M_{\text{jj}}^2}{Q^2 + M_{\text{X}}^2} \qquad (4.11)$$

where M_{X} is the invariant mass of the diffractive system produced in the interaction. M_{X} was reconstructed from the EFOs with the following formula

$$M_{\text{X}} = \sqrt{\left(\sum_{i=1}^{Nhad} E_i\right)^2 - \left(\sum_{i=1}^{Nhad} \vec{p}_i\right)^2} \qquad (4.12)$$

where the sums run over the total number of EFOs. M_{jj} is the invariant mass of the two highest-E_T jets. In the MC, even though the information on the parton momenta is available, this definition of $x_{I\!P}^{\text{obs}}$ and $z_{I\!P}^{\text{obs}}$ is used such to adopt a consistent reconstruction procedure over all the analysis.

4.2.7 Reconstruction of x_γ

The variable x_γ indicates the fraction of γ^* longitudinal momentum entering the hard subprocess. It was introduced in Eq. (2.57) with a definition that uses the four-momenta of partons participating to the interaction. Since these four-momenta are not directly measurable in the detector, an estimator for x_γ is used in the analysis,

x_γ^{obs}. The definition of x_γ^{obs} is

$$x_\gamma^{\text{obs}} = \frac{\sum_{j=1}^{2}(E_j - p_{Z,j})^{\text{LAB}}}{(E - p_Z)_{TOT}^{\text{LAB}}} \qquad (4.13)$$

In Eq. (4.13), all the quantities are measured in the laboratory reference frame. The sum in the numerator includes the two jets with the highest transverse energy as measured in the $\gamma^* p$ frame and the total $E - p_Z$ in the denominator is calculated considering only the hadronic system.

Chapter 5

Monte Carlo samples and theoretical calculations

Several theoretical models and simulations were used to extract the cross sections and to describe the data. Leading Order (LO) MC were used for the background estimation and the correction of the data for detector effects. The Next-To-Leading Order (NLO) calculations were obtained by modifying a program that normally could not be able to be used in the diffractive formalism. The features of the NLO program and the procedure carried out in order to adapt it to the diffractive case are also described here.

5.1 Monte Carlo simulation

The LO MC used for the simulation of the signal were RAPGAP and SATRAP . DJAN-GOH was used for the simulation of the background due to non-diffractive DIS. They all used parton shower models in order to mimic the effect of higher-order QCD terms. All the MC samples were produced applying a first loose jet selection in order to reduce the amount of data generated. A comprehensive list of all the MC samples used, together with their luminosities and the most important kinematical parameters is presented in Table 5.1.

5.1.1 Rapgap

The RAPGAP v2.08/18 [114] was used to simulate the diffractive signal. RAPGAP is based on the resolved-pomeron model. The "H1 fit2" dPDFs [115] were used for the samples generation. Although these dPDFs have been superseded by new and more refined fits, in the RAPGAP version used they were the most updated choice available. The pomeron flux adopted was the same Regge-motivated functional form used in the extraction of the dPDFs [115]:

$$f_{\mathbb{P}/p}(x_{\mathbb{P}}) = \int_{t_{cut}}^{t_{min}} \frac{e^{B_{\mathbb{P}}t}}{x_{\mathbb{P}}^{2\alpha_{\mathbb{P}}(t)-1}} dt \qquad (5.1)$$

The two parameters, $\alpha_{\mathbb{P}}$ and $B_{\mathbb{P}}$, are process-dependent quantities whose values were set to 0.25 and $4.0\,\mathrm{GeV}^{-2}$ respectively. The value of the pomeron flux is integrated over the squared four-momentum transferred at the proton vertex[1], t, since the scattered proton momentum is not measured with the Large Rapidity Gap method. The limit t_{min} is the maximum kinematically allowed value of t while $t_{cut} = -5$ GeV^2 is a limit set by the measurement apparatus.

The parton-shower simulation was based on the matrix-element parton shower (MEPS) model [116]. Three different RAPGAP samples were produced: two direct-photon samples (differing by the quarks produced in the hard scattering, one with only light quarks and the other with charm quarks only) and one resolved-photon sample (with all flavours allowed to be produced in the hard scattering). In resolved process simulation, the GRV-G-HO photon parton densities were used [73]. The three samples were then summed taking in account different absolute normalisations for the three cross sections as evaluated in Sect. 6.3. No sub-leading mesonic Regge trajectories were included in the generation, thus this sample is purely diffractive.

The QED radiative effects from initial-state (ISR) and final-state (FSR) radiation were simulated with the HERACLES program [117]. The hadronisation of partons into hadrons after the parton shower was modeled according to the Lund model [118] with JETSET 7.4 [119]. The charm quark fragmentation function used was the Peterson function [120] with the fragmentation free parameter set to $\epsilon_Q = 0.035$, as usual in heavy flavour analyses [121, 76].

5.1.2 Satrap

Another signal sample was generated with SATRAP [65]. SATRAP is based on the Golec-Biernat-Wüsthoff model of diffraction (see Sect. 2.2.4). The parameters of the model were determined from fits to the total $\gamma^* p$ cross section. The SATRAP generator is interfaced to RAPGAP , therefore this sample used the same programs and parameters of RAPGAP for the simulation of higher-order effects and the hadronisation. In the implementation of this sample only $q\bar{q}$ and $q\bar{q}g$ terms are considered. More complex contributions that would take in account any hadronic structure of the photon are not present with the result that this MC does not include any resolved photon contribution. Differently from the RAPGAP sample, the parton shower in SATRAP is based on the Colour Dipole Model (CDM) [122].

5.1.3 Djangoh

The background coming from non-diffractive dijet DIS events was estimated with an inclusive DIS sample simulated with the DJANGOH MC v1.1 [123]. This program used DJANGO as event generator and HERACLES for modelling the initial and final state QED radiation. The CTEQ4D proton PDFs [124] were used in the generation. The QCD cascade was simulated with the Colour Dipole Model as implemented in ARIADNE [125]. More informations about the generation of the sample used can be found in the Ref. [126].

---

[1]Notice that t is defined as a negative number.

Sample	Luminosity (pb^{-1})	Kinematic range	Notes
RAPGAP -LQ	842.592	$Q^2 > 2\,\mathrm{GeV}^2$, $y > 0.001$ $E^*_{\mathrm{T,jj}} > 2.5\,\mathrm{GeV}$, $-5.0 < \eta^*_{\mathrm{jj}} < 3.0$	Direct-photon, light quarks only Resolved pomeron model H1 fit2 dPDFs
RAPGAP -CC	793.586	$Q^2 > 2\,\mathrm{GeV}^2$, $y > 0.001$ $E^*_{\mathrm{T,jj}} > 2.5\,\mathrm{GeV}$, $-5.0 < \eta^*_{\mathrm{jj}} < 3.0$	Direct-photon, charm quark only Resolved pomeron model H1 fit2 dPDFs
RAPGAP -RES	1913.0	$Q^2 > 2\,\mathrm{GeV}^2$, $y > 0.001$ $E^*_{\mathrm{T,jj}} > 2.5\,\mathrm{GeV}$, $-5.0 < \eta^*_{\mathrm{jj}} < 3.0$	Resolved-photon GRV-GO-HO photon PDFs Resolved pomeron model H1 fit2 dPDFs
SATRAP	164.179	$Q^2 > 2\,\mathrm{GeV}^2$, $y > 0.001$ $E^*_{\mathrm{T,jj}} > 2.5\,\mathrm{GeV}$, $-5.0 < \eta^*_{\mathrm{jj}} < 3.0$	Saturation model
DJANGOH	13.117	$Q^2 > 3\,\mathrm{GeV}^2$	CTEQ4D proton PDFs

Table 5.1: *List of the MCs used in the presented analysis. The first column indicates the sample, the second column the generated luminosity of the sample, the third the kinematical range where the events were produced and the fourth the most relevant aspects of the samples.*

5.1.4 Detector simulation

After being generated, the MC events were passed to the detector simulation chain. The program MOZART was used to simulate the ZEUS detector. MOZART implemented by means of the GEANT 3.13 package the ZEUS geometry as well as the the response of the detector components to the particles passage. GEANT is a multi-purpose MC whose task is to simulate all the physics processes relevant to describe the energy losses and the multiple scattering of the particles passing through the detector. The trigger simulation was carried out with the program ZGANA which uses the output coming from MOZART. After the trigger simulation the physics quantities were reconstructed by exactly the same code used for the data.

5.2 NLO calculation

As introduced in Sect. 2.2.7, the validity of the QCD factorisation theorem in diffraction can be checked by comparing the measured cross section for production of diffractive dijets with the Next-to-Leading Order (NLO) calculation using the diffractive PDFs (dPDFs) extracted from the inclusive data. One of the main parts of the analysis presented in this thesis was therefore related to the calculation of the NLO prediction.

The program used for the calculation at the order α_S^2 was DISENT [127]. DISENT is a program able to perform calculation of jet production in DIS ep collisions both at order α_S and α_S^2. This program provides predictions only at the level of partons emerging from the hard interaction. DISENT works natively only in the non-diffractive case. On the other hand the model most commonly used to describe diffractive processes is the resolved pomeron model that relies on the proton vertex factorisation (see Sect. 2.2.2). As a consequence the dPDFs are factorised into a pomeron ($I\!\!P$) flux (depending only on $x_{I\!\!P}$and t)[2] and $I\!\!P$-PDFs (depending on Q^2 and $z_{I\!\!P}$, the fraction of the $I\!\!P$ longitudinal four-momentum taken by the parton entering the hard scattering subprocess). In order to obtain the NLO cross section for diffractive ep interactions, the program must calculate the NLO cross section for e-$I\!\!P$ collision, multiply it for the $I\!\!P$-flux and integrate over the $x_{I\!\!P}$and t kinematical range.

The DISENT program was modified such to carry out the calculation in the following way:

1. Divide the $x_{I\!\!P}$range into many intervals. In the calculation presented here, 150 different values of $x_{I\!\!P}$were considered, evenly spaced between $x_{I\!\!P,\min} = 0.0025$ and $x_{I\!\!P,\max} = 0.03$.

2. For each of these values, scale down the proton beam energy by a factor $x_{I\!\!P}$.

3. Replace the proton PDFs with the $I\!\!P$-PDFs.

4. Calculate the NLO cross section with DISENT.

5. Multiply it by the t-integrated pomeron flux calculated separately.

6. At this stage, the cross section given by the program is the NLO prediction for diffractive ep DIS at fixed $x_{I\!\!P}$. In order to have the cross section over the whole $x_{I\!\!P}$range, the steps above are repeated for all the $x_{I\!\!P}$values and the cross sections at different $x_{I\!\!P}$are summed. For a sufficiently fine $x_{I\!\!P}$binning this step approximates an integration over $x_{I\!\!P}$.

The calculation of the hard-process matrix element was performed in the $\overline{MS}$ scheme with five active flavours. Since only a limited number of orders are considered in the perturbative expansion of the matrix element, a scale dependence in the prediction is still present. The renormalisation scale, μ_{R}, was set equal to $E^*_{\mathrm{T,j1}}$, where $E^*_{\mathrm{T,j1}}$ is the transverse energy of the highest transverse energy jet in the event (the leading jet) as measured in the $\gamma^* p$ centre-of-mass frame. This choice was justified by the fact that the leading-jet transverse energy was the higher (thus dominant) scale for the most of the events. The factorisation scale was set to Q^2 [3]. The evolution of the strong coupling constant with the renormalisation scale was calculated with the QCD-NUM program [128]. QCDNUM wants as input the value of α_S at a scale $\mu_{\mathrm{R}} = M_Z$; the value set was $\alpha_S(M_Z) = 0.118$.

---

[2]In all the available dPDF sets, the parametrisation of the $I\!\!P$-flux is given after integrating over the kinematically allowed range in t, such that it depends only on $x_{I\!\!P}$.

[3]A more natural choice for the factorisation scale would have been $E^*_{\mathrm{T,j1}}$ but this could not be done because of limitations in the DISENT program.

The accuracy on the dPDFs fits is not at the level of the proton ones, due to limited precision of the measurement. Furthermore many different data sets can be used. These data sets are not always compatible over the entire kinematical range. For all these reasons the dPDFs parametrisation can give quite different results in certain kinematical regions. This can be seen in Fig. 5.1 where several dPDFs are compared as function of the reconstructed $z_{I\!P}$, $z_{I\!P}^{\mathrm{obs}}$, in different bins of Q^2. The calculation was repeated with different sets of dPDFs in order to discriminate among the dPDFs the ones providing a better description of the dijets data, given the validity of the QCD factorisation theorem for diffraction.

The following dPDFs were used:

- the ZEUS LPS+charm [53] - the result of an NLO DGLAP QCD fit to the inclusive diffractive structure functions measured by the ZEUS experiment with the leading proton spectrometer (LPS). In order to better constrain the dPDFs, measurements of D^* production cross section in diffractive DIS [75] were also included. The fit was restricted to the region $x_{I\!P} < 0.01$;

- the H1 2006 dPDFs [60] - the result of an NLO DGLAP QCD fit to a sample of inclusive diffractive structure functions measured by the H1 Collaboration. Two different parameterisations are available (Fit A and B) which differ in the parametrisation of the gluon distribution at the starting evolution scale. The data used as input to the fit were restricted to the region $Q^2 > 8.5\,\mathrm{GeV}^2$, $z_{I\!P}^{\mathrm{obs}} < 0.8$. Since the H1 measurements were not corrected for the contribution due to events where the proton dissociated into a low-mass state (*proton dissociation background*), in the comparison with the data the calculations were renormalised by a factor 0.87 [60];

- the Groys-Levy-Proskuryakov (GLP) dPDFs [130] - the result of an NLO DGLAP QCD fit to a sample of inclusive diffractive structure functions measured by the ZEUS Collaboration with the M_{X} method [63]. The data used in this fit were not corrected for proton dissociation background, therefore the obtained predicitons were scaled by a factor 0.70 ± 0.03 [63];

- the Martin-Ryskin-Watt 2006 (MRW 2006) dPDFs [131] - the result of a fit to the same data set as for the H1 2006 fit. Regge factorisation is assumed only at the input scale. The dPDFs are then evolved with an inhomogeneous evolution equation analogous to that for the photon PDFs. The inhomogeneous term accounts for the perturbative Pomeron-to-parton splitting. The inhomogenous terms in the evolution equation are additional to the usual NLO DGLAP terms and account for the perturbative Pomeron-to-parton splitting.

The only theoretical uncertainty considered was that coming from the limited order of the NLO calculations. The effect of the missing higher perturbative orders was estimated by varying μ_{R} by factors of 0.5 and 2. Uncertainties of more than 20% were obtained. Other possible sources of uncertainty not considered were the ones related to the dPDFs fits and the absence in DISENT of resolved-photon contributions that can affect the prediction at low values of x_γ.
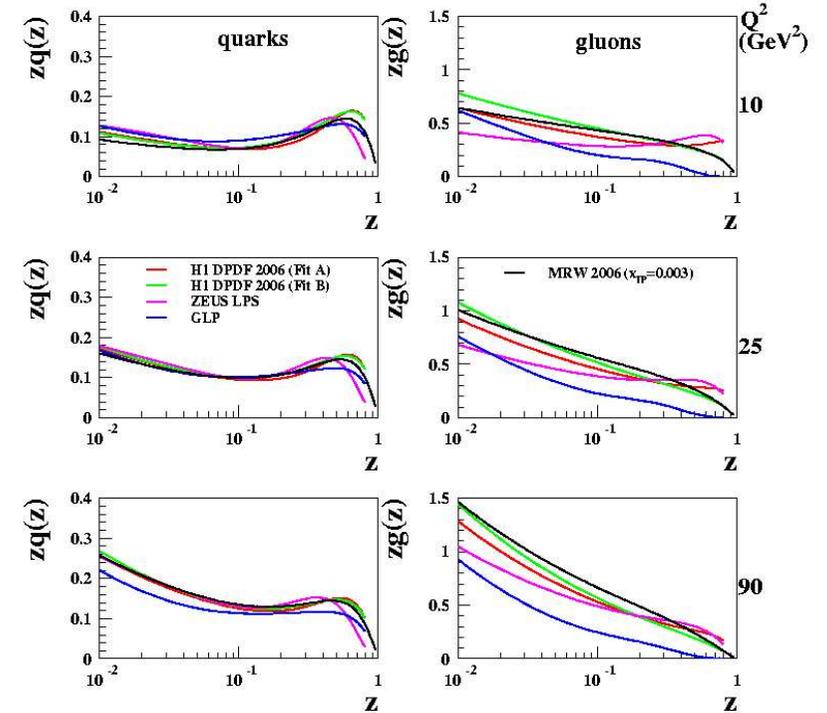
Figure 5.1: The $z_{I\!P}^{\mathrm{obs}}$ dependence of the $I\!P$-PDFs for different values of Q^2. The curves were produced at fixed $x_{I\!P} = 0.003$. The singlet (i.e. quark) and gluon contributions are shown separately. The uncertainties on the parametrisations are not shown.

5.2.1   Hadronisation corrections

The predictions given by the NLO program needed to be quoted at the hadron level (defined as the level of hadrons with a lifetime $\tau > 10\ ps$) in order to be compared to the measured cross section. The parton level differential cross section were corrected bin-by-bin to the hadron level by means of factors evaluated with RAPGAP . These hadronisation corrections take in account the effects of the hadronisation of the partons. The hadronisation correction for a given bin i of the differential cross section, C^i_{had}, is calculated doing the ratio of the hadron level cross section given by the MC, $\sigma^i_{\text{MC,HL}}$, over the parton level cross section given by the MC, $\sigma^i_{\text{MC,PL}}$:

$$C^i_{\text{had}} = \frac{\sigma^i_{\text{MC,HL}}}{\sigma^i_{MC,PL}}. \tag{5.2}$$

In the evaluation of the hadronisation corrections, the different RAPGAP contributions were summed up without the different normalisation weights (see Sect. 6.3). This was chosen for consistency with the NLO calculation where the different contributions are summed without reweighting.

The hadronisation corrections obtained are shown in Fig. 5.2 as a function of many variables. It can be noticed the strong rise at high $z^{\text{obs}}_{I\!\!P}$. This effect is a consequence of the interplay of the heavy mass of the charm quark with the low invariant mass of the hadronic system produced in the interaction. The heavy charm quark mass saturates the small phase space available such that the hadronisation model produces only two D mesons. The two exclusively produced mesons are then identified as jets by the k_T algorithm. Therefore the event that could have been generated with an arbitrary $z_{I\!\!P}$ is reconstructed, after the hadronisation, with $z^{\text{obs}}_{I\!\!P} \sim 1$. It can be noticed in the same figure that, if one doesn't consider the heavy quark contribution in RAPGAP, C_{had} is significantly smaller at high $z^{\text{obs}}_{I\!\!P}$. Since the cross section at high $z^{\text{obs}}_{I\!\!P}$ is small, the large correction there does not affect the corrections on the other variables.
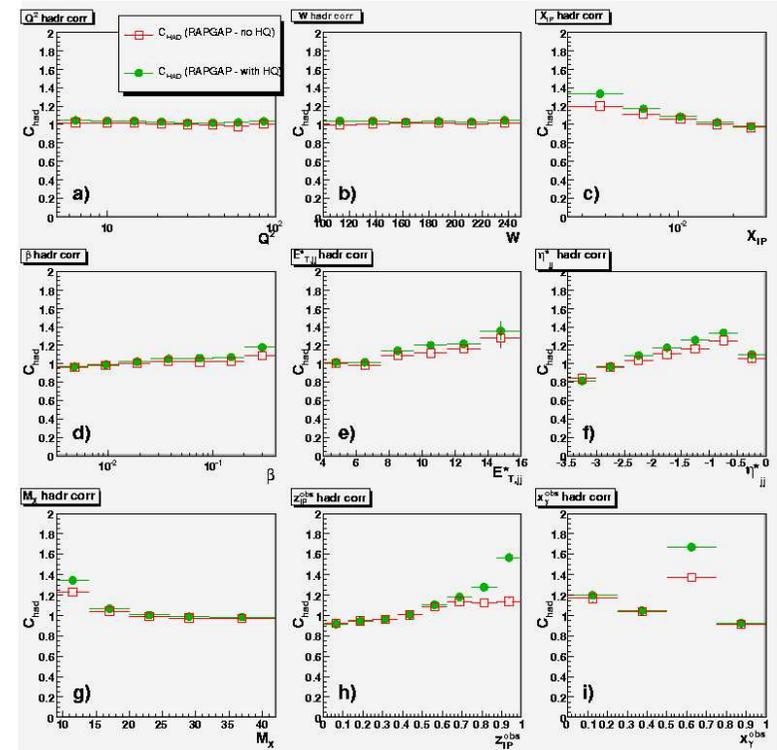
Figure 5.2: The hadronisation correction, C_{had}, calculated with the RAPGAP MC. The correction factors are shown as a function of (a) Q^2, (b) W, (c) $x_{I\!\!P}$, (d) β, (e) $E^*_{\text{T,jj}}$, (f) η^*_{jj}, (g) M_{X}, (h) $z^{\text{obs}}_{I\!\!P}$, (i) x^{obs}_γ. The full circles show the correction factors applied to the parton level NLO predictions obtained with DISENT. The open squares show the amount of the correction if one excludes the direct-photon charmed quark contribution in RAPGAP.

Chapter 6

Signal selection and background rejection

In order to extract a clean sample of diffractive dijet events in DIS, the data described in Sect. 4 were subject to a two-stage selection. First the data were filtered online by the ZEUS trigger. Two different trigger configurations were used to select the data presented here. The events passing this first online filter were then passed to an offline analysis program that applied the final selection. The selection asked for a well-reconstructed scattered electron, at least two jets with a minimum transverse energy emitted in a central pseudorapidity region of the detector and a rapidity gap in the event. The MC was subject to the same analysis chain. A comparison of the distributions coming from real data and MC are also shown here. The chapter describes also the strategies pursued in order to reject the main background sources for this analysis.

6.1 Signal selection

6.1.1 Online selection

Given the limited computing resources and bandwidth of the data transfer, it is impossible to store all the data coming from all the collisions. In order to reject events that did not have the characteristics of the physical signal desired, the data passed a three-stage trigger selection before being stored on tape and being analysed offline. At the first level of the trigger (FLT), some general requirements for a DIS selection were applied. For example, an isolated electromagnetic cluster in the calorimeter or the total electromagnetic energy above a threshold were the typical requirements, often in coincidence with a rough requirement on the quality of the tracks measured by the tracking chamber or a signal from the SRTD. Only the events passing the FLT were analised by the second level trigger (SLT). At this stage additional, more refined requirements were applied. The event was kept only if

- $\delta_{\mathrm{SLT}} > 30$ GeV, where δ_{SLT} is the total $E - p_Z$ measured at the SLT;

- one of the following was satisfied:

  - the electromagnetic energy in the rear (RCAL) or the barrel calorimeter (BCAL) was greater than 2.5 GeV

  - the electromagnetic or hadronic energy in the front calorimeter (FCAL) was greater than 10 GeV.

It is noted that until now no requirements on the diffractive nature of an event were applied. These were implemented at the third level trigger (TLT). At the TLT the requirements on the event were the following:

- the event had to pass the SLT requirements described above;

- $\delta_{\mathrm{TLT}} > 30$ GeV, where δ_{TLT} is the total $E - p_Z$ measured online at the TLT;

- a first rough but fast neural network algorithm was run in order to find a possible scattered electron. At least one candidate had to be found and its energy had to be higher than 4 GeV;

- if the same scattered electron candidate was found by the TLT in the RCAL, it had to be detected outside a rectangular region centered around the beam pipe (*trigger box cut*). The size of this box was 12×6 cm^2 (see Fig. 6.2);

- $E_{\mathrm{FPC,TLT}} < 20$ GeV, where $E_{\mathrm{FPC,TLT}}$ is the energy measured in the forward plug calorimeter online by the TLT. This acted as a diffractive rapidity gap selection since it was essentially a veto on the hadronic activity in the forward region of the detector.

The TLT logic described above (TLT-DIS06) was the one valid for the part of the data taking period when the positron beam was used (~ 61 pb^{-1}). When electrons were collided against protons (~ 3 pb^{-1}), the TLT logic was different, being a purely low-Q^2 DIS selection with $Q^2_{\mathrm{TLT}} > 2$ GeV^2. The FPC requirement was not applied.

6.1.2 DIS selection

A pure DIS sample was selected applying the following requirements on the events that passed the trigger selection:

- the SINISTRA neural network algorithm had to find at least one candidate with probability higher than 0.9. If the electron was in the CTD geometric acceptance, a track was required to be matched to the electromagnetic cluster. The SINISTRA probability is shown in Fig. 6.1 for events both before and after passing the DIS selection;

- the electron found in the RCAL had to lie outside a fiducial area centred around the beam pipe. The excluded area (*H*-shape box) was defined as follows [132]:

  - *box cut*, $-14 < X < 12$ cm and $-10 < Y < 10$ cm;

  - *cooling pipes*, four regions defined as
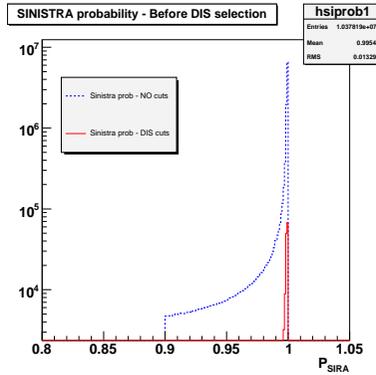
    * $-16 < X < -7$ cm and $4 < Y < 12$ cm;

Figure 6.1: SINISTRA *probability for data events before (dashed line) and after (solid line) the DIS selection.*

       ∗ $3 < X < 12$ cm and $4 < Y < 12$ cm;

       ∗ $-16 < X < -7$ cm and $-12 < Y < -4$ cm;

       ∗ $3 < X < 12$ cm and $-12 < Y < -4$ cm.

This cut rejects events where the electron was detected in regions of the RCAL with significant amount of inactive material difficult to be accurately described in the MC. This would affect the reconstruction of the electron four-momentum and consequently bias the measurement of the DIS kinematics. In Fig. 6.2 are shown the positions of the impact point on RCAL of the reconstructed scattered electrons for events passing the DIS selection. The profile of the H-shape box is clearly visible;

- the energy of the scattered electron had to be greater than 10 GeV. This selection guaranteed that the efficiency of the SINISTRA electron finder was sufficiently high (see Fig. 4.2);

- the Z-position of the vertex of the event had to be in the range $|Z_{\mathrm{VTX}}| < 50$ cm. This cut excluded events not originated from ep collisions, like events coming from beam-gas interactions.

After having applied these selection cuts, kinematic cuts on the virtuality of the γ^*, Q^2, and the γ^*p centre-of-mass energy, W, were applied. In this way the precise kinematic region where the ZEUS detector has a good sensitivity is defined. The cuts chosen were:

- $5 < Q^2_{\mathrm{DA}} < 100$ GeV^2;

- $100 < W_{\mathrm{DA}} < 250$ GeV;
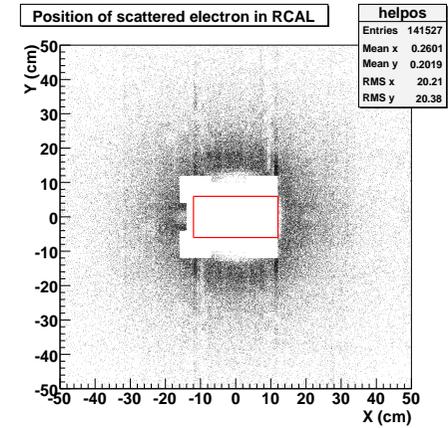
Figure 6.2: The position of the detected scattered electron on the RCAL surface. The red solid line represents the position and size of the box used at the TLT for the online selection.

where the index DA indicates that the two variables were reconstructed with the Double Angle method (see Sect. 4.2.4).

6.1.3   Jet selection

Jets were reconstructed at the detector level, i.e. from the measured four-momenta of the energy flow objects (EFOs), with the k_{T} algorithm described in Sect. 4.2.5. In the case of the MC samples, also the four-momenta of the particles before the hadronisation simulation and before the detector simulation are also used: we will refer to these jets as parton-level and hadron-level jets, respectively. In all the cases, the jets were reconstructed after having boosted the input four-momenta from the laboratory frame to the γ^*p rest frame. The input four-momenta were treated as massless. The jet-related quantities evaluated in this reference frame will be labeled with a star. The reconstructed jets are then boosted back to the laboratory frame (jet quantities in this frame are labeled with "LAB"). The jets are ordered in their transverse energy in the γ^*p rest frame, $E^*_{\mathrm{T,jet}}$, the first jet in the ordering being called the *leading jet*.

A clean dijet sample is obtained by imposing specific requirements on the jets reconstructed from the four-momenta. Before applying the cuts, the four-momenta of the jets were further corrected on a MC basis (see Appendix B). Only events with jets of a sufficiently high transverse energy to provide a hard scale needed for perturbative QCD calculation are selected. In order to reconstruct the jet properties in a precise way, selections on their directions were applied. The requirements on the jets were the following:
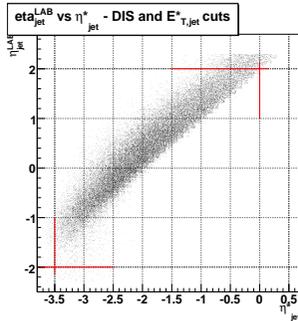
Figure 6.3: *The correlation of the jet pseudorapidity as measured in the laboratory, $\eta_{\text{jet}}^{\text{LAB}}$, and in the $\gamma^* p$ rest frame, η_{jet}^*, for the (upper plot) data and (lower plot) MC sample. The lines indicate the cuts on the jet pseudorapidity applied in the analysis.*

- At least two jets with

  - $E_{\text{T,jet}}^* > 4 \, \text{GeV}$;

  - $-2.0 < \eta_{\text{jet}}^{\text{LAB}} < 2.0$; this cut was applied in order to ensure a good containment in the CAL acceptance of the EFOs belonging to the jet;

  - $-3.5 < \eta_{\text{jet}}^* < 0.0$, that corresponds for the majority of the jets to the pseudorapidity range in the laboratory frame set by the previous cut, as visible in Fig. 6.3;

- $E_{\text{T,j1}}^* > 5 \, \text{GeV}$, where E_{T}^* is the transverse energy in the $\gamma^* p$ rest frame of the leading jet.

