Faculty of Mathematics and Physics
Charles University
Prague

Virtual photon structure from low Q^2 dijet production at HERA

Marek Taševský

Doctoral thesis
1999

Abstract

This thesis presents the study of manifestation of structure of the virtual photon in inclusive two-jet events in electron proton scattering. The analysis is based on the data corresponding to the luminosity $\mathcal{L} = 1$ pb^{-1} recorded by the H1 detector at *ep* collider HERA in 1995. The measurements were performed in the kinematic region $\sqrt{s_{ep}} \simeq 300\,\mathrm{GeV}$, $0.0 < Q^2 < 0.01\,\mathrm{GeV}^2$, $1.4 < Q^2 < 25\,\mathrm{GeV}^2$, $0.3 < y < 0.7$ and $E_T > 5\,\mathrm{GeV}$. The results were obtained using three approaches.

The virtuality dependence of the diparton and the inclusive dijet cross sections $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ was measured and compared to predictions of the leading order QCD event generators.

The virtuality dependence of the leading order effective parton distribution function of the virtual photon was extracted from the data and compared to existing parametrizations of the parton distribution functions of the photon.

The virtuality dependence of the inclusive dijet cross sections $\mathrm{d}\sigma/\mathrm{d}E_T$ and $\mathrm{d}\sigma/\mathrm{d}\eta$ was measured and compared to the next-to-leading order QCD calculations.

The data clearly show a suppression of the structure of the virtual photon with increasing virtuality. The leading order QCD models which do not introduce the virtual photon structure fail to describe the data. This conclusion was confirmed in the comparisons of the data with the next-to-leading order calculations. By comparing the data to the leading order QCD models and the next-to-leading order QCD calculations, both with the virtual photon structure included, a constraint to the parametrizations of the virtual photon structure is provided. In agreement with the direct comparisons in the framework of the effective parton distribution functions it is shown that the existing parametrizations of the virtual photon structure underestimate the data.

Contents

Chapter 1

Theoretical overview

1.1 Introduction

By means of collisions and decays of particles we learn about the composition of matter. The higher the energies of colliding particles, the deeper the structure we can see. For processes characterized by high momentum transfer (short distances) Quantum Chromodynamics (QCD) is applied with great success. It is, however, necessary to emphasize that in Quantum Field Theory (QFT) it is difficult to distinguish effects of the "structure" from those of "interactions". The fields in the lagrangian of the Standard Model (leptons, quarks and gauge bosons) represent fundamental particles, while the others are considered composite objects. For the latter concept of a parton distribution function (PDF) is used to describe their structure. This concept also appears successful for some fundamental particles, in particular the photon.

There are experiments making use of beams of real photons (e.g. NA-14 at CERN, E-687, E-691 at FERMILAB) but they are not suitable for the study of photon structure because of low energy and luminosity of the beams. However, one can exploit the fact that at electron-positron colliders (LEP, SLC, TRISTAN) and electron (positron)-proton collider at HERA the exchanged intermediate bosons between colliding particles are very often virtual photons. In Fig.1.1 a generalized picture of the interaction occurring at the above facilities is depicted.

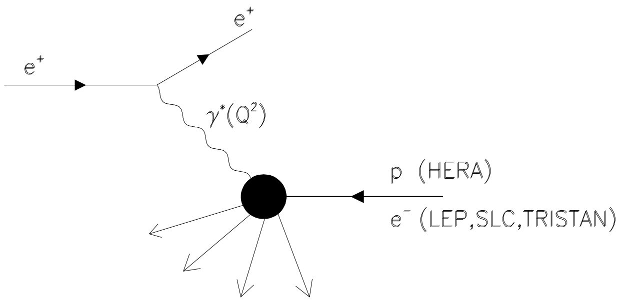


Figure 1.1: The general picture of the interaction occurring at colliders HERA, LEP, SLC and TRISTAN.

The spread of virtualities of exchanged photons is very large. At HERA it ranges from 10^{-8} to 10^4 GeV^2. Photons with virtualities smaller than about 10^{-2} GeV^2 are called *quasi-real* and interactions with them *photoproduction* (an average virtuality $\langle Q^2 \rangle \approx 10^{-3}$ GeV^2). In such events the structure of the photons is investigated using jets. On the contrary, photons of virtualities typically above few GeV^2 often serve as probes of the partonic content of the proton.

What has actually led to the introduction of the term "structure of the photon"? Classically (observed at large distances), photons do not interact with each other as they are neutral. In QFT, however, the photons can interact via their coupling to virtual charged fermion-antifermion pair, as shown in Fig.1.2. But the cross section in the lowest order of Quantum Electrodynamics (QED) is extremely small due to its proportionality to α^4 and for large centre-of-mass energies, S, it behaves like $1/S$.

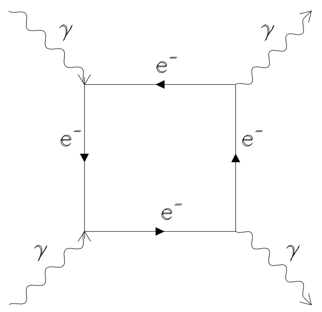
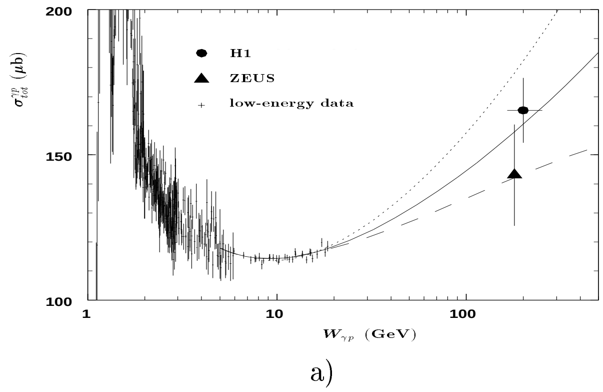


Figure 1.2: The interaction of two photons in the framework of QED.

And yet, there is experimental evidence that the photon in interaction with a hadron behaves in some respects like a hadron. One of the most important pieces of evidence concerns the energy dependence of the total cross section for γp interaction [1]. In Fig.1.3a we clearly observe that above roughly 10 GeV the cross section slowly rises with increasing mass of the hadronic system which is typical behaviour for hadron-hadron collisions. To explain this experimental observation the **VDM** (Vector Dominance Model [2]) was introduced in which the photon in soft processes (i.e. for low momentum transfers) behaves similarly to vector mesons with the same quantum numbers (e.g. ρ-meson) and $\sigma_{tot}(\gamma p)$ is proportional to $\sigma_{tot}(\rho p)$ (see Fig.1.3b).



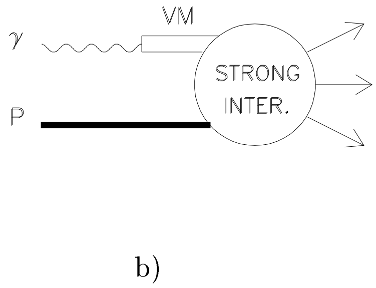


Figure 1.3: a) The total photoproduction cross section as function of the γp centre-of-mass energy $W_{\gamma p}$, b) the VDM diagram for $\gamma p \rightarrow$ hadrons.

Although in some aspects the photon behaves as a vector meson, it has, however, additional specific features which have become the subject of intensive theoretical and experimental investigation (for a review see [3]). The formalism of PDF, commonly used for hadrons, is also used to express the photon structure. Let us briefly recall the main historical facts leading to the introduction of the PDF of the proton.

1.2 Deep inelastic scattering and the proton structure

Inclusive cross section measurements have proven extremely fruitful for the understanding of the structure of matter. The relation of experimental results to theoretical concepts is most direct in the case of inclusive measurements where only the scattered lepton is measured and all possible hadronic final state configurations are summed over. In particular, experiments at SLAC carried out in fifties and sixties represent real milestones on the way to understanding the structure of the proton (see [4] for details).

The first milestone was made in the fifties in experiments where a beam of electrons was scattered elastically off various nuclei and protons. The results confirmed what was already well-known, namely that the nucleus is of finite size and composed of protons and neutrons. However, a non-trivial result was that the proton is not a point-like particle but has a finite size, of the order of 1 fm. This fact was deduced from the behaviour of the so called elastic formfactors.

The second milestone was made in the late sixties by systematic studies of properties of inelastic scattering of electrons on nucleons. The results that came from these experiments profoundly changed the picture of nucleon structure and directly led to the formulation of the Quark Parton Model (QPM) by Feynman, Bjorken and others.

1.2.1 Kinematics and basic notions

We will now briefly describe the basic notions of deep inelastic scattering (DIS) of charged (e^-, μ^-) as well as neutral (ν_e, ν_μ) leptons and their antiparticles on nucleons, in particular the proton:

$$l(k) + proton(P_p) \to l'(k') + X; l, l' = e, \mu, \nu_e, \nu_\mu \tag{1.1}$$

where X denotes any final state allowed by conservation laws and the letters in brackets stand for four-momenta of the corresponding particles. There are two types of DIS processes, called

- **neutral current** processes when $l' = l$. In these processes a (neutral) intermediate vector boson can be either the photon or Z boson, but in our kinematic range the latter contribution can be safely neglected.
- **charged current** processes when the electric charges of the initial and final leptons l, l' differ by one unit. In these processes the charged intermediate vector boson, W^+ or W^-, is exchanged.

In the following we will be interested only in electron proton scattering for which we have chosen this set of variables[1] (see Fig.1.4):

[1] m_e denotes lepton mass which can be often neglected, M_p denotes proton mass.

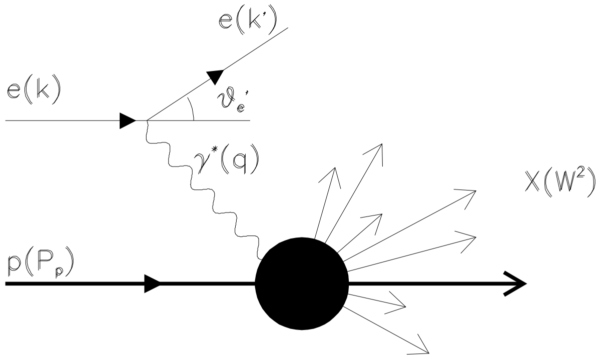


Figure 1.4: The deep inelastic scattering on the proton at HERA.

$$
\begin{aligned}
S &\equiv (k+P_p)^2 = M_p^2 + 2kP_p, && (1.2)\\
Q^2 &\equiv -q^2 \equiv -(k-k')^2 \doteq 2kk' = 4E_eE_e'\sin^2(\theta_e'/2), && (1.3)\\
y &\equiv \frac{qP_p}{kP_p} = \frac{E_{e,lab}-E'_{e,lab}}{E_{e,lab}} = \frac{\nu}{E_{e,lab}} = \frac{S-M_p^2}{S}\frac{1-\cos\theta_e'^*}{2}, && (1.4)\\
x &\equiv -\frac{q^2}{2P_pq} = -\frac{q^2}{2M_p\nu} = \frac{Q^2}{Q^2+W^2-M_p^2}, && (1.5)\\
W^2 &\equiv (q+P_p)^2 = \frac{Q^2(1-x)}{x} + M_p^2. && (1.6)
\end{aligned}
$$

There are two reference systems in which the process (1.1) is usually analyzed: the laboratory system, where the target proton is at rest and the centre-of-mass system (cms), where the colliding lepton and proton have equal but opposite momenta[2]. All the five quantities defined above are relativistic invariants. The variable S specifies the initial state of the colliding electron and proton, while the other three, x, y and q^2 together with S and the azimuthal angle ϕ_e' describe the final state electron in (1.1). The variable Q^2 defines the virtuality of the exchanged photon and W^2 is centre-of-mass energy of the hadronic system X produced by the absorption of the exchanged photon by the nucleon. The term "deep" means that *both* the invariants, Q^2 and P_pq, *are large* with respect to the proton mass, M_p

$$Q^2 \gg M_p^2; \quad \nu \gg M_p \tag{1.7}$$

but the ratio $x = Q^2/2P_pq$ in Eq.(1.5) stays finite. This defines the so called **Bjorken limit**. At HERA, where the beams are unpolarized and S fixed, the cross section of the process (1.1) depends on two variables only, for which one commonly takes (x, Q^2) or (x, y). In lowest order of QED the double differential cross section for DIS processes has the general form

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}y} = \frac{4\pi\alpha^2(2kP_p)}{Q^4}\left[\left(1-y-\frac{M_p^2xy}{S}\right)F_2(x,Q^2) + \frac{1}{2}y^2 2xF_1(x,Q^2)\right] \tag{1.8}$$

[2]The quantities calculated in cms are denoted with an asterisk.

The unknown functions $F_i(x, Q^2)$ are called **inelastic electromagnetic formfactors** or more often **structure functions** of the proton. The expression (1.8) follows from three fundamental properties of electromagnetic interactions: Lorentz invariance, gauge invariance and parity conservation. The famous result obtained at SLAC in the late sixties concerned the behaviour of the structure functions in the Bjorken limit (1.7). This result suggested that compared to elastic formfactors the structure functions are to a good approximation Q^2-*independent*! This phenomenon, called *structure function scaling*, actually led to the formulation of the **parton model**. It is based on three facts:

1. The scaling suggests that electrons scatter incoherently on essentially free, non-interacting charged parts of the proton, which Feynman named **partons**.
2. The variable x in Eq.(1.5) has a simple interpretation as a **fraction of the proton momentum carried by the struck parton**.
3. The **hadronization** stage of the scattering, i.e. the conversion of partons into observable hadrons occurs much later and does not influence the probability of the scattering of electrons on individual partons.

The fact that the quantity x spans the whole interval (0,1) was interpreted as an evidence that partons are distributed inside the proton with some non-trivial **distribution functions** $\mathbf{d_i(x)}$. The *inelastic* cross section (1.8) is then expressed in terms of the distribution functions of individual partons (neglecting target mass effects) as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}y} = \frac{4\pi\alpha^2}{Q^4} x(2kP_p) \sum_i e_i^2 d_i(x) \left[(1-y) + \frac{1}{2}y^2\right] \tag{1.9}$$

where[3] $\hat{s} = (k + \hat{p})^2 = (k + xP_p)^2 \doteq xS$ and $d_i(x)$ stand for the distribution function of the parton of flavour i with electric charge e_i (in units of the positron charge). Comparing Eq.(1.9) with the general form (1.8) we get:

- the structure functions in terms of the PDF:
$$F_2(x) = x \sum_i e_i^2 d_i(x). \tag{1.10}$$
- An important relation between the two, in principle independent, structure functions F_1, F_2 introduced in (1.8):
$$F_2(x) = 2xF_1(x). \tag{1.11}$$

This relation, confirmed experimentally to a great accuracy, provides evidence that charged partons carry spin 1/2. The measured structure functions F_1, F_2 are in fact, in contrast to the assumption in the QPM, also weakly dependent on Q^2. This phenomenon, called **scaling violation** can be seen in Fig.1.5 where one of the most recent measurements of Q^2-dependence of $F_2(x, Q^2)$ structure function in electromagnetic interactions is shown [5].

Scaling violations observed in data were explained in QCD, which slightly modifies the QPM, as described in Section 1.3.3. In QCD partons are identified with quarks and gluons and their

[3] The quantities with $\hat{}$ refer to the underlying parton level where parton will also mean the lepton and/or the photon.

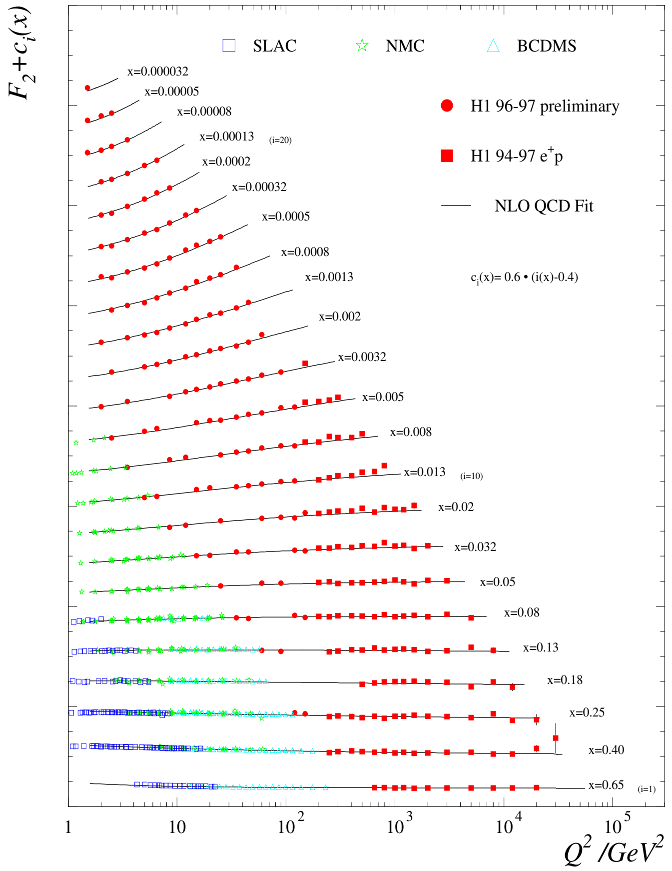


Figure 1.5: The Q^2-dependence of the proton structure function $F_2(x, Q^2)$ for several fixed values of x.

strong interactions are described by introducing an additional quantum number for partons, the colour charge. In QCD the "dressed" PDF of hadrons results from the resummation of multiple collinear parton emissions off the corresponding "bare" parton distributions. As a result of this resummation, the PDF acquires dependence on the *factorization scale* M. This scale dependence can be written in the form of coupled homogeneous Altarelli-Parisi evolution equations. They will be discussed later. In QCD the coupling of quarks and gluons is characterized by the renormalized colour coupling $\alpha_s(\mu)$, depending on the *renormalization scale* μ and satisfying the equation

$$\frac{\mathrm{d}\alpha_s(\mu)}{\mathrm{d}\ln\mu^2} \equiv \beta(\alpha_s(\mu)) = -\frac{\beta_0}{4\pi}\alpha_s^2(\mu) - \frac{\beta_1}{16\pi^2}\alpha_s^3(\mu) + \cdots, \tag{1.12}$$

where, in QCD with n_f massless quark flavours, the first two coefficients, $\beta_0 = 11 - 2n_f/3$ and $\beta_1 = 102 - 38n_f/3$, are unique, while all the higher order ones are ambiguous. The solution of Eq.(1.12) behaves like $1/\ln(\mu)$, i.e. increases at small scales (confinement) and vanishes at large scales (asymptotic freedom). This provides intuitive understanding of the fact that, on the one hand, the partons are bound in the hadron and, on the other hand, they behave like free particles when probed at short distances.

1.3 The photon structure

The experiments carried out in the last 15 years at PETRA, PEP and TRISTAN [6] and recently also at HERA and LEP [7], have convincingly shown that due to strong interactions between quarks the photon has a nontrivial structure.

1.3.1 Deep inelastic scattering on the photon

In close analogy to measurements of the proton structure at HERA, deep inelastic scattering on virtual photons, accessible at LEP, serves for investigating the structure of the photon (see Fig.1.6a and the upper rightmost diagram in Fig.1.7 where the proton is replaced by the photon). The DIS on the photon is characterized by the same set of variables as introduced in (1.2)–(1.6), replacing only

$$P_p \to P, \qquad M_P^2 \to P^2 = -p^2 \tag{1.13}$$

where P stands for the four-momentum of the target photon.

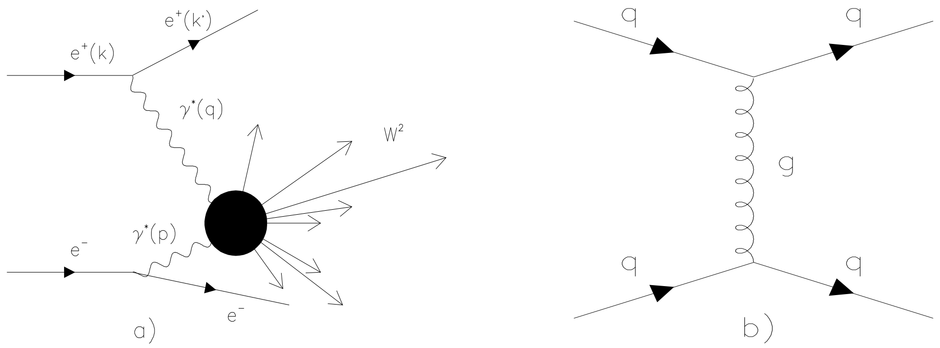


Figure 1.6: a) DIS on the photon at LEP, b) one of hard subprocesses of jet production in hadron-hadron collisions.

The formalism of PDF can be directly applied to the case of the real photon, but it makes sense to introduce also distribution functions of leptons in the photon which are fully calculable in QED. As shown in [8] the predictions of QED are in a good agreement with experimental data using e^+e^- or $\mu^+\mu^-$ final state.

In most experiments the partonic content of the photon, i.e. its PDF, can be measured via F_2^γ since due to a limited acceptance of an electron tagger y in (1.8) is small and consequently the structure function F_1^γ can be neglected.

1.3.2 Jet production

So far we have been considering only DIS on either the proton or the photon. In inclusive DIS processes there are rare cases where the jets are formed. It is worth recalling the definition of the hard scale.

Hard subprocesses in QCD

The hard subprocesses are characterized by hard scale, in other words by short distance or large momentum transfer. The information about the hard scale, given by kinematics of the process, helps in selecting the factorization as well as the renormalization scale which are parameters of theory. Very often both are put equal and proportional to the hard scale. For instance, in the measurement of proton structure function $F_2^p(x, Q^2)$, the hard scale is identified with Q^2 and the corresponding lowest order Feynman diagram is shown in the upper rightmost picture in Fig.1.7.

Another type of the hard subprocess occurs in jet production in hadron-hadron collisions. The jets come from partons undergoing the hard scattering. The hard scale in this case is taken to be the momentum transfer of this basic parton level subprocess. This scale is correlated with the transverse energy of the final state jets, E_T. An example of the lowest order Feynman diagram is shown in Fig.1.6b.

Jets can also be produced in *ep* collisions. We have two options for choice of hard scale, namely the virtuality Q^2 and the transverse momentum squared of jets, E_T^2. From the point of view of the relation between them we distinguish two kinematic regions.

1. $\mathbf{E_T^2 \gg Q^2}$: If Q^2 is close to zero we speak of photoproduction events (diagrams in the leftmost column in Fig.1.7) that are characterized by the electron being deflected by a very small angle from its original direction. If electron scatters through a larger angle with respect to its original direction compared to photoproduction, Q^2 is larger than few GeV^2. If, however, E_T^2 is much larger than the virtuality (diagrams in the middle column of Fig.1.7) the dynamics of these events is basically the same as in photoproduction. In both cases the scale is taken to be E_T^2. We distinguish two types of processes contributing to the cross section for jet production. In **direct photon** processes the photon enters hard scattering with the parton from the proton as a single object. When the virtuality of the photon is small compared to E_T^2, the photon has a time to develop its structure. Therefore in a fraction of cases the photon acts as a composite object whose structure is "resolved" by the parton from the proton. These cases are called **resolved photon** processes which differ from the previous ones by the presence of a so-called *photon remnant* that can be observed in the form of a broad jet [9].

2. $\mathbf{E_T^2 < Q^2}$: In this kinematic region, the photon lives too short a time to reveal its structure but can be used for extraction of quark or gluon distribution functions of the proton. For example, its gluon content (see lower rightmost diagram in Fig.1.7) was measured in [10]. Although here we have two scales, E_T^2 appears as the right choice.