6.1.4 Diffractive selection

Diffractive events are characterised by low values of $x_{I\!\!P}$ and by the presence of a LRG (see Sect. 2.2). The following selection criteria were applied[133]:

- $E_{\text{FPC}} < 1 \, \text{GeV}$, where E_{FPC} is the total energy in the FPC. The requirement of activity compatible with the noise level in the angular region covered by the FPC is equivalent to a rapidity-gap selection;

- $x_{I\!\!P}^{\text{obs}} < 0.03$ where $x_{I\!\!P}^{\text{obs}}$ is the reconstructed value of $x_{I\!\!P}$ (see Sect. 4.2.6). The cut on $x_{I\!\!P}^{\text{obs}}$ reduces the contribution of Reggeon exchange and other non-diffractive background.

The contamination of the non-diffractive background as a function of the applied diffractive selection cuts is shown in Fig. 6.4, through the distribution of η_{MAX}, where

Selection	Nr. Events after selection
Trigger	13764440
DIS cuts	147776
DIS+JETS cuts	37872
ALL cuts	5540

Table 6.1: *Number of events remaining after different stages of the signal selection.*

η_{MAX} is the pseudorapidity in the laboratory frame of the most forward EFO with energy higher than 400 MeV. Distributions of η_{MAX} are shown before and after applying cuts on E_{FPC} and $x_{I\!\!P}^{\text{obs}}$. The disagreement between the measured and the simulated distributions is the reason for not applying any explicit requirement on η_{MAX}, as was done in previous analyses [57, 75, 76]. After the E_{FPC} and $x_{I\!\!P}^{\text{obs}}$ cuts, the non-diffractive background from DJANGOH was estimated to be 2.4% of the total selected events and neglected in further analysis.

6.2 Trigger efficiency

It is important to evaluate reliably the fraction of events that have been lost because the trigger judged them wrongly to be background, although having all the characteristics for entering the physical signal sample selected in the way described above.

The definition of trigger efficiency, $\varepsilon_{\text{TRIGG}}$, is

$$\varepsilon_{\text{TRIGG}} = \frac{N_{\text{SEL}}^{\text{TRIGG}}}{N_{\text{SEL}}} \tag{6.1}$$

where N_{SEL} is the number of events in the starting sample that, independently of the trigger decision, pass the signal selection and $N_{\text{SEL}}^{\text{TRIGG}}$ is the subsample of N_{SEL} that passes also the three-level online filter[1]. In case of losses due to the trigger, one can correct offline the measured distributions according to the trigger efficiency. The absence of an unbiased (i.e. non-triggered) and high-statistics sample to be used as reference made the estimation of the trigger efficiency only from data samples problematic. The strategy used was to evaluate the efficiency of the trigger slot not in absolute way but relatively to another, more inclusive, filter. The inclusive low-Q^2 trigger slot, TLT-SPP15, was used as reference for this task. This slot has not only the advantage of being very inclusive and of having high statistics but also to be very well known. Its efficency is very high as it was tested by many previous DIS analyses [63]. The Eq.(6.1) was therefore modified as the following

$$\varepsilon_{\text{TRIGG}}^{\text{SPP15}} = \frac{N_{\text{SEL\&\&SPP15}}^{\text{TRIGG}}}{N_{\text{SEL\&\&SPP15}}} \tag{6.2}$$

where now the efficency of the diffractive trigger, $\varepsilon_{\text{TRIGG}}^{\text{SPP15}}$, is quoted relatively to the efficiency of the TLT-SPP15 slot and the events must not only pass the physics

---

[1] The trigger inefficiency is defined as $1 - \varepsilon_{\text{TRIGG}}$.
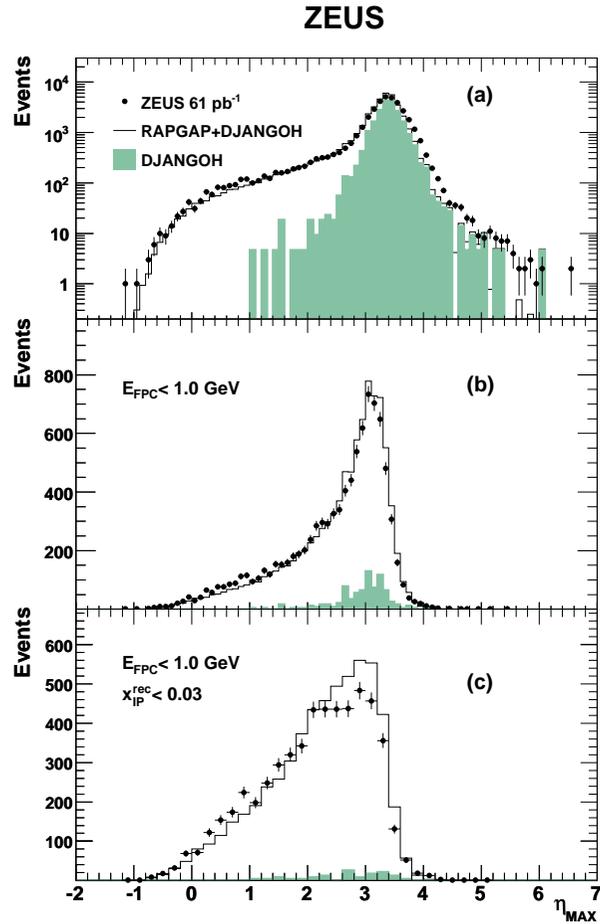
Figure 6.4: The measured η_{MAX} distribution (dots) (a) before diffractive selection, (b) after the E_{FPC} cut and (c) after adding the $x_{I\!\!P}^{\mathrm{obs}}$ cut. Also shown are area-normalised MC expectations obtained by fitting the relative amount of RAPGAP and DJANGOH to give the best description of the data before any diffractive selection.
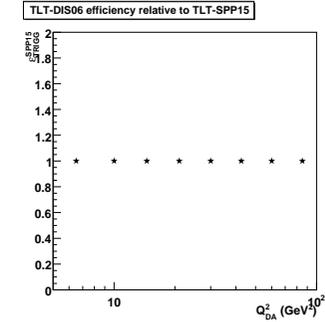
Figure 6.5: The efficiency of the diffractive trigger slot DIS06 relatively to the inclusive low-Q^2 trigger slot SPP15 as a function of Q^2.

selection but being also accepted by the reference trigger filter. In Fig. 6.5 the value of $\varepsilon_{\mathrm{TRIGG}}^{\mathrm{SPP15}}$ as a function of Q_{DA}^2 is shown. A full efficiency of the diffractive trigger compared to the inclusive one is observed over the entire range.

In order to estimate the absolute efficiency of the trigger as defined in Eq. (6.1), a MC study was performed with the RAPGAP sample available. The value of $\varepsilon_{\mathrm{TRIGG}}$ as a function of Q_{DA}^2 and $E_{T,j1}^*$ is shown in Fig. 6.6. The efficiency is always higher than 98 %. Given these results indicating a minimal loss of good physics events because of the trigger, it was decided to neglect in the next steps of the analysis any bias effect due to it.

6.3   Monte Carlo reweighting

The RAPGAP sample consists of three different subsamples corresponding to different physics processes (see Sect. 5.1.1): the direct-photon light-quark sample (LQ), the direct-photon charm-quark sample (CC) and a resolved-photon sample (RES). In order to have a complete MC sample, the three RAPGAP samples had to be summed up. This cannot be done in a straight forward way, because the relative contribution of direct and resolved processes to the total cross section is a priori unknown in Leading-Order calculations. The absolute normalisations for the different RAPGAP samples were evaluated from a fit to the data. As mentioned in Sect. 2.2.7, the variable most sensitive to the separation between direct and resolved processes is x_γ. The experimentally observable estimator of x_γ is x_γ^{obs}, introduced in Sect. 4.2.7. The data distribution as a function of x_γ^{obs} after all the selection cuts is shown in Fig. 6.7. The direct-photon events concentrate at high x_γ^{obs}, while the low-x_γ^{obs} part of the spectrum is composed of resolved-photon events and badly reconstructed direct-photon events. The three
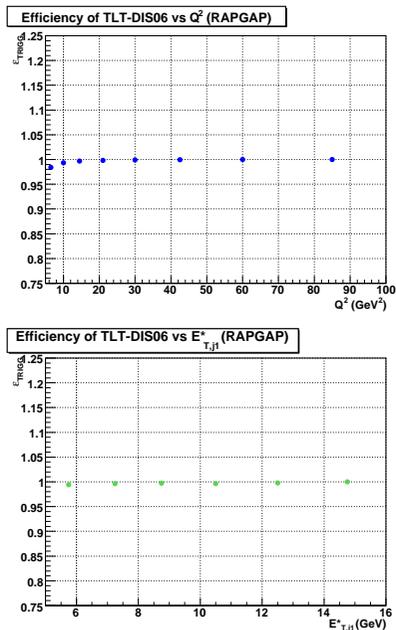
Parameter	Value from the fit
α_{LQ}	0.89
α_{CC}	0.89
α_{RES}	2.99

Table 6.2: The normalisation parameters of the single RAPGAP *subsamples. These parameters are used when the subsamples are summed together in order to obtain the total* RAPGAP *simulation.*
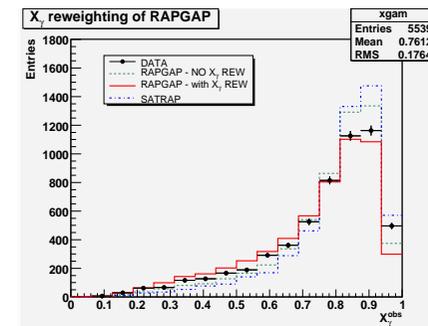
Figure 6.7: The x_γ distribution after all the selections for data (black dots) and the RAPGAP *MC before (dashed-dotted green line) and after (solid red line) the reweighting described in Sect.6.3. The* SATRAP *distribution is also shown as a dashed blue line.*

RAPGAP samples are summed up with different normalisations

$$N_{\mathrm{RAP}}^{\mathrm{TOT}} = \alpha_{\mathrm{LQ}} N_{\mathrm{RAP}}^{\mathrm{LQ}} + \alpha_{\mathrm{CC}} N_{\mathrm{RAP}}^{\mathrm{CC}} + \alpha_{\mathrm{RES}} N_{\mathrm{RAP}}^{\mathrm{RES}} \qquad (6.3)$$

The three normalisation coefficients are estimated such that the total-RAPGAP x_γ^{obs} distribution fits best the data one. The fit was performed via a χ^2 minimisation. The fit was performed over the entire x_γ^{obs} range after having normalised all the RAPGAP samples to the luminosity of the data.

Fig. 6.7 also shows the RAPGAP x_γ^{obs}-distributions obtained summing the different MC contributions with and without the reweighting coefficients. The description of the data by RAPGAP is improved. No significant distortion of the shapes was observed in the distributions of the other variables. In the following sections, we will refer to the total RAPGAP sample as the sample given by the sum of the three contributions as in Eq. (6.3) using the coefficients coming from the fit. The latter are summarised in Table 6.2. The same thing could not be done with SATRAP because it does not include any resolved-photon contribution in it (see Sect. 5.1.2).

Figure 6.6: Absolute efficiency of the diffractive trigger slot DIS06 estimated with the RAPGAP MC. The upper (lower) plot shows the efficiency as a function of Q_{DA}^2 ($E_{T,j1}^*$).

6.4   Control distributions

After all cuts, 5540 events of the originary data sample were selected. The number of events passing the different stages of the selection is summarised in Table 6.1. The distributions of the data after all the selection cuts are presented in Figs. 6.8-6.12. In the same plots the same MC distributions are compared to the data in order to test the level of the description of the simulation. The signal MCs, i.e. the SATRAP and the total RAPGAP samples, were renormalised to the total entries in the data. The background from non-diffractive DIS dijets as estimated with DJANGOH is also shown in the plots.

The control plots present the comparison between data and MC first in the global variables (Fig. 6.8), then for jet-related variables (Fig. 6.9 and 6.10), for diffraction-related variables (Fig. 6.11) and at the end for other quantities more linked to the detector performances (Fig. 6.12).

The general level of the agreement between data and MC is good. The distribution of the main kinematical variables (Q^2, W, jet variables) are fairly described by both the MCs. The phase space regions at low and high transverse energies of the jets is better described by RAPGAP , as it can be seen in Figs. 6.9 and 6.10. At medium values of the transverse energy of the jets, the two MC predict similar distributions. The SATRAP MC describes better the data distribution as a function of $z_{I\!P}^{\mathrm{obs}}$, in particular for $z_{I\!P}^{\mathrm{obs}} > 0.7$ (see Fig. 6.11c). The absence of a resolved-photon contribution in SATRAP affects its description of x_γ^{obs}, as it can be seen in Fig. 6.11d. Some discrepancies between the MC and the data distributions can be observed. The main one regards the description of the energy flow in the MC. In Fig. 6.8d it can be noted that the spectrum of the mass of the diffractive system produced in the interaction has less events at low masses while the intermediate and high masses regions are reasonably well described. The same effect can be seen in Fig. 6.12d where the number of EFOs in the event predicted by the MC is slightly shifted towards higher values compared to the data distribution. Consequently, the spectrum of *beta* (which is inversely related to M_X) is shifted to higher values in the MC compared to the data. However, the difference between MC and data was judged to be small and no further studies on it were carried out. In Fig. 6.12a, it can be noted that there is a remarkable difference between both the MCs and the data in the distribution as a function of the energy deposited in the FPC, E_{FPC}. This variable is very difficult to be simulated because involves non-perturbative processes like the fragmentation of the dissociated scattered proton. Anyway, the cut on E_{FPC} is applied at $E_{\mathrm{FPC}} < 1\,\mathrm{GeV}$, thus far from the region badly described at $0.1 < E_{\mathrm{FPC}} < 0.4$. Therefore this discrepancy was judged to be not important for the analysis and neglected.

6.5   Proton dissociation background

The process where the outgoing proton does not emerge intact from the diffractive interaction is called *proton dissociation* or *double dissociation* (see Fig. 6.13). The hadronic system coming from the dissociated proton, Y, has a low invariant mass, M_Y,
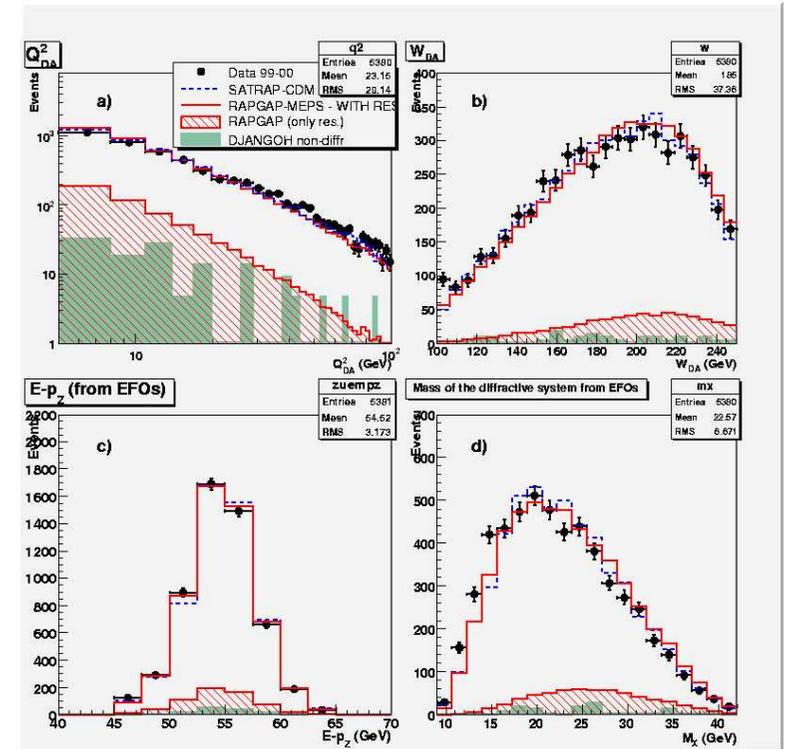

Figure 6.8: The control distributions as a function of (a) Q^2_{DA}, the virtuality of the exchanged γ^* measured with the Double Angle method, (b) W_{DA} the total energy available in the $\gamma^* p$ centre-of-mass system, (c) the total $E - p_Z$ and (d) M_X, the invariant mass of the diffractive system. The data entries are shown as dots, the statistical errors are shown as the error bars. The data are compared to the signal LO MC, RAPGAP (solid red line) and SATRAP (dashed blue line). The contribution from resolved-photon processes estimated with RAPGAP is shown as a hatched area. The non-diffractive DIS background estimated with DJANGOH is indicated by the dark solid area.
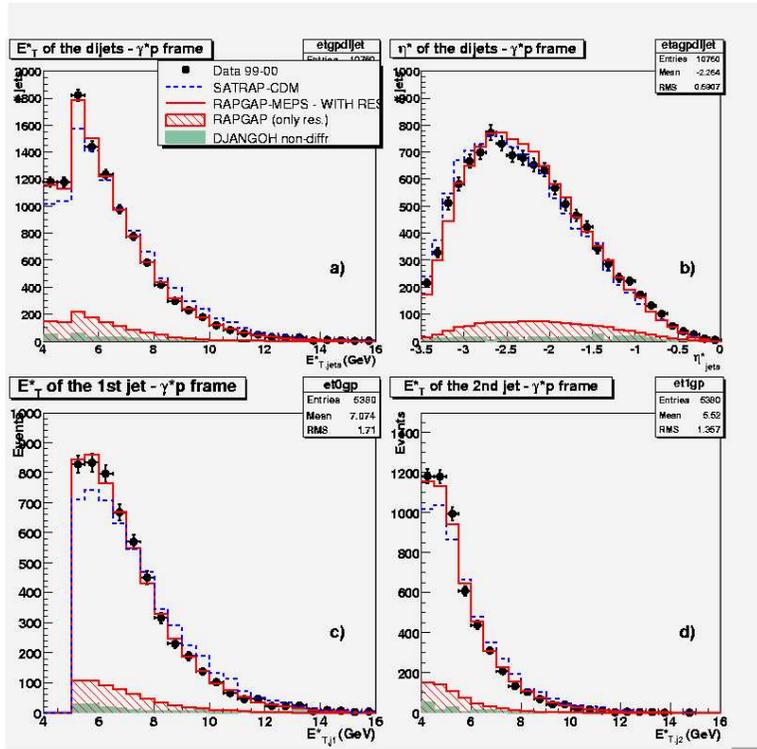
Figure 6.9: The control distributions as a function of (a) $E^*_{T,jj}$, the transverse energy of both the two jets with the highest transverse energy as measured in the γ^*p rest frame, (b) η^*_{jj}, the pseudorapidity of both the two jets with the highest transverse energy as measured in the γ^*p rest frame, (c)$E^*_{T,j1}$, the transverse energy of the highest transverse energy jet as measured in the γ^*p rest frame and (d) $E^*_{T,j2}$, the transverse energy of the next-to-highest transverse energy jet as measured in the γ^*p rest frame. Other details as in the caption of Fig. 6.8.
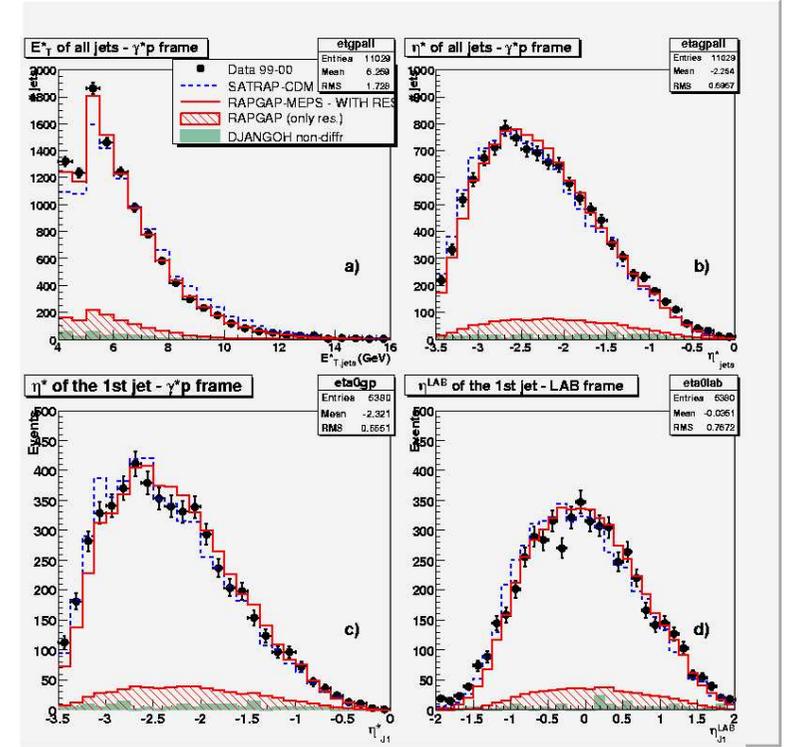
Figure 6.10: The control distributions as a function of (a) $E^*_{T,jet}$, the transverse energy of all the jets in the event as measured in the γ^*p rest frame, (b) η^*_{jet}, the pseudorapidity of all the jets in the event as measured in the γ^*p rest frame, (c)η^*_{j1}, the pseudorapidity of the highest transverse energy jet as measured in the γ^*p rest frame and (d) η^{LAB}_{j1}, the pseudorapidity of the highest transverse energy jet as measured in the laboratory frame. Other details as in the caption of Fig. 6.8.
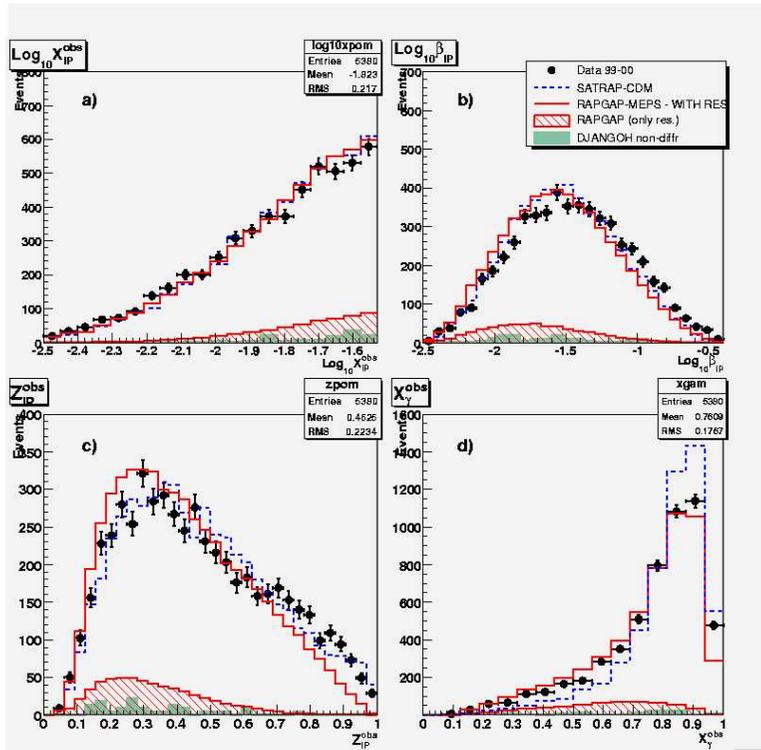
Figure 6.11: The control distributions as a function of (a) $Log_{10}x^{\mathrm{obs}}_{I\!\!P}$, (b) $Log_{10}\beta_{I\!\!P}$, (c) $z^{\mathrm{obs}}_{I\!\!P}$ and (d) $x^{\mathrm{obs}}_{\gamma}$. Other details as in the caption of Fig. 6.8.
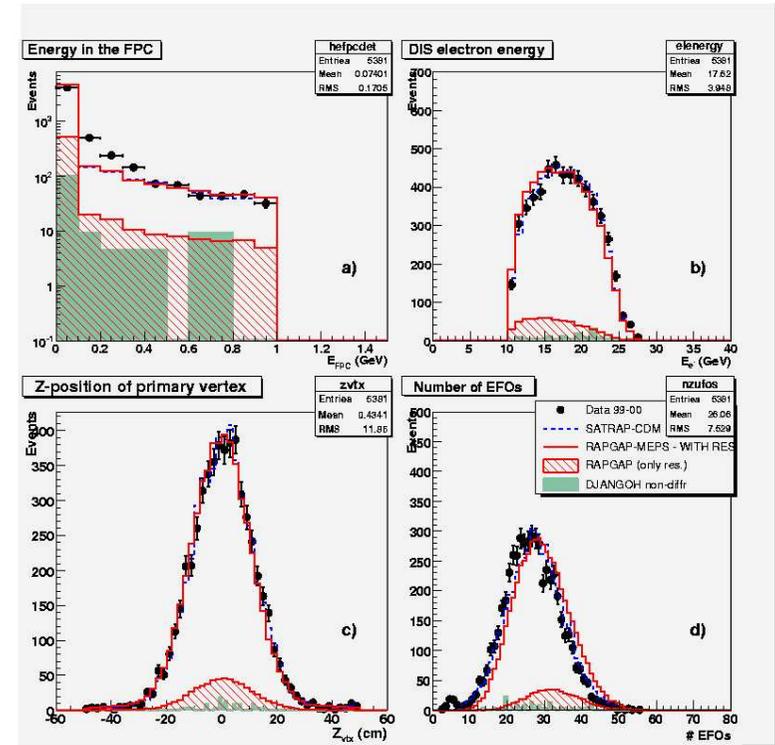

Figure 6.12: The control distributions as a function of (a) E_{FPC}, the energy measured in the FPC (b) $E_{e'}$, the energy of the scattered electron (c) Z_{VTX}, the position along the $Z-axis$ of the primary vertex and (d) the number of EFOs in the event. Other details as in the caption of Fig. 6.8.

typically below $2\,\mathrm{GeV}$. In this cases a LRG can be still observed. However, the proton-dissociative events are considered background and rejected from the signal sample, because the proton dissociation introduces a significant complication to the theoretical description of the event since the resolved-pomeron model (see Sect. 2.2.2) is no longer valid.

The experimental detection of this subset of events is challenging because, for sufficiently low values of M_Y, the resonant system escapes in the forward aperture of the detector and one cannot guess anything about the final state proton from the central detector only. Vague constraints on the values of M_Y can be set on kinematical and geometric acceptance basis. The FPC energy selection applied, $E_{\mathrm{FPC}} < 1\,\mathrm{GeV}$, restricts the values of the resonant system mass to approximately $M_Y \lesssim 2.3\,\mathrm{GeV}$. For M_Y below that value, a statistical subtraction from the selected signal sample is needed. The amount of proton dissociation background, f_{pdiss}, was estimated by previous analyses [75, 79] to be

$$f_{\mathrm{pdiss}} = (16 \pm 4)\%$$

This value was estimated with a MC study using the EPSOFT MC [134]. A more detailed description of the procedure can be found in Ref. [135]. The proton dissociation background could be estimated also through an experimental method by means of the ZEUS Leading Proton Spectrometer (LPS) [53]. Since the LPS tags diffractive events with only intact protons by construction, the amount of events with a dissociated proton can be obtained by comparing directly the LPS and the LRG inclusive measurements. By doing the ratio of the cross sections one obtains a value of f_{pdiss}

$$f_{\mathrm{pdiss}}^{\mathrm{LPS}} = (19 \pm 10)\%$$

The latter study showed that the proton dissociation background is indipendent of the other kinematical variables, as expected in the resolved pomeron model. Therefore the proton dissociation background was subtracted from the cross section independent of any variable. The compatibility between the two results obtained with the MC simulation and the LPS measurement and the smaller uncertainty of the former were the reasons for choosing the value of $(16 \pm 4)\%$ as estimate of f_{pdiss}.
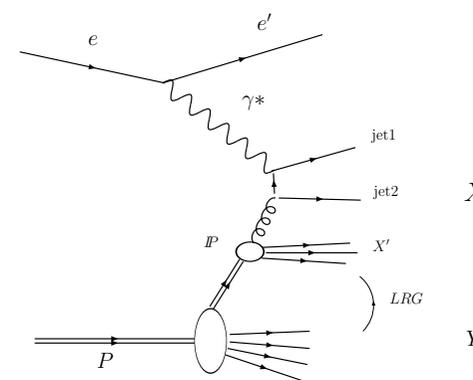

Figure 6.13: Schematic representation of proton dissociation in presence of diffractive DIS dijet production.

Chapter 7

Cross section measurement and discussion

This chapter presents the main results of the analysis: the extraction of the differential cross sections and the comparison of the Leading-Order (LO) and Next-To-Leading Order (NLO) theoretical predictions and the experimental measurements. The double differential cross sections, that can be used for fits to the dPDFs, are also shown. The corrections for detector acceptance and QED radiation are essential for an unbiased measurement of the physical process under study. The procedures for evaluating and implementing these corrections are described at the beginning of the chapter together with the error estimation.

7.1 Cross section extraction

The cross section for dijet production in diffractive DIS ep collision is measured at the hadron level, defined as the level of hadrons with a lifetime $\tau > 10 \, ps$. The detector level measurement was transported to the hadron level with correction factors that estimated the detector acceptance and efficiency. The measured cross section is corrected for QED effects and quoted at the QED Born level. The latter two corrections were both estimated on a MC basis as described in Sect. 7.1.1 and Sect. 7.1.2, respectively. The kinematic region where the cross section was measured is summarised in Table 7.1.