It is worth stressing that except for photoproduction the processes from both kinematic regions are included in the measurement of the proton structure function. Considering the proton, photon and partons massless in *ep* dijet production, we can define the fraction of the photon and proton four-momentum carried by the parton involved in binary hard scattering:

$$x_\gamma \equiv \frac{\sum_{i=1}^{2} E_{T,i}\, e^{-\eta_i}}{2E_\gamma}, \quad x_p \equiv \frac{\sum_{i=1}^{2} E_{T,i}\, e^{\eta_i}}{2E_p} \tag{1.14}$$

or equivalently

$$x_\gamma \equiv \frac{\sum_{i=1}^{2} (E - p_z)_i}{\sum_h (E_h - p_{z,h})}, \quad x_p \equiv \frac{\sum_{i=1}^{2} (E + p_z)_i}{\sum_h (E_h + p_{z,h})}, \tag{1.15}$$

where E_i, $p_{z,i}$, $E_{T,i}$ and η_i is the energy, longitudinal momentum, transverse energy and pseudorapidity of the jet respectively. E_h and $p_{z,h}$ is the energy and longitudinal momentum of a hadron, E_p and E_γ is the energy of the proton and photon in the $\gamma^* p$ cms. Using the polar scattering angle θ the pseudorapidity is defined as

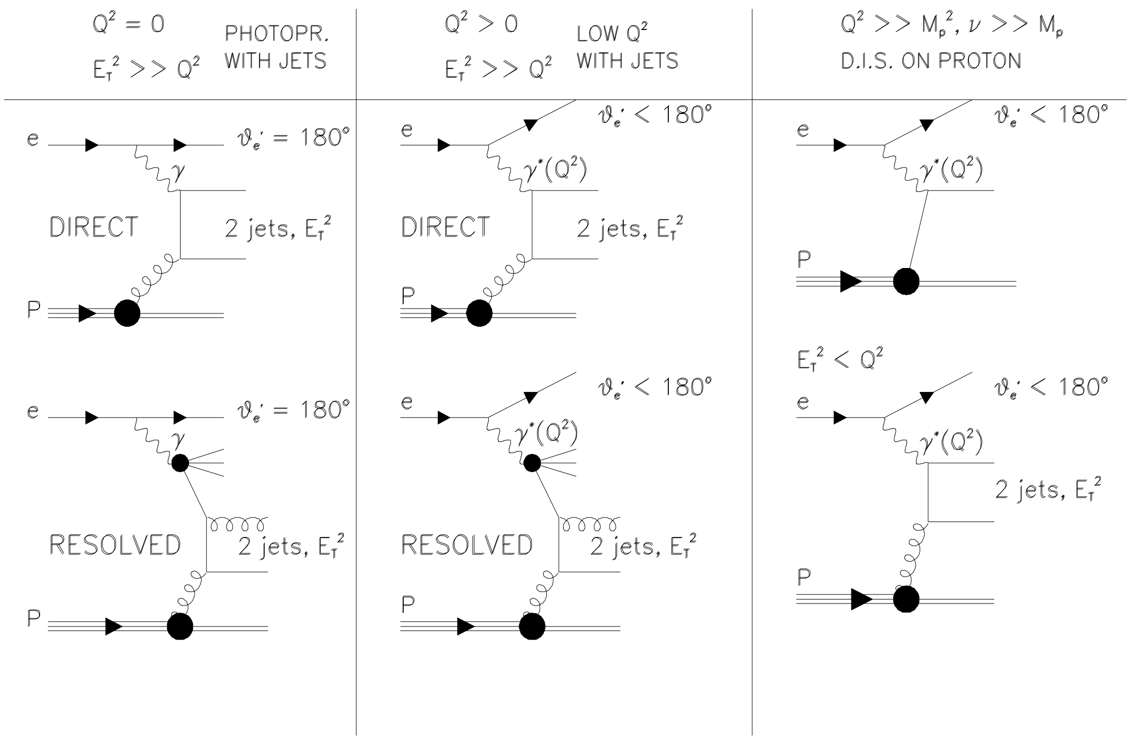


Figure 1.7: The different classes of events at HERA according to Q^2 and the its relation to the E_T^2. The angle of the scattered electron, θ_e', is related to the z-axis whose positive direction goes in the direction of the incident proton.

$$\eta \equiv -\ln(\tan\frac{\theta}{2}). \tag{1.16}$$

Note that under the above conditions all the formulae for the fraction of the four-momentum of the incident particle are equivalent. Also, the formulae for the matrix elements of the hard scatterings are simpler for the massless partons. If we take into account the masses of all the above particles, it will be useful to recall what follows from the constraints on the masses. If M_p and Q^2 are nonzero but negligible with regard to the total energy of the γ^*p cms then the two expressions for x_γ (x_p) in Eq.(1.14) and (1.15) are no longer identical but still very close to each other. Moreover, neither of them can be interpreted as the fraction of the four-momentum of the identical particle. Furthermore, if M_p and Q^2 are negligible with regard to the hard scale (which is generally smaller than the energy of the γ^*p cms) then expression for cross section can be approximated by the product of the PDF's and cross section of the hard scattering.

Note also that if the factorization scale M (defined as the maximum virtuality of the parton undergoing the parton emissions before entering the hard scattering) is small compared to the hard scale the concept of the PDF is justified. The latter condition restricts also the photon virtuality so the power corrections ($\sim Q^2/E_T^2$) can be neglected. In common practice, however, M is identified with the hard scale. This produces an additional contribution to the PDF with smaller M but this contribution is small due to $1/\tau$-dependence of the integrand. The smaller the Q^2, the smaller the relative contribution.

Events from the first two columns are the subject of the analysis presented in this thesis. Note that such processes, in contrast to DIS on the photon at LEP, involve cases with an incoming gluon from the photon even in LO approximation. However, as we are experimentally not able to distinguish individual subprocesses we can only measure an effective (i.e. some combination of quarks and gluons) parton distribution function of the photon. The gluon content is determined by subtracting the quark content (which is measured at LEP) from the total PDF.

1.3.3 PDF of the real photon

The factorization scale dependence of PDF of the photon is determined by the system of coupled inhomogeneous evolution equations

$$
\begin{aligned}
\frac{\mathrm{d}\Sigma(M^2)}{\mathrm{d}\ln M^2} &= \delta_\Sigma k_q + P_{qq}\otimes\Sigma + P_{qG}\otimes G, && (1.17)\\
\frac{\mathrm{d}G(M^2)}{\mathrm{d}\ln M^2} &= k_G + P_{Gq}\otimes\Sigma + P_{GG}\otimes G, && (1.18)\\
\frac{\mathrm{d}q_{\mathrm{NS}}(M^2)}{\mathrm{d}\ln M^2} &= \delta_{\mathrm{NS}} k_q + P_{\mathrm{NS}}\otimes q_{\mathrm{NS}}, && (1.19)
\end{aligned}
$$

where the convolution $\otimes$ is defined as

$$
(P\otimes q)(x) \equiv \int_x^1 \frac{\mathrm{d}y}{y} P(x/y) q(y) \tag{1.20}
$$

and the singlet and non-singlet quark distribution functions Σ and q_{NS} are given as

$$
\begin{aligned}
\Sigma(x,M) &\equiv \sum_{i=1}^{n_f} \left[q_i(x,M) + \overline{q}_i(x,M)\right] \equiv \sum_{i=1}^{n_f} q_i^+(x,M), && (1.21)\\
q_{\mathrm{NS}}(x,M) &\equiv \sum_{i=1}^{n_f} \left(e_i^2 - \langle e^2\rangle\right)\left(q_i(x,M) + \overline{q}_i(x,M)\right), && (1.22)
\end{aligned}
$$

$$
\delta_{\mathrm{NS}} \equiv 6n_f(\langle e^4\rangle - \langle e^2\rangle^2), \quad \delta_\Sigma \equiv 6n_f\langle e^2\rangle, \quad \langle e^n\rangle \equiv (\sum_{i=1}^{nf} e_i^n)/n_f \tag{1.23}
$$

Compared to the system of evolution equations, usually defined for quark, antiquark and gluon distribution functions, this notation has the advantage that the last equation is not coupled to the previous two. The main difference between the evolution equations for PDF of hadrons and those of the photon is the presence of the so called inhomogeneous splitting functions k_q and k_G which lead to a scale dependence characteristic of the photons and differing from that for hadrons.

To the lowest order in α the splitting functions $P_{ij}(x,M)$ and $k_i(x,M)$ are given as perturbation expansions in $\alpha_s(M)$:

$$
\begin{aligned}
k_q(x,M) &= \frac{\alpha}{2\pi}\left[k_q^{(0)}(x) + \frac{\alpha_s(M)}{2\pi}k_q^{(1)}(x) + \left(\frac{\alpha_s(M)}{2\pi}\right)^2 k_q^{(2)}(x) + \cdots\right], && (1.24)\\
k_G(x,M) &= \frac{\alpha}{2\pi}\left[\qquad\qquad \frac{\alpha_s(M)}{2\pi}k_G^{(1)}(x) + \left(\frac{\alpha_s(M)}{2\pi}\right)^2 k_G^{(2)}(x) + \cdots\right], && (1.25)\\
P_{ij}(x,M) &= \frac{\alpha_s(M)}{2\pi}P_{ij}^{(0)}(x) + \left(\frac{\alpha_s(M)}{2\pi}\right)^2 P_{ij}^{(1)}(x) + \cdots, && (1.26)
\end{aligned}
$$

where the leading order splitting functions $k_q^{(0)} = (x^2 + (1-x)^2)$ and $P_{ij}^{(0)}$ are *unique*[4], but the higher order splitting functions $k_q^{(j)}, k_G^{(j)}, P_{kl}^{(j)}, j \geq 1$ depend on the *factorization scheme* (FS)[5]. The general solution of the evolution equations (1.17) - (1.19) can be written as the sum of the general solution of the corresponding homogeneous equation, called the **hadronic** (HAD)[6] part and a particular solution of the full inhomogeneous equation. A subset of solutions of the full evolution equations resulting from the resummation of series of diagrams in Fig.1.8 and characterized by the value of the scale M_0 at which they vanish defines the so called **pointlike** (PL) parts of quark and gluon distribution functions. We can thus write quite generally ($D = q, \overline{q}, G$)

$$D(x, M^2) = D^{\mathrm{PL}}(x, M^2) + D^{\mathrm{HAD}}(x, M^2). \tag{1.27}$$

However, as there is an infinite number of pointlike solutions $q^{\mathrm{PL}}(x, M^2)$, which differ by terms satisfying the homogeneous evolution equation, the separation of quark and gluon distribution functions into their pointlike and hadronic parts is ambiguous and consequently these concepts have separately no unique physical meaning. To see the most important feature of the pointlike parts of quark and gluon distribution functions, we now consider in more detail the case of the non-singlet quark distribution function $q_{\mathrm{NS}}(x, M^2)$ which is explicitly defined via the series

$$q_{\mathrm{NS}}^{\mathrm{PL}}(x, M_0^2, M^2) \equiv \frac{\alpha}{2\pi} k_{\mathrm{NS}}^{(0)}(x) \int_{M_0^2}^{M^2} \frac{\mathrm{d}\tau}{\tau} + \int_x^1 \frac{\mathrm{d}y}{y} P_{qq}^{(0)}\left(\frac{x}{y}\right) \int_{M_0^2}^{M^2} \frac{\mathrm{d}\tau_1}{\tau_1} \frac{\alpha_s(\tau_1)}{2\pi} \frac{\alpha}{2\pi} k_{\mathrm{NS}}^{(0)}(y) \int_{M_0^2}^{\tau_1} \frac{\mathrm{d}\tau_2}{\tau_2} +$$

$$\int_x^1 \frac{\mathrm{d}y}{y} P_{qq}^{(0)}\left(\frac{x}{y}\right) \int_y^1 \frac{\mathrm{d}w}{w} P_{qq}^{(0)}\left(\frac{y}{w}\right) \int_{M_0^2}^{M^2} \frac{\mathrm{d}\tau_1}{\tau_1} \frac{\alpha_s(\tau_1)}{2\pi} \int_{M_0^2}^{\tau_1} \frac{\mathrm{d}\tau_2}{\tau_2} \frac{\alpha_s(\tau_2)}{2\pi} \frac{\alpha}{2\pi} k_{\mathrm{NS}}^{(0)}(w) \int_{M_0^2}^{\tau_2} \frac{\mathrm{d}\tau_3}{\tau_3} + \cdots, \tag{1.28}$$

where $k_{\mathrm{NS}}^{(0)}(x) = \delta_{\mathrm{NS}} k_q^{(0)}(x)$. In the leading-logarithmic approximation and in terms of moments such solutions are given explicitly as

$$q_{\mathrm{NS}}^{\mathrm{PL}}(n, M_0^2, M^2) = \frac{4\pi}{\alpha_s(M)} \left[1 - \left(\frac{\alpha_s(M)}{\alpha_s(M_0)} \right)^{1 - 2P_{qq}^{(0)}(n)/\beta_0} \right] a_{\mathrm{NS}}(n), \tag{1.29}$$

where

$$a_{\mathrm{NS}}(n) \equiv \frac{\alpha}{2\pi\beta_0} \frac{k_{\mathrm{NS}}^{(0)}(n)}{1 - 2P_{qq}^{(0)}(n)/\beta_0}. \tag{1.30}$$

and results from the resummation of the infinite series of diagrams in the upper part of Fig.1.8. The first term in Eq.(1.28), the photon splitting, corresponds to the first diagram in the upper part of Fig.1.8. As shown in [11, 12] the whole resummation substantially softens the x-dependence of $q_{\mathrm{NS}}^{\mathrm{PL}}(x, M^2)$ compared to the photon splitting alone, which is an important fact for comparisons of the data with next-to-leading order (NLO) calculations. The first term in Eq.(1.28) is proportional to $\ln(M^2/M_0^2)$ which dominates the scale dependence of the whole

[4]The **photon splitting function $\mathbf{k_q^{(0)}}$** determines the probability of finding the quark in the photon (or the quark in the gluon). It is derived from the same principles as the probability of finding the photon inside the electron, the so called *photon flux*.

[5]We can turn this statement around and consider any factorization scheme FS to be specified by the corresponding set of functions $k_j^{(i)}, P_{jk}^{(i)}, j, k = q, G; i \geq 1$.

[6]Sometimes also called "VDM part" because it is usually modelled by PDF of vector mesons.

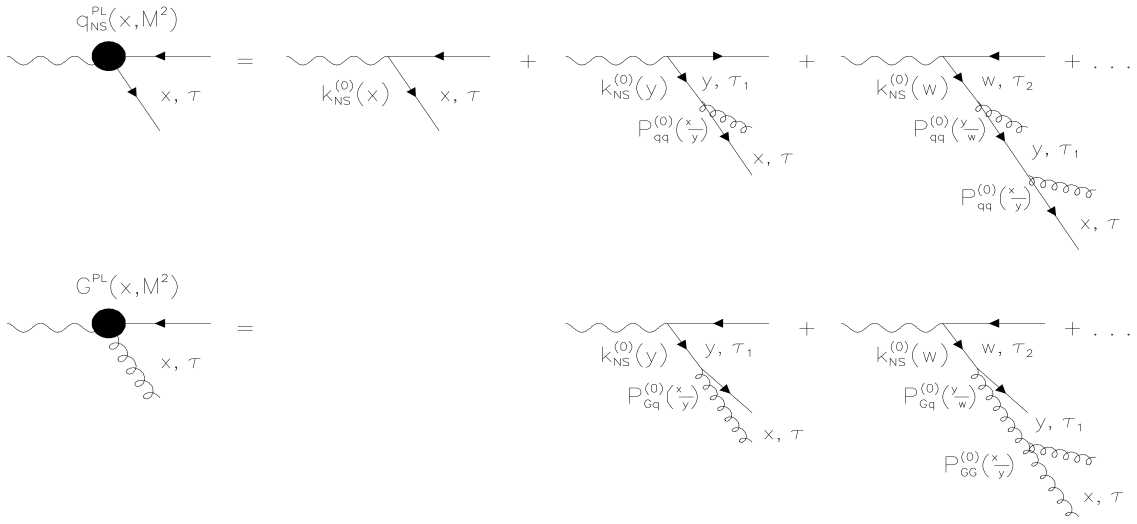


Figure 1.8: Diagrams defining the pointlike parts of non-singlet quark and gluon distribution functions.

pointlike solution. It has nothing to do with QCD and results exclusively from the basic QED $\gamma \to q\overline{q}$ splitting. After taking the derivative of $q^{\rm PL}_{\rm NS}(x,M^2)$ with respect to $\ln M^2$ we arrive at equation (1.19), the first term in Eq.(1.28) being responsible for the inhomogeneous term in Eq.(1.19). For $M/M_0 \to \infty$ the second term in brackets of (1.29) vanishes and therefore all the pointlike solutions share the same large M behaviour

$$q^{\rm PL}_{\rm NS}(x,M_0^2,M^2) \to \frac{4\pi}{\alpha_s(M)}a_{\rm NS}(x) \equiv q^{\rm AP}_{\rm NS}(x,M^2) \propto \ln\frac{M^2}{\Lambda^2}, \tag{1.31}$$

defining the so called *asymptotic pointlike* solution $q^{\rm AP}_{\rm NS}(x,M^2)$, originally derived by Witten [13]. Hence, thanks to the inhomogeneous term in the evolution equations the absolute value of the asymptotic pointlike solution is known without having to give the initial conditions.

The fact that $\alpha_s(M)$ appears in the denominator of (1.31) has been the source of claims (see, for instance, [14]) that $q(x,M) = \mathcal{O}(\alpha/\alpha_s)$. A closer inspection of (1.29) however, shows [15, 16] that this claim is misleading and $q(x,M^2) = \mathcal{O}(\alpha)$.

The arbitrariness in the choice of M_0 reflects the fact that as M_0 increases, fewer of the gluon radiation effects are included in the resummation (1.29) defining the pointlike part of quark distribution function $q^{\rm PL}$ and more included in the hadronic part. As we shall see in Chapter 6 (see also [11, 12]), the hadronic and pointlike parts have very different behaviour as functions of x and M.

The F_2^γ is calculated according to

$$\frac{1}{x}F_2^\gamma(x,Q^2) = q_{\rm NS}(M)\otimes C_q(Q/M) + \frac{\alpha}{2\pi}\delta_{\rm NS}C_\gamma +$$
$$\langle e^2\rangle\Sigma(M)\otimes C_q(Q/M) + \frac{\alpha}{2\pi}\langle e^2\rangle\delta_\Sigma C_\gamma + \langle e^2\rangle\frac{\alpha_s}{2\pi}G(M)\otimes C_G(Q/M) \tag{1.32}$$

where the coefficient functions C_q, C_G and C_γ can be expanded in powers of $\alpha_s(\mu)$. The renormalization scale μ is in principle independent of the factorization scale M. The coefficient functions C_q, C_G and C_γ are actually independent of μ because the μ-dependence of $\alpha_s(\mu)$ is cancelled by explicit dependence of $C_q^{(i)}$, $C_G^{(i)}$ and $C_\gamma^{(i)}, i \geq 2$ on μ [17]. On the other hand, both the PDF and the coefficient functions depend on the factorization scale M, but in such a correlated manner that physical quantities, like F_2^γ, are independent of M, provided expansions (1.24)–(1.26) are taken to all orders in $\alpha_s(M)$ and $\alpha_s(\mu)$.

Although the concept of PDF for the structure of the photon has been developed in analogy with that for the proton structure there are differences: while $F_2^\gamma(x, Q^2)$ rises with Q^2 for all x (see Fig.1.9, [3]), the Q^2-dependence of $F_2^p(x, Q^2)$ depends on x (see Fig.1.5). Another difference is that F_2^γ is large for high x, while F_2^p decreases at large x. Both these facts are due to the photon splitting term in the evolution equations for the photon which is absent in those for the proton.

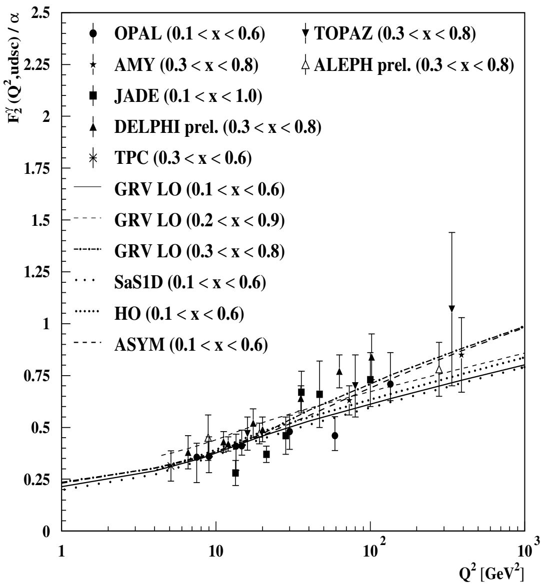


Figure 1.9: The photon structure function F_2^γ/α as a function of Q^2. Taken from [3].

1.3.4 PDF of the virtual photon

In the collider experiments [6, 7] the incoming electrons or positrons acted as sources of virtual photons, which were subsequently probed in hard collisions with partons from the other beam particle. Most of the previous studies concerned the region of very small photon virtualities Q^2, where the scattered electron goes close to the beam line. In HERA conditions this typically meant $\langle Q^2 \rangle \lesssim 1$ GeV^2. The transition between photoproduction and deep inelastic scattering has recently attracted considerable theoretical attention [18, 19, 20]. The point is that as its virtuality increases, the photon has less and less time to develop those aspects of its hadronic structure that depend on large distances and its interactions may therefore be easier calculable by perturbative means. This is expected to hold not only for hard processes like the jet production but also for total cross-sections [20]. The way the photon structure varies with its virtuality promises to hold important information on the interplay between perturbative and non-perturbative aspects of QCD.

Returning to solutions of evolution equations, it is a fact that one does not have to look for the general solution of evolution equations in the form described in Eq.(1.27). For example, GRV parametrizations [21] of the photon PDF cannot be divided to these two parts. On the other hand, Schuler and Sjöstrand (henceforth denoted by "SaS") [22, 23] offers parametrizations in which the contributions corresponding to VDM and pointlike parts can be obtained separately. A naive picture of the interaction of the photon with the parton from the proton could look like that in Fig.1.10.

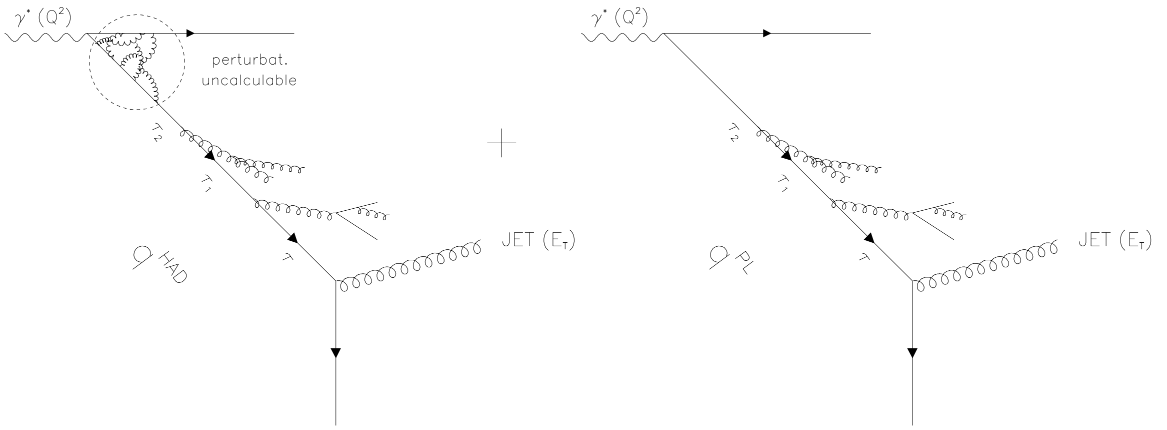


Figure 1.10: Two contributions corresponding to hadronic and pointlike parts of the quark PDF of the virtual photon interacting with the parton from the proton.