Kinematic region
$5 < Q^2 < 100 \, \text{GeV}^2$
$100 < W < 250 \, \text{GeV}$
$E_{\text{T,j1}}^* > 5.0 \, \text{GeV}, \quad -3.5 < \eta_{\text{j1}}^* < 0.0$
$E_{\text{T,j2}}^* > 4.0 \, \text{GeV}, \quad -3.5 < \eta_{\text{j2}}^* < 0.0$
$x_{\mathbb{P}} < 0.03$

Table 7.1: The kinematic region where the cross section was measured.

The values of the differential cross sections are averaged over the bin in which they are presented. For any variable κ, the cross section was determined as

$$\frac{d\sigma}{d\kappa} = C \, \frac{N_D(1 - f_{\text{pdiss}})}{\mathcal{L} \, \Delta\kappa}, \tag{7.1}$$

where N_D is the number of data events counted in a bin, C includes the effects of the acceptance and the QED correction factors as determined from MC, $\mathcal{L}$ is the integrated luminosity and $\Delta\kappa$ is the bin width. The main criteria for the choice of the bin width were: the resolution on the variable itself that limits the minimum bin size; the minimisation of migrations of events between different bins and the statistical significance of the measurement, both suggesting a large bin size; a high number of bins (i.e. a small bin size) in order to study in more detail the characteristics of the process.

The differential cross section for dijet production in diffractive DIS ep collision is presented as a function of the following variables:

- Q^2, the virtuality of γ^*, the photon exchanged between the colliding e and p. This variable defines the hard scale of a DIS interaction and allows the perturbative QCD (pQCD) description of the process;

- W, the energy of the γ^*p centre of mass. It defines the total energy available in the hard interaction;

- $x_{\mathbb{P}}^{\text{obs}}$, the reconstructed value of $x_{\mathbb{P}}$, as defined in Sect. 4.2.6. In the resolved pomeron model, it represents the fraction of the initial p momentum taken by the diffractive exchange, $\mathbb{P}$, probed by the γ^*. Therefore it is a variable that characterises the diffractive properties of the process;

- β, introduced in Sect. 2.2.3 that is connected to x_{Bj}, the x-Bjorken variable commonly used in standard DIS (see Sect. 2.1.3). The relation between β and x_{Bj} is $x_{\text{Bj}} = x_{\mathbb{P}} \cdot \beta$;

- $E_{\text{T,jj}}^*$, the transverse energies of the two jets with the highest transverse energy as measured in the γ^*p rest frame. Thus, in the Eq. (7.1) it contributes with two entries per event[1]. The transverse energy of the jets provides an additional hard scale in the process that often is higher than Q^2, therefore the pQCD theory has a special sensitivity to this variable.

- η_{jj}^*, the pseudorapidities of the two jets with the highest transverse energy as measured in the γ^*p rest frame. This is another jet variable very useful for testing the quality of the pQCD prediction, given also that the jet direction is experimentally measured with good precision;

---

[1]The use of $E_{\text{T,jj}}^*$ instead the single transverse energies of the jets was dictated by the convergence of the NLO calculation that is sensitive to the small difference in the E_{T}^* requirement between the first and the second jet. The definition of a more inclusive variable helped to cancel out divergences in the calculation.

- M_X, the invariant mass of the diffractive system produced in the $\gamma^* I\!P$ interaction;

- $z_{I\!P}^{\rm obs}$, the estimator of $z_{I\!P}$ (see Sect. 4.2.6). The latter is the fraction of the $I\!P$ momentum taken by the parton entering the hard sub-process. This is a very important variable because it is the one sensitive to the dPDFs (in a fashion similar to $x_{\rm Bj}$ for inclusive DIS analyses). The quality of different dPDF fits can be studied mainly as a function of this variable. The cross section as a function of this variable can also be used as input to QCD fits for the dPDFs themselves;

- $x_\gamma^{\rm obs}$, the estimator of x_γ (see Sect. 4.2.7). The variable x_γ is the fraction of γ^* longitudinal momentum entering the hard-subprocess. Although for DIS processes one expects only events with $x_\gamma^{\rm obs} = 1$ (direct-photon processes), in the low-Q^2 range studied in this analysis a contribution from $x_\gamma^{\rm obs} < 1$ (coming from resolved-photon processes and badly reconstructed direct-photon events) to the total cross section is still present. The most accepted models of QCD factorisation breaking in diffraction predict a suppression of the cross section only for the resolved-photon processes. In the diffractive photoproduction of dijets, the factorisation breaking is expected to exhibit a dependence on $x_\gamma^{\rm obs}$. It would therefore be interesting to measure the cross section as a function of the same quantity in DIS.

7.1.1   Acceptance calculation

Detector effects (like geometric acceptance, finite resolution, detection efficiency) bias the measurement such that the detector-level cross section can be significantly different to the hadron-level one. The size of these distortions is estimated on a MC basis and the detector-level measurement is corrected for it with the *bin-by-bin method*. The correction factors, C_f, for the i-th bin of any variable are defined as[2]

$$C_{f,i} = \frac{N_{\rm HAD,i}^{\rm MC}}{N_{\rm DET,i}^{\rm MC}} \qquad (7.2)$$

where $N_{\rm HAD,i}^{\rm MC}$ and $N_{\rm DET,i}^{\rm MC}$ indicate the number of events predicted by the MC in the i-th bin at the hadron level and the detector level, respectively. In the RAPGAP case, the total number of events was calculated by summing the contributions coming from the three subsamples after having renormalised them with the weights evaluated in Sect. 6.3.

The bin-by-bin method is a reliable way to estimate the C_f only if the MC describes the data distributions at the detector level. In fact, only in this case one can have a reasonable trust in the simulation of the detector implemented in the MC which is the base for an unbiased C_f determination. As it was presented in Sect. 6.4, the description of the MC was considered to be good for both RAPGAP and SATRAP . Since the detector simulation was exactly the same for both, any difference originated only by the different physical model used in the two MCs. None of the latter was

[2]Another variable with a physical meaning equivalent to C_f is the acceptance, α, defined as the inverse of C_f.

found to be significantly better. Because of this, it was decided to use for the cross section extraction the arithmetic mean of the two C_f. Fig. 7.1 shows the values of C_f for all the bins of all the variables considered for the differential cross sections. The C_f coming from the two MCs are quite similar. A typical value of $C_f \sim 1.25$ is observed with some trends as a function of variables like $E_{\rm T,jj}^*$ and $z_{I\!P}$.

In order to have a quantitative estimation of the detector effects, the stability and the purity are also presented. The *stability*, s, of any i-th bin is defined as

$$s = \frac{N_{\rm HAD\&\&DET}^{\rm MC}}{N_{\rm HAD}^{\rm MC}} \qquad (7.3)$$

where $N_{\rm HAD\&\&DET}^{\rm MC}$ is the number of events that are generated and reconstructed in the same i-th bin as estimated with the MC. The instability, defined as $1 - s$, quantifies the fraction of events for which the measurement was so biased to induce a migration of the value to another bin or even outside the kinematic phase space. A good measurement needs to have s as high as possible. The stability for RAPGAP and SATRAP is shown in Fig. 7.2.

The *purity*, p, of any i-th bin is defined as

$$p = \frac{N_{\rm HAD\&\&DET}^{\rm MC}}{N_{\rm DET}^{\rm MC}}. \qquad (7.4)$$

The purity quantifies the fraction of events detected in a bin that were actually generated in the same bin. The impurity, $1-p$, is again due to detector effects that alter the value of the variable and increment the number of entries collected in a bin with events coming from other bins or from outside the kinematic region. The p for RAPGAP and SATRAP is shown in Fig. 7.3. Notice that the correction factor is given by the ratio p/s.

The typical value for s is approximately 0.3 while for p is slightly below 0.4. The value of s is not particularly high and this can cause concern on the quality of the correction to the hadron level estimated with the MC. The reason for such a low value of s can be found in the migrations of events from one bin to another (which is an issue common to any analysis) and a low efficiency typical of the LRG method used for selecting the diffractive signal. The *efficiency* is defined in a way similar to the stability but dropping the requirement that the generated event has to be reconstructed in the same bin. Thus, the efficiency of any i-th bin is defined as

$$\varepsilon = \frac{N_{\rm HAD\&\&REC}^{\rm MC}}{N_{\rm HAD}^{\rm MC}} \qquad (7.5)$$

where $N_{\rm HAD\&\&REC}^{\rm MC}$ is the number of events that are generated in the i-th bin and selected at the detector level. The reconstructed value can be in any of the bins. Fig. 7.4 shows the efficiencies for the variables considered. The differences between Fig. 7.2 and 7.4 are due to the migrations of the reconstructed values of the variables from the bins where they were originally generated. The stabilities are typically 50% of the efficiencies. The low values of the efficiencies is related mainly to the rapidity gap selection which is applied only at the detector level. It can be noted in Fig. 7.4c
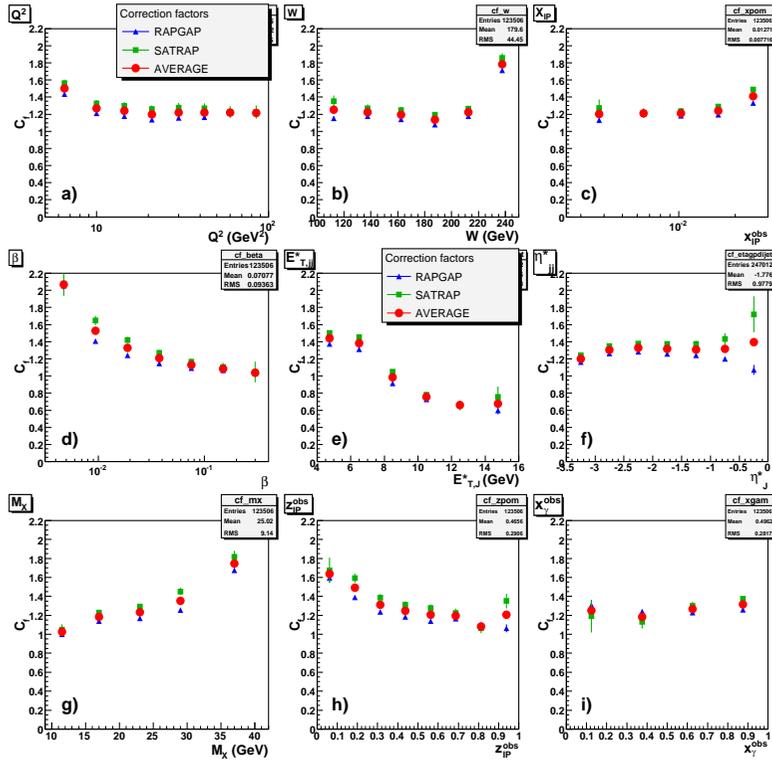
Figure 7.1: The correction factor, C_f, used to extract the cross sections at the hadron level. The values actually used in the analysis are represented by the red full circles, evaluated as the mean between the C_f predicted by RAPGAP and the SATRAP . The contributions from RAPGAP and SATRAP are shown separately as blue triangles and green squares, respectively. The error bars represent the statistical uncertainty on the estimation of C_f. The C_f is presented as a function of (a) Q^2, (b) W, (c) $x_{I\!\!P}^{\mathrm{obs}}$, (d) β, (e) $E_{\mathrm{T,jj}}^*$, (f) η_{jj}^*, (g) M_X, (h) $z_{I\!\!P}^{\mathrm{obs}}$ and (i) x_γ^{obs}. The variables are described more in detail in Sect. 7.1.

tha the efficiency drops at high $x_{I\!\!P}^{\mathrm{obs}}$. This is in fact the kinematic region most sensitive to the LRG selection, since $x_{I\!\!P}^{\mathrm{obs}}$ and the size of the rapidity gap are related. Thus, the values of the stability, which are low compared to many analyses, are the result of two effects which are individually under control.

7.1.2 QED radiative corrections

The emissions of a real photon from the incoming or outcoming electron are called *initial state radiation* (ISR) and *final state radiation* (FSR) respectively (see Fig.). These higher-order QED processes modify the four-momentum of the electron and alter as a consequence the reconstruction of the kinematics of the event. Since it is impossible experimentally to tag ISR and FSR events, the only possibility for taking them into account is to estimate the bias that they cause to the measurement with the MC and then correct the measured cross sections. The QED radiative corrections, C_{QED}, are the factors used to correct the cross sections back to the QED Born level. They are defined as

$$C_{\mathrm{QED}} = \frac{\left(\dfrac{d\sigma_{\mathrm{NOQED}}}{d\kappa}\right)}{\left(\dfrac{d\sigma_{\mathrm{QED}}}{d\kappa}\right)} \qquad (7.6)$$

where $\left(\dfrac{d\sigma_{\mathrm{NOQED}}}{d\kappa}\right)$ indicate the value in the i-th bin of differential cross section as a function of the generic variable κ as predicted by the MC without ISR and FSR. The same quantity but with the QED radiation processes allowed in the generation is indicated by $\left(\dfrac{d\sigma_{\mathrm{QED}}}{d\kappa}\right)$. The corrections for ISR and FSR were estimated with HERACLES through the SATRAP MC (see Sect. 5.1.1 and 5.1.2). The values of C_{QED} used in the analysis are shown in Fig. 7.5.

7.1.3 Systematic uncertainties

The uncertainties related to the experimental devices and techniques contribute to the systematic uncertainties. Several sources of systematic uncertainties have been checked. In general the systematic checks could be divided in two categories:

- *Experimental uncertainties.* Detector effects are unsmeared by means of the MC simulation. Nonetheless, a perfect description of the detector cannot be achieved and some characteristics and performances of the device are not well enough known and may be not well simulated. This causes an uncertainty on the result of the unfolding procedure. These uncertainties are taken in account by changing in the analysis some parameters that are believed to be the most critical and study the impact of the changes on the cross section measurement. The changes can be either on the real data or the MC side. The latter was chosen because of the higher statistics of the MC sample. The systematics evaluated in this way were:
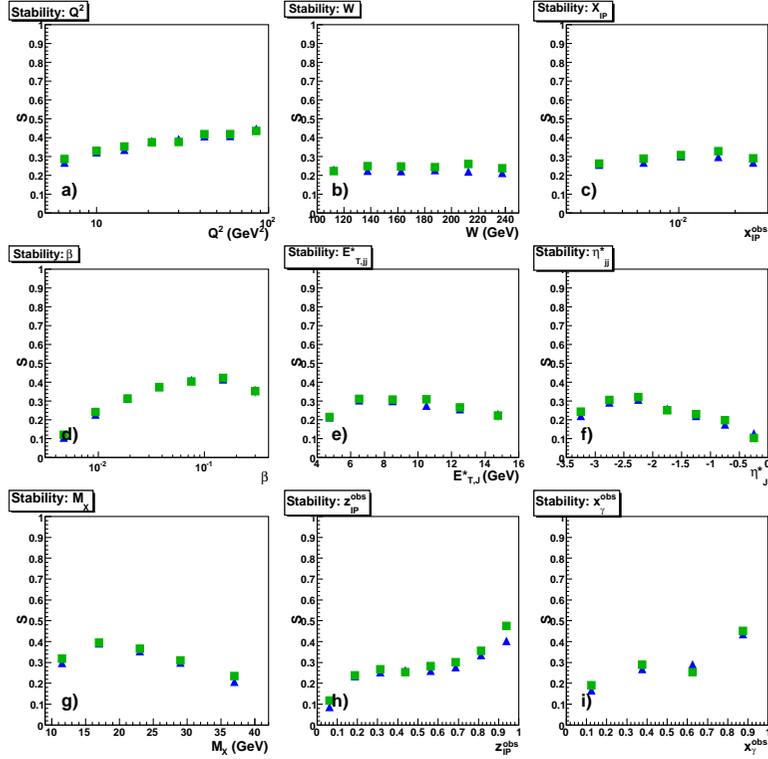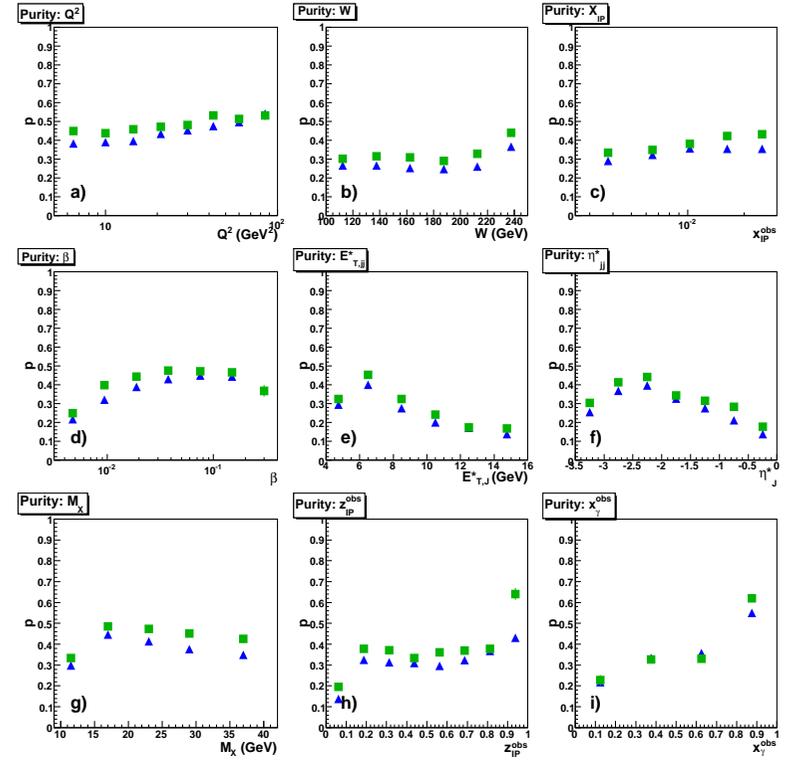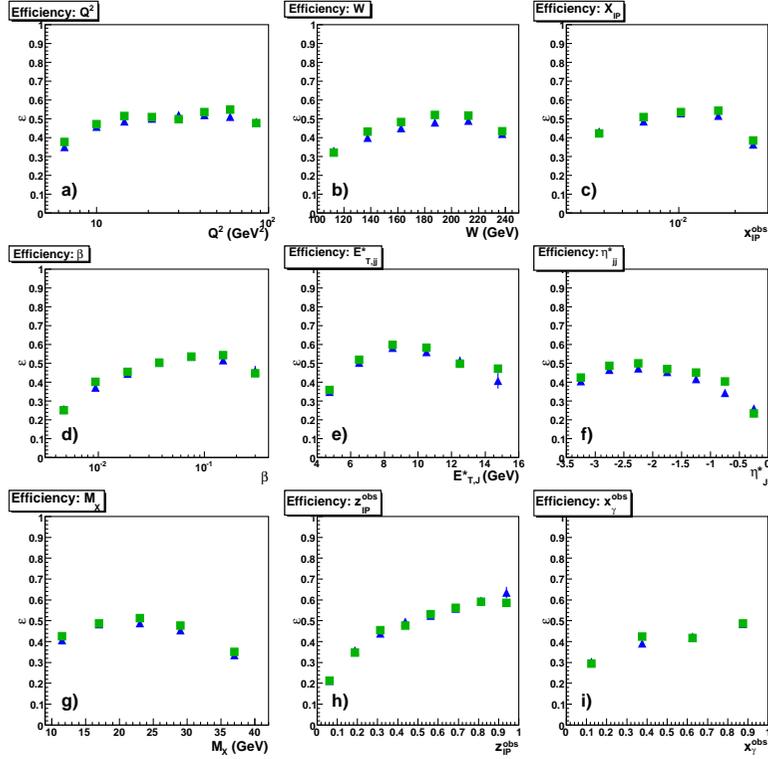
Figure 7.2: The stability, s, calculated with RAPGAP (blue triangles) and SATRAP (green squares). The error bars represent the statistical uncertainty on the estimation of s. The stability is presented as a function of (a) Q^2, (b) W, (c) $x_{I\!P}^{\mathrm{obs}}$, (d) β, (e) $E_{\mathrm{T,jj}}^*$, (f) η_{jj}^*, (g) M_{X}, (h) $z_{I\!P}^{\mathrm{obs}}$ and (i) $x_{\gamma}^{\mathrm{obs}}$. The variables are described more in detail in Sect. 7.1.

Figure 7.3: The purity, p, calculated with RAPGAP (blue triangles) and SATRAP (green squares). The error bars represent the statistical uncertainty on the estimation of p. The purity is presented as a function of (a) Q^2, (b) W, (c) $x_{I\!P}^{\mathrm{obs}}$, (d) β, (e) $E_{\mathrm{T,jj}}^*$, (f) η_{jj}^*, (g) M_{X}, (h) $z_{I\!P}^{\mathrm{obs}}$ and (i) $x_{\gamma}^{\mathrm{obs}}$. The variables are described more in detail in Sect. 7.1.

Figure 7.4: The efficiency, ε, calculated with RAPGAP (blue triangles) and SATRAP (green squares). The error bars represent the statistical uncertainty on the estimation of ε. The efficiency is presented as a function of (a) Q^2, (b) W, (c) $x_{\mathbb{P}}^{\mathrm{obs}}$, (d) β, (e) $E_{\mathrm{T,jj}}^*$, (f) η_{jj}^*, (g) M_{X}, (h) $z_{\mathbb{P}}^{\mathrm{obs}}$ and (i) $x_{\gamma}^{\mathrm{obs}}$. The variables are described more in detail in Sect. 7.1.
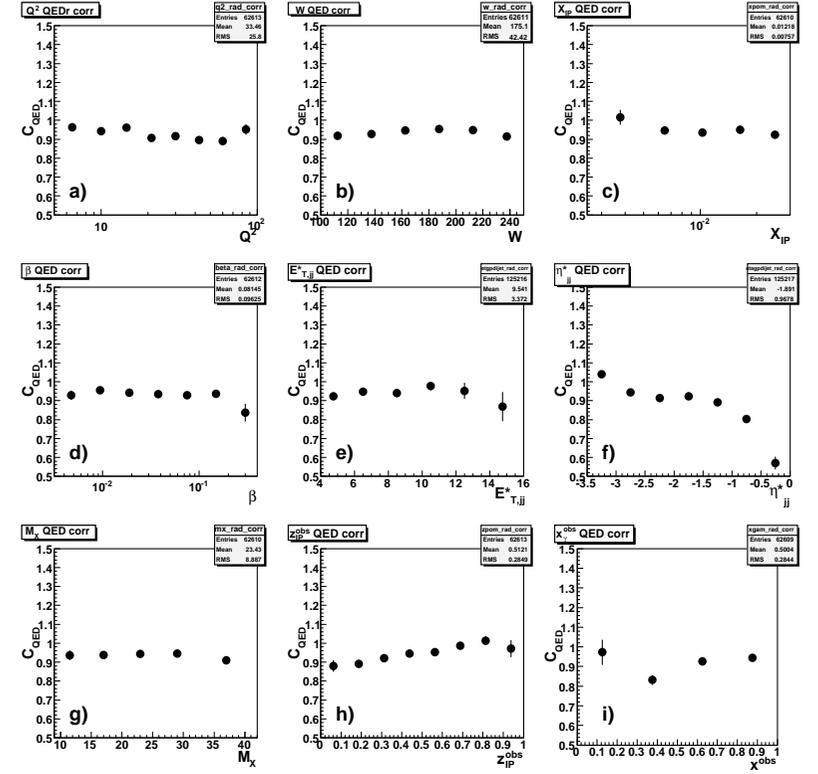
Figure 7.5: The QED radiative corrections, C_{QED}, used to correct the measured cross sections for higher order QED effects. The error bars represent the statistical uncertainty on the estimation of the corrections. The C_{QED} is presented as a function of (a) Q^2, (b) W, (c) $x_{\mathbb{P}}^{\mathrm{obs}}$, (d) β, (e) $E_{\mathrm{T,jj}}^*$, (f) η_{jj}^*, (g) M_{X}, (h) $z_{\mathbb{P}}^{\mathrm{obs}}$ and (i) $x_{\gamma}^{\mathrm{obs}}$. The variables are described more in detail in Sect. 7.1.

– the energy measured by the CAL was varied by $\pm 3\%$ in the MC to take into account the uncertainty on the CAL calibration, giving one of the largest uncertainties. Deviations from nominal cross section values were of the order of $\pm 5\%$, but reached $\sim 15\%$ in some bins;

– the energy scale of the scattered electron was varied in the MC by its uncertainty, $\pm 2\%$. The resulting variation of the cross sections was always below $\pm 3\%$;

– the position of the SRTD was changed in the MC by ± 2 mm in all directions to account for the uncertainty on its alignment. The change along the Z direction gave the largest effect and in a few bins caused a cross section variation of $\pm 2\%$;

– the model dependence of the acceptance corrections was estimated by using separately RAPGAP and SATRAP for unfolding the data. The variations from the central value (obtained using the average between RAPGAP and SATRAP were typically of the order of $\pm 5\%$ but reached $\sim \pm 10\%$ in some bins.

• *Measurement stability checks.* A bad description of the data distribution by the MC simulation could result in a biased estimation of the correction factors used for the unfolding. The migrations of events from and to the kinematical region may be wrongly reproduced. Thus, an estimation of this kind of effects was obtained by changing the cuts used in some of the selections described in Sect. 6.1. The cut variations applied were the following:

– the cut on the FPC energy was varied by ± 100 MeV in the MC;

– the cut on the scattered-positron energy was lowered from 10 to 8 GeV;

– the fiducial region for the electron selection was enlarged and reduced by 0.5 cm;

– the lower cut on $\delta = (E - p_Z)_{\rm TOT}$ was changed from 45 to 43 GeV.

The single contributions to the uncertainty coming from each systematic source can be seen in Appendix D.

The variations on the cross section induced by these cuts, with except of the ones coming from the calorimeter and electron energy, were summed in quadrature together in order to give the total systematic uncertainty. The uncertainties related to the calorimeter and electron energy scales caused a variation on the cross section that correlated many bins of the measurement. Thus, they were treated separately as correlated systematic uncertainties. They were summed in quadrature together with the uncertainty on the amount of proton dissociation background subtracted ($\pm 4\%$ for all the bins, see Sect. 6.5). Another source of correlated uncertainty was the error on the luminosity measurement ($\pm 2.25\%$) but this was not included neither in the plots nor in the tables.

As described in Sect. 6.1.4, in order to select a clean diffractive sample it was not necessary to apply an explicit large rapidity gap selection by means of a cut on the
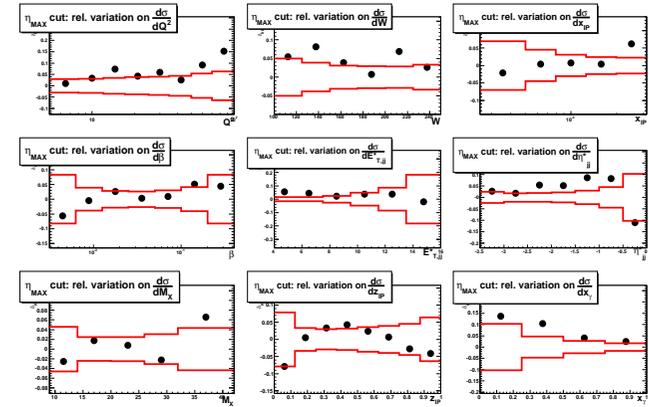
Figure 7.6: Relative variations on the single differential cross section induced by the introduction in the signal selection of a cut $\eta_{\rm maX} < 2.8$, where $\eta_{\rm MAX}$ is the pseudorapidity of the most forward EFO with an energy $E_{EFO} > 400$ MeV.

pseudorapidity of the most forward EFO ($\eta_{\rm MAX}$ cut). In order to present the results in a way compatible with previous measurements, the analysis was repeated with a cut on $\eta_{\rm MAX} < 2.8$. Only EFOs with an energy higher than 400 MeV were used for the $\eta_{\rm MAX}$ calculation. The latter requirement rejected fake EFOs originating from noise in the CAL or particles not coming from the primary vertex. With this additional cut, the number of selected events is reduced to 4012. The variation on the differential cross section is shown in Fig. 7.6. Notice that the changes of the differential cross section caused by the $\eta_{\rm MAX}$ cut can not be ascribed as systematic uncertainties since they are not coming from a not perfect extrapolation of the MC to the most forward $\eta_{\rm MAX}$ region where the MC does not describe the data (see Fig. 6.4). Using this wrong model for the extrapolation results in a wrong extraction of the cross section in the amount shown in Fig. 7.6. Conversely, the measurement described in Sect. 6 has no extrapolation and therefore it is safe from this kind of uncertainty.

7.2   Total cross section

The total cross section for the production of dijets in diffractive DIS in the kinematic region specified in Table 7.2 was measured to be

$$\sigma_{\rm TOT}^{\gamma^* p \to jjX'} = 91.8 \pm 1.2({\rm stat})\ ^{+3.3}_{-5.4}({\rm syst.})\ ^{+5.2}_{-3.8}({\rm corr.})\ {\rm pb}. \qquad (7.7)$$

In Table 7.3, the total measured cross section is compared to the NLO predictions using as input the different dPDFs described in Sect. 5.2 . All the results of the

calculation are compatible with data within the experimental and theoretical errors. However, the calculation using the MRW 2006 or the H1 2006 - Fit Bdescribe better the absolute normalisation of the data.

7.3   Comparison to Monte Carlo models

The single differential cross sections as a function of the variables listed in Sect. 7.1 predicted with the RAPGAP and SATRAP LO MC are compared to the measured values in Fig. 7.7 and 7.8. Since the MC predictions are not expected to describe the normalisation, the cross sections predicted by both MCs were normalised to the data area. The total correlated uncertainty is shown as a shaded band in the figures.

The $E^*_{T,jj}$ distribution is a steeply falling function as expected in pQCD (Fig. 7.8a) and the jets tend to populate the backward region (Fig. 7.8d). The most prominent features of the data are the rise of the cross section with $x^{obs}_{\mathbb{P}}$, the peak at $z^{obs}_{\mathbb{P}} \sim 0.3$ and the tail of the cross section at low x^{obs}_{γ} values. The requirement of two jets with high E_T suppresses the contribution of low values of x^{obs}_{γ}. The relatively low value of the peak position in the $z^{obs}_{\mathbb{P}}$ distribution indicates that in the majority of the events the dijet system is accompanied by additional hadronic activity. Most of the events are produced at large x^{obs}_{γ} as expected in DIS but a tail at low x^{obs}_{γ} indicates the presence of a small but not negligible resolved-photon contribution.

In general a good agreement between data and LO MC is observed. The two MC provide a very similar prediction. It should be stressed that this was expected theoretically. In fact, the dijet measurement selects very small dipoles and RAPGAP uses the k_T factorization scheme. It was demonstrated [136] that for small sizes of the dipole (i.e. high transverse energies of the jets) the two models are equivalent. The good agreement between the RAPGAP and SATRAP results confirms that because the input distribution to RAPGAP (dPDFs) is very different than the input distribution to SATRAP (gluon density from inclusive DIS data). Therefore the good agreement of the two approaches indicates the consistency of the QCD description at the leading order level. The main differences between the two MCs are a better description of the data by RAPGAP at high $E^*_{T,jj}$ and by SATRAP at high $z^{obs}_{\mathbb{P}}$ and the x^{obs}_{γ} differential cross section, where the inclusion of resolved-photon processes in RAPGAP improves the description of the data (Fig. 7.8d). The contribution of the resolved-photon processes to the total RAPGAP cross section was estimated to be 16%. SATRAP has no resolved-photon in it (that in the saturation model would be implemeted with processes of order higher than the $q\bar{q}$ and $q\bar{q}g$) and this MC doesn't describe the data cross section in the lowest x^{obs}_{γ} region.

7.4   Comparison to NLO QCD prediction

The comparison with NLO predictions is a fundamental step for this analysis. Oppositely to the LO case, the NLO calculation is expected to predict the absolute normalisation of the cross section, within the approximation of neglecting the higher orders of
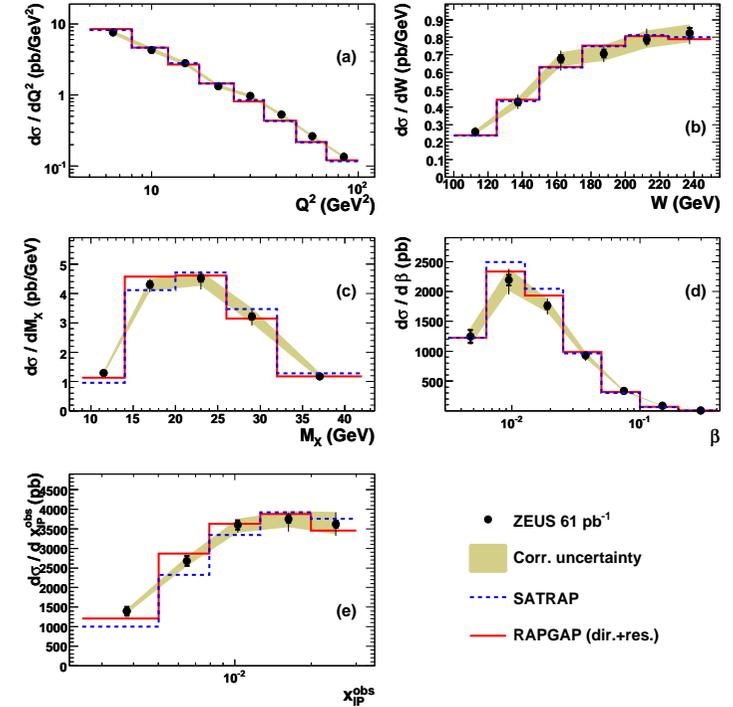

Figure 7.7: Measured differential cross section (dots) as a function of (a) Q^2, (b) W, (c) M_X, (d) β and (e) $x^{obs}_{\mathbb{P}}$. The inner error bars represent the statistical uncertainty and the outer error bars represent the statistical and uncorrelated systematic uncertainties added in quadrature. The shaded band represents the correlated uncertainty. For comparison the area-normalised predictions of the RAPGAP (solid lines) and the SATRAP (dashed lines) MC models are also shown.
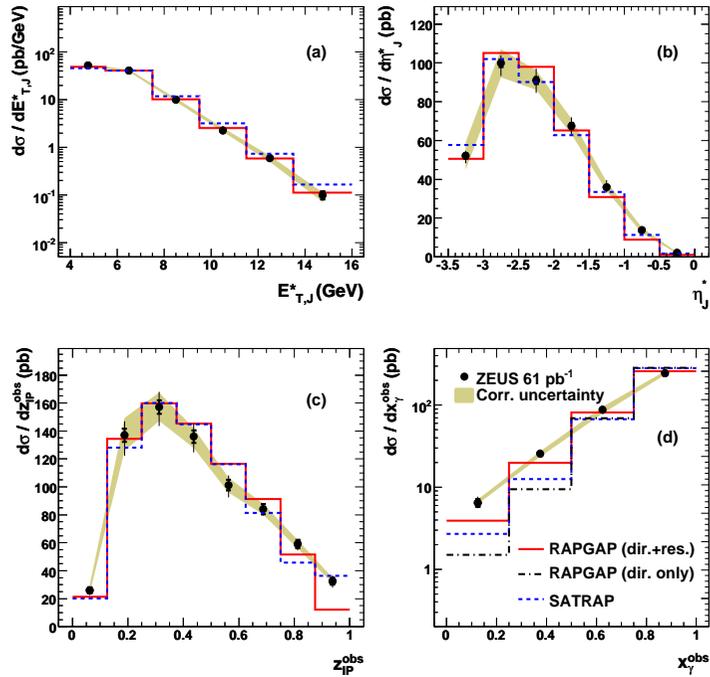
Figure 7.8: Measured differential cross section as a function of (a) $E^*_{T,jj}$, (b) η^*_{jj}, (c) $z^{obs}_{I\!P}$ and (d) x^{obs}_γ. The dashed-dotted line represents the area-normalised RAPGAP with only the direct photon contribution. Other details as in the caption of Fig. 7.7.

the perturbative serie. Therefore the comparison between NLO and data is the benchmark for testing the QCD factorisation theorem in diffraction with dijets in the DIS kinematical regime. On the other hand, if one assumes the validity of the factorisation theorem, the same comparison can be used for testing the accuracy of the different dPDFs used in the calculation.

The NLO predictions for the differential cross section are compared to the data in Figs. 7.9 and 7.10. The estimated theoretical uncertainties are shown only for the calculations using the ZEUS LPS+charm dPDFs and are similar for all the other calculations. For ease of comparison the ratios of data to the MRW 2006 prediction are presented in Figs. 7.11 and 7.12. The variation due to the choice of the dPDFs is displayed with respect to the MRW 2006 in the same figure. In general the shape of the measured cross section is described by the NLO calculations within the theoretical uncertainties. However, only the predictions using the MRW 2006 dPDFs and the H1 2006 − FitB are able to describe satisfactorily the data over the entire kinematic range. The differential cross section as a function of $z^{obs}_{I\!P}$ presented in Fig. 7.10d supports this statement very clearly. This quantity is the most sensitive to the choice of the dPDFs used in the calculation. The central values of the predictions using the MRW 2006 and the H1 2006 − FitB dPDFs describe very well the data over the whole range in $z^{obs}_{I\!P}$. Conversely, the calculations using the ZEUS LPS+charm and H1 2006 − FitA dPDFs exhibit a different trend and are incompatible with the data at high $z^{obs}_{I\!P}$, even considering the large theoretical uncertainties. The description of the x^{obs}_γ dependence is not reproduced by all the predictions, independently of the dPDFs. This is related to the NLO calculation rather than the dPDFs used since in DISENT there is not any resolved-photon contribution. The results presented in Sect. 7.3 showed instead that a contribution from resolved-photon processes is needed to obtain a good description at low x^{obs}_γ. Anyway the contribution to the total cross section of this particular phase space region is very small and the other variables are practically insensitive to this aspect of the calculation.

Any final statement about the factorisation theorem is limited by the large size of this theoretical uncertainty ($\sim 25\%$). The good agreement of the central values of some of the NLO curves presented supports the validity of the theorem, as suggested already in previous analyses [79, 80, 81]. There are no evidences of factorisation breaking in the production of dijets in diffractive DIS and that the factorisation theorem holds within an uncertainty of approximately 25%. This large theoretical uncertainty emphasizes the precision of the experimental measurement that can be very useful for future QCD fits as explained in the next section.

The calculation using the GLP fit dPDFs as input gave predictions clearly off in both normalisation and shape. This discrepancy with the measured data can be observed in Fig. 7.13 where a comparison as a function of $E^*_{T,jj}$ and $z^{obs}_{I\!P}$ is presented. The prediction underestimates the absolute normalisation of the cross section by a factor $0.3 - 0.4$ over the whole kinematic range. Because of this bad description of the data, the calculation using the GLP fit dPDFs was discarded and not studied in depth.

The sensitivity to the choice of the renormalisation scale is studied more in detail in Fig. 7.14. The differential cross section as a function of Q^2, $E^*_{T,jj}$ and $z^{obs}_{I\!P}$ is presented for two different choices of the renormalisation scale, μ_R. In one case the value of μ_R was

the default used through the whole analysis, i.e. the transverse energy in the $\gamma^* p$ of the leading jet, $\mu = E^*_{T,j1}$. The use of $E^*_{T,j1}$ was motivated by the choice to use a physical quantity as renormalisation scale. The default choice is compared to an expression taking into account also the virtuality of the exchanged photon, $\mu_R = \sqrt{E^{*\,2}_{T,j1} + Q^2}$. As expected, a better agreement with the data at higher Q^2 is observed, while at transverse energies of the jet already higher than 6 GeV there are no differences between the two choices for μ_R.
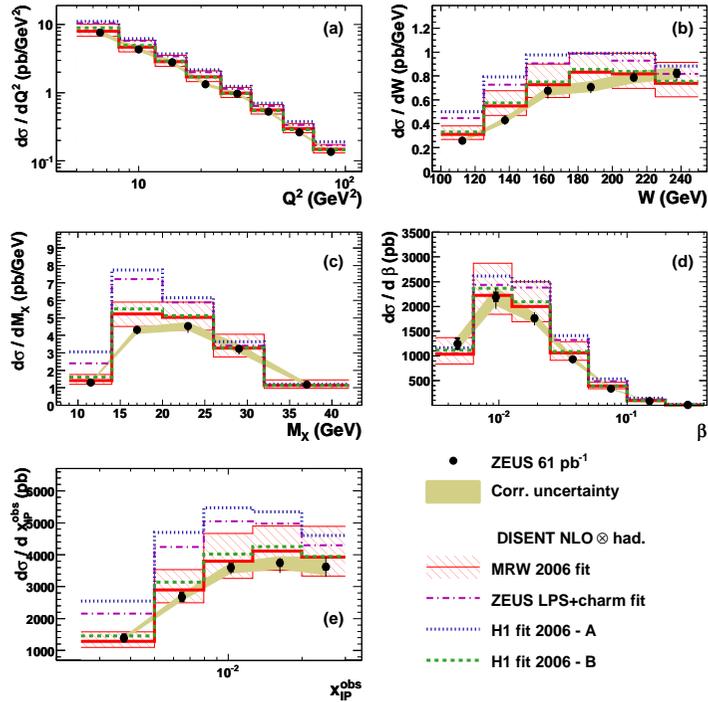
Figure 7.9: Measured differential cross section as a function of (a) Q^2, (b) W, (c) M_X, (d) β and (e) $x^{obs}_{I\!P}$ compared to the NLO predictions obtained using the available dPDFs, as indicated in the figure. The hatched area indicates the theoretical uncertainty of the predictions estimated using the MRW 2006 fit dPDFs. Other details as in the caption of Fig. 7.7.
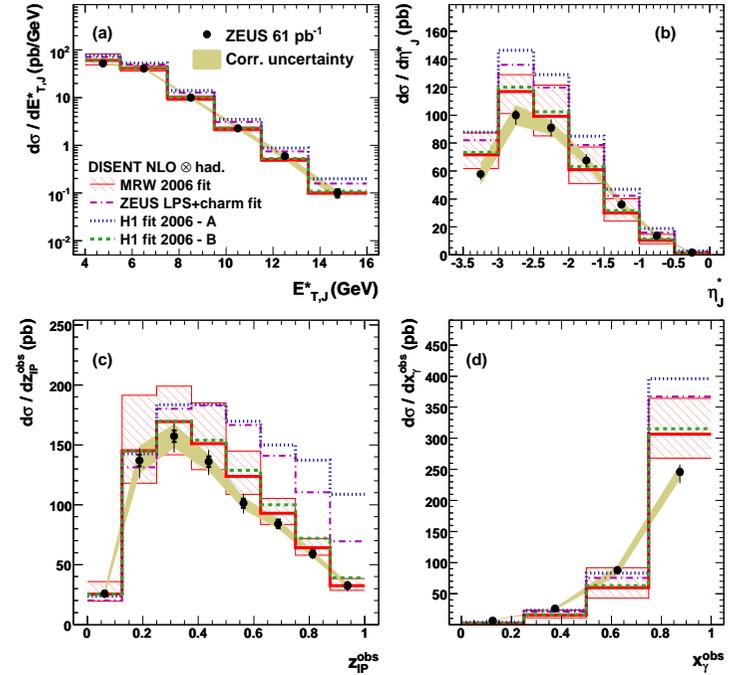

Figure 7.10: Measured differential cross section as a function of (a) $E^*_{T,jj}$, (b) η^*_{jj}, (c) $z^{obs}_{I\!P}$ and (d) x^{obs}_γ compared to the NLO prediction obtained using the available dPDFs. Other details as in the caption of Fig. 7.9.

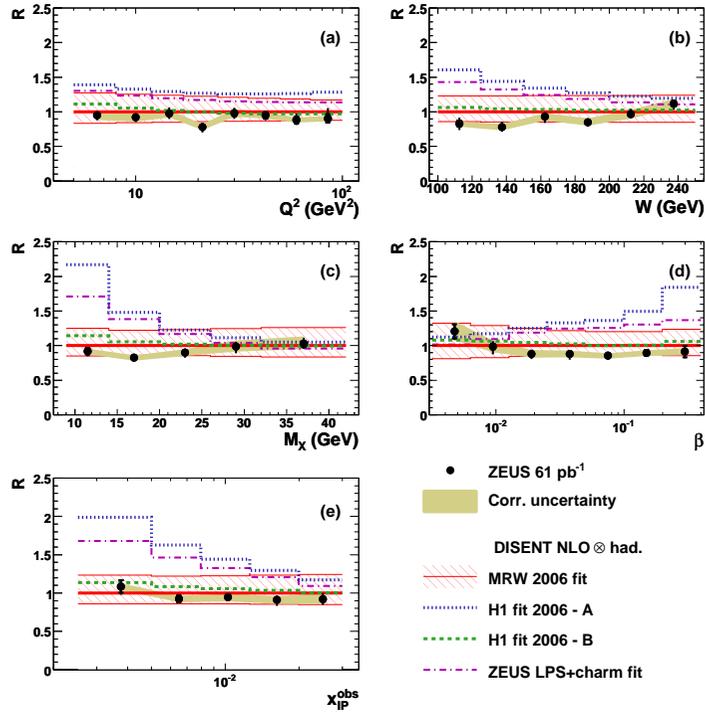*Figure 7.11: Ratio, R, of the data to the NLO prediction using the MRW 2006 dPDFs (dots) as function of (a) Q^2, (b) W, (c) M_X, (d) β and (e) $x_{I\!P}^{\mathrm{obs}}$. Also shown is the ratio of NLO calculations with other dPDFs to MRW 2006. Other details as in the caption of Fig. 7.9.*
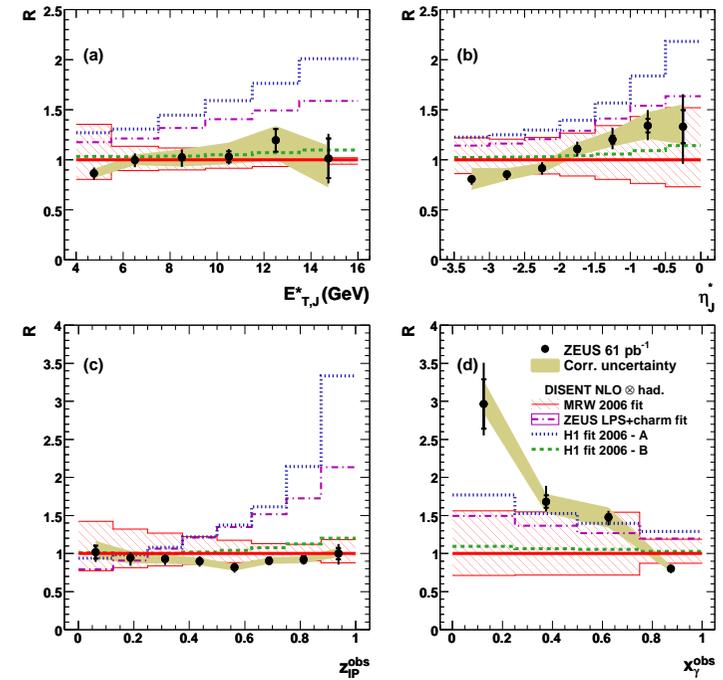

Figure 7.12: Ratio, R, of the data to the NLO prediction using the MRW 2006 dPDFs (dots) as function of (a) $E_{\mathrm{T,jj}}^*$, (b) η_{jj}^*, (c) $z_{I\!P}^{\mathrm{obs}}$ and (d) x_γ^{obs}. Other details as in the caption of Fig. 7.11.
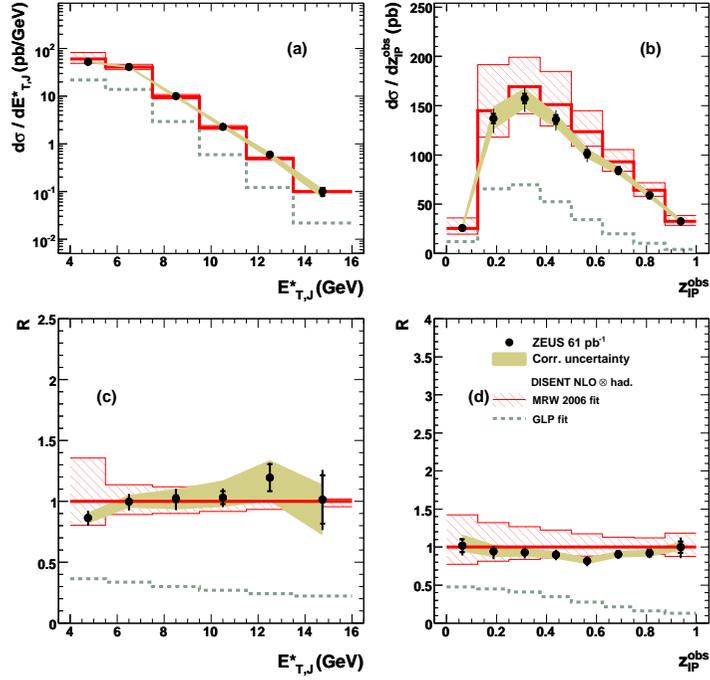
Figure 7.13: Comparison to the NLO prediction using the GLP fit dPDFs. The NLO prediction using the GLP fit dPDFs as input (dashed line) is compared to the measured cross section (dots) and the NLO prediction using the MRW 2006 dPDFs (solid line). The comparison is presented as a function of (a) $E^*_{T,jj}$ and (b) $z^{obs}_{I\!P}$. The ratio, R, of the data to the NLO prediction using the MRW 2006 dPDFs (dots) is shown as function of (c) $E^*_{T,jj}$ and (d) $z^{obs}_{I\!P}$. The ratio of the prediction using the GLP fit dPDFs to the prediction using the MRW 2006 dPDFs is presented as a dashed line. Other details as in the caption of Fig. 7.9.
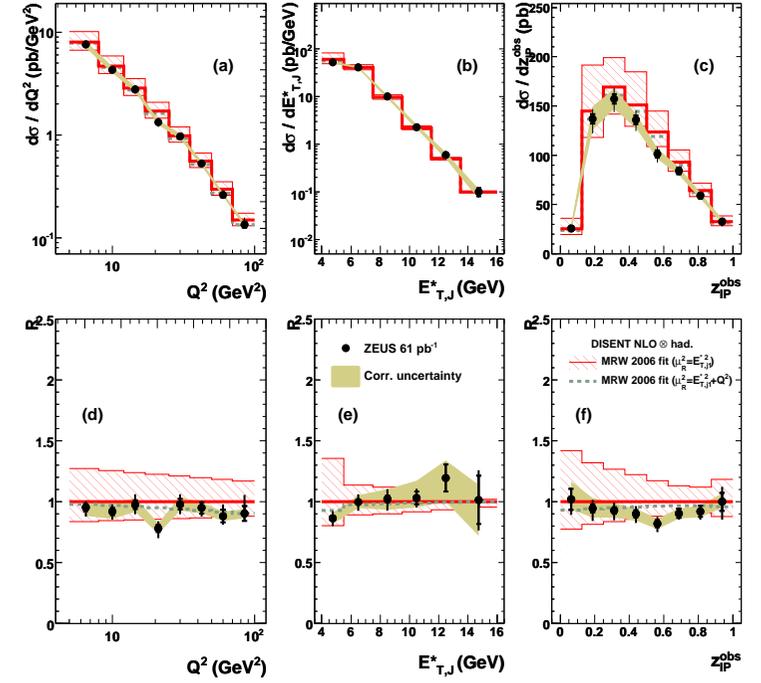

Figure 7.14: Comparison to the NLO prediction choosing a different renormalisation scale, μ_R. The NLO prediction calculated using $\mu^2_R = E^{*\ 2}_{T,j1} + Q^2$ (dashed line) is compared to the measured cross section (dots) and the NLO prediction using the default $\mu^2_R = E^{*\ 2}_{T,j1}$ (solid line). The comparison is presented as a function of (a) Q^2, (b) $E^*_{T,jj}$ and (c) $z^{obs}_{I\!P}$. The ratio, R, of the data to the NLO prediction using $\mu^2_R = E^{*\ 2}_{T,j1}$ (dots) is shown as function of (d) Q^2, (e) $E^*_{T,jj}$ and (f) $z^{obs}_{I\!P}$. The ratio of the prediction using $\mu^2_R = E^{*\ 2}_{T,j1} + Q^2$ to the prediction using $\mu^2_R = E^{*\ 2}_{T,j1}$ is presented as a dashed line. The NLO calculation was performed using always the MRW 2006 dPDFs as input. Other details as in the caption of Fig. 7.9.

7.5 Double differential cross section

Precious informations about the dynamics of the production of dijets in diffractive DIS can be extracted by measuring the double differential cross section. The double differential cross section is presented as a function of $z_{\mathbb{P}}^{\text{obs}}$ in different regions of Q^2 and $E_{\text{T,j1}}^*$. Such a detailed study is made possible only thanks to the high statistics of the data sample. The variables for the double differential cross sectins were chosen because of their particular relevance in the pQCD calculation: $z_{\mathbb{P}}^{\text{obs}}$ is the variable directly sensitive to the dPDFs while Q^2 and $E_{\text{T,j1}}^*$ are the two hard scales present in the process [3]. The control distributions are presented in Figs. 7.15 and 7.16. As pointed out describing the single differential distribution as a function of $z_{\mathbb{P}}^{\text{obs}}$ in Sect. 6.4, the data are better described by the SATRAP MC. Some differences between data and MC are visible, in particular at high $z_{\mathbb{P}}^{\text{obs}}$. Due to the sharp fall of the cross section at low $z_{\mathbb{P}}^{\text{obs}}$ (see Fig. 7.8c), the first bins in $z_{\mathbb{P}}^{\text{obs}}$ of the single differential measurement would have a low statistics. It was decided to merge it with the second bin. Thus, the double differential cross section has one bin less than the single differential one. The factors applied to correct the data for the detector acceptance and resolution are shown in Fig. 7.17. Values similar to the single differential C_f were observed (see Fig. 7.1h). The strong trend as a function of $E_{\text{T,j1}}^*$ is consistent with the one observed in Fig. 7.1e. The stabilities and purities for the double differentila cross sections are presented in Figs. 7.18 and 7.19. The values are very low ($\lesssim 15\%$) over the entire range of the measurement. However, it has to be kept in mind the low efficiency ($\sim 40-50\%$) of the measuremnt, as discussed in Sect. 7.1.1 and the satisfactory good description of the data distributions by the MC. Thus, the correction factors were still judged to be reliable for correcting the measuremnt to the hadron level. Furthermore, any possible bias in the correction factors is taken into account in the systematic uncertainty, where the correction factor is changed by evaluating it with either RAPGAP or SATRAP rather than the average of the two.

The measured double differential cross section as a function of $z_{\mathbb{P}}^{\text{obs}}$ in different regions of Q^2 and $E_{\text{T,j1}}^*$ is shown in Figs.7.20 and 7.21, respectively. The NLO predictions are compared to the data in the same figures. As it was done for the single differential cross section, the ratios of data to the MRW 2006 prediction are presented in Figs. 7.22 and 7.23 together with the ratios of the NLO predictions using the other dPDFs to the MRW 2006 NLO calculation. The same considerations made in Sect. 7.4 are valid: the NLO calculation using the MRW 2006 and the H1 $2006 -$ FitB dPDFs describes the data over the entire kinematical range much better than the NLO curves using the other dPDFs. It can be noted that the level of the NLO description at fixed $z_{\mathbb{P}}^{\text{obs}}$ stays approximately constant as a function of the hard scale used in the calculation. This is a good evidence that the QCD description adopted in the analysis works. The QCD theory is able to predict the evolution of the dPDFs as a function of the scale once that the $z_{\mathbb{P}}^{\text{obs}}$ dependence is given at an initial scale. The fact that the shape of the cross section as a function of $z_{\mathbb{P}}^{\text{obs}}$ is described by the NLO calculation at different

---

[3]The measured points were presented as a function of either Q^2 or $E_{\text{T,j1}}^*$ and not combinations of them, e.g. $Q^2 + (E_{\text{T,j1}}^*)^2$, because it was preferred to present the cross sections in terms of physical quantities.

scales is a remarkable evidence that the DGLAP evolution works in the context under examination.

Under the assumption of the validity of the QCD factorisation theorem, these data could be included in future QCD fits to the dPDFs themselves, together with the inclusive DIS data. This technique, already exploited in previous analyses both of the proton [83] and diffractive [81] PDF, has been proved to reduce significantly the uncertainty on the gluon parton densities. This can be intuitively understood from the data presented here: the major difference between the H1 $2006 -$ FitB and Fit A is in the gluon dPDF and these data have a clear discriminating power between the two fits. Therefore these data have a significant potential to further constrain the gluon dPDF.
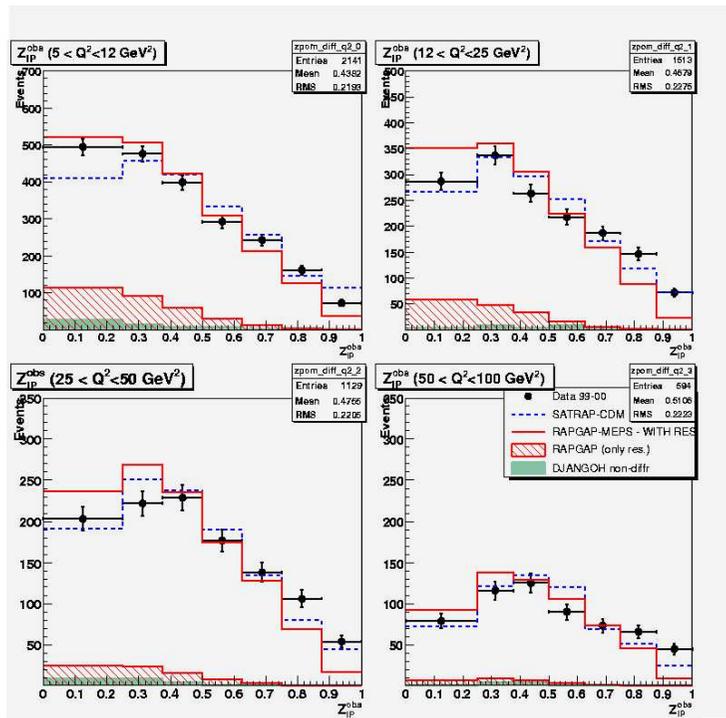
Figure 7.15: The control distributions as a function of $z_{I\!P}^{\rm obs}$ in different regions of Q^2. The data entries are shown as dots, the statistical errors are shown as the error bars. The data are compared to the signal LO MC, RAPGAP (solid red line) and SATRAP (dashed blue line). The contribution from resolved-photon processes estimated with RAPGAP is shown as a hatched area. The non-diffractive DIS background estimated with DJANGOH is indicated by the dark solid area.
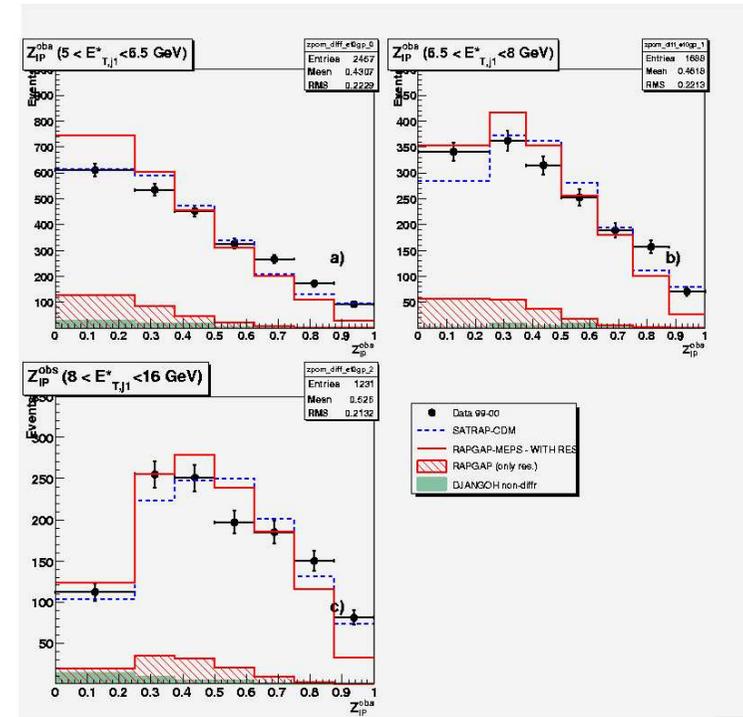

Figure 7.16: The control distributions as a function of $z_{I\!P}^{\rm obs}$ in different regions of $E_{T,j1}^*$ (dots). Other details as in the caption of Fig. 7.15.
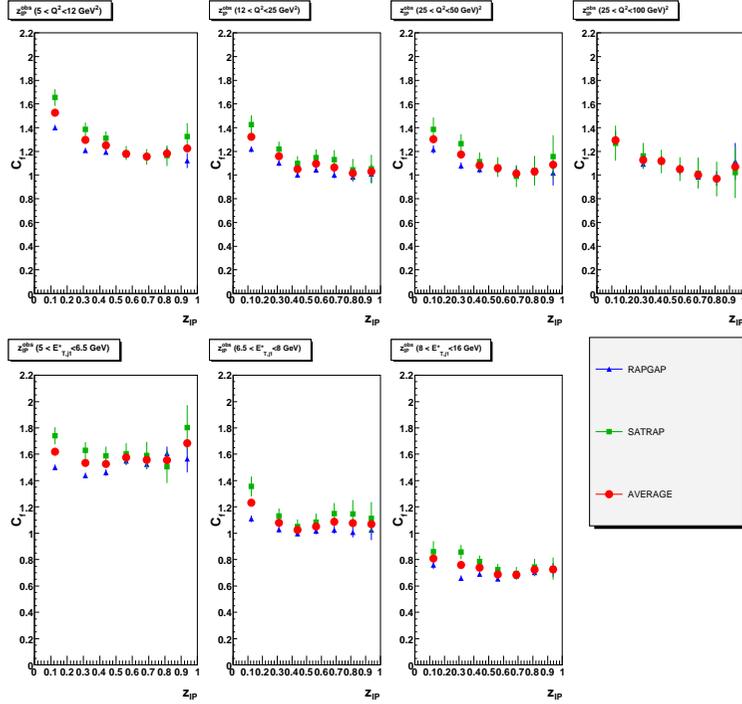
Figure 7.17: The correction factor, C_f, used to extract the double differential cross sections at the hadron level. The values actually used in the analysis are represented by the red full circles, evaluated as the mean between the C_f predicted by RAPGAP and the SATRAP . The contributions from RAPGAP and SATRAP are shown separately as blue triangles and green squares, respectively. The error bars represent the statistical uncertainty on the estimation of C_f. The upper (lower) row presents the C_f as a function of $z_{I\!P}^{\mathrm{obs}}$ in different bins of Q^2 $(E_{\mathrm{T},\mathrm{j}1}^{*})$.