Fig.1.10 shows a building up of the quark distribution function of the photon before entering the hard subprocess with another quark from the proton. Two parts of the quark distribution function, as defined in Eq.(1.27), are shown: the hadronic one (left) and the pointlike one (right). The incoming quark of virtuality τ_{i+1} emits a quark with virtuality τ_i and a real gluon. In the collinear kinematics, i.e. for strongly ordered virtualities τ, τ_1, τ_2 in Fig.1.10 we have

$$\frac{m_q^2}{1-x} + xQ^2 \ll \cdots \ll |\tau_2| \ll |\tau_1| \ll |\tau| \leq M^2. \tag{1.33}$$

Then we should specify the scale μ to be used as an argument in $\alpha_s(\mu)$ in the vertices along the ladder. This task is not trivial and its solution is based on the incorporation of radiative corrections (see [24]). The "best" scale μ^2 should be identified with the highest virtuality of all partons interacting in this vertex. This implies that the scales of α_s are different at different vertices along the ladder and increase as we move towards the hard subprocess. These two features, the strong ordering and choice of the scale in the vertices, are crucial for derivation of the dressed PDF of the hadron or the pointlike part of the photon presented in Eq.(1.28).

By emitting gluons, the virtuality of incoming quark increases towards the hard scale which is determined by the E_T of the produced (gluon) jet. At the beginning small virtual corrections occur in the hadronic part arising from gluon loops connecting the quark with remnant antiquark from the photon. For small virtualities $\alpha_s(\tau)$ is large confining the emitted gluons in a bound state which cannot be calculated perturbatively and is modelled by VDM ansatz. It vanishes rapidly with increasing Q^2 (as $m_\rho^2/(m_\rho^2+Q^2)^2$, see [22]). At higher virtualities perturbative QCD can be applied. The gluon emission starts almost parallel to the original quark leg and successively the gluons are emitted with increasing opening angle. The emissions are still absorbed in the photon PDF. Once the opening angle is so large that the emission can no longer be considered parallel, the quark enters the hard interaction since its virtuality reached the hard scale.

The pointlike part, in contrast to the hadronic one, is calculable in the collinear kinematics using the resummation procedure described above. It does not contain non-perturbative effects as does the hadronic part and the photon splitting is calculable from QED.

In the following a short review of existing parametrizations for PDF of the virtual photon is given.

- Drees and Godbole [25] introduced a simple suppression factor

$$L(Q^2, M^2, \omega^2) \equiv \frac{\ln((M^2+\omega^2)/(Q^2+\omega^2))}{\ln((M^2+\omega^2)/\omega^2)}, \tag{1.34}$$

  which multiplies all quark distribution functions of the real photon, while those of the gluon are multiplied by L^2. In the simplest interpretation of this formula free parameter ω governing the virtuality dependence, is assumed to be independent of x. Consequently, the Q^2-dependence is also the same for all x. This simple approach can be applied to any existing parametrization of the real photon and the parameter ω can be tuned to the data.

- Recently, SaS constructed four LO parametrizations starting from general arguments relying on dispersion relation in Q^2 [26]. They differ in choice of initial scale and the way they are extracted from data on F_2^γ. All five quark flavours are taken into account. The SaS parametrizations predict a specific form for the x-dependence of the virtuality suppression which is clearly distinguishable from that of (1.34). The demonstration of this difference seen in effective PDF is studied in Chapter 5. The PDF for the virtual photon in [22] was found as follows. For the real photon evolution equations were used with some initial conditions given by experimental measurements. The value of the PDF of the real photon at the required scale then served as input into the dispersion relations for the PDF written as a function of Q^2. The resulting PDF at $Q^2 > 0$ differ from the solutions found using the evolution equations by terms of the power corrections type which increase as the virtuality approaches the hard scale from below.

- In another approach Glück, Reya and Stratman [27] introduced the virtuality dependence by specifying the initial conditions imposed on the solutions of the inhomogeneous evolution equations for PDF of the photon at $Q^2 = M^2$. Only the light quarks u, d, s and the gluon are included as intrinsic partons. Consequently the GRS effective PDF of the photon are not directly comparable with those of SaS. Moreover as the standard MC event generators, including HERWIG and RAPGAP used in our analysis, do not contain the $2 \to 3$ matrix elements required for the consistency of the GRS approach, they are not compared to our results.

Existing measurements of the virtual photon structure

The available experimental data on the structure of the virtual photon are still scarce. The first ever measurement of $F_2^\gamma(x, M^2, Q^2)$ was made by the PLUTO experiment [6] in the middle of the eighties for $Q^2 \leq 0.8\,\mathrm{GeV}^2$ and $M^2 = 5\,\mathrm{GeV}^2$. At ZEUS [28] the Q^2-dependence of the ratio $\sigma^{\mathrm{res}}/\sigma^{\mathrm{dir}}$ was measured in the region $0.1 \leq Q^2 \leq 0.7\,\mathrm{GeV}^2$. This result indicates that already for quite small virtualities of the photon the resolved contribution drops relative to the direct one by a factor of about 2.

There are three recent analyses by the H1 collaboration bringing results on the virtual photon structure. In [29] the jet production in the region of virtualities $Q^2 \leq 49\,\mathrm{GeV}^2$ has been investigated using the inclusive single jet sample. This analysis clearly demonstrated that models in which the photon only couples directly to the partons of the hard scattering process fail to describe the data. On the other hand models which include the resolved component of the photon are in good agreement with the data. In [30] the NLO calculations have been compared to the data for Q^2-dependence of dijet rates in the region $5 \leq Q^2 \leq 100$ GeV^2. One concludes that the data require the resolved photon contribution up to $Q^2 \lesssim 20$ GeV^2. In [31] an effective PDF of the virtual photon has been extracted for the first time as a function of x_γ, E_T^2 and Q^2. The effective PDF was measured to be approximately independent of E_T^2 and suppressed with increasing Q^2.

1.3.5 Transverse vs. longitudinal virtual photons

In the Equivalent Photon Approximation (EPA) the incoming leptons act as sources of fluxes of transverse and longitudinal photons, given as [32]

$$f^T_{\gamma/e}(y, Q^2) = \frac{\alpha}{2\pi}\left[\frac{1+(1-y)^2}{y}\frac{1}{Q^2} - \frac{2m_e^2 y}{Q^4}\right] \tag{1.35}$$

$$f^L_{\gamma/e}(y, Q^2) = \frac{\alpha}{2\pi}\frac{2(1-y)}{y}\frac{1}{Q^2} \tag{1.36}$$

The transverse and longitudinal fluxes thus coincide for $y = 0$, while at $y = 1$, f_L^γ vanishes. Note that while for $Q^2 \gg m_e^2$ the second term in (1.35) is negligible with respect to the leading $1/Q^2$ one, close to $Q^2_{\mathrm{min}} = m_e^2 y^2/(1-y)$, their ratio is finite and approaches $2(1-y)/(1+(1-y)^2)$.

Similar to the transverse virtual photon, the photon splitting term in the quark distribution function of the longitudinal photon can be calculated within QED. The analysis of the vertex $\gamma(Q^2) \to q\overline{q}$ in collinear kinematics yields [16]:

$$q_L^{\gamma^*}(x) = 6\frac{\alpha}{\pi}e_q^2 x(1-x). \tag{1.37}$$

Neglecting longitudinal photons is a good approximation for large $y \to 1$, where the flux vanishes, as well as for very small virtualities. The estimates made in [11, 12] using SaS1D for the transverse photon PDF (as they do not exist for the longitudinal photons) suggest that the contribution of γ_L^* to the effective PDF is not negligible for $Q^2 \gtrsim 1\,\mathrm{GeV}^2$.

In this analysis the contribution of γ_L^* is assumed to be small, however, the question of its importance in the jet production is certainly an exciting task which can be solved in the near future with higher statistics of the data at HERA.

Chapter 2

Monte Carlo event generators

Software programs that model physics processes are called *event generators*. They are based on Monte Carlo (MC) technique and are used mainly for two purposes: to compare theoretical expectations with the data and to correct the data for detector effects. In MC generators we distinguish three levels:

1. **Parton level**. The hard scattering binary process based on leading-order (LO) matrix element is generated. The properties of outgoing partons are established. QCD initial and final state parton radiations are switched on[1].

2. **Hadron level**. The hadronization mechanism as well as the model for additional transverse activity[2] is applied on all created partons.

3. **Detector level**. The detector response is simulated [33] using GEANT package [34] on all produced particles, some hadrons decay. The information at this moment is of the same form as that from the detector during data taking. In both cases, such "raw data" then pass through the same chain of reconstruction programs. By this procedure the geometrical acceptance and finite resolution for various quantities is simulated.

Using MC event samples we are able to determine the correlations between quantities on different levels which enables us to correct the data to a required level.

In the time when this analysis started only two Monte Carlo event generators involving virtual photons in the direct and resolved photon processes were available: HERWIG [35] and RAPGAP [36]. Both of them (HERWIG, version 5.9 and RAPGAP, version 2.6) have been used for correction of the data. Currently also PYTHIA [37] from version 6.0 and PHOJET [38] from version 1.06 include the virtual photons in both channels.

HERWIG is a general purpose MC event generator, applicable to a wide range of hard processes in all conceivable collisions. It has a sophisticated treatment of parton showers, based on the idea of colour coherence, but only rudimentary capabilities to simulate soft processes that do not accompany some basic hard scattering process. Hadronization is done via decay of colourless clusters, formed during hard scattering and parton shower stages.

[1] In the case of extraction of the photon PDF, making sense only in LO approximation, the data are corrected to the level of LO hard parton scattering only, thus without initial and final state radiations.

[2] In PYTHIA the model for the additional transverse activity is called *multiple parton interactions* (MI) and is connected with the parton level, though it affects the jets from hadrons.

RAPGAP, formulated originally for the simulation of diffractive processes, has recently been extended to several other processes in ep collisions, including the jet production. It combines standard hard scattering and parton showers with LUND string fragmentation using JETSET. As option, it offers colour dipole model as alternative model for parton cascading.

For the purpose of analyzing jets in our kinematic conditions and comparing their properties with different event generators, it is useful to keep in mind how they treat

Photon fluxes. In resolved photon processes, both MC generators use Equivalent Photon Approximation (EPA) in which the incoming lepton acts as a source of fluxes of transverse and longitudinal virtual photons. The expressions for fluxes were given in (1.35)–(1.36) in Chapter 1.

In versions used in this analysis neither of the models included the contribution of longitudinal photons. The applicability of EPA, which is crucial for any investigation of virtual photon structure, depends on the process and kinematic region studied. It can be tested only in the direct channel, where its results can be compared with those obtained from the exact matrix elements (see below), and for the transverse photons only. We studied the accuracy of EPA in two regions of y, namely $0.2 < y < 0.8$ (corresponding roughly to our actual cuts) and $0.2 \leq y$. We observed that the accuracy depends sensitively on y and deteriorates rapidly at low y but in the kinematic region covered by this analysis the EPA is a very good approximation of exact matrix elements. We expect the meaningfulness of EPA (with transverse photons only) in the resolved channel to decrease with decreasing x_γ, reflecting a similar dependence on y in the direct channel. Nevertheless it gives us some confidence to use it in the whole region of virtualities studied in this thesis. Moreover, we emphasize that a complete analysis of the structure of virtual photon requires the inclusion of longitudinal photons [11, 12, 39], principally because the Q^2-dependent part of the contribution of the transverse photon is of the same order as the contribution of the longitudinal part.

Direct photon processes. Both generators produce events according to the exact matrix elements of $2 \to 3$ subprocesses

$$e + q \to e + q + G \tag{2.1}$$

$$e + G \to e + q + \overline{q} \tag{2.2}$$

In some generators the direct processes are generated combining EPA with matrix elements corresponding to binary subprocesses $\gamma^* q \to Gq$ and $\gamma^* G \to q\overline{q}$. This approximation is reliable at small virtualities but may lead to large discrepancies as $Q^2 \to E_T^2$.

Strong coupling α_s. In its standard version HERWIG always employs the NLO formula for α_s. There is nothing inconsistent on that, but one must be aware of this difference when comparing results of HERWIG with those of other MC event generators. For the purpose of extraction of effective PDF it is, however, preferable to use the LO form of α_s, which, moreover, is a default in RAPGAP. In α_s formula, the default values of Λ were taken, namely 0.18 GeV in HERWIG and 0.25 GeV in RAPGAP.

Scale. Although in principle the factorization scale (M) and renormalization scale (μ) are independent of each other, we identified them, i.e we set $\mu = M$. In HERWIG the common scale is defined as

$$\mu_{\mathrm{H}}^2 \equiv \frac{2\hat{s}\,\hat{t}\,\hat{u}}{\hat{s}^2+\hat{t}^2+\hat{u}^2} = \frac{4\hat{p}_T^2}{3+\cos^2\hat{\theta}}, \tag{2.3}$$

where $\hat{s}, \hat{t}, \hat{u}$ are Lorentz invariant Mandelstam variables and $\hat{p}_T, \hat{\theta}$ are the transverse momentum and scattering angle of the outgoing partons in diparton cms where we also have

$$\hat{p}_T^2 = \frac{1}{4} s_{ep} y x_\gamma x_p \sin^2\frac{\hat{\theta}}{2} \qquad \text{and} \tag{2.4}$$

$$\cos\hat{\theta} = \tanh\frac{\Delta\hat{\eta}}{2}. \tag{2.5}$$

This means that for values of the scale in HERWIG we have $\mu_{\mathrm{H}}^2 \in \; <\hat{p}_T^2, 4/3\hat{p}_T^2>$, which is close to the values of the scale defined in RAPGAP as

$$\mu_{\mathrm{R}}^2 \equiv p_T^2 + 4m_p^2, \tag{2.6}$$

where m_p stands for the parton mass.

Parton cascading. As mentioned, HERWIG has a very elaborate treatment of cascading using coherent parton showers with angular ordering on both incoming and outgoing parton legs and for both proton and photon side. RAPGAP includes LUND string fragmentation for the initial as well as final parton showers on both proton and photon side.

PDF of the proton and virtual photon. In MC event samples used to correct the data to hadron level GRV-HO parametrization for the proton as well as photon were used. For consistency in the extraction of the effective photon PDF all MC event samples used for correction of the data to the level of hard scattering were generated with LO parametrizations of the proton as well as photon PDF. The virtuality suppression of the photon PDF was modelled by

- Drees-Godbole approach or
- using SaS1D and SaS2D which have the virtuality dependence built in in a more sophisticated way.

Both approaches are described in Chapter 1.

It is also important to specify the way heavy quarks c and b are treated. There are principally two ways of including the contribution of c and b quarks to hard scattering process. In one advocated by the GRV group, c, b are not considered as "intrinsic" content of hadrons or photons and the $c\overline{c}$ final states are generated by processes involving photons, gluons and light quarks in initial state. This procedure avoids the complications stemming from heavy quark mass thresholds, but requires the inclusion of more complicated matrix elements instead of simpler ones used in the alternative approach, in which heavy quarks are considered intrinsic above their respective mass thresholds. This second approach, adopted by MRS [40] and CTEQ [41] groups, is a reasonable approximation provided the hard scale M is sufficiently above the heavy quark masses, $M \gg m_h$. In opposite situation, i.e. when $M \ll m_h$, heavy quarks decouple and we may ignore them. In the kinematic range studied in this analysis the approach in which the number of effectively massless quarks $n_f = 4$ should be a good approximation and we shall adopt it in the following.

Additional transverse energy activity. It is well known that properties of jets with moderate E_T are significantly influenced by the additional soft particles, accompanying the basic parton level hard scattering but not originating from it. It is therefore important that the Monte Carlo used describes correctly the E_t-flow also outside the jets. The necessity of adding this transverse energy activity is documented e.g. in [42]. In HERWIG these soft particles are optionally generated via the mechanism called *soft underlying event* (SUE). The basic result of this mechanism is a redistribution of energy and longitudinal momentum of particles coming from the decay of the beam remnant clusters. In HERWIG without the SUE, these remnant clusters have typically large relative longitudinal momenta and, consequently, their decay particles are separated by large gap in rapidity. Most of these particles go into the beam tube and therefore do not influence the jet properties. The essence of the SUE is to allow these beam clusters to "collide" in a manner typical for soft hadronic (eg. pp) collisions, thereby producing many soft particles occupying the central region between the original beam remnant clusters, where the hard jets are usually located. In events with SUE the jet characteristics of low E_T jets may be significantly distorted by the accompanying soft particles. The fraction of events in which the hard scattering is accompanied by the SUE is given by the parameter PRSOF. We recall that there is no such additional mechanism in RAPGAP. For certain quantities the mechanism of the SUE has similar consequences as the multi-parton interaction option of PYTHIA, but it differs from the MI in two crucial aspects: the SUE does not change the cross-section of the basic hard scattering process, and it is a soft collision, not, as in the MI, yet another hard parton scattering described by perturbative QCD.

$\mathbf{p_T^{min}}$. The parameter $p_T^{\min}$ in both HERWIG and RAPGAP specifies the minimal transverse momentum of the basic parton level hard scattering process in γ^*p cms. To avoid distorting the jet E_T spectra close to the minimum transverse energy of jets used in the analysis ($E_T > 5$ GeV is required, see Chapter 4), $p_T^{\min}$ must be chosen somewhat below it. We took $p_T^{\min} = 3\,\mathrm{GeV}$ in both HERWIG and RAPGAP for the low Q^2 sample and $p_T^{\min} = 2.5\,\mathrm{GeV}$ in HERWIG for the photoproduction simulation. The stability of the HERWIG results with respect to the variation of $p_T^{\min}$ is investigated in detail in [39] and depends sensitively on the value of parameter PRSOF. For large values of PRSOF, the jet cross sections become so sensitive to the choice of $p_T^{\min}$, that it prevents any meaningful theoretical prediction [39]. It is therefore crucial that the comparison of HERWIG results to data leads to a small value of PRSOF, somewhere in the region $0.1 - 0.2$ (will be shown in Chapter 5), where HERWIG results are still only slightly sensitive to the choice of $p_T^{\min}$.

intrinsic $\mathbf{p_T}$. The value of this parameter gives the width of the intrinsic p_T-distribution of valence partons in incoming hadrons. In both MC generators the Gaussian distributions ($\mathrm{d}N/\mathrm{d}p_T^2 \propto exp(-p_T^2/\kappa^2)$) were used with $\kappa = 0.7\,\mathrm{GeV}$ in HERWIG and $\kappa = 0.44\,\mathrm{GeV}$ in RAPGAP in case of protons, $\kappa = 0$ in case of photons. The values of these parameters affect the p_T of the proton and photon remnant jet [9] which were not studied here.

Chapter 3

HERA and H1

The following chapter briefly describes the *ep* collider HERA and one of the four detectors operating at HERA, H1. In the description of the H1 detector [33] emphasis is put on those subdetectors which are relevant for this analysis.

3.1 HERA

HERA (**H**adron **E**lektron **R**ing **A**nlage) is the first ever constructed electron proton storage-ring, situated at the DESY (**D**eutsches **E**lektronen **SY**nchrotron) laboratory in Hamburg, Germany. DESY activities are focused on two areas:

- the investigation of the fundamental properties of matter in particle physics.
- the use of synchrotron radiation in surface physics, chemistry, molecular biology, geophysics and medicine.

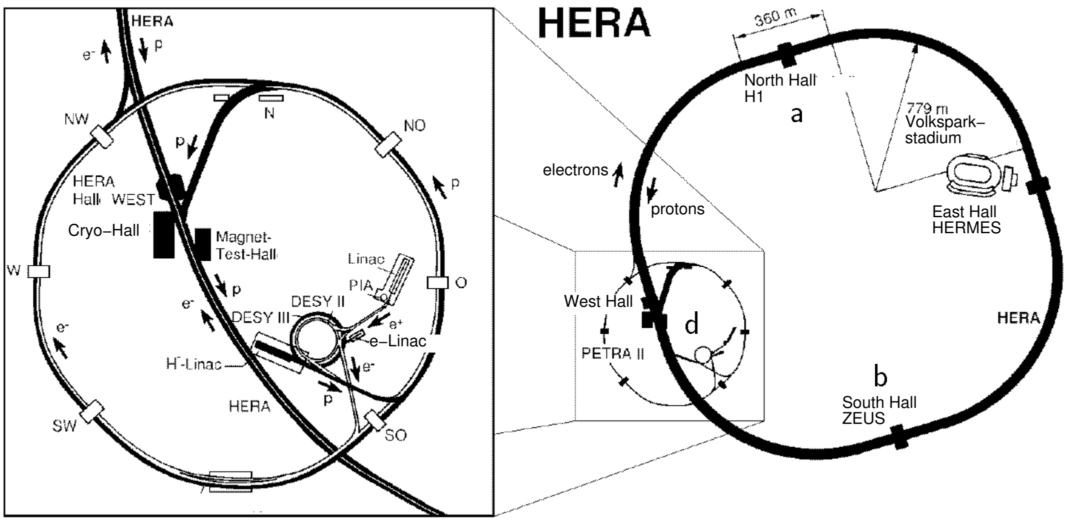


Figure 3.1: The layout of the accelerator complex HERA (right) together with the major experiments: H1 (a), ZEUS (b), HERMES (c), HERA-B (d). The left figure presents a zoomed view of the HERA pre-accelerator chain.

With various smaller systems (see Fig.3.1), electrons (or positrons)[1] and protons are first pre-accelerated and then injected into two separate rings of 6.3 km circumference where they are accelerated to their final energy of 27.6 GeV and 820 GeV in 1995. Finally the electrons and protons are brought to head-on collisions in two places where big detectors were built (H1 in the North Hall, ZEUS in the South Hall). Since 1995 HERMES experiment (West Hall) is operated where the origin of nucleon spin is investigated using collisions of polarized electrons with a polarized gas target (hydrogen, deuterium, ^{3}He or ^{4}He). In recently commissioned fixed target experiment, HERA-B (East Hall), the CP violation is investigated in B-meson systems using collisions of halo of the proton beam with tungsten wires.

The beams of particles collide in the form of bunches in 96 ns bunch crossing intervals (frequency 10.4 MHz) which corresponds to a maximum of 210 bunches in the ring. In the 1995 running period, typically 174 electron and 174 proton bunches were filled. In addition, there were 15 unpaired electron and 6 unpaired proton bunches. The bunches with no colliding partner, called pilot bunches, are an important tool to check the beam induced background.

3.2 The H1 Detector

Since the centre of mass from HERA collisions is boosted along the proton direction the H1 detector is considerably more massive and highly segmented in this direction. The coordinate system convention used by H1 collaboration and in this analysis is the following: the centre corresponds to the *ep* interaction point, the positive z-axis points towards the incident proton direction (hereafter referred to as the "forward") and the positive x-axis points horizontally towards the centre of HERA.

Figure 3.2 gives an overview of the H1 detector [33]. Starting the description from the interaction point, silicon trackers surround the beam pipe $\boxed{1}$ in central and backward region (not shown). The central $\boxed{2}$ and forward $\boxed{3}$ tracking systems contain different layers of drift chambers and trigger proportional chambers. The Liquid Argon (LAr) electromagnetic $\boxed{4}$ and hadronic $\boxed{5}$ calorimeters mounted in Liquid Argon cryostat, surrounds the central trackers. A super-conducting coil $\boxed{6}$ of radius 3 m is placed outside the LAr calorimeter and provides an uniform magnetic field of 1.2 T. This field is compensated by another super-conducting coil $\boxed{7}$ in order not to influence the HERA accelerator. The iron return yoke $\boxed{10}$ of the magnet is laminated and filled with limited streamer tubes. The small fraction of hadronic energy leaking out of the back of the calorimeter is registered here and muon tracks are found. Muon identification further benefits from additional chambers $\boxed{9}$ inside and outside the iron. The muon tracks in the forward direction are analyzed in a supplementary toroidal magnet $\boxed{11}$ sandwiched between drift chambers. The SpaCal electromagnetic and hadronic calorimeters $\boxed{12}$ are located in the backward direction behind the trackers. In the forward direction, a gap between the LAr calorimeter and the beam pipe is covered by the PLUG $\boxed{13}$ calorimeter. The whole detector is shielded by the concrete shielding $\boxed{14}$. The electron tagger at position $z = -33$ m upstream the interaction point (in HERA tunnel) detects electrons with a very small scattering angle inducing a photoproduction event and, taken in coincidence with a corresponding photon detector at $z = -103$ m upstream the interaction point, monitors the luminosity by the bremsstrahlung process. Two scintillator walls in backward direction are installed to recognize background produced by the proton beam upstream the H1 detector.

[1]In the following and throughout this thesis the incident and scattered lepton is referred to as electron.

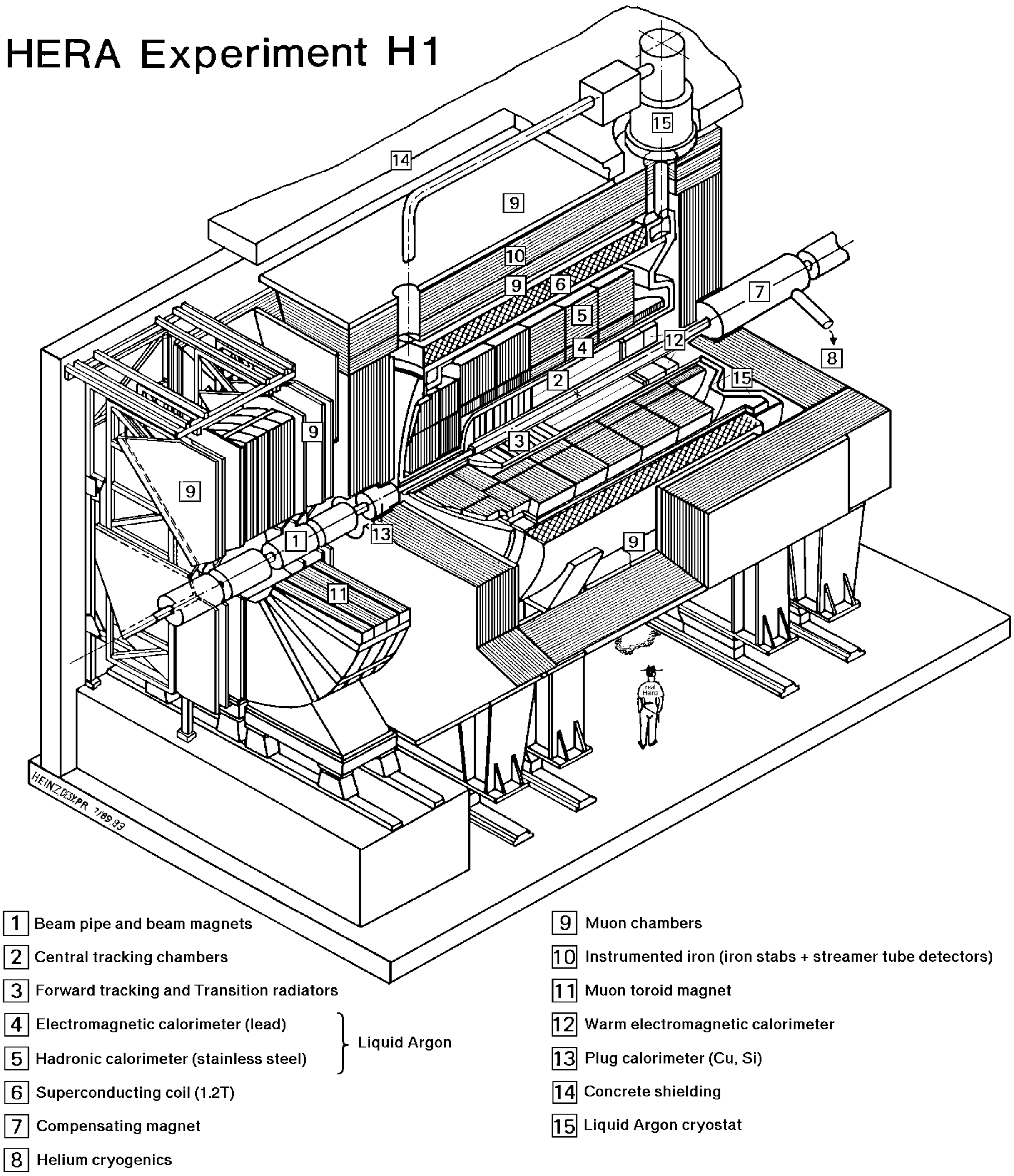


Figure 3.2: The H1 detector with its main components.

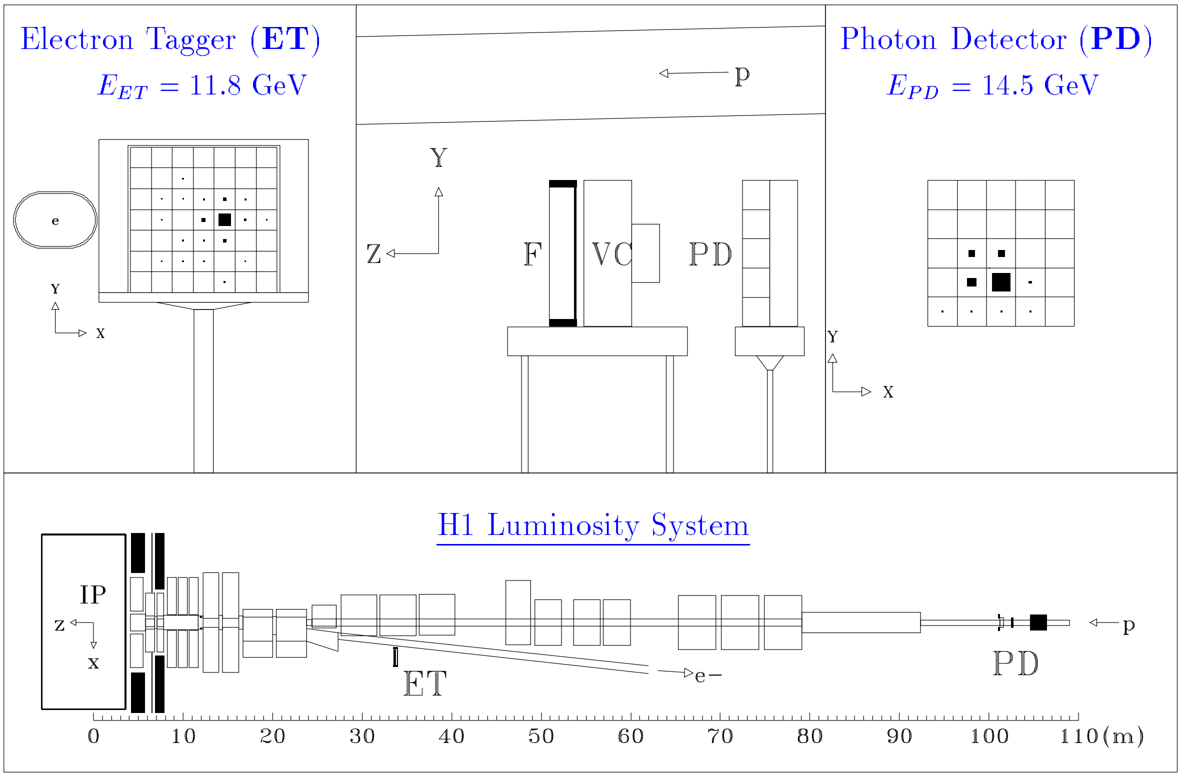


Figure 3.3: The bremsstrahlung event measured by the H1 luminosity system. In the lower part the complete luminosity system is shown with the electron tagger (ET) and the photon detector (PD). The upper central picture shows a zoom of the photon detection system consisting of PD, the filter (F) and veto counter (VC).

3.2.1 Luminosity system

The luminosity system is a multipurpose device. Its main task is a fast relative luminosity measurement. In addition it provides electron beam monitoring for the HERA machine, absolute luminosity measurement in the interaction region, tagging of the photoproduction events and energy measurement for small angle scattered electrons and photons from initial state radiation. The luminosity is determined from the rate of the Bethe-Heitler events $ep \to ep\gamma$ having a large and precisely calculable cross section [43]. The main source of background is bremsstrahlung from the residual gas in the beam pipe, $eA \to eA\gamma$, with even a larger cross section. The background is measured using electron pilot bunches. The luminosity monitor detects scattered electron and outgoing photons in coincidence. It contains therefore two arms: the electron tagger (ET) and the photon detector (PD). The general view of the luminosity system is shown in Fig.3.3.

Scattered electrons are deflected by a set of quadrupoles and a bending magnet. They pass an exit window at $z = -27.3\,\mathrm{m}$ and hit ET at $z = -33.4\,\mathrm{m}$.

The bremsstrahlung photons leave the beam pipe through the photon exit window at $z = -92.3\,\mathrm{m}$, where the beam pipe bends upward, and hit PD at $z = -102.9\,\mathrm{m}$. A Pb filter (F) followed by a water Čerenkov counter (VC) protects the detector from the high synchrotron

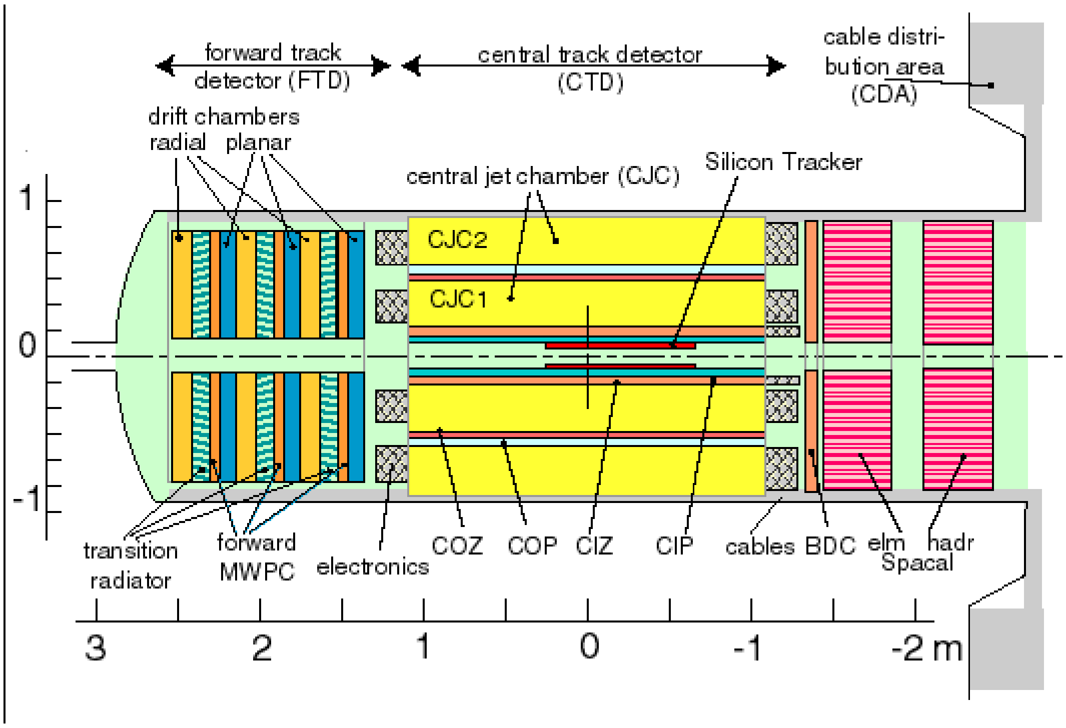


Figure 3.4: The side view of H1 tracking system.

flux. The aperture of PD is $\lesssim 0.45\,\text{mrad}$, limited by the inner diameter of lead collimator. This corresponds to an acceptance of ≈ 98 % for bremsstrahlung photons which varies only slightly with changes of the beam optics.

Besides the luminosity measurements the system is used to tag quasi-real photoproduction events with $Q^2 < 0.01\,\text{GeV}^2$ in the energy interval $0.2 < y < 0.8$. In this case the trigger requires only the energy in ET, while PD and VC act as a veto. This suppresses proton induced background substantially.

The major contribution to the systematic error in the absolute luminosity comes from the dependence of the system acceptance on possible variations of the electron beam angle in the interaction region. This beam tilt is measured on-line with a precision of 0.01 mrad by reconstructing the average position of the photon spot at PD. The final absolute values of the integrated luminosity, used for physics analyses, are calculated off-line from a detailed understanding of the acceptance of the luminosity detectors. The total systematic error from the luminosity system in 1995 was 1.07% [44].

3.2.2 Trackers

The H1 tracking system (see Fig.3.4) has been designed in order to provide triggering, reconstruction and momentum measurement of charged particle tracks. Particle trajectories are bent in the magnetic field. To achieve good track reconstruction efficiency over a maximum angular range, two mechanically distinct detector modules were constructed, the central tracker (CTD) and forward tracker (FTD). Each is optimized for tracking and triggering in its angular region.

Central tracker

- **Drift chambers**
  Track reconstruction in the central region is based on the two large concentric jet chambers, CJC1 and CJC2, covering the polar angle range $15^\circ < \theta < 165^\circ$. The chambers have wires strung parallel to the beam axes. The inner ring consists of 30 segments with 24 sense wires each, while the outer ring is divided into 60 segments equipped with 32 sense wires. A space point is measured with a resolution of $170\,\mu$m in the drift coordinate ($r\phi$-plane) and, by charge division measurements, with a resolution $\sim 1\%$ of the wire length in z. The parameters of tracks are obtained from up to 56 points on track, which are derived from the drift time measurements on different sense wires. Through the magnetic field, only tracks with transverse momentum $p_T \gtrsim 0.15\,$GeV will pass through both chambers. The momentum resolution achieved are $\sigma_p/p^2 \lesssim 0.01\,\text{GeV}^{-1}$ and $\sigma_\phi = 21\,$mrad.

  Two thin drift chambers, the central inner (CIZ) and central outer (COZ) z-chambers complement the measurement of charged track momenta in the central chambers. The CIZ fits inside CJC1 and COZ in between CJC1 and CJC2. In these chambers the drift direction is parallel to the beam axis with the wires strung on polygonal support structures arund the axis. The CIZ (COZ) comprises 15 (24) rings of 12 (9) cm length in z-direction with 4 layers of sense wires in each ring. The internal z-resolution is about $250\,\mu$m at 1 cm drift distance. The $r\phi$-coordinate is measured using also the charge division method with a precision of 1-2% of 2π.

  The vertex fitted tracks recognized in the CJC are supplied with the z-chamber information using a combined fit-reconstruction procedure. The precision of the vertex position determination is thereby improved to 2 mm. Due to z-chambers the θ-measurement accuracy for individual tracks is improved from about 10 mrad to 1 mrad.

- **Proportional chambers**
  The inner multiwire proportional chamber (CIP) is placed closest to the beam pipe. The angular coverage of the chamber is $9^\circ < \theta < 171^\circ$. The chamber is composed of 60 sectors in z-direction and of 8 sectors in ϕ.

  The central outer proportional chamber (COP) has angular acceptance of $25^\circ < \theta < 155^\circ$ and consists of 18 sectors in z-direction and of 16 sectors in ϕ.

  Both chambers provide a fast timing signal with a time resolution better than 96 ns separation time between consecutive HERA bunch crossings. A "ray" combination of pads hit in CIP with pads hits in COP and in the forward proportional chambers (FWPC) is used to trigger charged particle trajectories originating from the interaction region.

Forward tracker

FTD provides an accurate measurement of trajectories of charged particles in the forward direction, $5^\circ < \theta < 30^\circ$, track information on individual particles within jets by means of transition radiation detection and a fast forward track trigger. It consists of three supermodules. Each of them includes three different orientations of planar wire drift chambers designed to provide accurate θ-measurements, FWPC for fast triggering, a passive transition radiator and a radial wire drift chamber which provides accurate $r\phi$-information.

3.2.3 Liquid argon calorimeter

The liquid argon calorimeter surrounds the tracker and covers the polar angular range between 4° and 153°. The LAr calorimeter is placed inside the large H1 coil. The segmentation along the beam axis is in eight "wheels". Each of the six barrel wheels is segmented in ϕ into eight identical "stacks". The two forward wheels are each mechanically assembled as two half rings. Each stack in turn is divided into an electromagnetic section with lead as absorber material, followed by a hadronic section made of welded stainless steel absorber plates with independent read-out cells inserted between the plates.

The total thickness of the electromagnetic part varies between 20 and 30 radiation lengths for electrons and 1.0–1.4 interaction lengths for hadrons. The total depth of the calorimeter varies between 4.5 and 8.0 nuclear interaction lengths, depending on polar angle.

The energy resolution for electrons $\sigma(E)/E$ is about $10\%/\sqrt{E}$ (E in GeV) with a constant term below 1%. The energy resolution for hadrons is about $50\%/\sqrt{E}$ with a constant term below 2%. Typical noise levels in the read-out channels are between 10 and 30 MeV. The LAr calorimeter is non compensating. The charge output for hadrons is about 30% smaller than that for electrons. A special software weighting technique has to be applied to get the proper hadronic energy scale for hadrons.

Both the absolute energy scale and the resolution of the LAr calorimeter have been verified using test beams and H1 data. By comparing the measured track momenta of electrons and positrons with the corresponding energy deposits in the LAr calorimeter, the electromagnetic energy scale is presently known with 3% uncertainty. By studying p_T-balance between the scattered electron and the hadronic final state particles, the hadronic energy scale is known with an uncertainty of 4%.

3.2.4 Backward detectors

In the 1994-1995 winter shutdown the upgrade of the backward part of the H1 detector has been performed. The backward proportional chamber (BPC) was replaced by backward drift chamber (BDC) and the backward electromagnetic calorimeter by a lead-fiber "spaghetti" calorimeter (SpaCal). Both BDC and SpaCal are crucial for measurement of the scattered electron in this analysis.

BDC

The BDC is subdivided into 8 octants consisting of 4 double layers. The signal wires are strung in polygons around the beam axis in order to optimize a resolution of the polar scattering angle θ which is particularly relevant for the reconstruction of the event kinematics. The double layers are rotated by 11.25° to obtain some measurement of ϕ. Each signal wire is contained in a separate cathode cell. The wires in different planes of the double layers are shifted by half the cell size in order to resolve a right-left ambiguity of the drift origin. The wire spacing is 1 cm for the inner 16 cells in the octant and 3 cm for the outer part. A transition region between the two zones contains cells with 2 cm width. The angular coverage of the BDC is between 153° and 177.5°. The spatial resolution for individual hits is 0.3 mm leading to a θ-resolution better than 0.5 mrad if no showering occured in the material between the vertex and the BDC. The BDC track finding efficiency was studied in [45]. It amounts to 98% in the inner and 96% in the outer region.

SpaCal

The SpaCal calorimeter [46] comprises the electromagnetic and hadronic section.

The electromagnetic part of the SpaCal consists of 1192 cells with an active volume of $4.05 \times 4.05 \times 25\,\mathrm{cm}^3$ each. The cells are made of grooved lead plates filled with scintillating fibers with a diameter of 0.5 mm. The scintillating light of each cell is converted into an electric pulse using photomultiplier tubes (PMT). The active length of the electromagnetic SpaCal corresponds to 27.47 radiation lengths and one nuclear interaction length. The angular coverage of the calorimeter is between 153° and 177.8°.

The use of PMT and an electronic chain with low noise level (3 MeV) permits very low trigger thresholds and a reliable reconstruction of small energy deposits. The PMT provide a time resolution of about 1 ns allowing to reduce the non-ep background on the trigger level.

The energy resolution for electrons was measured in test beam experiments [47, 48] and parametrized as $7.5\%/\sqrt{E}$ with a constant term 2%. The spatial resolution of the calorimeter in the transverse plane was measured to be 3.4 mm.

The hadronic part of the SpaCal comprises 136 cells of $12 \times 12 \times 12\,\mathrm{cm}^3$ providing one nuclear interaction length. The fibres are of the same type as in the electromagnetic section but have a diameter of 1 mm. The energy resolution for hadrons is $\sim 30\%/\sqrt{E}$.

3.2.5 Time-of-flight (TOF) system

The TOF system consists of several scintillation counters, namely the Backward TOF (BTOF, at $z = -275$ cm), Forward TOF (FTOF, at $z \approx 790$ cm), the Plug TOF (PTOF, at $z \approx 540$ cm) and the veto walls (at $z \approx -390$ cm). The time resolution is at the level of ns and allows to reject beam induced background arriving out-of-time in the H1 detector. For this purpose the TOF counters are operated in veto mode on the first trigger level L1. The TOF system supplements the time sensitive electron trigger of the SpaCal discussed below.

3.2.6 Trigger system

The trigger system presently consists of four levels of event filtering, L1, L2, L4 and L5. In 1995 L2 was not yet operational. It was commissioned in 1996.

The first trigger level L1 [49] is a dead time free system providing trigger decisions after $2.5\,\mu$s. During this time the full event information is stored in pipelines. The L1 trigger conditions (subtriggers) are based on simple logical combinations of the trigger signals (trigger elements) from the different detector components. Due to limited capacity of the Read Out System the rates of some subtriggers have to be down-scaled. If the subtrigger conditions are fulfilled the "raw" bit is set. If the event is actually selected by this subtrigger the "actual" bit is set. The L1 accepts the event if at least one subtrigger has an actual bit set. In order to establish the trigger efficiencies from the data a set of monitor triggers is installed with loose (minimum bias) trigger conditions.

The L2 validates the L1 decision using more complex algorithm within $20\,\mu$s. Two kinds of second level trigger techniques are used in H1: neural network and topological triggers. At a positive trigger decision the read-out of the event is started. The full event information is transferred to the trigger level L4.

The L4 is a software filter farm consisting of 30 parallel processors[2]. Additional background suppression, using reconstructed event information, is performed here. About 20% of the events pass the L4 selection cuts and are written to tape. The output event rate is technically limited to about 20 Hz or 1.5 MBytes/s. For monitoring purposes, one out of 100 rejected events is written to a separate data file.

At L5, finally, the full event reconstruction is performed. Calibration and correction constants are applied to the data. The calibrated events are classified according to their physics properties. Unclassified events are rejected. Similar to L4 1% of the rejected events are stored for monitoring. Classified events are written on the so called Data Summary Tapes (DST) which are normally used as a basis of the further physics analysis.

[2]Since 1997 36 processors are active.

Chapter 4

Data selection

4.1 Introduction

The analysis is based on data recorded in 1995 with nominal z-position of the ep interaction point. In the same year a set of data was recorded in which the z-position was shifted by $+70\,\mathrm{cm}$. The later data were analyzed and the results of F_2^p measurement published in [50]. In presented analysis the data from two classes of ep processes were investigated: the *photoproduction data* in which the scattered electron is detected in the electron tagger and the *low Q^2 data* with the electron caught in the SpaCal. Still limited statistics of the year 1995 required considerate choice of the kinematic bounds in which we are going to study the structure of the photon. In general, the procedure leading to physics results was the same for both samples however in some aspects different treatment was adopted, namely in triggers, in the reduction of the background and in taking the acceptance of electron detecting device into account. Therefore we will deal with each data sample separately.