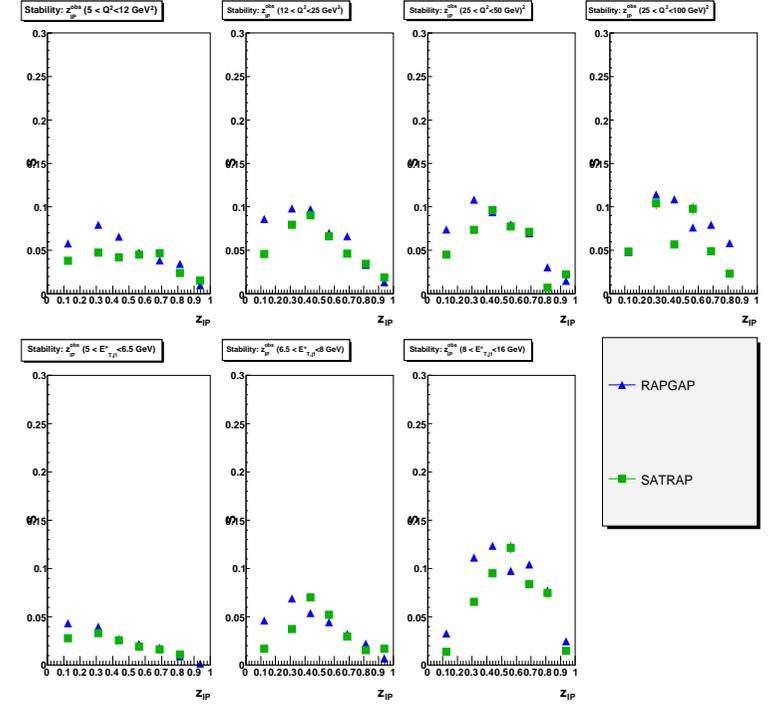
Figure 7.18: The stability, s, of the bins of the double differential cross section estimated with RAPGAP (blue triangles) and SATRAP (green squares). The error bars represent the statistical uncertainty on the estimation of s. The upper (lower) row presents s as a function of $z_{I\!P}^{\mathrm{obs}}$ in different bins of Q^2 $(E_{\mathrm{T},\mathrm{j}1}^{*})$.
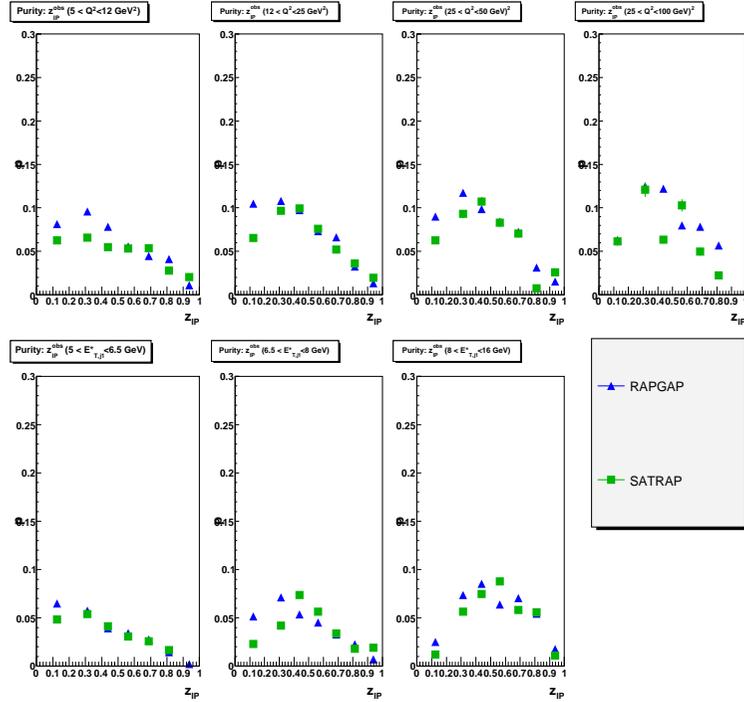
Figure 7.19: The purity, p, of the bins of the double differential cross section estimated with RAPGAP *(blue triangles) and* SATRAP *(green squares). The error bars represent the statistical uncertainty on the estimation of p. The upper (lower) row presents p as a function of $z_{I\!P}^{\rm obs}$ in different bins of Q^2 ($E_{\rm T,j1}^*$).*
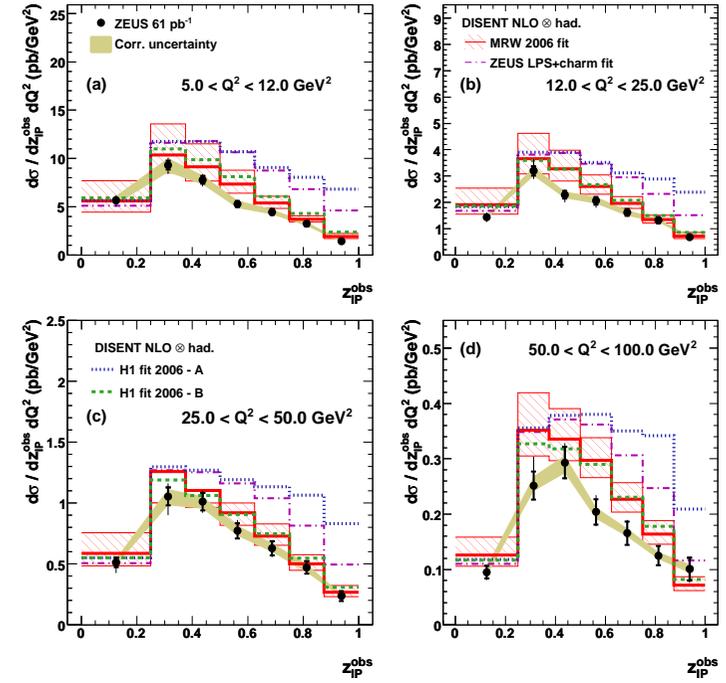

Figure 7.20: Measured differential cross section as a function of $z_{I\!P}^{\rm obs}$ in different regions of Q^2 (dots). Other details as in the caption of Fig. 7.9.
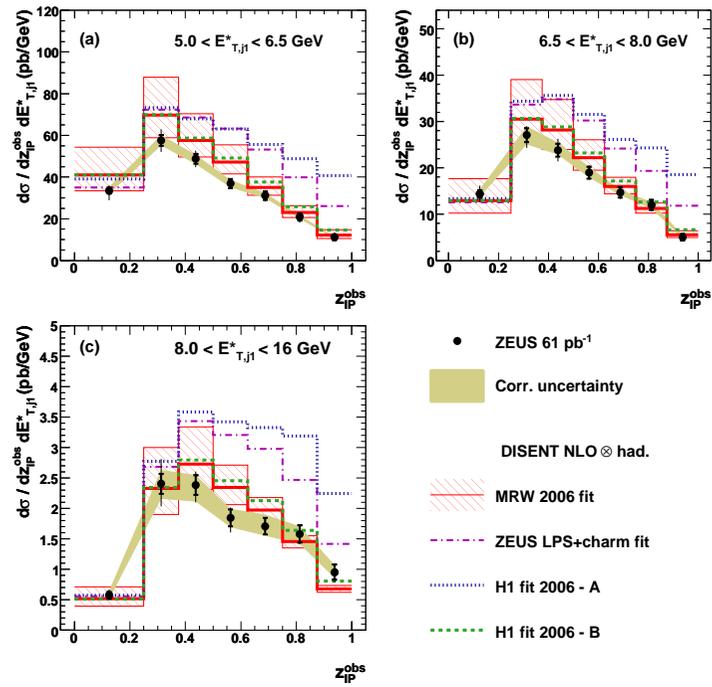
Figure 7.21: Measured differential cross section as a function of $z_{I\!P}^{\mathrm{obs}}$ in different regions of $E_{T,j1}^{*}$ (dots). Other details as in the caption of Fig. 7.9.

Figure 7.22: Ratio, R, of the data to the NLO prediction using the ZEUS LPS+charm dPDFs (dots) as function of $z_{I\!P}^{\mathrm{obs}}$ in different regions of Q^2. Other details as in the caption of Fig. 7.11.
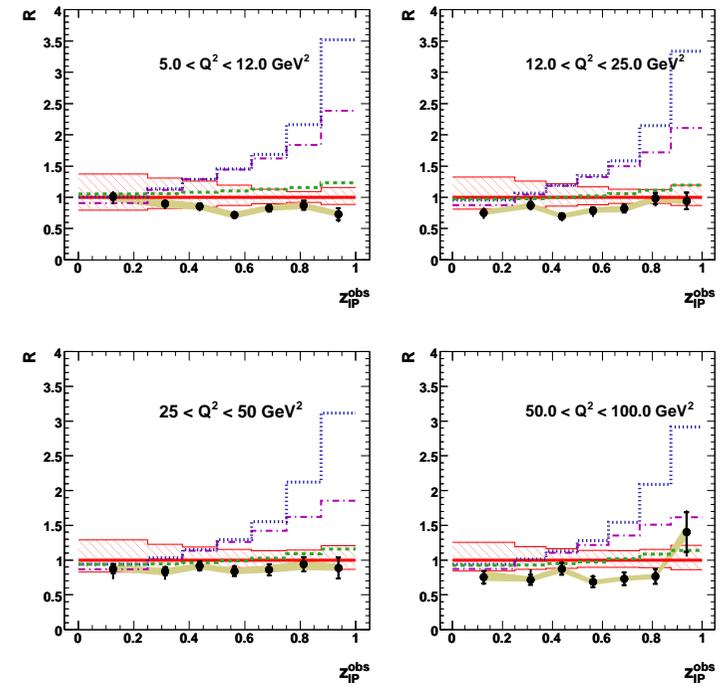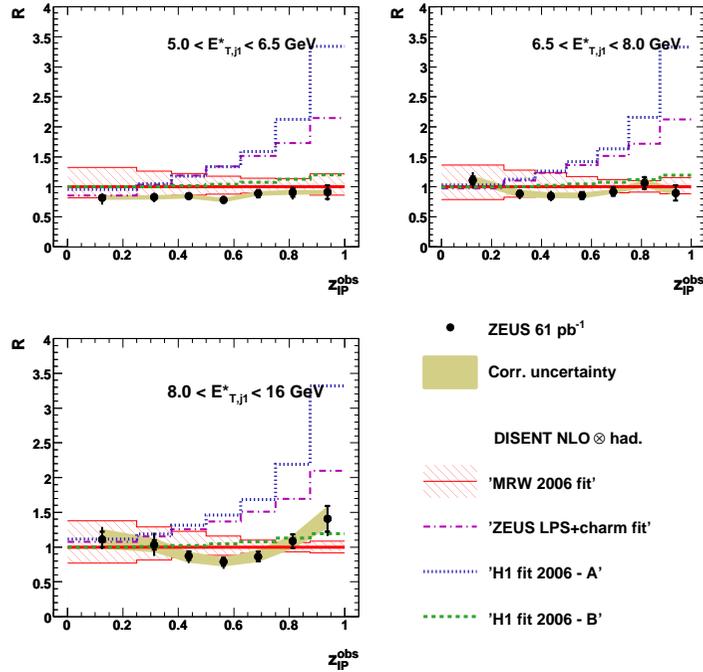
Figure 7.23: *Ratio, R, of the data to the NLO prediction using the ZEUS LPS+charm dPDFs (dots) as function of $z_{I\!P}^{obs}$ in different regions of $E_{T,j1}^*$. Other details as in the caption of Fig. 7.11.*

Kinematic region
$5 < Q^2 < 100\,\text{GeV}^2$, $100 < W < 250\,\text{GeV}$
$E_{\text{T,j1}}^* > 5\,\text{GeV}$, $-3.5 < \eta_{\text{j1}}^* < 0.0$
$E_{\text{T,j2}}^* > 4\,\text{GeV}$, $-3.5 < \eta_{\text{j2}}^* < 0.0$
$x_{I\!P} < 0.03$

Table 7.2: *The kinematic region of definition of the cross section measured in this analysis. The cross section was defined for at least two jets reconstructed in the γ^*p rest frame. The leading jet had to have a transverse energy in the γ^*p rest frame higher the $5\,GeV$ while the jet with the second highest E_T^* had to have a transverse energy higher than $4\,GeV$. All the jets had to lie in the pseudorapidity range $-3.5 < \eta_{jet}^* < 0$ as measured in the γ^*p rest frame. The virtuality of the photon, Q^2, and the total γ^*p centre-of-mass energy, W, had to be $5 < Q^2 < 100\,GeV^2$ and $100 < W < 250\,GeV$, respectively. The value of $x_{I\!P}$ had to be lower than 0.03.*

	σ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	δ_{theor} (pb)	Δ_{DIFFR} (pb)
Data	91.5	1.2	$+3.3$ -5.4	$+5.2$ -3.8	$-$	$+4.1$
ZEUS LPS+charm	120.3	$-$	$-$	$-$	$+29.4$ -18.3	$-$
H1 2006 - Fit A	130.2	$-$	$-$	$-$	$+31.2$ -19.9	$-$
H1 2006 - Fit B	102.5	$-$	$-$	$-$	$+24.7$ -15.6	$-$
MRW 2006	99.3	$-$	$-$	$-$	$+23.4$ -14.7	$-$

Table 7.3: *Total cross section for the production of diffractive dijets compared to expectations of NLO calculations using various dPDFs as indicated in the Table. The kinematic range of the measurement is specified in Table 7.2. The statistical, δ_{stat}, uncorrelated systematic, δ_{syst}, and energy scale uncertainties, δ_{ES}, are quoted separately. The theoretical uncertainty on the NLO calculations, δ_{theor}, is quoted in the sixth column. The difference with the measured cross section with and without η_{MAX} cut, Δ_{DIFFR}, is presented in the last column. The uncertainties on the proton dissociation subtraction and the luminosity measurement are not presented in the table.*

Chapter 8

Conclusions and outlook

This thesis presented an analysis of the production of dijets in diffractive Deep Inelastic Scattering (DDIS). The measurements were performed with the data collected by the ZEUS detector at the HERA collider during the data-taking period 1999-2000 for a total integrated luminosity of $61.3\,\mathrm{pb}^{-1}$. The single- and double-differential cross sections for the production of dijets in diffractive DIS have been measured with the ZEUS detector in the kinematic region $5 < Q^2 < 100$ GeV^2, $100 < W < 250$ GeV and $x_{I\!\!P} < 0.03$, requiring at least two jets with $E^*_{T,\mathrm{jet}} > 4$ GeV in the pseudorapidity region $-3.5 < \eta^*_{\mathrm{jet}} < 0.0$ and the highest-E^*_T jet with $E^*_{T,\mathrm{j1}} > 5$ GeV.

By requiring dijets in the DIS regime, it is possible to study diffraction in presence of a hard scale (either the virtuality of the photon exchanged between the electron and the proton or the transverse energy of the jets). This allows to use the perturbative theory for describing the process. Furthermore, the main channel for producing dijets is via boson-gluon fusion, i.e. this process is directly sensitive to the gluon content of the diffractive PDFs. It has been demonstrated in previous analyses that gluon-initiated processes constitute the majority of the diffractive interactions. Because of these peculiar properties, diffractive dijet production is an excellent benchmark for the theory.

Experimental features of the analysis. This was the first analysis of this kind carried out at ZEUS. The DIS kinematics were reconstructed with the double angle method exploiting the informations of the scattered electron and the hadronic final state as measured with the high-resolution uranium calorimeter. The jets were reconstructed in the $\gamma^* p$ rest frame from energy flow objects with the longitudinally-invariant k_T clustering algorithm. The diffractive events were tagged by requiring the presence of a large rapidity gap (LRG) in the direction of the scattered proton. This method provides a clear experimental signature for diffraction and a final sample with high statistics. In this analysis a particularly relevant role was played by the forward plug calorimeter (FPC) subdetector. The LRG requirement was imposed by vetoing events with an energy deposit in the FPC higher than the signal fluctuations. No explicit cut on the pseudorapidity of the particles produced in the interaction was applied. The experimental method pursued provided increased statistics and required smaller unfolding corrections. In this way the impact of the simulation of processes like the fragmen-

tation of the partons produced and the showering of the particles in the calorimeter, poorly described by the MC, was reduced. All these experimental techniques resulted in a precise measurement, to date the most accurate of its kind.

The total, single and double differential cross section were extracted. The total cross section was measured to be

$$\sigma^D_{\mathrm{TOT}}(ep \to epj_1j_2X') = 91.5 \pm 1.2(stat.) \, ^{+3.3}_{-5.4}(syst.) \, ^{+6.4}_{-5.3}(corr.) \, \mathrm{pb}$$

The differential cross section as a function of $E^*_{\mathrm{T,jj}}$ exhibited a steep fall off as expected from QCD. The requirement of two jets with a minimum E^*_T imposes a kinematical constraint on M_X biasing the distributions towards higher values of M_X and $x^{\mathrm{obs}}_{I\!\!P}$. A peak at $z^{\mathrm{obs}}_{I\!\!P} \sim 0.3$ in the differential cross section was observed. This indicates that in the most of the events a large hadronic activity is present outside the dijet system. A tail at low values of x^{obs}_γ is also observed.

The double differential cross section was presented both in bins of $z^{\mathrm{obs}}_{I\!\!P}$ and Q^2 and in bins of $z^{\mathrm{obs}}_{I\!\!P}$ and $E^*_{\mathrm{T,j1}}$. The measurement of double differential cross sections provides at the same time a more detailed analysis of the process and a valuable input for theoretical models.

Comparison to LO Monte Carlos. The measured cross sections were compared to two LO MCs, RAPGAP and SATRAP . Although they were based on different theoretical approaches, the two MCs provided similar predictions, as expected theoretically for high E^*_T jets. The agreement in shape with the experimental distributions was good over the whole kinematical range. The introduction in RAPGAP of resolved-photon processes improved the description at low x^{obs}_γ.

Comparison to NLO calculation. The improvement in the recent years of the theoretical and experimental understanding of diffraction has given the possibility to develop more refined Next-To-Leading Order (NLO) QCD fits and calculations that can now be tested. Recent analyses from ZEUS and H1 supported the validity of the QCD factorisation theorem in DIS. Under this assumption, the precise measurement presented here can be exploited for putting more stringent constraints on the theory that still has not the same accuracy as for the inclusive scattering. The DISENT program was used to calculate the NLO predictions for the cross section. Since this program was natively written for standard non-diffractive DIS interactions, it had to be modified in order to accept the dPDFs. This was proposed in a previous analysis by the H1 collaboration and was done for the first time at ZEUS.

The NLO calculation was performed with different dPDFs available at the time of the analysis. Significant differences between the central values of the predictions were observed. The absolute normalisation of the measured cross section was described by the NLO calculation, supporting the validity of QCD factorisation in diffractive DIS. However, the large theoretical uncertainties did not allow to make any definitive statement about it. This aspect, compared to the precision of the measured data, suggests that these data have a strong constraining power on the theoretical predictions. The central values of the calculation using two of the available dPDFs, the H1 2006 $-$ FitB

and MRW 2006 fit, provided a good description of the data both in normalisation and shape over the entire kinematical range. The predictions using the ZEUS LPS+charm and H1 2006 − FitA dPDFs were higher than the data in normalisation and did not describe the shape of the data distributions in some regions of the phase space, especially at high $z_{I\!P}^{\rm obs}$.

Outlook. These data can be very useful for developing a more refined theoretical description of diffraction. It has been shown that the dijets data can be included in a NLO QCD fit together with the inclusive data improving significantly the accuracy of the parton densities, especially the gluon contribution. The double differential cross section was presented in such a way to include it into such a combined fit. This would be the most natural next step for this analysis. However, the dijet data are rich of informations that have still to be extracted. The azimuthal asymmetries between the scattered electron and the high $E_{\rm T}^*$ jets can reveal insights on the ratio between the yields of diffractive dijets with longitudinally and transversely polarised virtual photons. Repeating the analysis without the LRG requirement but tagging the scattered proton with the ZEUS Leading Proton Spectrometer would give a measurement that, although not competitive in precision with the present one, would provide a richer amount of information like the t-distribution for this set of events. An extension to higher values of the transverse energy of the jets would be welcomed because of the reduced theoretical uncertainties in that kinematical region. It could be achieved with the high statistics of the HERA II data sample. However, modifications of the detector apparatus may reduce the advantages of this analysis: in order to allocate the magnets that increased the luminosity, the rear beam hole in the calorimeter was enlarged and the FPC was taken away, affecting in this way the detection of both low Q^2 (i.e. high statistics) and diffractive events. A study would be needed in order to determine the feasibility of such an analysis.

The theory of diffraction can be significantly improved by the data presented in this thesis. Any progress in the dPDFs accuracy and in the understanding of soft rescatterings will be of great support to the rich diffractive physics program at the LHC.

The study of dijets is a precious tool for understanding diffractive physics and including it in the QCD framework. The use of QCD theoretical methods in diffraction has become available in the recent years and dijets are a powerful benchmark for testing and improving them. Big improvements have been achieved on the way of promoting diffractive physics as a respectable branch of QCD, as shown in this thesis, nonetheless a lot of work has still to be done and diffraction has still a lot to teach us about the nature of the proton and the strong interaction in general.

Appendix A

ZEUS Coordinate system

The ZEUS coordinate system is shown in Fig. A.1. It is a right-handed Cartesian system with the $Z-$axis pointing along the proton beam direction. The origin of the reference system is located at the nominal interaction point. The standard polar coordinates are used to define the θ and ϕ angles and the distance $\rho = \sqrt{x^2 + y^2}$.
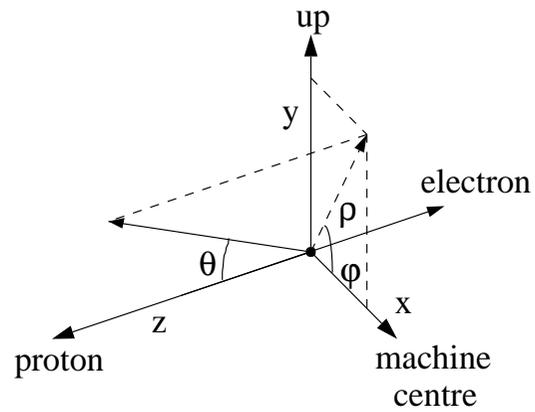
Figure A.1: The ZEUS coordinate system.

Appendix B

Jet energy corrections

The energy flow objects (EFOs) used for reconstructing the jets were corrected for the inactive material in the detector as described in Sect. 4.2.3. The description of the inactive material proved to be reliable because of the good description of several measured quantities by the MC (the detector-level transverse energy of the jets, for instance, see Fig. 6.8-6.12). However, the correlation with the same quantity at the hadron-level is affected by that. This because the energy losses biased the jet reconstruction. The transverse energies of the jets were therefore corrected in both MC and data with some factors estimated with the RAPGAP MC described in Sect. 5.1.1. The evaluation of the jet energy corrections was carried out through the following steps:

- A DIS jet sample was selected by imposing the DIS selections described in Sect. 6.1.2 and cuts on jets looser than the ones used in the final analysis:

  - $E_{\mathrm{T,jet}}^* > 2.5\,\mathrm{GeV}$, where $E_{\mathrm{T,jet}}^*$ was the transverse energy of the jet in the $\gamma^* p$ rest frame;

  - $-3.5 < \eta_{\mathrm{jet}}^* < 0.0$, where η_{jet}^* was the pseudorapidity of the jet in the $\gamma^* p$ rest frame;

  - $-2.0 < \eta_{\mathrm{jet}}^{\mathrm{LAB}} < 2.0$, where $\eta_{\mathrm{jet}}^{\mathrm{LAB}}$ was the pseudorapidity of the jet in the laboratory frame.

  The hadron selection applied was exactly the same of the final analysis. For the evaluation of the jet energy corrections, no diffractive cuts were applied, neither at the hadron nor at the detector level. The latter choice was motivated by gaining statistics for this kind of analysis that is supposed to be insensitive to the diffractive requirement.

- The detector-level jets were matched to the hadron-level jets in the $\eta - \phi$ plane in the $\gamma * p$ rest frame. The distance between a detector-level and a hadron-level jet, ΔR was defined as

$$\Delta R = \sqrt{(\eta_{\mathrm{jet,det}}^* - \eta_{\mathrm{jet,had}}^*)^2 + (\phi_{\mathrm{jet,det}}^* - \phi_{\mathrm{jet,had}}^*)^2} \qquad (B.1)$$

where the subscripts *det* and *had* refer to the detector- and hadron-level quantities, respectively. A detector-level jet was matched to the closest hadron-level jet; in any case it had to be $\Delta R < 1$ in order to match the detector-level jet;

- the correlation in $E_{\mathrm{T,jet}}^*$ between the matched pairs of detector- and hadron-level jets was plotted in bins of η_{jet}^*;

- a fit in the form

$$E_{\mathrm{T,jet}}^{*,\mathrm{det}} = P0 + P1 * E_{\mathrm{T,jet}}^{*,\mathrm{had}} \qquad (B.2)$$

was performed for each bin in η_{jet}^*;

- the Eq. B.2 was inverted and the new corrected transverse energy of the jet at the detector level, $E_{\mathrm{T,jet}}^{*,\mathrm{det}}(corr)$ was evaluated according to

$$E_{\mathrm{T,jet}}^{*,\mathrm{det}}(corr) = \frac{E_{\mathrm{T,jet}}^{*,\mathrm{det}} - P0}{P1} \qquad (B.3)$$

The function used to correct the detector-level transverse energy, $E_{\mathrm{T,jet}}^{*,\mathrm{det}}$, depended linearly by $E_{\mathrm{T,jet}}^{*,\mathrm{det}}$ itself. The two parameters of the linear correction changed according to the pseudorapidity in the $\gamma * p$ rest frame. This allowed to take more carefully in account the different amounts of inactive material in the detector.