The event selection is based on the identification of the scattered electron and the presence of jets. In Sections 4.2.1,4.3.1 the event selection of global character is mentioned, the Sections 4.2.2 and 4.3.2 describe the electron identification in different devices and in the Section 4.2.3 the jet algorithm used is described together with delimitation of the space in the detector where the jets are searched.

4.2 Low $\mathbf{Q^2}$ sample $(1.4 < \mathbf{Q^2} < 25\,\mathbf{GeV^2})$

4.2.1 Run selection

The early time of the data taking in 1995 was stigmatized by various hardware problems connected with installation of the new backward detector components the BDC and the SpaCal. Some read-out errors and particularly the SpaCal trigger inefficiencies disabled to use the data with the SpaCal information. In 1995 reliable functionality of the SpaCal had been found in the runs recorded between 20th of October and the end of the 1995 run period (27th of November). In addition the following criteria were imposed:

- runs are of good or medium quality (according to H1 ep – event classification)
- all components of H1 detector relevant for this analysis are functional.
  If at least one of the HV alarm bits for CJC, CIZ, COZ, BDC, SpaCal, LAr, luminosity system and TOF–counters is on, the luminosity is corrected accordingly.

4.2.2 Electron identification

The electron is identified in the SpaCal using a number of cuts which were carefully tuned to minimize the background. The detailed studies leading to these cuts and their explanations are given in [51]. Slightly modified cuts for shifted vertex data can be found in [50] and their closer explanation in [52].

The selection criteria for the electron identification are:

- The cluster in the electromagnetic part of the SpaCal with the highest energy deposit is selected as the electron candidate. The minimum energy required for the electron candidate is 8 GeV.

- The signal from the electron candidate lies within a window of $8.5\,\mathrm{ns} \leq t \leq 16.5\,\mathrm{ns}$ from the time expected for the genuine ep interactions. This cut is designed to reject out-of-time proton background which, due to different flight path lengths, arrives sooner in the SpaCal.

- The cluster radius R_{cl}, describing the transverse size of the shower caused by the electron candidate, is smaller than 3.5 cm. The energy $E_{e,had}$ in the Hadronic SpaCal behind the electron candidate cluster, estimating the longitudinal size of the shower, is smaller than 0.5 GeV. This cut is motivated by the fact that electromagnetic showers in general are completely contained in the electromagnetic SpaCal. An energy leakage into the beam pipe is suppressed by a cut on the energy in the SpaCal veto layer of $E_{veto} < 1\,\mathrm{GeV}$.

- At least one reconstructed track is found in the BDC with number of hits greater than 4 (the maximum number of hits is 8). Each track in the BDC is characterized by the coordinates of its central point and a direction vector. This point is extrapolated using a straight line from the vertex to the SpaCal transverse plane in the z-coordinate given by the coordinates of the electron candidate cluster. The distance of this track in the SpaCal plane from the electron candidate cluster is smaller than 2 cm in radial and 2.5 cm in azimuthal direction.

- The radial distance of the electron candidate cluster from the beam line is greater than 8.7 cm. The main goal of this cut is to minimize energy leakage into the beam pipe. In addition the spatial resolutions of both the SpaCal and the BDC significantly deteriorate very close to the beam pipe due to edge effects.

- $E - p_z \equiv \sum(E_h - P_{z,h}) > 35$ GeV, where E_h and $P_{z,h}$ are the energy and longitudinal momentum of a particle[1]. The sum is over all energy deposits measured with the calorimeters supported by tracking information (see below CLIM method) excluding the electron tagger and the photon detector. As shown in [51], this cut is powerful to reject photoproduction background.

- The z-coordinate of the interaction vertex lies within 35 cm of the nominal position.

[1] If not explicitly specified the term *particle* means common name for combined objects on detector level (see description of CLIM method below) or partons generated in Monte Carlo event generator or hadrons coming from them.

The efficiency of the electron finder procedure was estimated using event generator and was found to be less than 2% [51]. Once the characteristics of the selected electron are fully determined we calculate the virtuality, Q_e^2, and the inelasticity, y_e, both by electron method, i.e. relying on the information on the scattered electron only.

4.2.3 Jet algorithm and jet selection

The cone algorithm used here is applied to the particles transformed to the $\gamma^* p$ cms with positive z-axis along the direction of the incident proton. The recombination of particles is done in the E_T-recombination scheme which is based on quantities: transverse energy E_T, pseudorapidity η and azimuthal angle ϕ. The axis of a jet is defined [53] by

$$E_{T,jet} = \sum_i E_{T,i}, \qquad \eta_{jet} = \frac{\sum_i E_{T,i}\,\eta_i}{\sum_i E_{T,i}}, \qquad \phi_{jet} = \frac{\sum_i E_{T,i}\,\phi_i}{\sum_i E_{T,i}} \tag{4.1}$$

where the sums run over all particles assigned to the jet. The resulting jet cross sections are dependent on how the algorithm treats the choice of jet initiators and configurations of overlapping jet cones. It has repeatedly been pointed out that many definitions of cone algorithms are not infrared and/or collinear safe [54, 55].

In this analysis we use the definition implemented in PXCONE [56] which does not suffer from problems discussed in [54, 55]. This definition which corresponds closely to the Snowmass proposal [57] and to the algorithm used in the CDF experiment [58] is also used by the OPAL collaboration [59].

Jets are defined by the particles i within a cone of radius $R_c = 1$ (or a half-angle R_c) in pseudorapidity and azimuth space ($R_c^2 = (\Delta\eta)^2 + (\Delta\phi)^2$) around the jet axis. The parameter setting chosen for this analysis is $\epsilon = 3\,\mathrm{GeV}$ and $ovlim = 0.75$. The algorithm operates as follows:

1. Each particle in the event is considered as a seed of a jet, i.e. its momentum is taken as the axis of a cone of half-angle R_c.
2. All particles lying within this cone are found and their momenta are summed.
3. If the momentum sum and the cone axis do not coincide, the momentum sum is taken to define a new cone axis and step 2 is repeated.
4. When a stable solution has been reached, the particles in the cone are deemed to constitute a "proto-jet". Unless the set of particles forming the proto-jet is identical to one already found, the new proto-jet is added to a list.
5. At this stage two narrow proto-jets may be found separated by an angle between R_c and $2R_c$. To permit this configuration to be found as a single jet, the steps 2 to 4 are performed for all midpoints of pairs of proto-jets as possible initial cone axes. This leads to infrared safety of the procedure [55].
6. Particles from proto-jets with transverse energy $E_{T,jet} < \epsilon$ are removed from the list if they are not contained in other proto-jet with $E_{T,jet} > \epsilon$.
7. If two proto-jets have some particle in common, the overlap fraction f is computed as the ratio of the total energy of the particles common to both proto-jets to the energy of the lower energy proto-jet.

(a) If $f > ovlim$ the higher energy proto-jet remains unchanged and the lower energy proto-jet is eliminated. If multiple overlaps occur, the highest energy pairs are treated first.

(b) If $f < ovlim$ each particle in the overlap region is assigned to the proto-jet to whose axis it lies closest in (η,ϕ). The jet momenta are recalculated using Eq.(4.1). Thus it may happen that some particles within a jet have a distance larger than R_c.

8. Proto-jets with transverse energy $E_{T,jet} < \epsilon$ are again rejected and the remaining ones are called jets.

In this analysis the events are accepted if at least two highest E_T jets satisfy conditions

- $E_T \geq 5\,\mathrm{GeV}$.
- $-3 \leq \eta \leq 0$.

For such jets, the response of the LAr calorimeter (acceptance, resolution) is well understood.

4.2.4 Clusters and tracks

Clusters and tracks are basic elements we work with on the detector level. All the methods used to calibrate the cell energy in the electromagnetic part of the SpaCal are described in great detail in [51]. The treatment of 1995 data in analyses of the proton structure functions is described in [50, 51, 52]. It was found that some corrections were needed to apply, e.g. a finer energy recalibration of the electromagnetic SpaCal cells of the distance smaller than 35 cm from the beam line, cell energy calibration of the insert[2] and last but not least the calibration of the SpaCal for hadrons.

Based on good electron identification and measurement of hadron flow in the LAr calorimeter and SpaCal this analysis requires a careful treatment of cell information all over the H1 detector. Therefore we adopted all mentioned steps from analysis [51] which we name "ELAN procedure". Because of bigger emphasize on hadron flow measurement in this analysis we will concentrate in the following on the effect of the ELAN procedure on the hadron flow into the SpaCal, i.e. essentially on the effect of the SpaCal calibration for hadrons.

The energy scale corrections for hadrons are applied to the SpaCal cells which do not belong to the cluster of the selected electron candidate in the electromagnetic part of the SpaCal. The corrections were obtained by demanding that the inelasticity measured with the hadronic final state, y_{JB}, is on average balanced with y_e reconstructed from the scattered electron. By this requirement we want to ensure not only that the event kinematics is consistent between different reconstruction methods[3] but also that the energy originating from hadronic final state is well balanced with the energy deposited by the scattered electron.

Three quantities defined in expressions (4.2) served to investigate how the above procedure improved the global balance in the event.

$$\begin{gathered} p^e_{T,x} + p^h_{T,x} \\ p^e_{T,y} + p^h_{T,y} \end{gathered} \tag{4.2}$$

[2] We impose a cut on the energy in the veto layer in the electron identification procedure.

[3] Although we use only electron method it is of course desirable to have this consistency.

$$(E - p_z)_{b.i.} - (E - p_z)_{a.i.} = 2E_e - E'_e(1 - cos\theta'_e) - (E - p_z)$$

The quantities $p_{T,x}$, $p_{T,y}$ are constructed as follows. First we define the transverse momentum vector of the scattered electron in the (x,y)-plane of the laboratory system. The (x,y)-frame is then rotated so that the scattered electron points to the x-axis in this new frame. The quantities $p^e_{T,x}$ ($p^h_{T,x}$) then stand for x-component of the transverse momentum vector of the electron (hadron) system in the new frame and similarly $p^e_{T,y}$ ($p^h_{T,y}$) are projections onto y-axis. The third quantity represents the difference between ($E - p_z$) quantities of initial state (before interaction) and final state (after interaction).

Note that $p_{T,x}$ is very sensitive and actually the only one quantity convenient for the investigation of the balance between the electron and hadron system in the transverse plane (compare e.g. p^h_T/p^e_T where mere magnitudes of vectors are exploited) since it takes properly into account also the directions of these transverse momentum vectors. It is especially important to use this quantity for the low Q^2 events where the errors in the measurement of the hadron system vector, being apparently comparable with small p_T of the electron, may alter the original direction of the vector significantly. As $p^e_{T,y} = 0$ by definition, $p_{T,y}$ quantity is kind of a measure of the symmetry in the distribution of the p_T-vector of the hadron system around the x-axis.

From the following it will be obvious that the ELAN procedure did not give a satisfactory global balance when used for a small subsample of deep inelastic events with hard jets (see Fig.4.1). Therefore we used additional procedure, so called “CLIM method”.

CLIM method

The purpose of this approach, currently used in H1 analyses (e.g. [31, 60, 61]) is to compensate for losses in calorimetric measurements of hadrons. The aim is to reach an energy balance between the scattered electron and hadron system. Although these losses are supposed to be corrected by MC simulation the correction factors may be quite large which results eventually in large uncertainties in the corrected data. A natural consequence of the compensation, or the improved balance, is an improvement in resolution of kinematic variables. This method combines all calorimetric clusters and tracks reconstructed in the central drift chamber. The momentum vector of a cluster is determined from its energy deposit in the calorimeter and its direction oriented from the interaction vertex to the cluster centre of gravity. The position of the vertex was corrected for a tilt of the beam line with respect to the central line of the central drift chamber[4]. The momentum p_{lim} of each central track with momentum p_{ch} was included with the value p_{lim} =min(p_{ch}, 400 MeV). The direction of p_{lim} vector coincides with the direction of p_{ch} vector. The value 400 MeV was determined as that value for which the means and half-widths of the distributions of the quantities in (4.2) correspond to the best balanced sample of events. The effect of above procedures is documented in Fig.4.1. The first column from the left shows the starting distribution where no corrections were applied, the second column corresponds to application of the ELAN procedure resulting in a slight but perceptible improvement in the mean values and the third column represents an implementation of both the ELAN and CLIM methods reaching almost zero mean values and conserving nearly the same half-widths. A reference event sample on which the comparisons were performed was restricted by the cuts defined so far after the ELAN and CLIM procedures were implemented.

Globally we see a rather small influence of the ELAN procedure. This is due essentially to three facts. Firstly, the final tuning of calibration constants for the SpaCal including the insert

[4]This line defines the z–axis of the laboratory system.

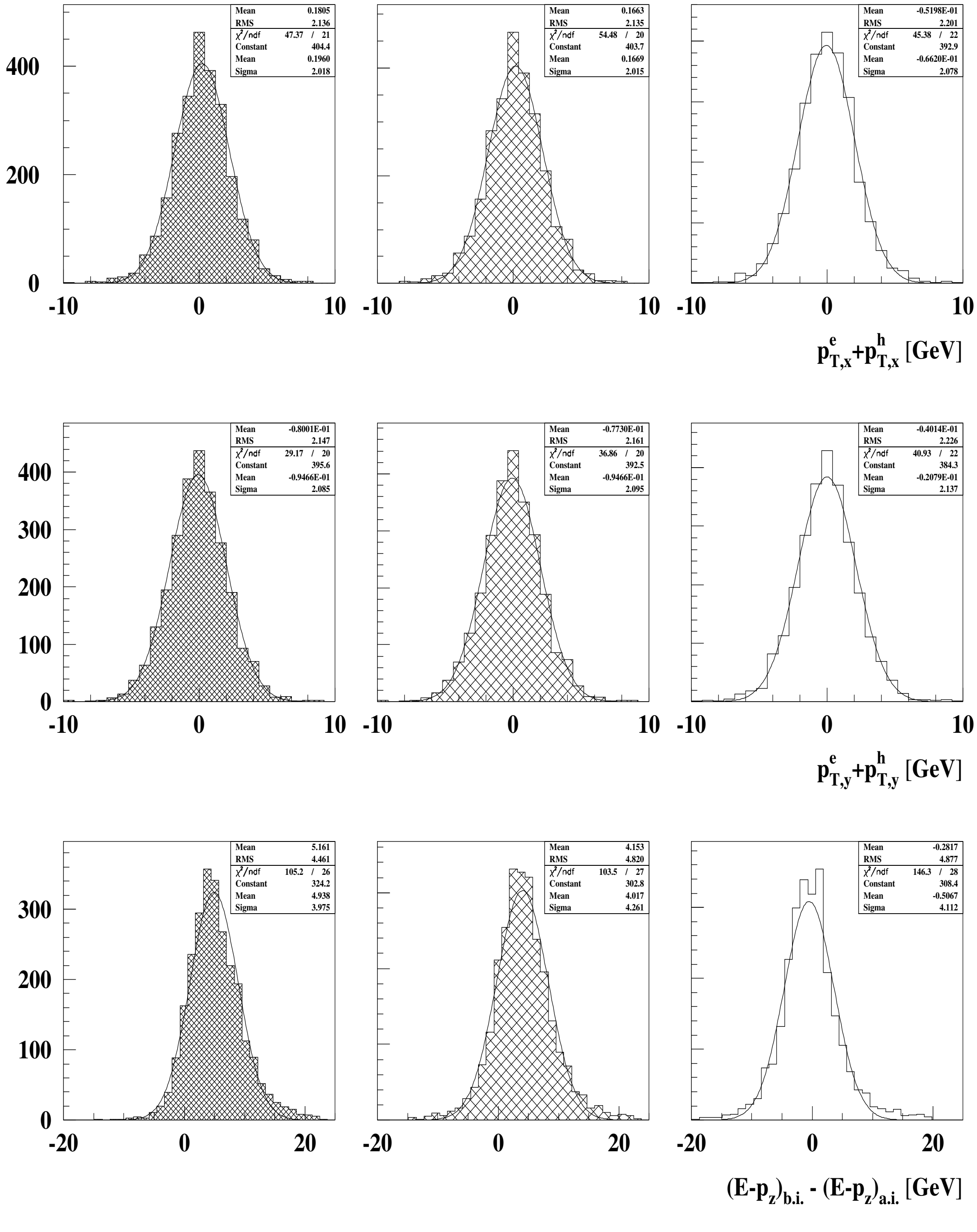


Figure 4.1: The global balance of the full event sample using three quantities described in the text. The first, second and third columns correspond in sequence to no corrections, application of the the ELAN procedure and application of both the ELAN and CLIM procedures.

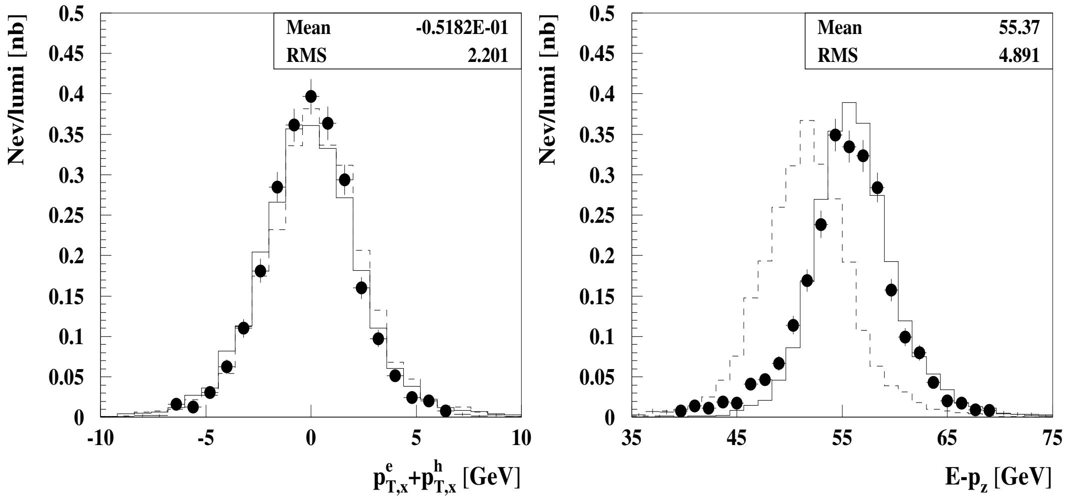


Figure 4.2: The distributions of $p_{T,x}$ and $(E-p_z)$ quantities. The full points (full histograms) represent the data (HERWIG 5.9) after implementation of the ELAN and CLIM procedures, the dashed histogram represents the prediction of HERWIG 5.9 without any corrections. The mean and RMS values are given for the data after application of the the ELAN and CLIM procedures.

represent small corrections important in analyses of the proton structure function. Secondly, and mainly, the calibration for hadrons involved in the ELAN procedure concerns only the electromagnetic part of the SpaCal. The SpaCal itself, in comparison with LAr calorimeter, fills a small part of space and, in particular, it is placed at a position where the hadron flow is seldom seen. The backward hadron flow is caused largely by a hadronized photon remnant which is present in the resolved photon processes. Their contribution to the total jet cross section appreciably decreases with increasing Q^2. Thirdly, this calibration affects a relatively small number of clusters because the backward hadron flow is low compared to photoproduction (see Fig.5.6 and compare also Fig.5.4 with Fig.5.7).

The same treatment was also applied to the MC event sample. Fig.4.2 shows that the MC model (HERWIG 5.9 for illustration) describes the data satisfactorily and that the quantity $(E - p_z)$ peaks at the required value, i.e. approximately at $2E_e$, in both the data and the MC model. The small disagreement between the data and HERWIG for lower values of $E-p_z$ is due to QED radiation that occurs in the data but it is not implemented in HERWIG (see Section 5.4.2 about the influence of these radiations).

As stated earlier, both procedures are to compensate for losses in measurements of hadrons so we should observe some improvement in resolution. The comparisons are shown in Fig.4.3 for two ranges of virtuality (1.4–8 and 8–50 GeV^2). In both of them an evident improvement is achieved in the jet pseudorapidity, ϕ-angle, x_γ and $(E-p_z)$ quantities. The lack of improvement in resolution of the jet E_T remained not understood.

In the following we will exploit the $p_{T,x}$ quantity (Eq.(4.2)) for one more check. Let us suppose—contrary to the assumption made so far—that the hadron system is well measured. Then an independence of the mean $p_{T,x}$ of Q^2 (and better $p_{T,x} = 0$ for all Q^2) would prove that the p_T^e as well as Q^2 are determined precisely. The Fig.4.4 shows that this is within errors the case up to $Q^2 = 50\,\mathrm{GeV}^2$.

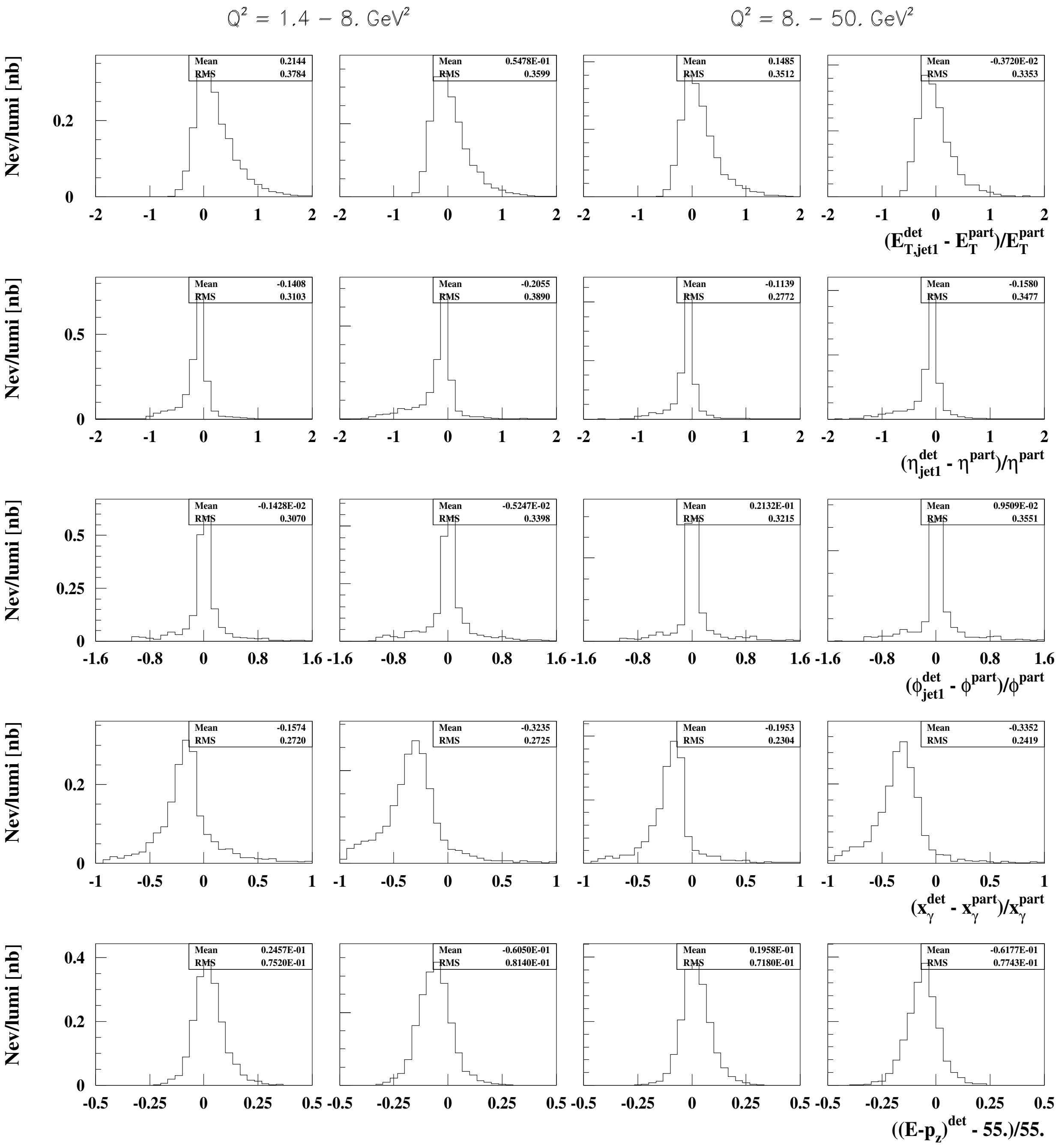


Figure 4.3: The resolution defined by the difference between the quantities on the detector level and the hard scattering level. The variables studied with a help of HERWIG 5.9 are the jet transverse energy, pseudorapidity and angle in azimuthal plane and x_γ and $(E-p_z)$ quantities. The first column corresponds to the sample after applying both the ELAN and CLIM procedures, the second column shows the same without any corrections, both for the range $Q^2 = 1.4 - 8\,\mathrm{GeV}^2$, the third and fourth column was produced analogously for the range $Q^2 = 8 - 50\,\mathrm{GeV}^2$.

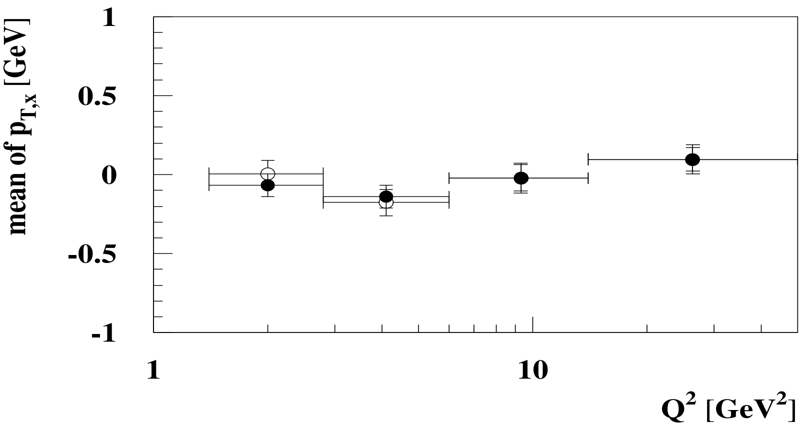


Figure 4.4: The dependence of the mean $p_{T,x}$ on Q^2 in the data. All the cuts as well as the ELAN and CLIM procedures were applied. The full points correspond to the values from Gaussian fit, the open ones to the values from histograms.

4.2.5 Trigger selection and event weights

An increasing luminosity delivered by HERA gives rise to increasing trigger rates in the H1 trigger system (described in Section 3.2.6). To keep the data recorded by H1 to a reasonable size a rejection or down-scaling of events has to be applied. The down-scaling factors which are applied on trigger levels 1–4 may vary from run to run[5]. In this analysis we adopted an approach from [62] based on the requirement that the spread of event weights (calculated from the down-scaling factors) in a selected event sample is as small as possible. This implies that two events of the same type (the same pattern of raw subtriggers was set on) should be assigned the same weight independently of the run in which it occurred or the actual down-scaled subtriggers. We can imagine for example the situation when a certain pattern of raw subtriggers contributes a great deal to the cross section in some run due to high run luminosity and a high down-scaling factor of one subtrigger whereas in another run with low luminosity the same pattern gives negligible contribution because the down-scaling factor is low. This imbalance of the significance level between the events of the same kind can be countervailed or smoothed by averaging the event weights over the run luminosities.

In the following the procedure for trigger selection and event weight calculation is described.

The available number of events in the data is found by applying the cuts except for the trigger condition. In order to calculate the event weight one has to know which actual subtrigger bit was set on in a single event. The spectrum of all actual subtriggers contributing to the selected event sample is shown in Fig.4.5. The whole spectrum (128 bits) is truncated as the remaining bits do not contribute at all. We see that the dominant contribution comes from subtriggers s0, s2 and s56. In selecting the set of subtriggers we have to make a compromise between the desires for high statistics and low systematic error of the trigger efficiency. Investigations as to inefficiencies and down-scaling factors led us to the selection of the following set of subtriggers:

$$s2, s3, s4, s12, s53, s56, s60. \tag{4.3}$$

[5] In 1995 only trigger level L1 was down-scaled.

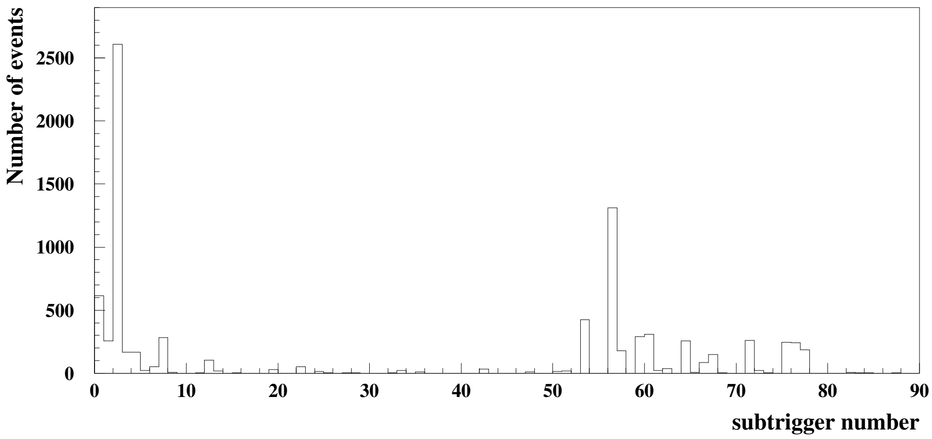


Figure 4.5: Spectrum of all subtriggers contributing to the event sample selected by all the cuts except for triggers.

The s0 subtrigger was eventually excluded due to high down-scaling factors. Slightly simplified definitions of the subtriggers read as follows:

"SpaCal" subtriggers:

s2 = (IET > 1) ∧ (ZVTX_t0 ∨ DCrϕ_T0) ∧ (¬ RZ_veto) ∧ (FB_TOF)
s3 = (IET > 2) ∧ (FBP_TOF)
s4 = (IET > 2) ∧ (¬ SPCLe_IET_cen) ∧ (FBP_TOF)
s12 = (IET > 1) ∧ (LU_PD) ∧ (¬ SPCLe_ATOF) ∧ (FBP_TOF)

"Jetty" subtriggers:

s53 = (DCrϕ_Tc ∧ DCrϕ_TNeg ∧ DCrϕ_THig) ∧ (LAr_BW) ∧ (¬ RZ_veto) ∧ (FB_TOF)
s56 = (DCrϕ_TNeg) ∧ (LAr_BR) ∧ (¬ RZ_veto) ∧ (FBP_TOF)
s60 = (DCrϕ_Ta) ∧ (LAr_EW) ∧ (¬ RZ_veto) ∧ (FBP_TOF)

Individual trigger elements use the information of the following subdetectors:

(ZVTX_t0): at least one track found by the z-vertex trigger.
The z-vertex trigger uses the information from the MWPCs to determine the event vertex position along the beam axis. The event vertex is found by (geometrical) interconnection of every pair of signals from the cathode pads of MWPCs with straight lines in the (r,z)-plane. The intersection of this straight line with the z-axis leads to an entry in a 16 bin wide histogram (-44 cm $< z <$ 44 cm). The vertex is found by searching for a peak in the histogram. This trigger element asks for at least one entry in the histogram.

(DCrϕ): The DCrϕ trigger uses the information from two central $r\phi$-drift chambers to find tracks of charged particles originating from the beam axis. A track in the (r,ϕ)-plane translates to a special bit-pattern (r is given by the position of the sense wire, ϕ is reflected by the bin of the drift time measured by signal wires). The trigger searches for bit-patterns that agree with predefined masks corresponding to track candidates with $p_T > 400\,\mathrm{MeV}$. The trigger elements include information on the transverse momentum, multiplicity and the sign of curvature. In order to reduce the background sufficiently the DC$r\phi$ trigger is only sensitive to tracks whose distance to the beam axis is small ($|DCA| < 2\,\mathrm{cm}$).

- (DC$r\phi$_T0): at least one track candidate in CJC at the time t0
- (DC$r\phi$_Tc): at least three track candidates in CJC
- (DC$r\phi$_TNeg): at least one track candidate in CJC with negative curvature
- (DC$r\phi$_THig): at least one track candidate in CJC with $p_T > 800\,\mathrm{MeV}$
- (DC$r\phi$_Ta): at least one track candidate in CJC

($\neg$ RZ_veto): This trigger element is based on fast histogramming of hits in the z-chambers CIZ and COZ. It gives an estimate of the z-position of the event vertex. Events with too many tracks pointing outside the interaction region are rejected.

(FBP_TOF): In addition to the intrinsic TOF condition of the SpaCal IET trigger, all subtriggers require the absence of out-of-time background signals. This condition is realized using a background veto signal from TOF counters FTOF, BTOF and PTOF and the Veto Wall.

($\neg$ SPCLe_IET_cen): no activity in the central IET bit

($\neg$ SPCLe_ATOF): no activity in the electromagnetic SpaCal at wrong time

(LU_PD): set if an energy in the photon detector is in certain range

(LAr_BW, LAr_BR, LAr_EW): these trigger elements are set if a sum over all Big Towers deposits in central barrel of LAr calorimeter and SpaCal is above threshold, if Big Tower deposit in barrel is above threshold and validated by a track in MWPC, if a sum over all Big Towers weighted by a profile is above threshold, respectively.

For each of the SpaCal subtriggers a minimum energy deposit in the SpaCal is required. IET>0, IET>1 and IET>2 refer to different thresholds of the SpaCal IET trigger, providing full trigger efficiency at 5 GeV, 8 GeV and $\sim$23 GeV, respectively [63]. The event weight is then calculated according to the formulae [62]

$$\omega_j = \frac{\sum_{k=1}^{N_{runs}} L_k}{\sum_{k=1}^{N_{runs}} L_k P_{jk}}, \quad P_{jk} = 1 - \prod_{i=1}^{N_{subtr}} (1 - \frac{r_{ij}}{d_{ik}}) \tag{4.4}$$

In Eq.(4.4) L_k is the integrated luminosity of run k and P_{jk} is the probability that at least one of N_{subtr} subtriggers triggers the event j in run k for which we need to know the down-scaling factors d_{ik} of subtriggers i in run k. r_{ij} is equal to 1 if raw subtrigger i is set in the event j. This procedure gives the spectrum of the event weights shown in Fig.4.6.

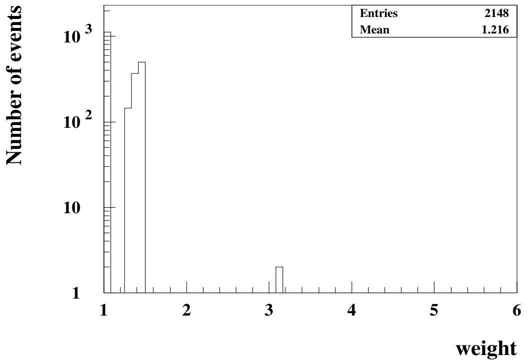


Figure 4.6: Spectrum of the event weights calculated for set of subtriggers (4.3) using formulae (4.4).

4.2.6 Trigger efficiency

The "jetty" subtriggers were used only for more effective exploitation of available statistics. They were not taken into account unless one of the raw "SpaCal" subtriggers was set on. Therefore the efficiency was determined only for the "SpaCal" part of subtriggers. The efficiency can be assessed with the help of either a monitor trigger or a trigger completely independent of the set of tested subtriggers. The monitor trigger has looser trigger conditions (lower energy threshold etc.) than the set of subtriggers that are tested. In the sample of events that we selected applying *all* the cuts (i.e. including the trigger conditions) the tested subtriggers and consequently the monitor raw triggers were set on. This means that in events fulfilling all the criteria *but* the trigger conditions the requirement that the monitor trigger triggers the events (which defines the trigger sample) should be identical to the requirement that the tested subtriggers do so if the efficiency of the latter is just 100%. The efficiency is then given as the ratio of the number of events in the trigger sample where one of the raw subtriggers s2, s3, s4, s12 was on and the total number of events in the trigger sample.

$$\epsilon_{trigger} = \frac{n(\mathrm{Tr\_indep}_{ACTUAL} \wedge \mathrm{Tr\_tested}_{RAW})}{n(\mathrm{Tr\_indep}_{ACTUAL})} \qquad (4.5)$$

where the abbreviation *indep* means either independent or monitor trigger. The monitor trigger for s2, s3, s4, s12 was trigger s0 whose definition is

$$\mathsf{s0} \ = (\mathsf{IET} > 1) \wedge (\mathsf{FBP\_TOF}).$$

Its efficiency is 100% [51]. The found trigger efficiency of the subtrigger set studied is shown in Fig.4.7 where its dependence on various quantities is plotted. All the distributions correspond to the virtuality region studied here, $1.4 < Q^2 < 25\,\mathrm{GeV}^2$, while the Q^2 spectrum itself is extended up to $Q^2 = 50\,\mathrm{GeV}^2$. Observed dependences are within statistical errors fairly constant in all of the distributions. This fact together with numerical values in the ratio in Eq.(4.5) ($365/374 = 0.98 \pm 0.01$) leads us to the conclusion that the inefficiency of the set of subtriggers s2, s3, s4, s12 is a global number +3% which will be included to the total systematic error.

The L1 subtriggers are able to trigger on genuine *ep* collisions with high efficiency and at the same time to reject background processes effectively. The main background processes are explained in the next section where also their relics in the analysis are estimated.

The loss of events due to the software filters L4 and L5 has been checked and found negligible.

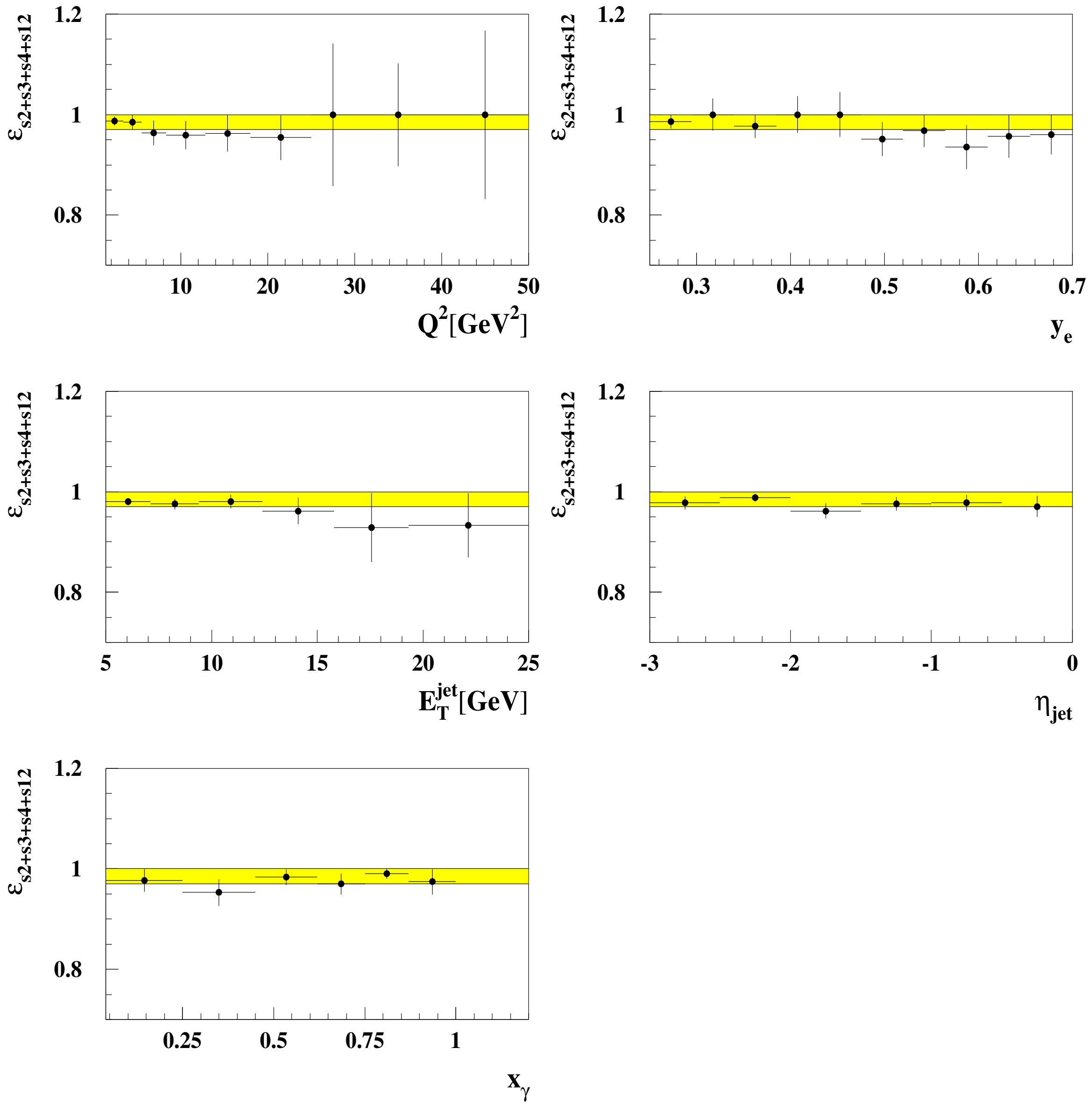


Figure 4.7: The trigger efficiency of the set of subtriggers s2, s3, s4 and s12 whose definitions are described in the text. The shaded region shows a 3% - band of inefficiency. All the distributions correspond to the virtuality range $1.4 < Q^2 < 25\,\mathrm{GeV}^2$, the Q^2 spectrum itself is extended up to $Q^2 = 50\,\mathrm{GeV}^2$.

4.2.7 Background

Non-ep background

We distinguish several types of the background induced by non-ep interactions.

PROTON INDUCED BACKGROUND: the beam protons may collide with molecules from the residual gas within the beam pipe (proton-gas background) or they may hit the walls or other structures of the beam pipe (proton-wall background). Both processes can either occur within the H1 detector or upstream, i.e. behind the time-of-flight system.

ELECTRON INDUCED BACKGROUND: the main electron induced background stems from synchrotron radiation of the electron beam as well as from bremsstrahlung on the residual gas. Both background processes occur within and outside the H1 detector and may produce a signal in the detector or the luminosity system.

Another possible source of this background is the interaction of the electron with residual gas. It may take place within or outside the H1 detector.

COSMIC MUONS: events are rejected if two tracks of opposite charge can be fitted to a single track and the total number of tracks is less than 7.

HALO MUONS: halo muons cross the detector horizontally and deposit a large transverse energy flow. Halo muons are rejected by looking for horizontal energy bands (i.e. fixed r and ϕ) in the calorimeters.

COHERENT NOISE: some events with large transverse energy are due to coherent (electronic) noise in the LAr calorimeter. These events are rejected by searching for a typical (channel related) pattern of distributed energy.

An efficient reduction of the background from non-ep interactions is provided by the minimum electron energy and the vertex requirement. The residual electron (proton)-induced background was estimated by analyzing the so-called electron (proton) pilot bunch events in which these pilot bunches do not have their colliding partners. The contamination is given by

$$bg = \frac{c^{tot}}{c^{pilot}} \frac{N^{tot}}{N^{pilot}} \tag{4.6}$$

The ratio c^{tot}/c^{pilot}, corresponding to currents in the respective bunches, is given in [44] where for electron (proton) bunches it is 12.37 (31.57) during the whole run period. N^{tot} is number of events after applying all the cuts, N^{pilot} is number of events surviving the same cuts in the electron (proton) pilot bunch sample. We found no event in our full data sample (amounting to 2148 events) which would have been caused by electron or proton pilot bunch. In order to make more reliable estimate we repeated the calculation using a sample with more statistics, namely having applied only the cuts $1.4 < Q^2 < 50\,\mathrm{GeV}^2$ and $0.25 < y < 0.7$. In the sample of 37532 events $bg_{electron} = 0.4 \pm 0.1\%$ and $bg_{proton} = 0.2 \pm 0.1\%$. Thus the non-$ep$ background in the full sample is considered negligible.

Photoproduction background

In the background events considered to come from photoproduction processes the scattered electron escapes the main detector along the beam pipe and the electron signal may be faked by the energy deposition in the SpaCal associated with the hadronic final state coming from the fragmentation of the photon remnant.

The photoproduction background was estimated by applying all the cuts in two Monte Carlo event samples

$$\begin{array}{r l}\text{PHOJET:} & \mathcal{L} = 0.15\,\mathrm{pb}^{-1},\ Q^2 < 1.00\,\mathrm{GeV}^2,\ 0.25 < y < 1.00 \\ \text{HERWIG:} & \mathcal{L} \approx 1\,\mathrm{pb}^{-1},\ Q^2 < 0.01\,\mathrm{GeV}^2,\ 0.18 < y < 0.75,\ p_T^{min} = 2.5\,\mathrm{GeV}\end{array}$$

where in HERWIG, in contrast to PHOJET, predominantly hard jets are produced. These files were not generated in the full range of y but they still should be considered representative as

1. the minimum energy cut for the electron candidate cluster ensures that events with low y are not able to fake electrons in the SpaCal (the cases with $y < 0.25$ would have to be caused by a hadronic activity concentrated in one SpaCal cluster of energy $E'_e > 20.7$ GeV!)

2. in general, for DIS processes the region $y > 0.7$ is more dangerous. There is more energy available for photon remnant pointing backward. But as the acceptance of the electron tagger is satisfactory only up to $y = 0.7$ the influence of the photoproduction background has to be estimated from data (see method based on the rejection of events with negative CJC tracks [64]). It is however reasonable to believe that the requirement on the $E - p_z$ quantity and in particular the presence of two hard jets reduces the contamination by the photoproduction background to negligible level in the high y region.

The estimates based on the number of events which passed all the analysis selection criteria give 0.1% contamination in PHOJET sample and 0% in HERWIG sample.

We realize of course that there are some cases where genuine low energy electrons entering the SpaCal under small angles may be misidentified by a collimated deposit of the hadrons from photon remnant but the numerous cuts on electron candidate cluster reject all such cases.

Thus we can conclude that the background resulting from photoproduction events is negligible.

Diffractive background

Diffractive events, characterized by a large rapidity gap, are excluded from the analysis. For detailed conditions see [60].

4.3 Photoproduction sample ($\mathbf{Q^2 < 10^{-2}\ GeV^2}$)

4.3.1 Run selection

The same components of the H1 detector as for the *low Q^2 sample* were used. The key device is the electron tagger as it serves as an identifier and one of the main parts of the trigger for photoproduction events. We analysed the same run period as in the case of the low Q^2 events where the runs were already carefully surveyed and checked. The statistics available is still several times higher compared to the *low Q^2 sample* due to rising *ep* cross section with decreasing Q^2.

4.3.2 Electron identification

The efficiency of the electron identification is given by the electron tagger acceptance.