The correlation between hadron- and detector-level jet transverse energies before the corrections is shown in Fig. B.1. After the corrections, the same correlations look like in Fig. B.2. The parameters $P0$ and $P1$ obtained from the linear fit B.2 and used for correcting the jets are shown in Fig. B.3 as a function of η_{jet}^*.
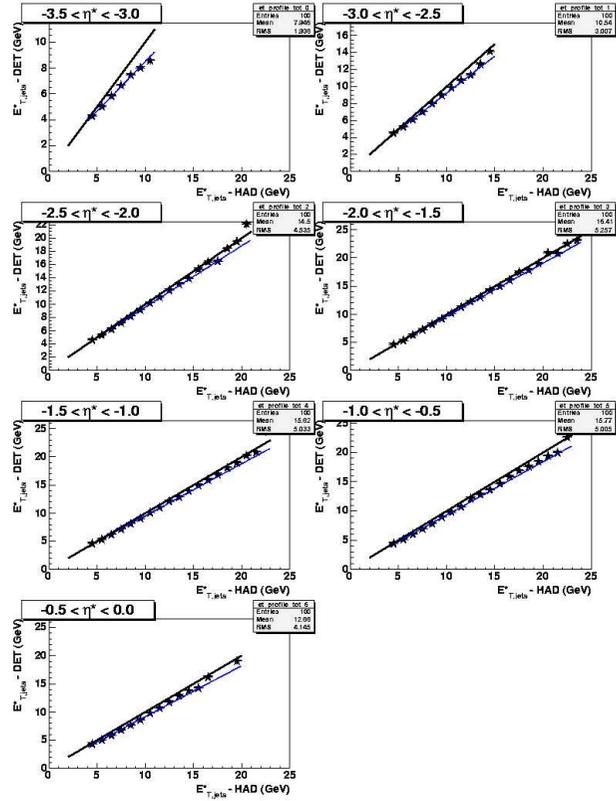
Figure B.1: The correlation of the $E^*_{\mathrm{T,jet}}$ for MC jets reconstructed at the hadron and detector level before the jet energy corrections. The RAPGAP sample described in Sect. 5.1.1 was used. The correlation is shown in bins of pseudorapidity of the jet as measured in the γ^*p rest frame. The thin blue line shows the result of a linear fit to the correlation. The thick black line indicates the 45° line.
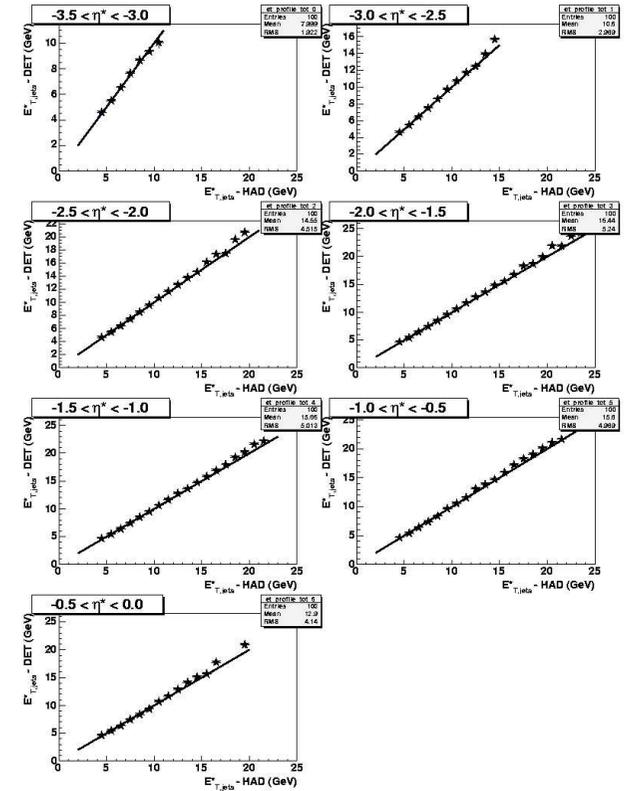

Figure B.2: The correlation of the $E^*_{\mathrm{T,jet}}$ for MC jets reconstructed at the hadron and detector level after the jet energy corrections. The RAPGAP sample described in Sect. 5.1.1 was used. The correlation is shown in bins of pseudorapidity of the jet as measured in the laboratory frame. The thin blue line shows the result of a linear fit to the correlation. The thick black line indicates the 45° line.
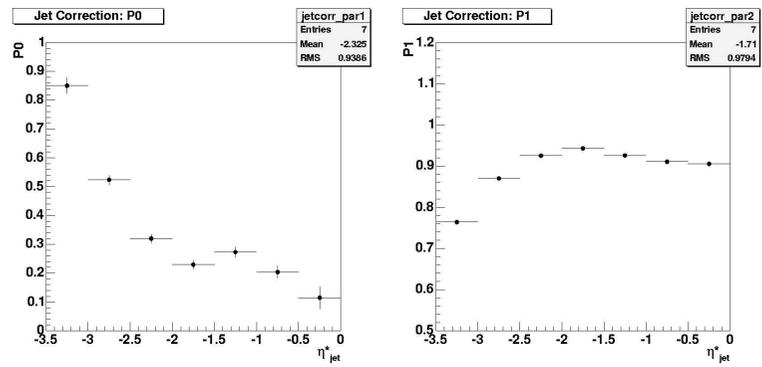
Figure B.3: The parameters (left) P0 and (right) P1 used for the jet corrections as a function of the jet pseudorapidity in the $\gamma^* p$ rest frame, η^*_{jet}.

Appendix C

Resolutions

The resolution indicates the accuracy of the reconstruction at the detector-level of the hadron-level value of a variable. It was estimated with the RAPGAP MC (see Sect.5.1.1). The resolutions are presented as a function of the variables used in the cross section extraction (see Sect.7) in Figs. C.1–C.3. The plots show for each variable κ

- the correlation between the hadron- and detector-level values for all the events;

- the distribution for the events of the residual, r_κ, defined as

$$r_\kappa = \frac{d_\kappa - h_\kappa}{h_\kappa}$$

where d_κ and h_κ indicate the detector-level and hadron-level values of the variable κ for each event;

- the mean and the full width at half maximum (FWHM) of the residual distributions for each bin of each variable.

Eventual selections on the considered variable were not applied in order to not bias the resolutions. The resolution on η_{j1}^* has been calculated only on matched jets, i.e. only detector-level jets matched in the $\eta - \phi$ plane to hadron level jets entered in the resolution calculation. The matching procedure is the same described in Appendix B.
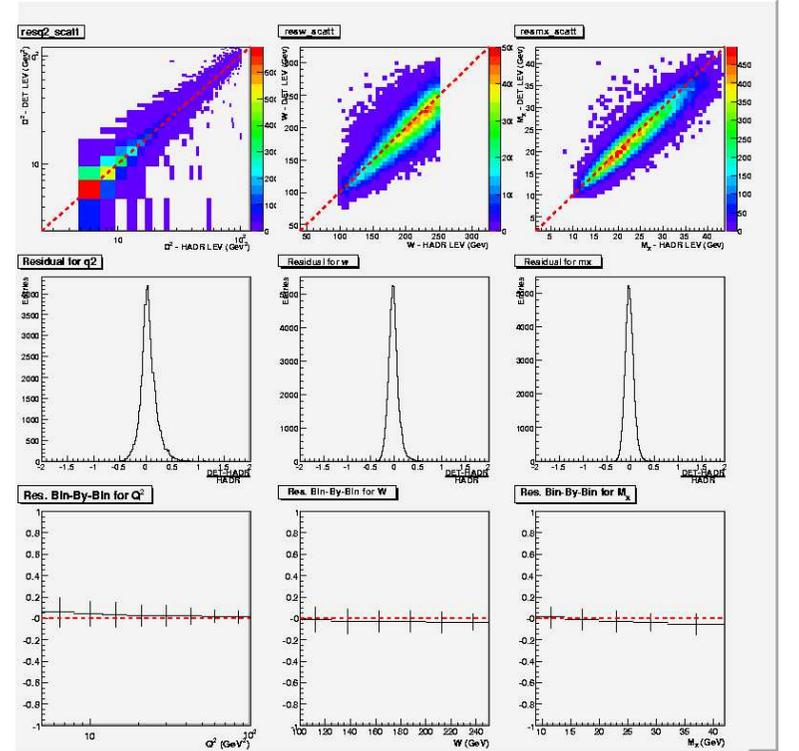

Figure C.1: Resolution on (left column) Q^2, (central column) W and (right column) M_X. The first row indicates the correlation between the generated ($X-$axis) and reconstructed ($Y-$axis) values of each variable. The middle row the total distribution of the residuals for each variable. The bottom row presents the distribution of the residuals for each bin in the generated value of the variables. In the latter plot, the points indicate the mean of the distributions while the rror bars the Half Width at Half Maximum of the distributions.
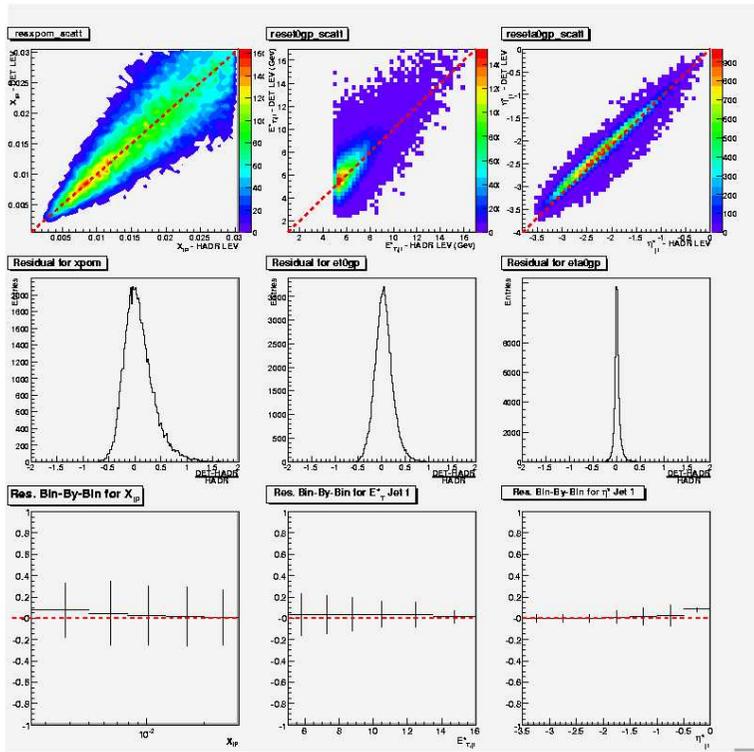
Figure C.2: Resolution on (left column) $x_{I\!\!P}$, (central column) the transverse energy of the jet with highest transverse energy in the $\gamma^* p$ frame, $E_{T,j1}^*$ and (right column) the pseudorapidity of the jet with highest transverse energy in the $\gamma^* p$ frame, η_{j1}^*. Other details in the caption of Fig. C.1.
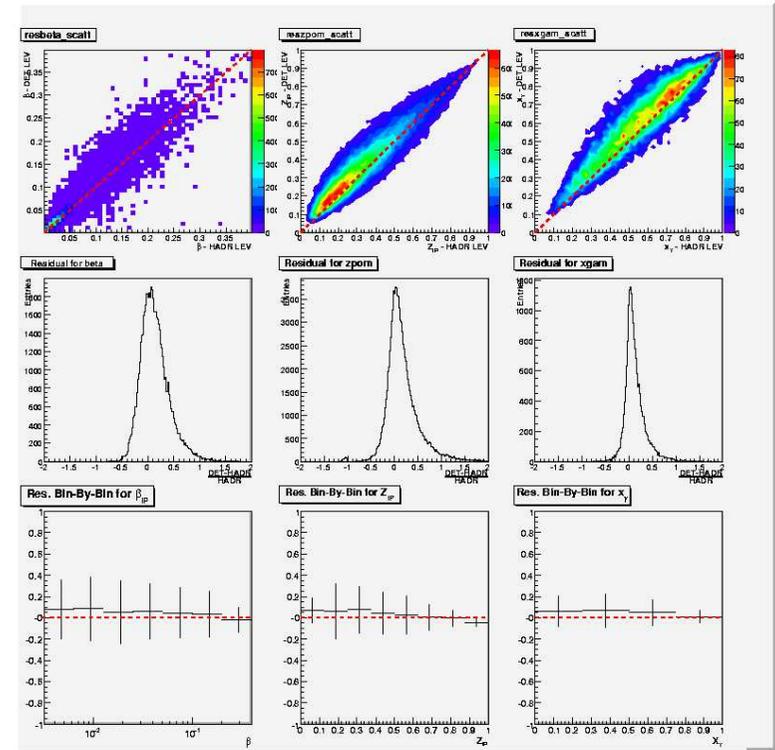

Figure C.3: Resolution on (left column) β, (central column) $z_{I\!\!P}^{\mathrm{obs}}$ and (right column) x_γ^{obs}. Other details in the caption of Fig. C.1.

Appendix D

Systematics

The uncertainties related to the experimental devices and techniques contribute to the systematic uncertainties. Several sources of systematic uncertainties have been checked. A more detailed description of the systematic checks and their treatment can be found in Sect. 7.1.3. In the following pages, the single contributions coming from each of them are presented. The contributions can beFor ease of presentation in the figures, an identifying number was associated to each of them. They are listed as follows:

ID1 the energy scale of the calorimeter was decreased by 3% only in the Monte Carlo (MC) simulation;

ID2 the energy scale of the calorimeter was increased by 3% only in the MC simulation;

ID3 the energy scale of the DIS scattered electron was decreased by 2% only in the MC simulation;

ID4 the energy scale of the DIS scattered electron was increased by 2% only in the MC simulation;

ID5 the energy scale of the FPC was decreased by 10% only in the MC simulation;

ID6 the energy scale of the FPC was increased by 10% only in the MC simulation;

ID7 the size of the fiducial area outside which the DIS electron had to lie was shrinked by 0.5 cm;

ID8 the size of the fiducial area outside which the DIS electron had to lie was enlarged by 0.5 cm;

ID9 the selection on the energy of the scattered electron was decreased from 10 GeV to 8 GeV;

ID10 the lower threshold of the selection on the total $E - p_Z$ was changed from 45 GeV to 43 GeV;

ID11 the position along the $Z-$axis of the SRTD was shifted by 2 mm away from the interaction point only in the MC;

ID12 the position along the $Z-$axis of the SRTD was shifted by 2 mm towards the interaction point only in the MC;

ID13 the factors used to correct for the detector smearing where calculated using only the RAPGAP MC;

ID14 the factors used to correct for the detector smearing where calculated using only the SATRAP MC;

In Figs. D.1-D.9, the relative uncertainties introduced by each of the systematic uncertainties listed above are presented for each variable separately. In Figs. D.10-D.23 the contribution of each systematic is presented for all the variables in one plot.
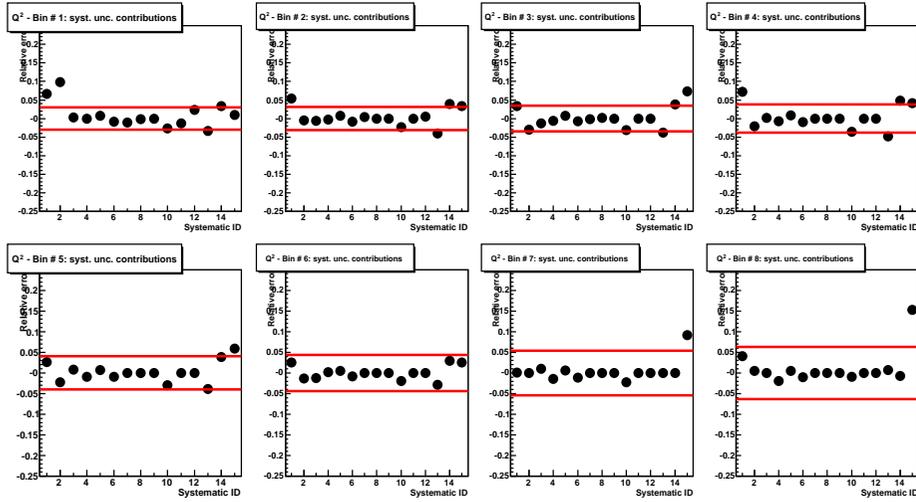
Figure D.1: The relative error introduced by the systematic uncertainties for Q^2. Every different pad considers only one bin of the variable. Every pad shows the single contributions from all the systematic sources. The solid line indicates the size of the statistical error.

Figure D.3: The relative error introduced by the systematic uncertainties for $x_{I\!P}$. Other details as in the caption of Fig. D.1.
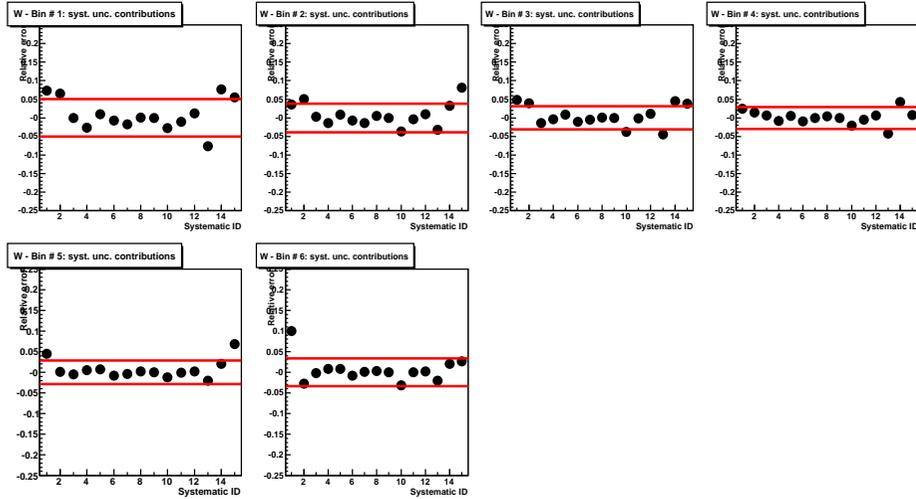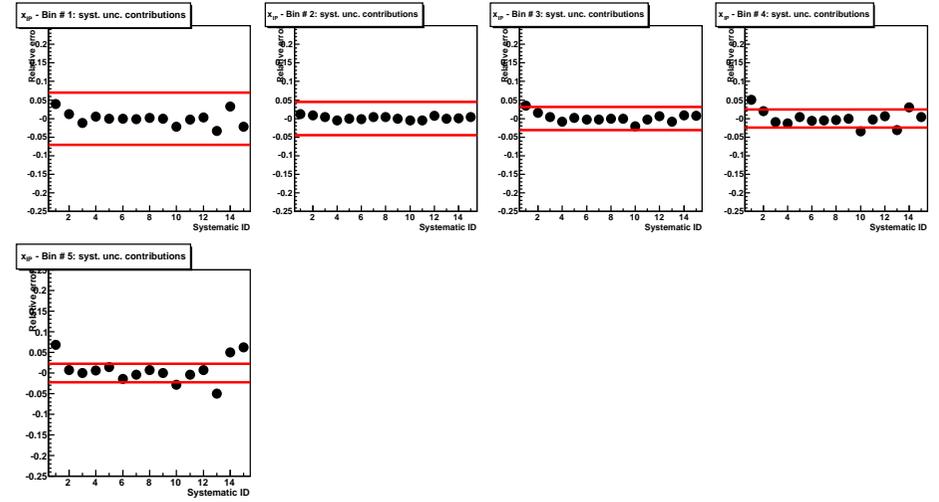

Figure D.2: The relative error introduced by the systematic uncertainties for W. Other details as in the caption of Fig. D.1.
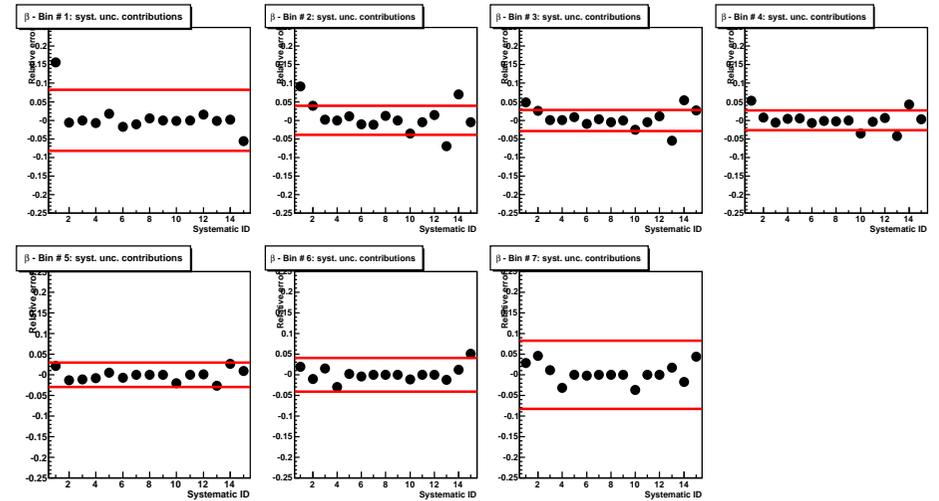
Figure D.4: The relative error introduced by the systematic uncertainties for β. Other details as in the caption of Fig. D.1.

Figure D.5: The relative error introduced by the systematic uncertainties for $E^*_{\mathrm{T,jj}}$. Other details as in the caption of Fig. D.1.

Figure D.7: The relative error introduced by the systematic uncertainties for M_{X}. Other details as in the caption of Fig. D.1.

Figure D.6: The relative error introduced by the systematic uncertainties for η^*_{jj}. Other details as in the caption of Fig. D.1.
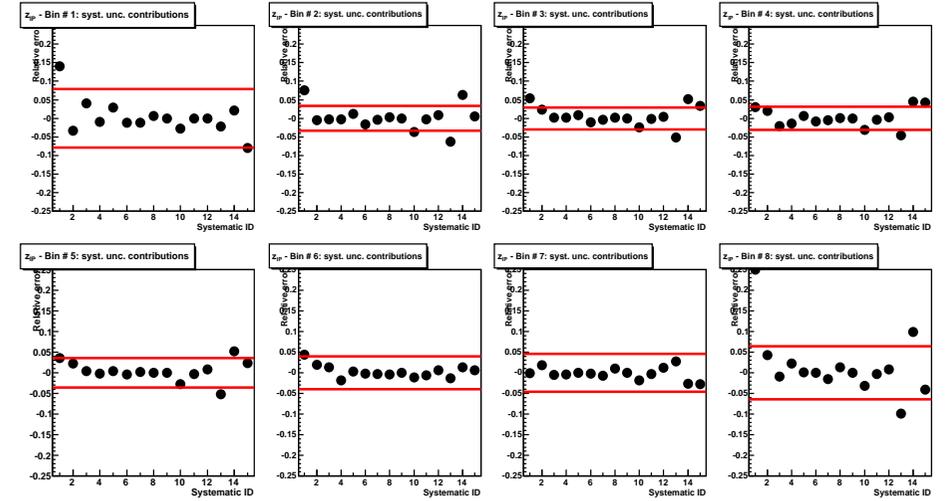
Figure D.8: The relative error introduced by the systematic uncertainties for $z^{\mathrm{obs}}_{I\!\!P}$. Other details as in the caption of Fig. D.1.
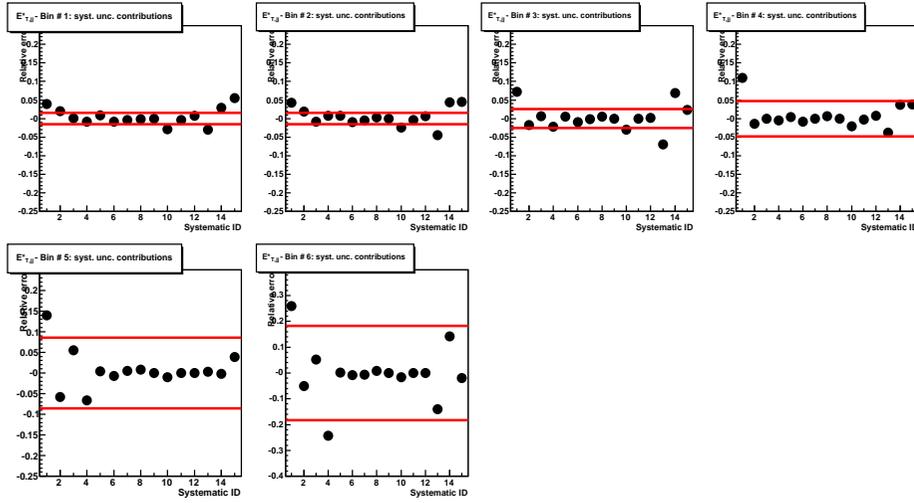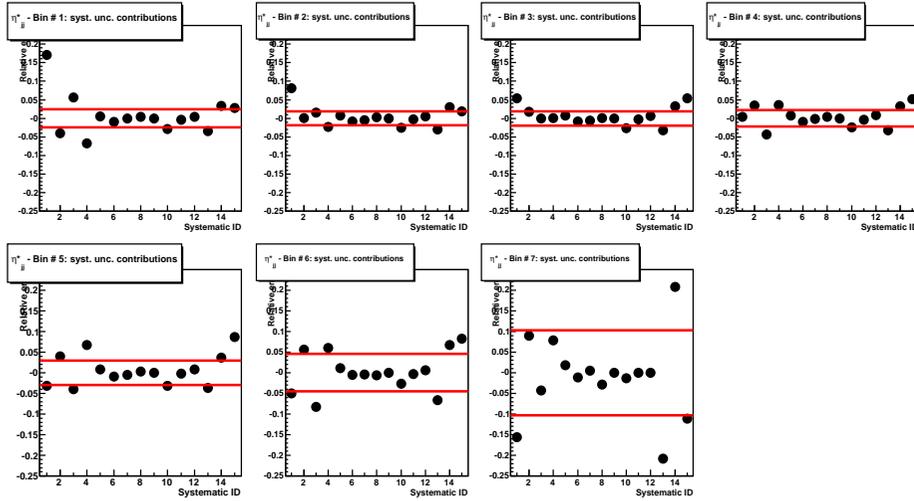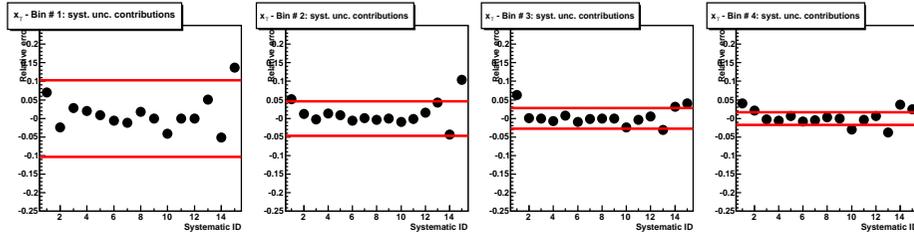
Figure D.9: The relative error introduced by the systematic uncertainties for x_γ^{obs}. Other details as in the caption of Fig. D.1.