- For good containment of the electromagnetic shower in the electron tagger we restrict the x-coordinate of the shower centre to $|x_{etag}| < 6.5\,\mathrm{cm}$ (boundaries at $x_{etag}^{bound} = \pm 7.7\,\mathrm{cm}$, Moliere radius $R_M = 2.1\,\mathrm{cm}$).
- The electron tagger "sees" the electron if this is scattered through the polar angle $\theta' < 5\,\mathrm{mrad}$ and its energy lies in the range $5.5 < E_{etag} < 22\,\mathrm{GeV}$. This corresponds to $Q^2 \lesssim 0.01\,\mathrm{GeV}^2$. The acceptance of the detector depends on y and Q^2 and also on the position of the electron beam. The dependence on Q^2 can be neglected, the dependence on y is taken into account in this analysis. Due to limited acceptance of the electron tagger the energy of the scattered electron has to be restricted to the range $8 \lesssim E_{etag} \lesssim 19\,\mathrm{GeV}$ which corresponds to the range $0.3 < y_{etag} < 0.7$. This cut also reduces the possible background due to random coincidences of proton and electron induced background events within the same bunch crossing (see Section 4.3.6).
- The z-coordinate of the interaction vertex lies within 35 cm of the nominal position.

4.3.3 Clusters and tracks

Both methods used in the *low Q^2 sample* (named ELAN and CLIM) were also made use of for *photoproduction data*. Here the ELAN procedure has a rather bigger influence on the resulting backward hadron flow compared to the *low Q^2 data* especially for the two facts mentioned in the end of the Section 4.2.4. Indeed, the rate of resolved photon processes in the dijet sample falls with increasing Q^2 which in photoproduction results in a sizable cross section in the low x_γ region allowing the photon remnant to carry away a big part of the incoming photon momentum. Nevertheless, for reasons cited in Section 4.2.4 the requirement on the balance between the electron and hadron system forced the introduction of the CLIM procedure here too. The only quantity usable for a study of the balance is the third one in expressions (4.2) since the others lose their meaning when the scattered electron has a negligible transverse momentum. The distributions of the third quantity from (4.2) and $E - p_z$ after application of the ELAN and CLIM procedures are plotted in Fig.4.8. Mean values close to the required ones were achieved for $p_{lim} = 350\,\mathrm{MeV}$ in the data and $p_{lim} = 400\,\mathrm{MeV}$ in the MC event sample together with smaller correction factors involved in the ELAN procedure. From comparison with Fig.4.2 it is seen that QED radiations in photoproduction play a negligible role.

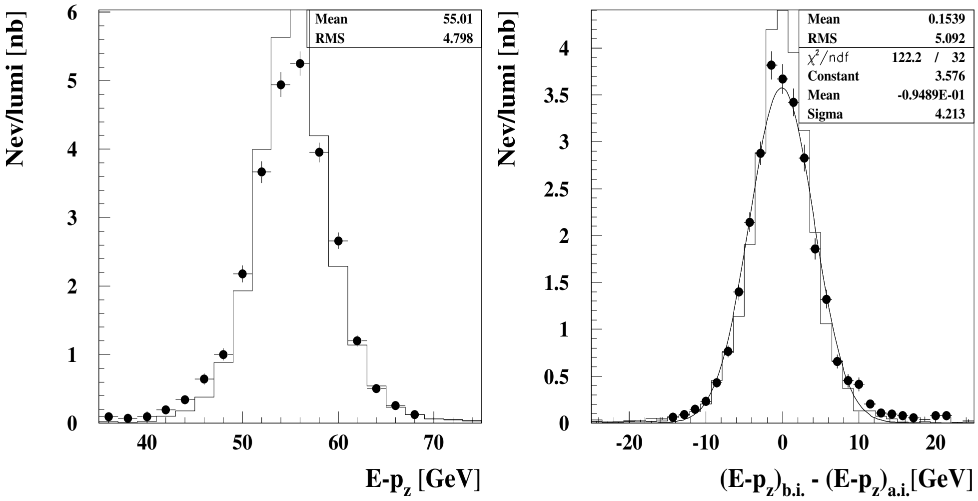


Figure 4.8: The distribution of the third quantity from (4.2) and $E-p_z$. The full points (full histogram) represent the data (HERWIG 5.9) after implementation of the ELAN and CLIM procedures. The mean and RMS values are given for the data.

As a consequence of the implementation of the ELAN and CLIM procedures an improved resolution in many quantities was achieved (not shown).

Just as in the *low Q^2 sample*, events are accepted if at least two highest E_T jets satisfy the conditions

- $E_T \geq 5\,\mathrm{GeV}$.
- $-3 \leq \eta \leq 0$.

4.3.4 Trigger selection and event weights

The set of subtriggers used for photoproduction events was found by the procedure described in subsection 4.2.5. A dominant contribution to the full event sample comes from subtriggers s80, s82 and s83. Because of the high down-scale factors of s80 the final set of subtriggers used for the analysis of photoproduction events was

$$s82, s83. \tag{4.7}$$

Simplified definitions look like this:

s82 $=$ (eTAG) $\wedge$ (LAr_BR) $\wedge$ (DCRϕ_TNeg) $\wedge$ (ZVTX_t0) $\wedge$ ($\neg$ SPCLe_E_tot) $\wedge$ (FB_TOF)
s83 $=$ (eTAG) $\wedge$ (DCRϕ_TNeg) $\wedge$ (ZVTX_mul>1) $\wedge$ (ZVTX_t0) $\wedge$ ($\neg$ SPCLe_E_tot) $\wedge$ (FB_TOF)

The meanings of single trigger elements are as follows:

(eTAG): set if the following conditions are fulfilled in coincidence:

- Energy measured in the electron tagger with $E_{etag} > 4\,\mathrm{GeV}$.
- No energy measured in the photon detector.

- No signal from the water Čerenkov counter.

The latter two requirements reduce triggers due to bremsstrahlung events.

(SPCLe_E_tot): the total energy in the electromagnetic SpaCal is in a certain range

(ZVTX_mul>1): there are at least two tracks in MWPC

The other symbols are explained in the Section 4.2.5. The spectrum of the event weights for the set (4.7) calculated according to formulae 4.4 is shown in Fig.4.9.

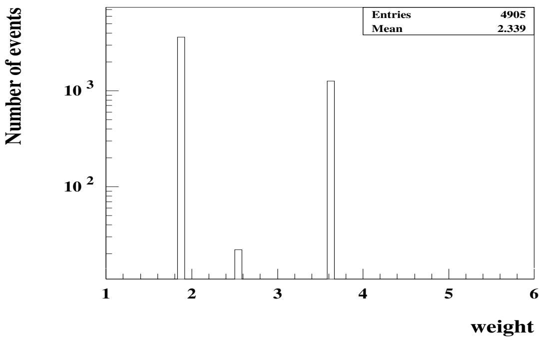


Figure 4.9: Spectrum of the event weights calculated for set of subtriggers (4.7) using formulae (4.4).

4.3.5 Trigger efficiency

The monitor trigger for subtriggers s82 and s83 was s80 whose definition is

$$\mathsf{s80} \;=\; (\mathsf{eTAG}) \wedge (\mathsf{ZVTX\_t0}) \wedge (\mathsf{B\_TOF}).$$

The efficiency of the subtrigger s80 is 100% over the entire y-region studied here [65]. The trigger efficiency of the set s82, s83 calculated according to Eq.(4.5) as depending on various quantities is shown in Fig.4.10. Due to the high down-scale factors of s80 the sample of events on which the efficiency could be tested was rather small. Within statistical errors we find no dependence of the efficiency on any of the quantities shown. The relation (4.5) expressed in numbers ($118/119 = 0.99 \pm 0.01$) gives us a global inefficiency in the trigger sample (defined in subsection 4.2.6). And similarly to the low Q^2 *sample* we conclude that the inefficiency of the set of subtriggers s82, s83 is a global number +2% which will be included in the total systematic error.

4.3.6 Background

For an estimate of the electron-induced background the procedure with electron pilot bunches described in Section 4.2.7 can be used. Following the Eq.(4.6) we again put 12.37 for c^{tot}/c^{pilot} as given in [44]. In 4905 events which survived all the cuts we found no event caused by the electron pilot bunch. In order to study this problem on higher statistics we loosened the cuts to the following conditions: $0.25 < y_{etag} < 0.7$, $E_T^{jet} > 5\,\mathrm{GeV}$ and the presence of actual subtrigger s82 or s83. In 15173 events 7 were caused by the electron pilot bunch which leads to the final estimate $bg_{electron} = 0.05 \pm 0.02\%$.

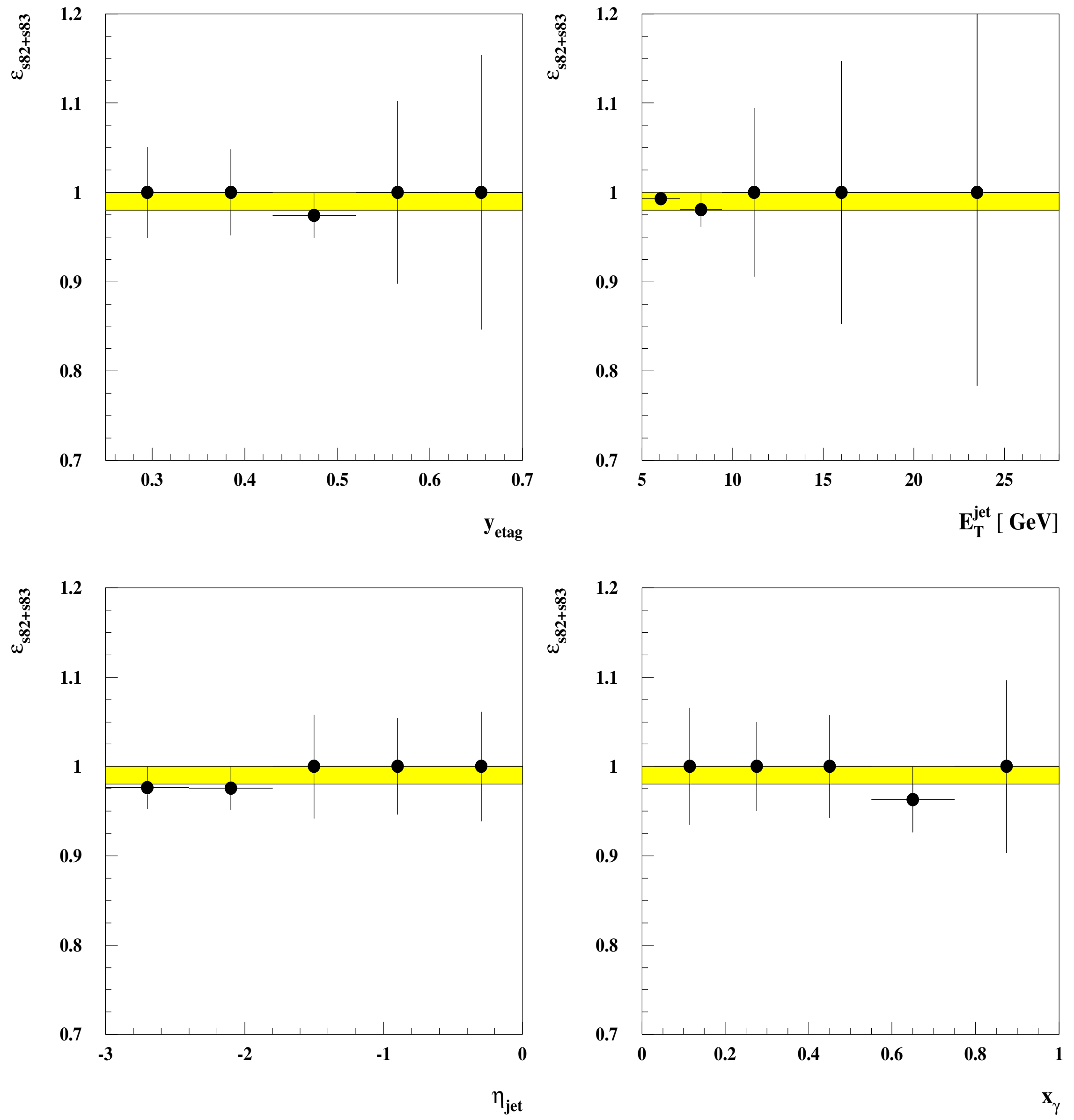


Figure 4.10: The trigger efficiency of the set of subtriggers s82, s83 in photoproduction whose definitions are described in the text. The shaded region shows a 2% - band of inefficiency.

The background from the interactions of the proton beam with the residual gas in random coincidence with a signal in the electron tagger is expected to be more significant. This background cannot be studied using events from the proton pilot bunches since our selections required a scattered electron and therefore events from proton pilot bunches are excluded. An estimate of this background can be done by analyzing the energy spectrum of the scattered electron. According to studies carried out in [66, 67] this background does not exceed the level of 1% which is due to a rejection of the events where a photon with an energy $E_\gamma > 2\,\mathrm{GeV}$ was detected in the photon tagger and due to the upper bound of the electron energy.

As in the *low Q^2 sample* the diffractive events were excluded from the *photoproduction data*.

4.4 Summary of the event selection cuts

In the Table 4.1 all the cuts explained in this chapter are summarized.

	cut #	**Low $\mathbf{Q^2}$ data** $1.4 < Q^2 < 25\,\mathrm{GeV}^2$	**Photopr. data** $Q^2 < 10^{-2}\,\mathrm{GeV}^2$
E	1	$0.25 < y_e < 0.7$	$0.3 < y_{etag} < 0.7$
l	2	$8.5\,\mathrm{ns} < t < 16.5\,\mathrm{ns}$	$\lvert x_{etag}\rvert < 6.5\,\mathrm{cm}$
e	3	$R_{cl} < 3.5\,\mathrm{cm}$	$E_\gamma < 2\,\mathrm{GeV}$
c	4	$E'_{e,had} < 0.5\,\mathrm{GeV}$	
t	5	$E_{veto} < 1\,\mathrm{GeV}$	
r	6	$\lvert\Delta R\rvert_{Sp-BDC} < 2\,\mathrm{cm}$	
o	7	$\lvert R\Delta\phi\rvert_{Sp-BDC} < 2.5\,\mathrm{cm}$	
n	8	$\lvert R_{(x,y)}\rvert > 8.7\,\mathrm{cm}$	
	9	$35 < E - p_z < 75\,\mathrm{GeV}$	
	10	$\lvert z_{vtx}\rvert < 35\,\mathrm{cm}$	
	11	$E_{T,jet} > 5\,\mathrm{GeV}$	
J	12	$-3 < \eta_{jet1,2} < 0$	
e	13	$\lvert\phi_{jet1} - \phi_{jet2}\rvert > 2$	
t	14	$E_{T,jet1} > 7$ GeV, $E_{T,jet2} > 5\,\mathrm{GeV}$	
s	15	$-2.5 < \eta_{jet1,2} < 0$	
Tr	16	s2, s3, s4, s12, s53, s56, s60	s82, s83

Table 4.1: The event selection cuts for the *low Q^2 data* (left column) and *photoproduction data* (right column) as described in the text. The cuts in rows without columns are common for both data samples.

In the following two different analyses will be presented using both data samples. In the first one, the data will be compared to Monte Carlo event generators and the effective PDF of the virtual photon will be extracted. For this, the selection cuts # 1—13 and 16 for the *low Q^2 data* will be imposed, leading to 2148 events, for photoproduction the cuts # 1—3, 10—13 and 16 will be used leaving 4905 events with us. In the second analysis, the data will be compared to NLO calculations. In the *low Q^2 data* the cuts # 1—10, 14—16 will be applied, which 1226 events will survive whereas in photoproduction the conditions # 1—3, 10 and 14—16 will be required leading to 2827 events. The cuts #13—15 will be discussed later.

Chapter 5

Comparison of the data with Monte Carlo generators

5.1 Introduction

The use of Monte Carlo event generators with LO matrix elements is twofold. Firstly, they can be used for comparisons with the data whereby we can learn much about interaction mechanism of studied processes. Secondly, they serve to correct (unfold) the data from detector to hadron or parton level which enables us to compare them with predictions of *different* LO event generators without a need to simulate the detector response. Such corrected data can be of course compared with NLO parton level calculations as well if we take the hadronization corrections into account.

The ideas put into LO event generators as well as NLO calculations reflect our current knowledge about the whole interaction mechanism. The Monte Carlo event generators have only leading order matrix elements supplemented by parton showers but they also contain non-perturbative physics; the NLO calculations include exact NLO matrix elements and α_s however they are for apparent difficulties essentially free of any non-perturbative ideas.

Whenever we want to make a comparison of uncorrected data with event generators on the detector level we have to bear in mind that distributions of quantities will be dependent on the detector acceptance and resolution effects which inhibit their interpretation and exploitation by physicists who are not in general familiar with the particular detector.

Should we attempt to make physics conclusions from comparisons with LO event generators on any level we have to be aware of (at least) these two caveats:

- It is sometimes possible to move the generator predictions towards required values by adjusting the generator parameters, including parameters used in the hadronization mechanism[1]. Adjusting the LO event generator predictions to fit the data possible NLO effects can be hidden.
- Occasionally it happens that two different generator parameters have similar or exactly opposite effects leading to a strong correlation between the parameters.

It is true that for the purpose of correction of the data any generator with any underlying physics may suffice. If we want to unfold a distribution of a given quantity from the detector to the hadron level we have to check that some correlation between values on the detector and

[1]The last remark is irrelevant for the comparisons on the parton level.

the hadron level exists or shows some functional dependence. For practical reasons (smaller number of iterations in unfolding procedure, smaller systematic errors) it is, however, better if the predictions of the MC generator are in a reasonable agreement with the data.

As the aim of this analysis is to provide information on the virtual photon structure the data have been unfolded to hadron level (which enables us to compare it with any LO event generator and NLO calculations) and to the hard scattering level (which enables us to extract the photon PDF and to compare it with any parameterisation). This chapter is thus devoted to a detailed check of whether the LO event generators used are able to describe the data and to a description of a procedure for coping with several free parameters in generation. We will also mention what was revealed by detailed studies leading to this procedure.

The most important parameters for the generation of samples used to correct for detector effects are summarized in Table 5.1.

model	sample #	process	$Q^2[\mathrm{GeV}^2]$	γ^* - PDF	p_T^{min}	N_{events}	$\mathcal{L}$ [pb^{-1}]
HERWIG 5.9	1	DIR	0.0 — 0.01	-	2.5	100 k	0.64
	2	RES	0.0 — 0.01	GRV-HO	2.5	100 k	0.08
	3	DIR	0.5 — 100.0	-	3.0	61 k	2.50
	4	RES	0.5 — 100.0	GRV-HO	3.0	73 k	2.19
RAPGAP 2.6	5	DIR	1.0 — 50.0	-	3.0	100 k	3.05
	6	RES	1.0 — 50.0	GRV-HO	3.0	100 k	3.63

Table 5.1: Characteristics of Monte Carlo event samples used for correction of the data to hadron level and for making predictions on the detector level. GRV-HO parametrization for the proton PDF was set throughout. For photoproduction the events were generated via weighting in p_T. The luminosity was calculated from all generated events and cross section modified to describe the data.

From first comparisons it was obvious that the predictions for the cross sections from HERWIG were too low to describe the data. To show this a spectrum of x_γ from the uncorrected data in the range of virtuality $8 < Q^2 < 25\,\mathrm{GeV}^2$ is compared with HERWIG results for direct processes (sample 3) in Fig.5.1. Assuming a negligible contribution from resolved photon processes in the region where direct ones dominate (i.e. close to $x_\gamma = 1$) for these values of Q^2 the peak in HERWIG should match that in the data. What we see however is a sizable discrepancy between the data and HERWIG. A simple scaling of HERWIG prediction by a factor 1.7 brings the peak from MC generator very close to the data peak.

The situation is rather more complicated concerning the resolved photon processes in HERWIG as the results depend on two parameters, namely ω (defining the Q^2-suppression of the photon structure, see Eq.(1.34)) and PRSOF (specifying the amount of SUE - see discussion in Chapter 2). Since we cannot rule out a possible multiplicative factor for the cross section we arrive at the conclusion that the HERWIG prediction depends on four parameters in this very specific region of phase space. We will name them $\{c_{res}, c_{dir}, \omega, \mathrm{PRSOF}\}$ where c_{res} and c_{dir} stand for multiplicative factors for the cross section of resolved and direct photon processes respectively.

To disentangle this issue we will exploit the *photoproduction data* in which the ω parameter drops out and thus the parameters c_{res} and PRSOF can be established with higher accuracy. We will assume that the same multiplicative factors hold for both the *photoproduction* and *low Q^2 samples.*

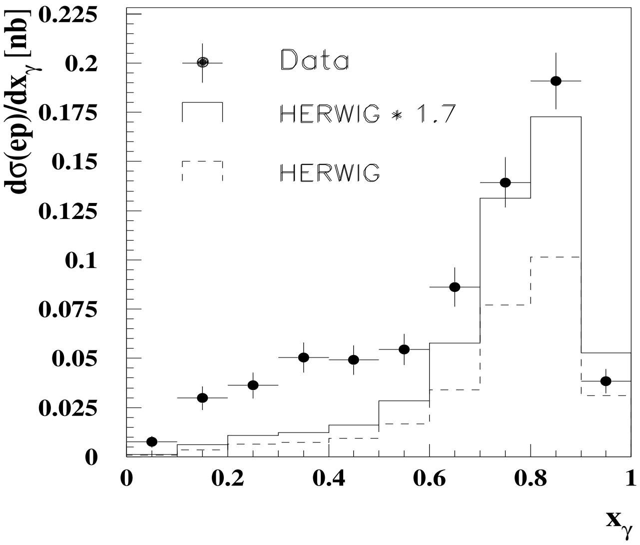


Figure 5.1: An evidence of low cross sections in HERWIG 5.9 for direct processes: a peak by $x_\gamma = 1$ predicted by HERWIG (dashed histogram) does not match the peak in the data (full points). The full histogram represents the cross section in HERWIG multiplied by a factor 1.7. All for $8 < Q^2 < 25\,\mathrm{GeV}^2$.

5.2 Fixing the free parameters on detector level

5.2.1 HERWIG

- Step 1 [photoproduction]:
  By varying c_{res}, c_{dir} and PRSOF parameters in samples 1 and 2 we arrived at several combinations for which the description of the *photoproduction data* on detector level is satisfactory:

$$\begin{aligned} &\{c_{res} = 2.1, \mathrm{PRSOF} = 0.10\}, \\ &\{c_{res} = 1.7, \mathrm{PRSOF} = 0.25\}, \\ &\{c_{res} = 1.5, \mathrm{PRSOF} = 0.40\}. \end{aligned} \tag{5.1}$$

  The influence of c_{dir} variation on the results is rather small. We found that c_{res} and c_{dir} satisfy

$$c_{res} - 0.3 \;\lesssim\; c_{dir} \;\lesssim\; c_{res} + 0.3$$

  Of course, any intermediate value of PRSOF between those above would result in some intermediate value of c_{res} (c_{dir}). However, configurations with PRSOF = 0.0 or PRSOF > 0.5 are no longer able to reproduce the shapes of distributions in the data. We made one simplifying, nevertheless important assumption:

$$c_{res} = c_{dir}.$$

- Step 2 [photoproduction]:
  We turn to the determination of the parameter PRSOF in photoproduction. The quantities most sensitive to the amount of SUE are *jet pedestals* and *energy flow outside the jets.*

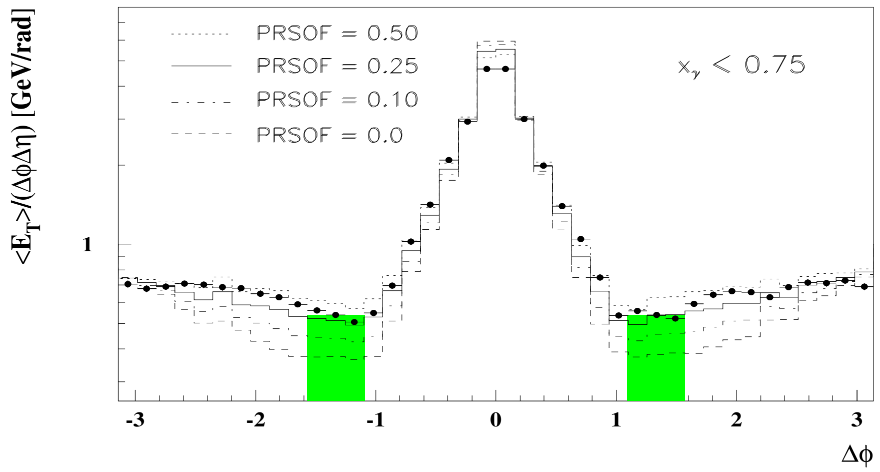


Figure 5.2: The transverse energy jet profile of the highest E_T jet as a function of the $\Delta\phi$ integrated over $|\Delta\eta| \leq 1$ for events with $x_\gamma < 0.75$. The *photoproduction data* are represented by full points with statistical errors only and also shown are the predictions of HERWIG for different values of PRSOF parameter. The hatched bands show the regions where the jet pedestals were calculated from.