*Figure D.10: The relative error introduced by the systematic **ID1** uncertainties for the following variables: Q^2, W, $x_{I\!\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_{X}, $z_{I\!\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

*Figure D.11: The relative error introduced by the systematic **ID2** uncertainties for the following variables: Q^2, W, $x_{I\!\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_{X}, $z_{I\!\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

*Figure D.12: The relative error introduced by the systematic **ID3** uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

*Figure D.14: The relative error introduced by the systematic **ID5** uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

*Figure D.13: The relative error introduced by the systematic **ID4** uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*
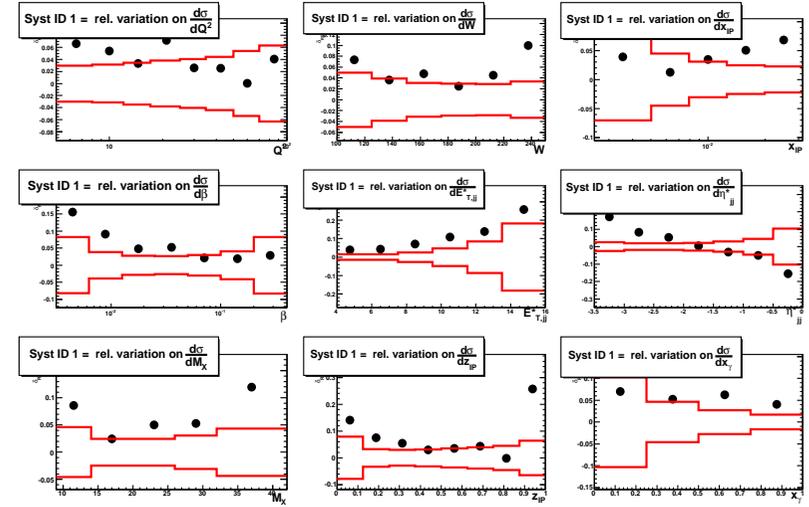

*Figure D.15: The relative error introduced by the systematic **ID6** uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*
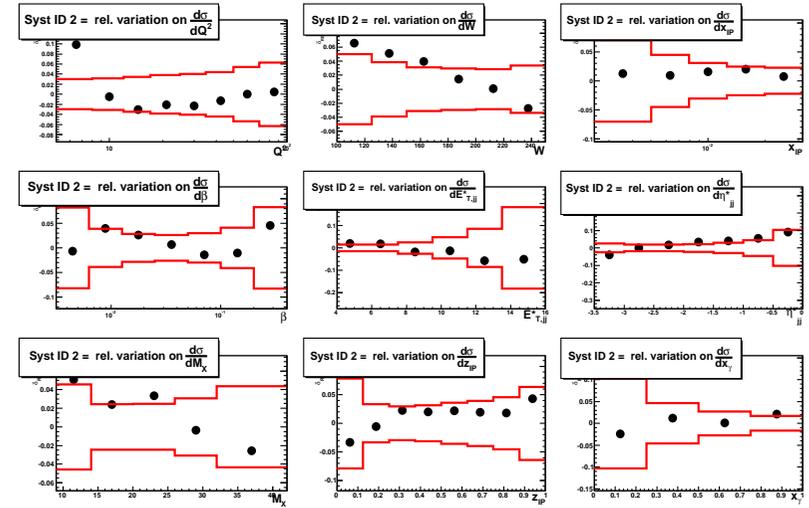
*Figure D.16: The relative error introduced by the systematic **ID7** uncertainties for the following variables: Q^2, W, $x^{\mathrm{obs}}_{I\!P}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z^{\mathrm{obs}}_{I\!P}$ and $x^{\mathrm{obs}}_{\gamma}$. The solid line indicates the size of the statistical error.*

*Figure D.18: The relative error introduced by the systematic **ID9** uncertainties for the following variables: Q^2, W, $x^{\mathrm{obs}}_{I\!P}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z^{\mathrm{obs}}_{I\!P}$ and $x^{\mathrm{obs}}_{\gamma}$. The solid line indicates the size of the statistical error.*
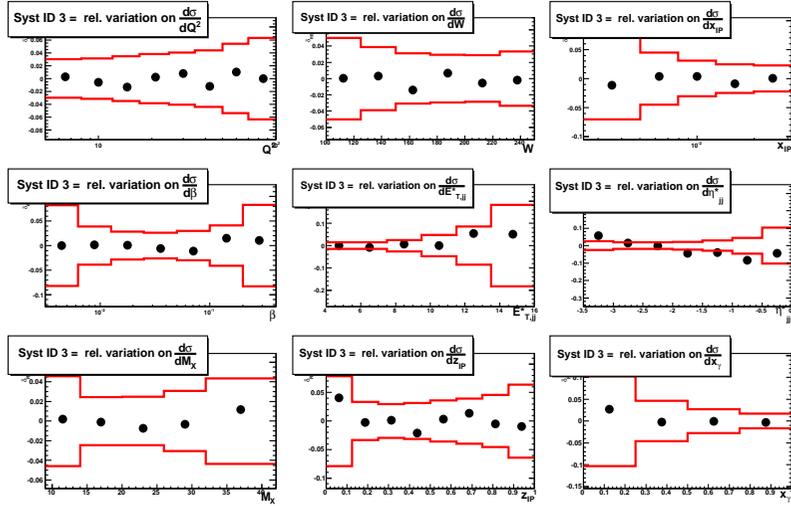

*Figure D.17: The relative error introduced by the systematic **ID8** uncertainties for the following variables: Q^2, W, $x^{\mathrm{obs}}_{I\!P}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z^{\mathrm{obs}}_{I\!P}$ and $x^{\mathrm{obs}}_{\gamma}$. The solid line indicates the size of the statistical error.*
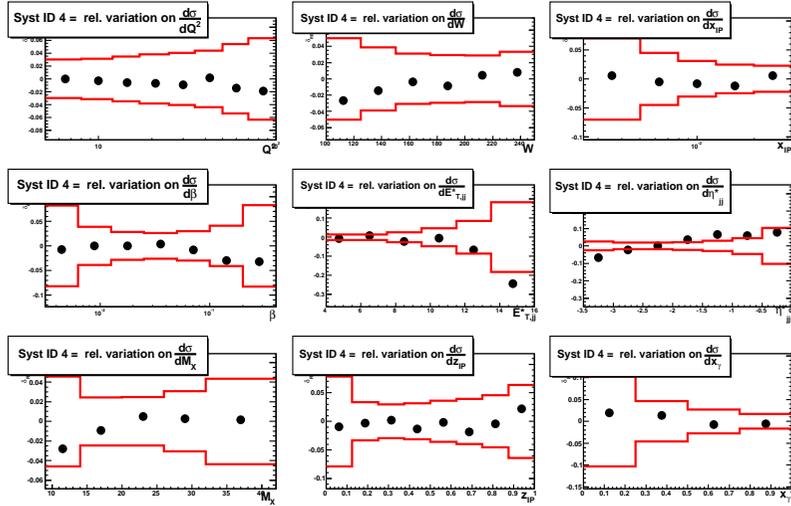

*Figure D.19: The relative error introduced by the systematic **ID10** uncertainties for the following variables: Q^2, W, $x^{\mathrm{obs}}_{I\!P}$, β, $E^*_{\mathrm{T,jj}}$, η^*_{jj}, M_{X}, $z^{\mathrm{obs}}_{I\!P}$ and $x^{\mathrm{obs}}_{\gamma}$. The solid line indicates the size of the statistical error.*
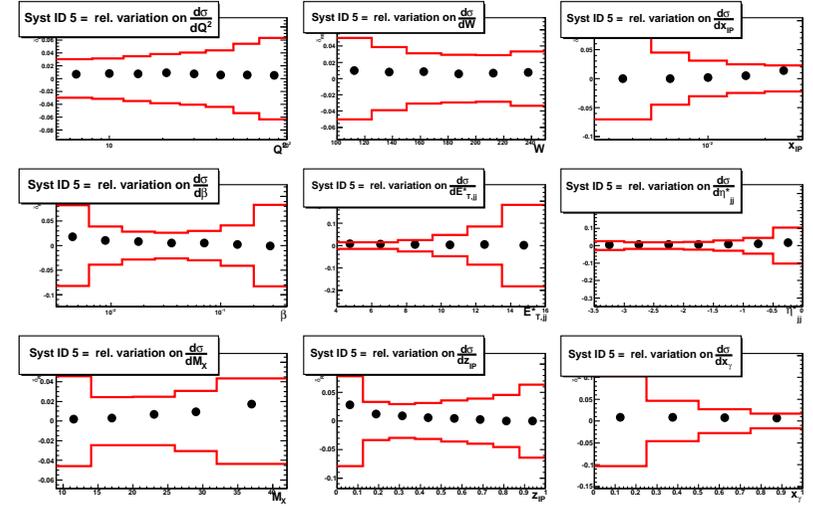
Figure D.20: The relative error introduced by the systematic **ID11** *uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_X, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

Figure D.22: The relative error introduced by the systematic **ID13** *uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_X, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*
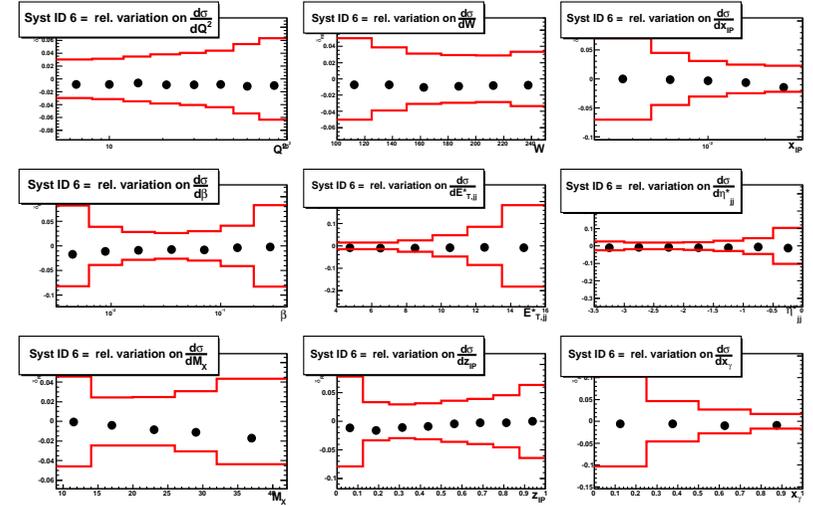

Figure D.21: The relative error introduced by the systematic **ID12** *uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_X, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*
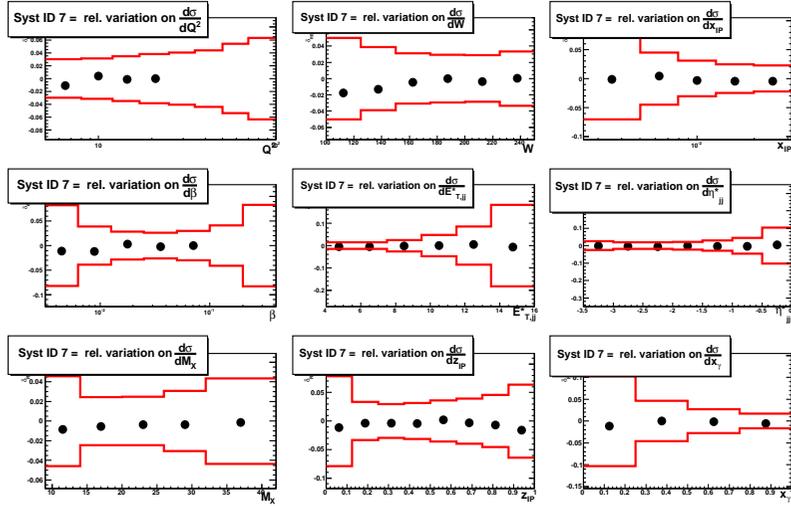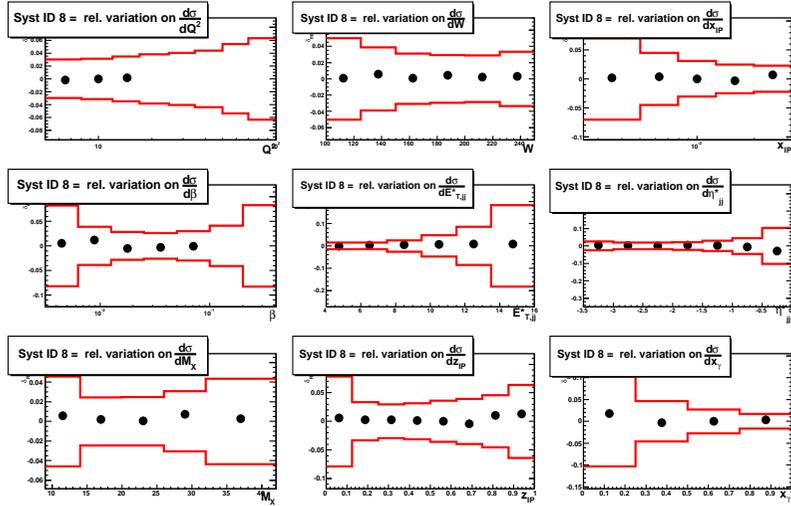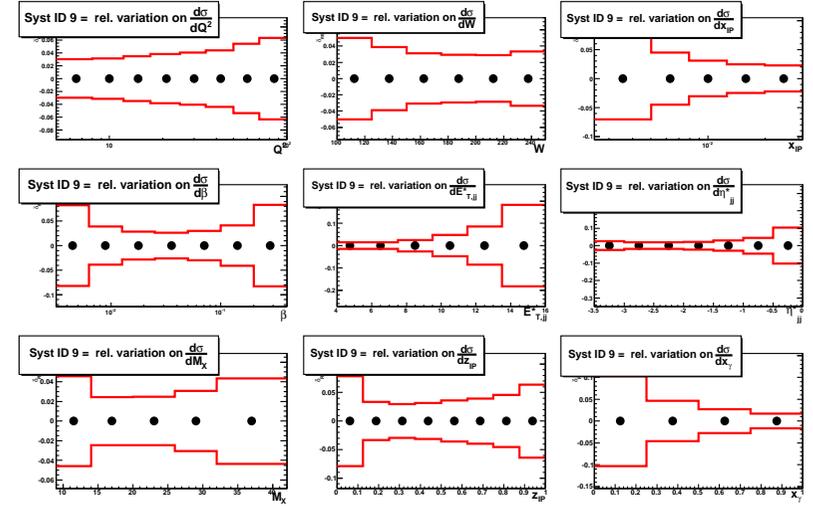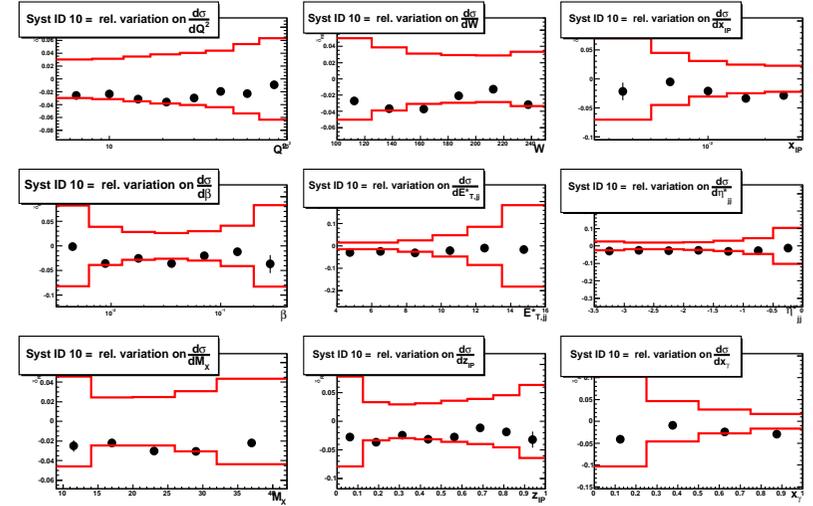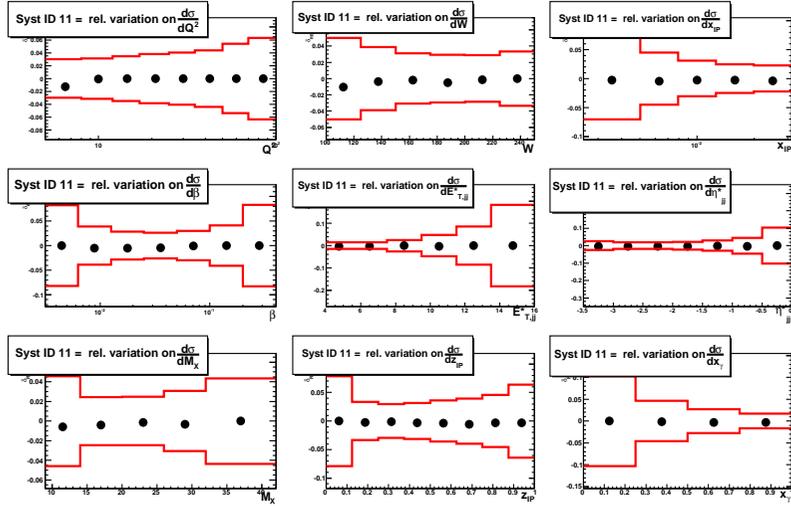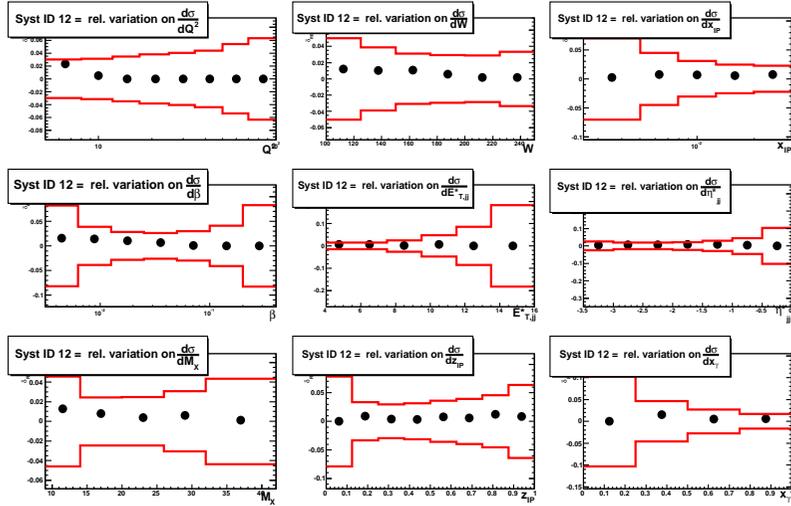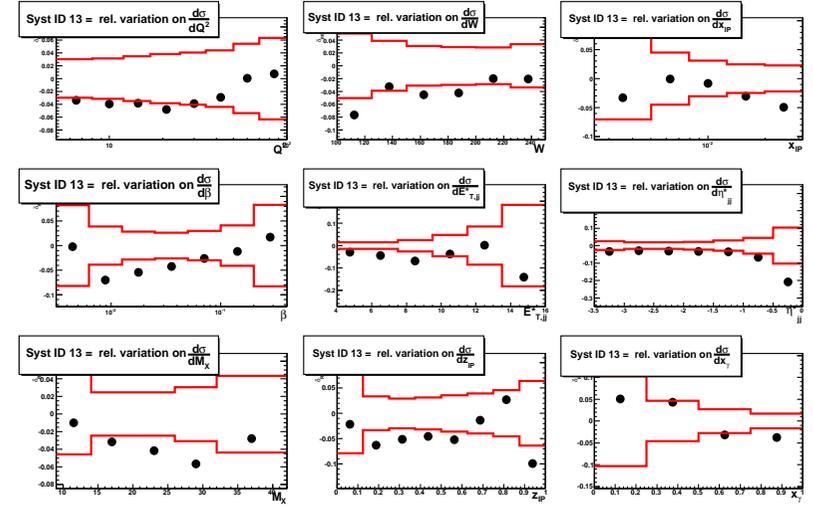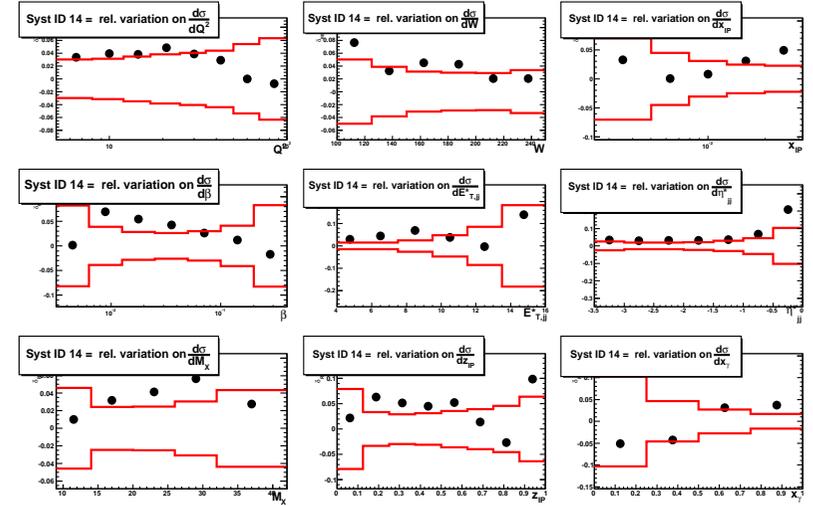

Figure D.23: The relative error introduced by the systematic **ID14** *uncertainties for the following variables: Q^2, W, $x_{I\!P}^{\mathrm{obs}}$, β, $E_{\mathrm{T,jj}}^*$, η_{jj}^*, M_X, $z_{I\!P}^{\mathrm{obs}}$ and x_γ^{obs}. The solid line indicates the size of the statistical error.*

Appendix E

Numerical values of the measured differential cross section

Q^2 bin (GeV^2)	$d\sigma/dQ^2$ (pb/GeV^2)	δ_{stat} (pb/GeV^2)	δ_{syst} (pb/GeV^2)	δ_{ES} (pb/GeV^2)	Δ_{DIFFR} (pb/GeV^2)
5, 8	7.4	$\pm$ 0.2	$^{+0.3}_{-0.5}$	$^{+0.5}_{-0.5}$	0.1
8, 12	4.2	$\pm$ 0.13	$^{+0.2}_{-0.3}$	$^{+0.3}_{-0.3}$	0.1
12, 17	2.7	$\pm$ 0.1	$^{+0.2}_{-0.2}$	$^{+0.2}_{-0.2}$	0.2
17, 25	1.28	$\pm$ 0.05	$^{+0.08}_{-0.12}$	$^{+0.08}_{-0.07}$	0.05
25, 35	0.94	$\pm$ 0.04	$^{+0.07}_{-0.07}$	$^{+0.06}_{-0.05}$	0.06
35, 50	0.52	$\pm$ 0.02	$^{+0.02}_{-0.03}$	$^{+0.03}_{-0.03}$	0.01
50, 70	0.26	$\pm$ 0.01	$^{+0.02}_{-0.01}$	$^{+0.01}_{-0.01}$	0.02
70, 100	0.133	$\pm$ 0.008	$^{+0.020}_{-0.003}$	$^{+0.005}_{-0.006}$	0.020

Table E.1: Values of the differential cross section as a function of Q^2 for the production of diffractive dijets. The range over which the cross section is averaged is given in the first column. The statistical, δ_{stat}, uncorrelated systematic, δ_{syst}, and energy scale uncertainties, δ_{ES}, are quoted separately. The theoretical uncertainty on the NLO calculations, δ_{theor}, is quoted in the sixth column. The difference with the measured cross section with and without η_{MAX} cut, Δ_{DIFFR}, is presented in the last column. The uncertainties on the proton dissociation subtraction and the luminosity measurement are not presented in the table.

W bin (GeV)	$d\sigma/dW$ (pb/ GeV)	δ_{stat} (pb/ GeV)	δ_{syst} (pb/ GeV)	δ_{ES} (pb/ GeV)	Δ_{DIFFR} (pb/ GeV)
100, 125	0.25	$\pm$ 0.01	$^{+0.02}_{-0.03}$	$^{+0.01}_{-0.01}$	0.01
125, 150	0.42	$\pm$ 0.02	$^{+0.04}_{-0.03}$	$^{+0.02}_{-0.03}$	0.03
150, 175	0.65	$\pm$ 0.02	$^{+0.04}_{-0.06}$	$^{+0.04}_{-0.04}$	0.03
175, 200	0.69	$\pm$ 0.02	$^{+0.03}_{-0.04}$	$^{+0.05}_{-0.04}$	0.01
200, 225	0.77	$\pm$ 0.02	$^{+0.06}_{-0.03}$	$^{+0.05}_{-0.05}$	0.05
225, 250	0.80	$\pm$ 0.03	$^{+0.03}_{-0.06}$	$^{+0.05}_{-0.05}$	0.02

Table E.2: Values of the differential cross section as a function of W. Other details as in the caption of Table E.1.

M_X bin (GeV)	$d\sigma/dM_X$ (pb/ GeV)	δ_{stat} (pb/ GeV)	δ_{syst} (pb/ GeV)	δ_{ES} (pb/ GeV)	Δ_{DIFFR} (pb/ GeV)
9, 14	1.25	$\pm$ 0.06	$^{+0.02}_{-0.08}$	$^{+0.05}_{-0.06}$	-0.03
14, 20	4.2	$\pm$ 0.1	$^{+0.2}_{-0.2}$	$^{+0.2}_{-0.2}$	0.1
20, 26	4.4	$\pm$ 0.1	$^{+0.2}_{-0.4}$	$^{+0.2}_{-0.2}$	0.0
26, 32	3.1	$\pm$ 0.1	$^{+0.2}_{-0.3}$	$^{+0.3}_{-0.2}$	-0.1
32, 42	1.15	$\pm$ 0.05	$^{+0.09}_{-0.06}$	$^{+0.12}_{-0.09}$	0.08

Table E.3: Values of the differential cross sections with respect to M_X. Other details as in the caption of Table E.1.

β bin ($\times 10^{-2}$)	$d\sigma/d\beta$ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	Δ_{DIFFR} (pb)
0.32, 0.63	1238	$\pm$ 102	$^{+30}_{-76}$	$^{+150}_{-137}$	-73
0.63, 1.26	2110	$\pm$ 82	$^{+152}_{-219}$	$^{+195}_{-175}$	-11
1.26, 2.51	1713	$\pm$ 48	$^{+106}_{-131}$	$^{+110}_{-107}$	45
2.51, 5.01	894	$\pm$ 24	$^{+39}_{-80}$	$^{+53}_{-49}$	3
5.01, 10.00	324	$\pm$ 10	$^{+9}_{-18}$	$^{+14}_{-17}$	4
10.00, 19.95	84	$\pm$ 3	$^{+4}_{-3}$	$^{+4}_{-4}$	4
19.95, 39.81	8.8	$\pm$ 0.7	$^{+0.4}_{-0.5}$	$^{+0.5}_{-0.5}$	0.4

Table E.4: Values of the differential cross sections with respect to β. Other details as in the caption of Table E.1.

η_{jj}^* bin	$d\sigma/d\eta_{\mathrm{jj}}^*$ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	Δ_{DIFFR} (pb)
-3.5, -3	56.1	$\pm$ 1.4	$^{+2.5}_{-3.7}$	$^{+7.7}_{-7.6}$	1.6
-3, -2.5	97.1	$\pm$ 1.8	$^{+3.6}_{-6.1}$	$^{+7.1}_{-7.0}$	1.8
-2.5, -2	88.4	$\pm$ 1.7	$^{+5.7}_{-5.9}$	$^{+5.0}_{-4.8}$	4.8
-2, -1.5	65.5	$\pm$ 1.5	$^{+4.1}_{-4.1}$	$^{+3.7}_{-4.0}$	3.4
-1.5, -1	34.9	$\pm$ 1.1	$^{+3.3}_{-2.6}$	$^{+2.7}_{-2.0}$	3.0
-1, -0.5	13.4	$\pm$ 0.6	$^{+1.4}_{-1.3}$	$^{+1.4}_{-1.3}$	1.1
-0.5, 0	1.7	$\pm$ 0.2	$^{+0.4}_{-0.4}$	$^{+0.3}_{-0.2}$	-0.2

Table E.7: Values of the differential cross sections with respect to η_{jj}^*. Other details as in the caption of Table E.1.

$x_{\mathbb{P}}^{\mathrm{obs}}$ bin ($\times 10^{-2}$)	$d\sigma/dx_{\mathbb{P}}^{\mathrm{obs}}$ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	Δ_{DIFFR} (pb)
0.25, 0.50	3131	$\pm$ 220	$^{+103}_{-191}$	$^{+127}_{-137}$	-67
0.50, 0.79	6099	$\pm$ 274	$^{+51}_{-89}$	$^{+247}_{-302}$	25
0.79, 1.26	8105	$\pm$ 250	$^{+108}_{-361}$	$^{+363}_{-434}$	63
1.26, 1.99	8329	$\pm$ 205	$^{+266}_{-676}$	$^{+455}_{-456}$	35
1.99, 3.00	8070	$\pm$ 182	$^{+656}_{-651}$	$^{+695}_{-582}$	503

Table E.5: Values of the differential cross sections with respect to $x_{\mathbb{P}}^{\mathrm{obs}}$. Other details as in the caption of Table E.1.

$z_{\mathbb{P}}^{\mathrm{obs}}$	$d\sigma/dz_{\mathbb{P}}^{\mathrm{obs}}$ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	Δ_{DIFFR} (pb)
0, 0.125	25.4	$\pm$ 2.0	$^{+0.9}_{-2.4}$	$^{+3.79}_{-2.12}$	-2.08
0.125, 0.25	131.6	$\pm$ 4.4	$^{+8.6}_{-13.5}$	$^{+12.29}_{-10.80}$	0.68
0.25, 0.375	152.3	$\pm$ 4.5	$^{+9.4}_{-12.3}$	$^{+9.9}_{-9.6}$	5.1
0.375, 0.5	132.0	$\pm$ 4.1	$^{+8.2}_{-10.0}$	$^{+6.0}_{-8.0}$	5.6
0.5, 0.625	98.3	$\pm$ 3.5	$^{+5.7}_{-7.6}$	$^{+5.1}_{-5.2}$	2.3
0.625, 0.75	82.9	$\pm$ 3.3	$^{+1.4}_{-2.9}$	$^{+3.9}_{-3.9}$	0.5
0.75, 0.875	57.8	$\pm$ 2.6	$^{+1.7}_{-3.2}$	$^{+2.9}_{-3.0}$	-1.6
0.875, 1	31.5	$\pm$ 2.0	$^{+3.1}_{-4.0}$	$^{+2.2}_{-1.5}$	-1.3

Table E.8: Values of the differential cross sections with respect to $z_{\mathbb{P}}^{\mathrm{obs}}$. Other details as in the caption of Table E.1.

$E_{\mathrm{T,jj}}^*$ bin (GeV)	$d\sigma/dE_{\mathrm{T,jj}}^*$ (pb/ GeV)	δ_{stat} (pb/ GeV)	δ_{syst} (pb/ GeV)	δ_{ES} (pb/ GeV)	Δ_{DIFFR} (pb/ GeV)
4, 5.5	50.9	$\pm$ 0.8	$^{+3.2}_{-3.7}$	$^{+2.5}_{-2.9}$	2.8
5.5, 7.5	39.8	$\pm$ 0.6	$^{+2.6}_{-2.8}$	$^{+2.3}_{-2.0}$	1.8
7.5, 9.5	9.7	$\pm$ 0.3	$^{+0.7}_{-0.9}$	$^{+0.8}_{-0.9}$	0.2
9.5, 11.5	2.2	$\pm$ 0.1	$^{+0.1}_{-0.1}$	$^{+0.3}_{-0.2}$	0.1
11.5, 13.5	0.59	$\pm$ 0.05	$^{+0.02}_{-0.01}$	$^{+0.07}_{-0.10}$	0.02
13.5, 16	0.10	$\pm$ 0.02	$^{+0.02}_{-0.02}$	$^{+0.01}_{-0.03}$	0.00

Table E.6: Values of the differential cross sections with respect to $E_{\mathrm{T,jj}}^*$. Other details as in the caption of Table E.1.

$x_{\gamma}^{\mathrm{obs}}$ bin	$d\sigma/dx_{\gamma}^{\mathrm{obs}}$ (pb)	δ_{stat} (pb)	δ_{syst} (pb)	δ_{ES} (pb)	Δ_{DIFFR} (pb)
0, 0.25	6.3	$\pm$ 0.6	$^{+0.9}_{-0.5}$	$^{+0.7}_{-0.3}$	0.9
0.25, 0.5	25.2	$\pm$ 1.2	$^{+2.9}_{-1.5}$	$^{+1.5}_{-2.2}$	2.6
0.5, 0.75	85.2	$\pm$ 2.3	$^{+4.4}_{-5.5}$	$^{+7.6}_{-8.0}$	3.4
0.75, 1	238.3	$\pm$ 4.0	$^{+10.9}_{-17.1}$	$^{+12.3}_{-11.6}$	5.8

Table E.9: Values of the differential cross sections with respect to $x_{\gamma}^{\mathrm{obs}}$. Other details as in the caption of Table E.1.

$z_{I\!P}^{\rm obs}$ bin	$d\sigma/dz_{I\!P}^{\rm obs}dE_{T,j1}^*$ (pb/ GeV)	$\delta_{\rm stat}$ (pb/ GeV)	$\delta_{\rm syst}$ (pb/ GeV)	$\delta_{\rm ES}$ (pb/ GeV)	$\Delta_{\rm DIFFR}$ (pb/ GeV)
$5.0 < E_{T,j1}^* < 6.5\,{\rm GeV}(< E_{T,j1}^* >= 5.7\,{\rm GeV})$					
0, 0.25	31.8	$\pm$ 1.3	$^{+2.1}_{-4.1}$	$^{+1.9}_{-1.3}$	-1.2
0.25, 0.375	55.8	$\pm$ 2.4	$^{+4.4}_{-4.8}$	$^{+2.5}_{-2.4}$	3.1
0.375, 0.5	47.5	$\pm$ 2.2	$^{+3.0}_{-3.0}$	$^{+1.9}_{-2.4}$	2.5
0.5, 0.625	36.2	$\pm$ 2.0	$^{+1.1}_{-1.7}$	$^{+1.6}_{-1.9}$	1.0
0.625, 0.75	30.5	$\pm$ 1.9	$^{+0.3}_{-2.0}$	$^{+1.4}_{-1.3}$	-0.8
0.75, 0.875	20.3	$\pm$ 1.5	$^{+1.3}_{-2.1}$	$^{+0.9}_{-0.9}$	-1.2
0.875, 1	10.6	$\pm$ 1.1	$^{+0.2}_{-1.0}$	$^{+0.4}_{-0.5}$	-0.4
$6.5 < E_{T,j1}^* < 8.0\,{\rm GeV}(< E_{T,j1}^* >= 7.2\,{\rm GeV})$					
0, 0.25	14.1	$\pm$ 0.8	$^{+1.5}_{-1.6}$	$^{+1.0}_{-0.6}$	0.2
0.25, 0.375	26.2	$\pm$ 1.4	$^{+1.2}_{-2.3}$	$^{+1.9}_{-2.0}$	-0.9
0.375, 0.5	22.9	$\pm$ 1.3	$^{+2.0}_{-1.8}$	$^{+1.7}_{-0.9}$	1.9
0.5, 0.625	18.4	$\pm$ 1.2	$^{+0.9}_{-1.0}$	$^{+1.2}_{-1.0}$	0.7
0.625, 0.75	14.5	$\pm$ 1.1	$^{+0.7}_{-0.8}$	$^{+0.6}_{-0.6}$	0.1
0.75, 0.875	11.8	$\pm$ 0.9	$^{+0.8}_{-0.9}$	$^{+0.6}_{-0.8}$	0.0
0.875, 1	5.0	$\pm$ 0.6	$^{+0.1}_{-0.2}$	$^{+0.5}_{-0.2}$	-0.1
$8.0 < E_{T,j1}^* < 16.0\,{\rm GeV}(< E_{T,j1}^* >= 9.7\,{\rm GeV})$					
0, 0.25	0.56	$\pm$ 0.05	$^{+0.07}_{-0.04}$	$^{+0.08}_{-0.07}$	0.06
0.25, 0.375	2.3	$\pm$ 0.1	$^{+0.4}_{-0.3}$	$^{+0.2}_{-0.2}$	0.2
0.375, 0.5	2.3	$\pm$ 0.1	$^{+0.2}_{-0.2}$	$^{+0.2}_{-0.3}$	-0.1
0.5, 0.625	1.8	$\pm$ 0.1	$^{+0.1}_{-0.2}$	$^{+0.1}_{-0.2}$	0.0
0.625, 0.75	1.7	$\pm$ 0.1	$^{+0.1}_{-0.0}$	$^{+0.2}_{-0.2}$	0.1
0.75, 0.875	1.5	$\pm$ 0.1	$^{+0.0}_{-0.1}$	$^{+0.1}_{-0.2}$	0.0
0.875, 1	0.9	$\pm$ 0.1	$^{+0.0}_{-0.1}$	$^{+0.1}_{-0.1}$	-0.1

Table E.10: Values of the double differential cross sections with respect to $z_{I\!P}^{\rm obs}$ in bins of $E_{T,\rm j1}^*$. Other details as in the caption of Table E.1.

$z_{I\!P}^{\rm obs}$ bin	$d\sigma/dz_{I\!P}^{\rm obs}dQ^2$ (pb/ GeV^2)	$\delta_{\rm stat}$ (pb/ GeV^2)	$\delta_{\rm syst}$ (pb/ GeV^2)	$\delta_{\rm ES}$ (pb/ GeV^2)	$\Delta_{\rm DIFFR}$ (pb/ GeV^2)
0, 0.25	5.5	$\pm$ 0.2	$^{+0.4}_{-0.6}$	$^{+0.4}_{-0.4}$	0.0
0.25, 0.375	9.0	$\pm$ 0.4	$^{+0.6}_{-0.8}$	$^{+0.7}_{-0.7}$	-0.1
0.375, 0.5	7.6	$\pm$ 0.4	$^{+0.4}_{-0.5}$	$^{+0.5}_{-0.5}$	0.3
0.5, 0.625	5.1	$\pm$ 0.3	$^{+0.1}_{-0.2}$	$^{+0.4}_{-0.4}$	0.1
0.625, 0.75	4.4	$\pm$ 0.3	$^{+0.1}_{-0.2}$	$^{+0.3}_{-0.2}$	0.1
0.75, 0.875	3.2	$\pm$ 0.3	$^{+0.1}_{-0.2}$	$^{+0.2}_{-0.2}$	-0.1
0.875, 1	1.4	$\pm$ 0.2	$^{+0.1}_{-0.2}$	$^{+0.1}_{-0.1}$	-0.1
$12 < Q^2 < 25\,{\rm GeV}^2(< Q^2 >= 17.2\,{\rm GeV}^2)$					
0, 0.25	1.38	$\pm$ 0.08	$^{+0.13}_{-0.16}$	$^{+0.09}_{-0.07}$	0.07
0.25, 0.375	3.1	$\pm$ 0.2	$^{+0.4}_{-0.2}$	$^{+0.2}_{-0.2}$	0.4
0.375, 0.5	2.2	$\pm$ 0.1	$^{+0.1}_{-0.2}$	$^{+0.1}_{-0.1}$	0.1
0.5, 0.625	2.0	$\pm$ 0.1	$^{+0.1}_{-0.2}$	$^{+0.1}_{-0.1}$	0.0
0.625, 0.75	1.6	$\pm$ 0.1	$^{+0.1}_{-0.1}$	$^{+0.1}_{-0.1}$	0.0
0.75, 0.875	1.3	$\pm$ 0.1	$^{+0.0}_{-0.1}$	$^{+0.1}_{-0.1}$	-0.1
0.875, 1	0.66	$\pm$ 0.08	$^{+0.01}_{-0.04}$	$^{+0.05}_{-0.03}$	-0.01

Table E.11: Values of the double differential cross sections with respect to $z_{I\!P}^{\rm obs}$ in bins of $E_{T,\rm j1}^*$. Other details as in the caption of Table E.1.

$z_{I\!\!P}^{\rm obs}$ bin	$d\sigma/dz_{I\!\!P}^{\rm obs}dQ^2$ (pb/ GeV^2)	$\delta_{\rm stat}$ (pb/ GeV^2)	$\delta_{\rm syst}$ (pb/ GeV^2)	$\delta_{\rm ES}$ (pb/ GeV^2)	$\Delta_{\rm DIFFR}$ (pb/ GeV^2)
$25 < Q^2 < 50\,{\rm GeV}^2 (< Q^2 >= 35.2\,{\rm GeV}^2)$					
0, 0.25	0.49	$\pm$ 0.03	$^{+0.03}_{-0.07}$	$^{+0.03}_{-0.02}$	-0.06
0.25, 0.375	1.00	$\pm$ 0.07	$^{+0.07}_{-0.13}$	$^{+0.06}_{-0.07}$	-0.03
0.375, 0.5	0.99	$\pm$ 0.07	$^{+0.06}_{-0.06}$	$^{+0.04}_{-0.06}$	0.05
0.5, 0.625	0.76	$\pm$ 0.06	$^{+0.06}_{-0.02}$	$^{+0.04}_{-0.04}$	0.06
0.625, 0.75	0.62	$\pm$ 0.05	$^{+0.03}_{-0.04}$	$^{+0.04}_{-0.03}$	-0.01
0.75, 0.875	0.47	$\pm$ 0.05	$^{+0.02}_{-0.02}$	$^{+0.03}_{-0.04}$	0.00
0.875, 1	0.23	$\pm$ 0.03	$^{+0.01}_{-0.01}$	$^{+0.02}_{-0.01}$	0.00
$50 < Q^2 < 100\,{\rm GeV}^2 (< Q^2 >= 69.5\,{\rm GeV}^2)$					
0, 0.25	0.09	$\pm$ 0.01	$^{+0.00}_{-0.01}$	$^{+0.01}_{-0.01}$	0.00
0.25, 0.375	0.25	$\pm$ 0.02	$^{+0.05}_{-0.01}$	$^{+0.02}_{-0.01}$	0.05
0.375, 0.5	0.29	$\pm$ 0.03	$^{+0.02}_{-0.01}$	$^{+0.01}_{-0.02}$	0.02
0.5, 0.625	0.20	$\pm$ 0.02	$^{+0.02}_{-0.01}$	$^{+0.01}_{-0.01}$	0.02
0.625, 0.75	0.16	$\pm$ 0.02	$^{+0.01}_{-0.00}$	$^{+0.01}_{-0.01}$	0.01
0.75, 0.875	0.12	$\pm$ 0.02	$^{+0.00}_{-0.00}$	$^{+0.01}_{-0.01}$	0.00
0.875, 1	0.10	$\pm$ 0.01	$^{+0.01}_{-0.01}$	$^{+0.01}_{-0.01}$	0.00

Table E.12: Values of the double differential cross sections with respect to $z_{I\!\!P}^{\rm obs}$ in bins of $E_{T,\rm j1}^*$. Other details as in the caption of Table E.1.

Bibliography

[1] S.L. Glashow, Nucl. Phys. **20**, 579 (1961);
S. Weinberg, Phys. Lett. **19**, 1264 (1967);
A. Salam, *Elementary particle theory*, ed. N. Svartholm (1968).

[2] F. Halzen and A. Martin, *Quarks and Leptons: an introductory course in Modern Particle Physics*, John Wiley & sons, Inc. (1984).

[3] E. Leader and E. Predazzi, *An Introduction to Gauge Theories and Modern Particle Physics*, Cambridge University Press (1996).

[4] M. Gell-Mann, Phys. Lett. **8**, 214 (1964).

[5] P.W. Higgs, Phys. Rev. Lett. **12**, 132 (1964);
P.W. Higgs, Phys. **145**, 1156 (1966).

[6] F. Gianotti, Collider Physics: LHC, ATL-CONF-2000-001.

[7] R. Feynman, Phys. Rev. Lett. **23**, 1415 (1969).

[8] R. Devenish and A. Cooper–Sarkar, *Deep Inelastic Scattering*, Oxford University Press (2004).

[9] E.D. Bloom et al., Phys. Rev. Lett. **23**, 930 (1969);
M. Breidenbach et al., Phys. Rev. Lett. **23**, 935 (1969).

[10] G. Sterman et al., *Handbook of pQCD*; available at
http://www.phys.psu.edu/~cteq/handbook/v1.1/handbook.pdf .

[11] B. Delamotte, Am. J. Phys. **72**, 170 (2004);
e-Print: http://arxiv.org/abs/hep-th/0212049 .

[12] W.-M. Yao et al., *Review of Particle Physics*, J. Phys. **G 33**, 1 (2006).

[13] C. Glasman,Proceedings of the XIII International Workshop on Deep Inelastic Scattering, Madison, USA (2005);
e-Print: http://arxiv.org/abs/hep-ex/0506035 .

[14] J.C. Collins, D.E. Soper and G. Sterman, *Perturbative Quantum Chromodynamics*, pag. 1, ed. A.H. Mueller, World Scientific Singapore.

[15] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972);
L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 96 (1975);
Y.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[16] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (2007);
G. Altarelli, Nucl. Phys **B 81**, 1 (1981).

[17] ZEUS Coll., S. Chekanov et al., Phys. Review **D 67**, 012007 (2003).

[18] Y. Balitzki and L.N. Lipatov, Phys. Rev. Lett. **28**, 822 (1978);
E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Phys.Rev. Lett. **44**, 443 (1976);
E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Phys.Rev. Lett. **45**, 199 (1977).

[19] J.R. Forshaw, G. Kerley and G.Shaw, Phys. Rev. **D 60**, 074012 (1999);
K. Golec-Biernat, Acta Phys. Pol. **33**, 2771 (2002).

[20] K. Golec-Biernat and M. Wüsthoff, Phys. Rev. **D 59**, 014017 (1999).

[21] L.V. Gribov, E.M. Levin and M.G. Ryskin, Phys. Rep. **100**, 1 (1983).

[22] K. Golec-Biernat, J. Phys. **G 28**, 1057 (2002).

[23] J.E. Huth et al., *Proc. of the 1990 DPF Summer Study on Physics, Colorado, ed. E.L. Berger*, (World Scientific, Singapore, 1992) 134.

[24] UA1 Coll., G. Arnison et al., Phys. Lett. **B 123**, 115 (1983).

[25] M.H. Seymour, CERN-TH/95-176.

[26] G.C. Blazey et al., e-Print: `hep-ex/0005012` .

[27] G.P. Salam and G. Soyez, JHEP **5**, 086 (2007).

[28] JADE Coll., W. Bartel et al., Z. Phys. **C 33**, 23 (1986);
JADE Coll., S. Bethke et al., Phys. Lett **B 213**, 235 (1988).

[29] S. Catani, Yu.L. Dokshitzer and B.R. Webber, Phys. Lett. **B 285**, 291 (1992), CERN-TH-6473-92;
S. Catani, Yu.L. Dokshitzer and B.R. Webber, Nucl. Phys. **B 406**, 187 (1993), CERN-TH-6775-93.

[30] S. Catani and M.H. Seymour, Nucl. Phys. **B 485**, 291 (1997);
Erratum-ibid., **B 510**, 503 (1997).

[31] Mark-J Coll., D.P. Barber et al., Phys. Rev. Lett. **43**, 830 (1979);
TASSO Coll., R. Brandelik et al., Phys. Lett. **B 86**, 243 (1979);
PLUTO Coll., C. Berger et al., Phys. Lett. **B 86**, 418 (1979).

[32] ZEUS Coll., S. Chekanov et al., Phys. Lett. **B 551**, 226 (2003).

[33] Z. Nagy and Z. Trocsanyi, Nucl. Phys. Proc. Suppl. **74**, 44 (1999).

[34] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 44**, 183 (2005).

[35] ZEUS Coll., J.Breitweg et al., Eur. Phys. J. **C 8**, 367 (1999).

[36] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C42**, 1 (2005).

[37] ZEUS Coll., S. Chekanov et al., Nucl. Phys. **B 765**, 1 (2007).

[38] CDF Coll., A. Abulencia et al., Phys. Rev. Lett. 96, 122001 (2006).

[39] L.D. Landau and I.Y. Pomeranchuk, Zu. Eksper. Teor. Fiz **24**, 505 (1953);
     E. Feinberg and I.Y. Pomeranchuk, Nuovo Cimento Suppl. **3**, 652 (1956);
     A.I. Akhiezer and I.Y. Pomeranchuk, Uspekhi Fiz. Nauk. **65**, 593 (1958);
     V.N. Gribov, Sov. Phys. JETP **29**, 377 (1969).

[40] T. Regge, Nuovo Cimento **14**,951 (1959);
     T. Regge, Nuovo Cimento **18**,947 (1959).

[41] P.D.B. Collins, *An Introduction to Regge Theory and High Energy Physics*, Cambridge University Press (1977).

[42] M. Arneodo and M. Diehl, Preprint DESY-05-214 (2005); e-Print: `http://arxiv.org/abs/hep-ph/0511047` .

[43] E. Predazzi, *Lectures given at Hadrons VI*, Florianopolis, Brazil (1998);
     e-Print: `http://arxiv.org/abs/hep-ph/9809454` .

[44] V.N. Gribov, Sov. Phys. JETP **14**, 478 (1962).

[45] A. Donnachie and P.V. Landshoff, Phys. Lett. **B 185**, 403 (1987);
     A. Donnachie and P.V. Landshoff, Phys. Lett. **B 296**, 227 (1993);
     A. Donnachie and P.V. Landshoff, Phys. Lett. **B 437**, 408 (1998);

[46] G. Ingelman and P.E. Schlein, Phys. Lett.**B 152**, 256 (1985).

[47] A. Donnachie and P.V. Landshoff, Phys. Lett. **B 191**, 309 (1987).

[48] UA8 Coll., R. Bonino et al., Phys. Lett. **B 211**, 239 (1988);
     UA8 Coll., A. Brandt et al., Phys. Lett. **B 297**, 417 (1992).

[49] UA8 Coll., A. Brandt et al., Nucl. Inst. Meth. **A 327**, 412 (1993).

[50] ZEUS Coll., M. Derrick et al., Z. Phys. **C 73**, 253 (1997).

[51] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 25**, 169 (2002).

[52] ZEUS Coll., S. Chekanov et al., Nucl. Phys.**B 658**, 3 (2003).

[53] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 38**, 43 (2004).

[54] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 6**, 587 (1999).

[55] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 48**, 749 (2006).

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

[57] ZEUS Coll., M. Derrick et al., Phys. Lett. **B 356**, 129 (1995).

[58] H1 Coll., T. Ahmed et al., Nucl. Phys. **B 429**, 477 (1994).

[59] H1 Coll., T. Ahmed et al., Phys. Lett. **B 348**, 681 (1995).

[60] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 48**, 715 (2006).

[61] K. Golec-Biernat, J. Kwiecinski and A. Szczurek, Phys. Rev. **D 56**, 3955 (1997).

[62] ZEUS Coll., M. Derrick et al., Z. Phys. **C 70**, 391 (1996).

[63] ZEUS Coll., S. Chekanov et al., Nucl. Phys. **B 713**, 3 (2005).

[64] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 30**, 1 (2003).

[65] K. Golec-Biernat and W. Wüsthoff, Phys. Rev. **D 60**, 114023 (1999);
     K. Golec-Biernat and W. Wüsthoff, Eur. Phys. J. **C 20**, 313 (2001).

[66] J. Bartels, J. Phys. **G 28**, 1045 (2002).

[67] J.C. Collins, Phys. Rev. **D 57**, 3051 (1998);
     J.C. Collins, J. Phys. **G 28**, 1069 (2002).

[68] CDF Coll., T. Affolder et al., Phys. Rev. Lett. **84**, 5043 (2000).

[69] V.A. Khoze, A.D. Martin and M.G. Ryskin, Phys. Lett. **B 502** (2001).
     A.B. Kaidalov et al., Phys. Lett.**B 559**, 235 (2003)

[70] M. Klasen and G. Kramer, Eur. Phys. J. **C 38**, 93 (2004).

[71] V.A. Khoze, A.D. Martin and M.G. Ryskin, Eur. Phys. J. **C 18**, 167 (2000).

[72] J.J. Sakurai, Phys. Rev. Lett. **22**, 981 (1969).

[73] M. Glück, E. Reya and A. Vogt, Phys. Rev. **D 45**, 3986 (1992);
     M. Glück, E. Reya and A. Vogt, Phys. Rev. **D 46**, 1973 (1992).

[74] A. Kaidalov et al., Eur. Phys. J. **C 21**, 521 (2001);
     A. Kaidalov et al., Phys.Lett. **B 567**, 61 (2003);

[75] ZEUS Coll., S. Chekanov et al., Nucl. Phys. **B 672**, 3 (2003).

[76] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 51**, 301 (2007).

[77] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 50**, 1 (2007).

[78] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 52**, 813 (2007).

[79] ZEUS Coll., S. Chekanov et al., *Subm. to Eur. Phys. J.*, DESY-07-161 (2007).

[80] H1 Coll., A. Aktas et al., Eur. Phys. J. **C 51**, 549 (2007).

[81] H1 Coll., A. Aktas et al., *Subm. to JHEP*; e-Print `arXiv:0708.3217` [hep-ex].

[82] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 23**, 13 (2002).

[83] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 42**, 1 (2005).

[84] S. Alekhin et al., *Proceedings of "HERA and the LHC: A Workshop on the implications of HERA for LHC physics", Part B*, CERN-2005-014, 2004-2005, Eds. A. De Roeck and H. Jung (2005); e-Print: `hep-ph/0601013` .

[85] ZEUS Coll., U. Holm et al., *The ZEUS Detector, Status Report 1993* (unpublished), DESY (1993).

[86] ZEUS Coll., B. Foster et al., Nucl. Inst. Meth. **A 338**, 254 (1994).

[87] M. Derrick et al., Nucl. Inst. Meth. **A 309**, 77 (1991).

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

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

[90] ZEUS Coll., A. Bamberger et al., FPC group, Nucl. Inst. Meth. **A 450**, 235 (2000).

[91] F. Goebel, *Measurement of the Diffractive Contribution to the DIS Cross Section Using the ZEUS Forward Plug Calorimeter*, Ph.D. Thesis, University of Hamburg, Report DESY-THESIS-01-049 (2001).

[92] A. Dwurazny et al., Nucl. Inst. Meth. **A 277**, 176 (1989).

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

[94] H. Bethe and W. Heitler, Proc. Roy. Soc. **A 146**, 83 (1934).

[95] J. Andruszków, "First measurement of HERA luminosity by ZEUS lumi monitor," DESY-92-066.

[96] K. Piotrzkowski, Nucl. Inst. Meth. **B 119**, 253 (1996).

[97] W.H. Smith, K. Tokushuku and L.W. Wiggers, *Proc. Computing in High Energy Physics (CHEP), Annecy, France, Sept. 1992*, C. Verkerk and W. Wojcik (eds.), p. 222. CERN, Geneva, Switzerland (1992). Also in preprint DESY 92-150B.

[98] S. Silverstein et al., Nucl. Inst. Meth. **A 360**, 322 (1995).

[99] A. Quadt et al., Nucl. Inst. Meth. **A 438**, 472 (1999).

[100] H. Uijterwaal, *Global Second Level Trigger for ZEUS*, Ph.D. Thesis, University of Amsterdam, 1992.

[101] U. Behrens et al., *Status of the ZEUS Eventbuilder*, Internal ZEUS-Note 92-054, DESY (1992).

[102] S.M. Fisher and P. Palazzi, *ADAMO Reference Manual for Version 3.3*, CERN-ECP;
available at `http://adamo.web.cern.ch/Adamo/refmanual/Document.html` .

[103] D. C. Bailey et al., *presented at International Conference on Open Bus Systems 92*, Zurich, Switzerland, 13-15 Oct 1992.

[104] G.F. Hartner, VCTRACK(3.7/04): Offline Output Information, Internal ZEUS-Note 97-064 (1997);
G.F. Hartner, VCTRACK Briefing: Program and Math, Internal ZEUS-Note 98-058 (1998).

[105] H. Abramowicz, A. Caldwell, R. Sinkus, Nucl. Inst. Meth. **A 365**, 508 (1995);
R. Sinkus and T. Voss, Nucl. Inst. Meth **A 391**, 360 (1997).

[106] T. Doeker, A. Frey and M. Nakao, *E*lectron Position Reconstruction - Update of the ELECPO routines, Internal ZEUS-Note 94-123 (1994);
Ch. Amelung, *Electron Position Reconstruction in ZEUS: Further Update of the ELECPO package (Based on 1995 data)*, Internal ZEUS-Note 96-093 (1996).

[107] G. Briskin, *Diffractive dissociation in ep deep inelastic scattering*, Ph.D. Thesis, University of Tel Aviv, Report DESY-THESIS-1998-036 (1998).

[108] N. Tuning, *Proton structure functions at HERA*, Ph.D. Thesis, University of Amsterdam (2001);
N. Tuning, *ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction.*, Internal ZEUS-Note 01-021 (2001).

[109] M. Turcato, *Measurement of beauty photoproduction at HERA*, Ph.D. Thesis, University of Padova, Report DESY-THESIS-03-039 (2003).

[110] R. Brun et al., preprint CERN-DD/EE/84-1, CERN, 1987.

[111] M. Wing, *Proceedings for "10th International Conference on calorimetry in High Energy Physics"*, CALTECH, Pasadena, USA (2002); e-Print: `hep-ex/0206036` .

[112] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 23**, 615 (2002).

[113] S. Bentvelsen, J. Engelen and P. Kooijman, *Proc. Workshop on Physics at HERA*, W. Buchmuller and G. Ingelman (ed.), Vol.1, p.23, DESY, Hamburg, Germany (1992).

[114] H. Jung, *The RAPGAP Monte Carlo for Deep Inelastic Scattering version 2.00/18* (2001); available at `http://www.desy.de/~jung/rapgap/` .

[115] H1 Coll., C. Adloff et al., Z. Phys. **C 76**, 613 (1997).

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

[117] A. Kwiatowski, H. Spiesberger and H.-J. Möhring, Comp. Phys. Comm. **69**, 155 (1992).

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

[119] M. Bengtsson and T. Sjöstrand, Comp. Phys. Comm. **46**, 43 (1987); T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994).

[120] C. Peterson, D. Schlatter, I. Schmitt and P. Zerwas, Phys. Rev. **D 27**, 105 (1983).

[121] ZEUS Coll., S. Chekanov et al., Phys. Lett. **B 649**, 111 (2007).

[122] G. Gustafson and U. Pettersson, Nucl. Phys. **B 306**, 746 (1988).

[123] H. Spiesberger, *HERACLES and DJANGOH: Event Generation for ep Interactions at HERA Including Radiative Processes*, (1998); available at `http://wwwthep.physik.uni-mainz.de/~hspiesb/djangoh/djangoh.html` .

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

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

[126] H. Lim, Ph.D. Thesis, The Graudate School Kiungpook National University, (2002).

[127] S. Catani and M. Seymour, Nucl. Phys. **B 485**, 291 (1997); Erratum-ibid. **B 510**, 503 (1997).

[128] Unpublished. Informations available at `http://www.nikhef.nl/~h24/qcdnum/`.

[129] ZEUS Coll., S. Chekanov et al., Phys. Rev. **D 69**, 012004 (2004).

[130] M. Groys, A. Levy and A.Proskuryakov, in *Proceedings of "HERA and the LHC: a workshop on the implications of HERA for LHC physics"*, CERN-2005-014, 2004-2005, Eds. A. De Roeck and H. Jung, p. 499 (2005); e-Print: `hep-ph/0601013` .

[131] A.D. Martin, M.G. Ryskin and G. Watt, Phys. Lett. **B 644**, 131 (2006).

[132] T. Yamashita, Ph.D. Thesis, University of Tokyo, (2001) (unpublished).

[133] T. Tawara, Ph.D. Thesis, University of Tokyo, (2007) (unpublished).

[134] M. Kasprzak, Internal ZEUS note 95-069 (1995).

[135] S. Kagawa, Ph.D. Thesis, University of Tokyo, (2004) (unpublished).

[136] J. Bartels, K. Golec-Biernat, H. Kowalski, Phys. Rev. **D66**, 014001 (2002).

Acknowledgments

This thesis is the result of three years of work in ZEUS, a period that was sometimes tough, sometimes funny, other times very rewarding. At the end I can say that it was a great life experience that taught me plenty of things about myself and the others. All of this could not be possible without the presence and help of many people and I want to thank all of them. I would like to mention some of them in particular.

I want to thank firstly my supervisors, Prof. Dr. R. Klanner and Prof. Dr. P. Schleper, for their fundamental support and inspiration. They helped me and guided me with their impressive experience and humanity.

I worked for more than three years in the CMS/ZEUS group of the Hamburg university and there I met many interesting and nice people. I want to thank T. Schörner-Sadenius, J. Sztuk and M. Turcato for all the times that they found the time to help me and answer to my questions. The students of the group form a nice and active crew and we spent together a lot of time and lunches at the extraordinary DESY-Kantine. I will keep a very nice memory of that. A mention in particular to E. Butz, my first room mate and student of italian selected expressions, and to D. Contarato who helped me to enter in the DESY and Hamburg environments when I was just at the start, stranded in a place where everybody was speaking german.

I want to thank Y. Yamazaki and T. Tawara who shared with me the workload and the pleasure of the diffractive dijets in DIS. I received a special guidance and an unvaluable human support from M. Arneodo, A. Bruni and M. Ruspa, I will always be grateful to them. Many thanks also to the coordinators of the diffractive group and the ZEUS management that followed the developing of my analysis: H. Abramowicz, K. Borras, J. Cole, E. Gallo, A. Levy and all the others.

I liked a lot to spend time with the italian group of ZEUS. I owe them at least a ton of coffee and I hope to meet them again in the future. Unfortunately these acknowledgments are becoming too long and I cannot mention all of them. However, a special thank goes to A. Montanari for all the times he hosted me in his office when I was homeless in building 1. The students at DESY that I met are too many to be mentioned one by one: I want to greet in particular R. Santamarta, J. Grebenyuk, J. Malka and N. D'Ascenzo.

Outside DESY, this Ph.D. was an extraordinary experience thanks also to my friends in Hamburg. I met so many nice people that it is impossible to salute them all. But a special mention must go to Patricia who shared the flat with me for three years denoting an incredible patience. I spent a lot of evenings and dinners chatting and drinking with Marco, his fellowship was extremely lovely and I thank him for that (and also for all the desserts that he offered to me). Thomas (aka Resi) and Carsten

191

are wonderful persons with whom I shared a lot of nights and beers in Sankt Pauli: they are two real good friends and I learned a lot from them. I trust that time and distance will not separate us.

Although far from Hamburg, my friends in Torino were always supporting me and remembering me: a big thanks to Marco, Paolo, Massimo, Fede, Stefano and Dario. It is nice to see that the kilometers did not decrease our friendship. I want to thank in particular Zelda who gave me a lot of support and suggestions during the years.

A very special kiss goes to Gemma, who is more beautiful than a lily. She contributed to make the last months unforgettable, I thank her a lot.

And at the end, last but not least, my family. My parents, Marisa and Massimo, and my brother Paolo were always close to me and ready to help with their unique warmness. I owe them a lot if I reached this goal. A final acknowledgment to my aunt Milena and my grandparents, Alfredo and Rosa, Nerino and Teresa. Some of them were not able to share with me the joy of this achievement but I keep them always in my heart.