– **The jet pedestals**

  The jet profiles are usually defined as a transverse energy flow per event plotted as a function of the distance of a particle from the jet axis, $\Delta\phi = \phi_{particle} - \phi_{jet}$. The particles are required to lie in the region $|\Delta\eta| = |\eta_{particle} - \eta_{jet}| \leq 1$. The term *pedestal* comes from the pictures of this kind and refers to the transverse energy plateau which can be seen just outside the jet core in Fig.5.2. The level of transverse energy rises again further from the jet core because of energy associated with the second high E_T jet.

  The jet pedestals were calculated from the transverse energy flow per unit area $\Delta\phi\Delta\eta$ outside the highest E_T jet and subsequently projected onto the $\Delta\phi$ axis. The width of the $\Delta\phi$ interval considered was 0.47 rad and taken from jet radius $R = 1.10$. In Fig.5.2 the jet profile for *photoproduction data* is shown together with prediction of HERWIG for several values of PRSOF parameter. The hatched bands indicate the regions where the jet pedestals were calculated from. The study was carried out for events with $x_\gamma < 0.75$ to approach as much as possible the resolved photon processes for which SUE was generated. The comparison of the pedestals in the data and MC suggests that a preferred value of PRSOF is 0.25 - 0.30.

– **The energy flow outside the jets**

  This definition of SUE was adopted from [42]. The transverse energy of particles is summed in some (η, ϕ)-region and then normalized to the unit area of (η, ϕ)-space and averaged over all events in the sample. The energy deposited around the axes of the two highest E_T jets within distance $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 1.3$ is excluded from the energy summation. The region of interest is depicted in Fig.5.3. The condition $R > 1.3$ ensures that we really look at the energy flow not connected with the hard jets.

  In Fig.5.4 the SUE as a function of x_γ in the *photoproduction sample* is shown together with prediction of HERWIG for different values of the parameter PRSOF. The study was performed in two regions of pseudorapidity of the particles. We

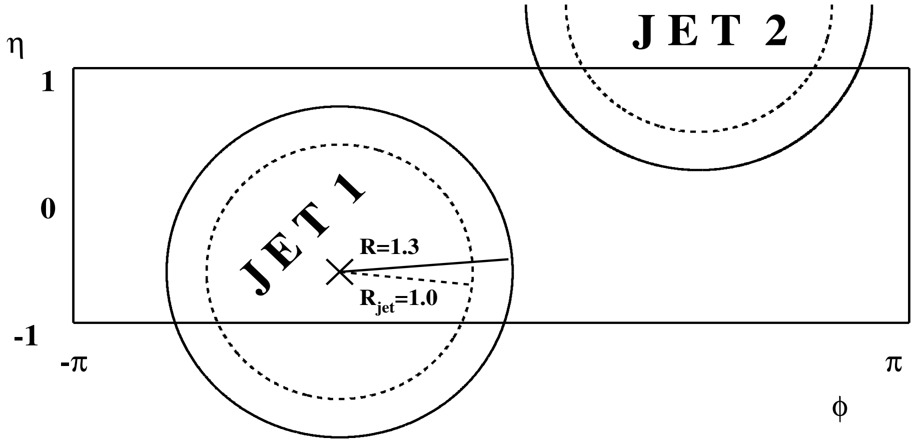


Figure 5.3: The illustration of the region where the transverse energy flow outside the jets is calculated.

conclude that HERWIG with PRSOF = 0.25 describes the data best.
Within available statistics of the *photoproduction data* and the MC samples we fix

$$\text{PRSOF} = 0.25 \pm 0.05.$$

and select the following combination:

$$\{c_{res} = c_{dir} = 1.7, \text{PRSOF} = 0.25\} \tag{5.2}$$

for description of the *photoproduction data* (all control plots will be shown later).

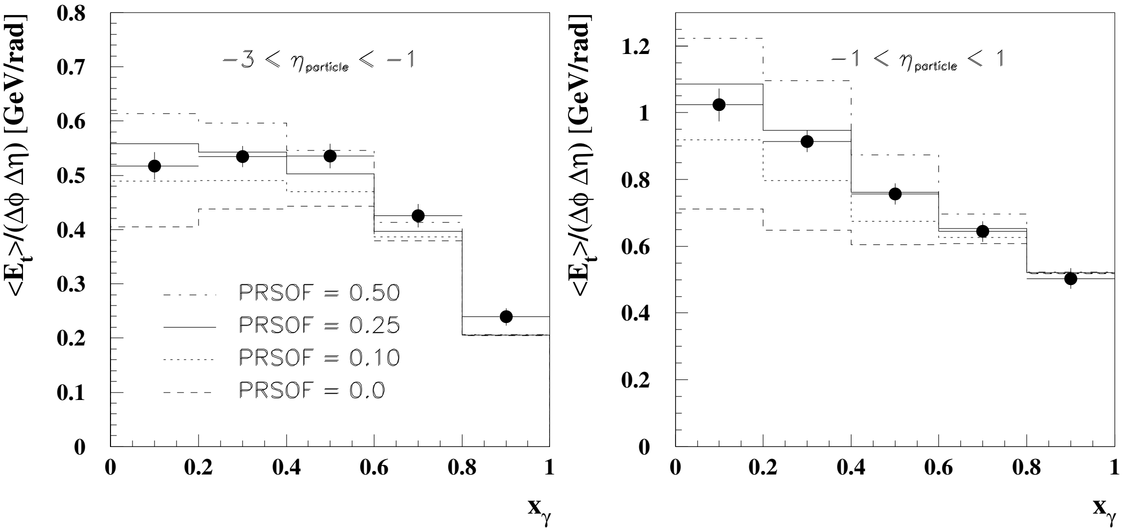


Figure 5.4: The transverse energy flow outside two highest E_T jets per unit area, $\Delta\phi\Delta\eta$, as a function x_γ. The *photoproduction data* are shown by full points with statistical errors only, different histograms correspond to indicated values of PRSOF parameter in HERWIG. The quantity was studied for two regions of pseudorapidity of particles.

- Step 3 [low Q^2]:
  As mentioned earlier, for HERWIG in the low Q^2 region (sample 4), we adopted the value of c_{res} determined in HERWIG *photoproduction* (*sample* 2). It is gratifying that the value of c_{dir} established independently in the *low* Q^2 *sample* 3 (see Fig.5.1) is consistent with c_{dir} in photoproduction. In the same way as in photoproduction we will establish the amount of SUE from the jet pedestals and the energy flow outside the jets. Both of these are ω-independent as confirmed in Fig.5.5 where the energy flow outside the jets is shown for several values of ω.

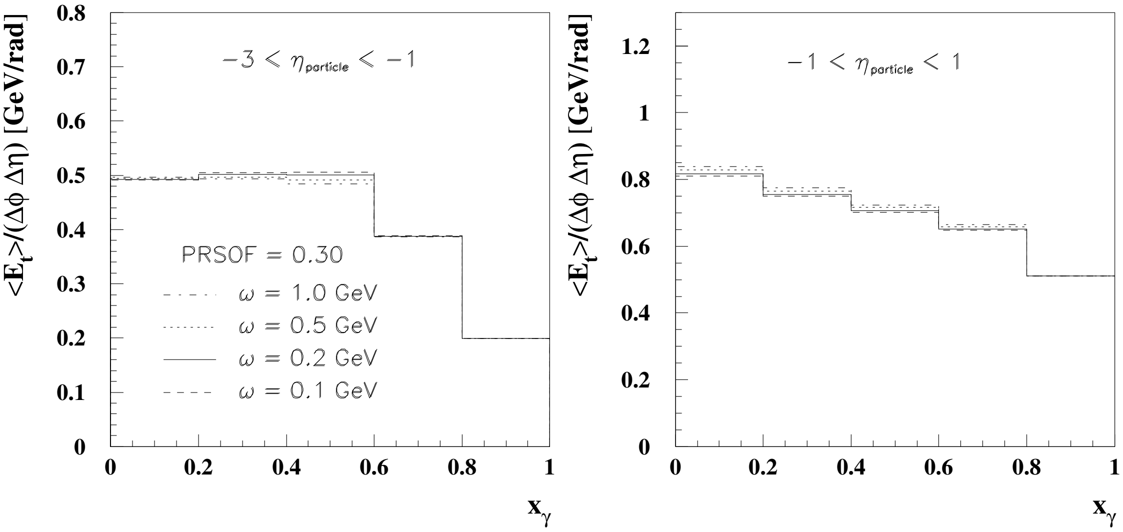


Figure 5.5: The transverse energy flow outside two highest E_T jets per unit area, $\Delta\phi\Delta\eta$, as a function of x_γ. The different histograms correspond to indicated values of ω parameter in HERWIG (PRSOF = 0.30 throughout). The quantity was studied for two regions of pseudorapidity of particles and integrated over $1.4 < Q^2 < 25\,\mathrm{GeV}^2$.

The situation is now similar to that in Step 2. Therefore we produced the same plots as in photoproduction to estimate the value of PRSOF parameter. In Fig.5.6 the Q^2-dependence of the jet pedestals up to $Q^2 = 50\,\mathrm{GeV}^2$ is shown. The photoproduction point for which the pedestals were taken from Fig.5.2 is also shown. This study was carried out for events with $x_\gamma < 0.75$ for the reason mentioned in Step 2. The data points present clear evidence of a decrease of the jet pedestal with increasing Q^2. The *low* Q^2 *data* with their statistical errors do not allow a precise determination of the value of PRSOF nevertheless the points prefer a value around 0.20.

Fig.5.7 shows the x_γ-dependence of the transverse energy flow outside the two highest E_T jets integrated over the region $1.4 < Q^2 < 25\,\mathrm{GeV}^2$. While the backward region demonstrates the same behaviour as in photoproduction, in the forward region the decrease of the flow with increasing x_γ is hardly perceptible. On the basis of the comparison of the two quantities and within statistical errors of the data and HERWIG we can say that

$$\mathrm{PRSOF} = 0.15 \pm 0.05.$$

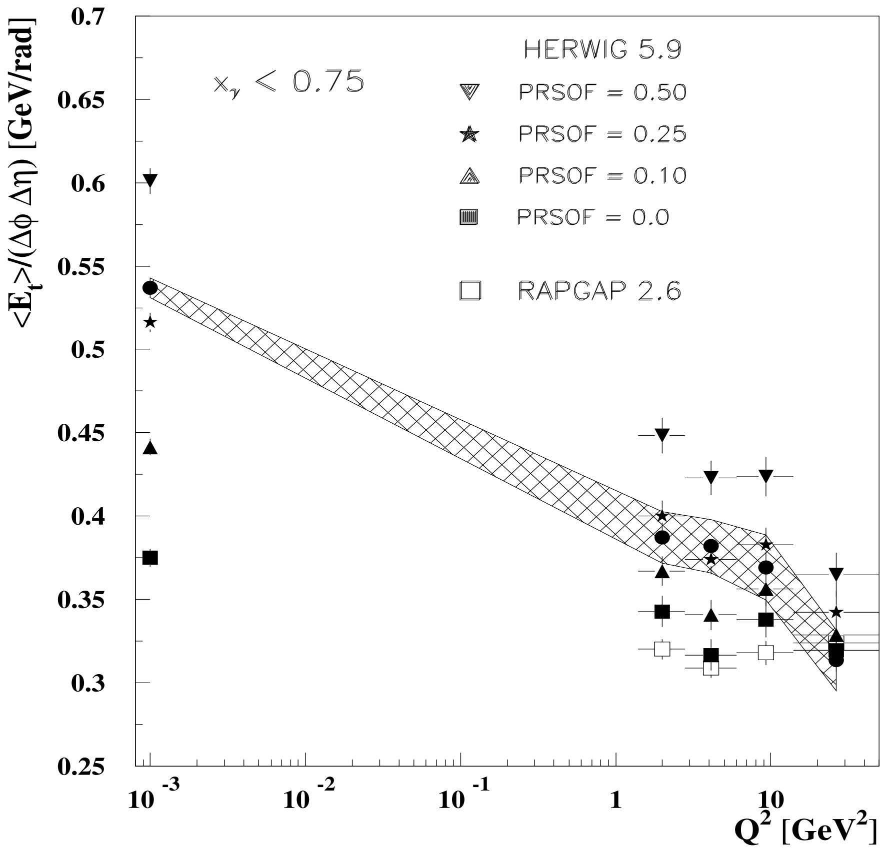


Figure 5.6: The transverse energy flow per unit area in $\Delta\phi\Delta\eta$ outside the highest E_T jet calculated from the jet profile distribution projected onto $\Delta\phi$ axis (the jet pedestals). The left most points plotted at $Q^2 = 1.10^{-3}$ GeV^2 correspond to photoproduction. The different closed symbols show the dependence of $< E_t >$ on the value of PRSOF parameter in HERWIG. Open points give RAPGAP prediction. The experimental data are shown by a hatched band; its width is given by the statistical error.

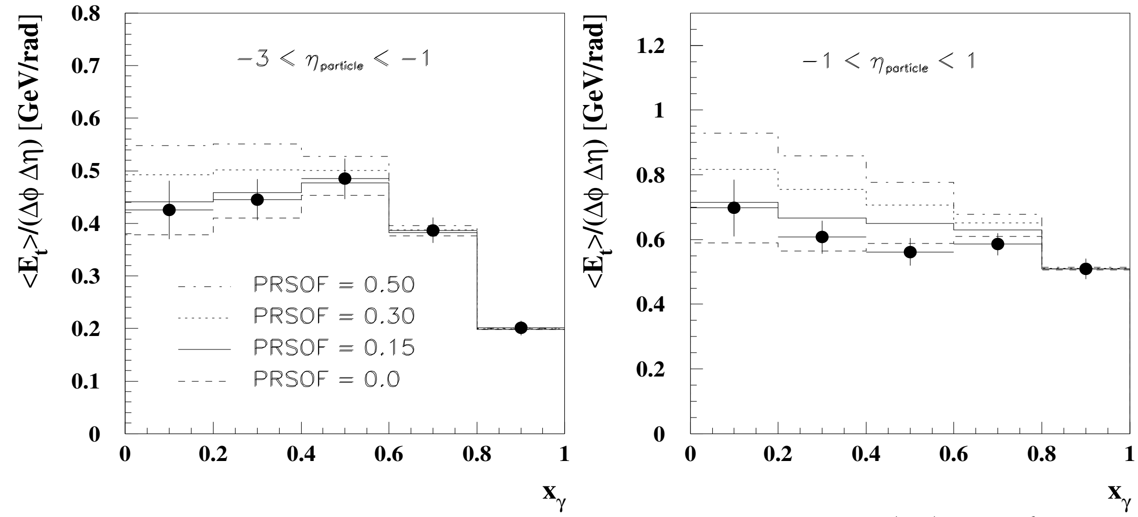


Figure 5.7: The transverse energy flow outside two highest E_T jets per unit area, $\Delta\phi\Delta\eta$, as a function of x_γ. The data, integrated over $1.4 < Q^2 < 25\,\mathrm{GeV}^2$, are shown by full points with statistical errors only, different histograms correspond to indicated values of PRSOF parameter in HERWIG. The quantity was studied for two regions of pseudorapidity of particles.

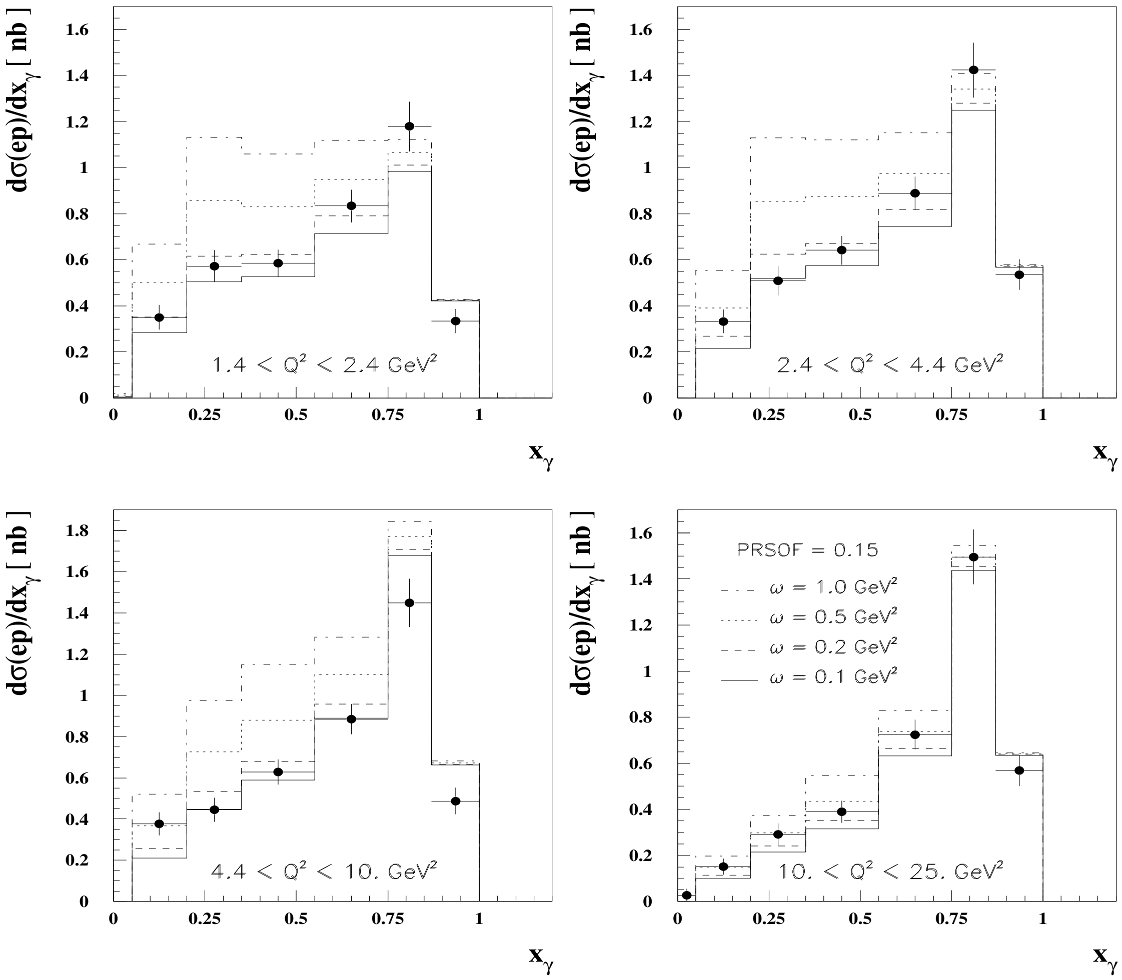


Figure 5.8: The inclusive dijet *ep* cross sections as a function of x_γ, in four Q^2 intervals (points) compared to HERWIG multiplied by 1.7 with PRSOF = 0.15. The dashed-dotted, dotted, dashed and full histograms correspond to HERWIG with $\omega = 1.0$, 0.5, 0.2 and 0.1 GeV resp.

Having three parameters c_{res}, c_{dir} and PRSOF determined the last one to be estimated in HERWIG is ω. Recalling the definition of Drees-Godbole suppression factor in resolved photon processes (1.34) we find x_γ-distribution as the most sensitive to ω. In Fig.5.8 its ω-dependence is shown for four Q^2 intervals. The three remaining parameters were set to the values just established. From the picture it is obvious that the choice $\omega = 0.2\,\text{GeV}$ describes the data most satisfactorily. This finding was also confirmed in many other distributions (will be shown later).

5.2.2 RAPGAP

As stated in Chapter 3 there is no possibility in RAPGAP of generating the additional soft underlying energy whose presence in the data was proved with help of HERWIG. A missing source of SUE in RAPGAP is clearly demonstrated in Fig.5.6 where RAPGAP symbols lie even below those of HERWIG with PRSOF = 0.0. The same relation to HERWIG is fully confirmed in Fig.5.9. The lack of SUE is also seen in distributions of other variables. This is then reflected in additional multiplicative factors for the cross sections of direct as well as resolved photon processes which have to be applied if one wants to achieve a satisfactory description of both the

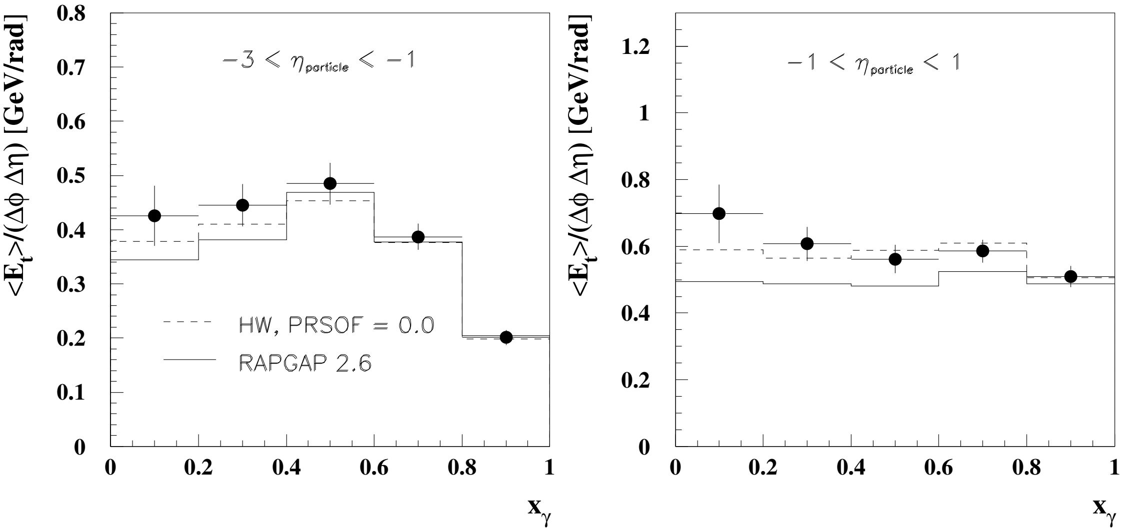


Figure 5.9: The transverse energy flow outside two highest E_T jets per unit area, $\Delta\phi\Delta\eta$, as a function of x_γ. The data, integrated over $1.4 < Q^2 < 25\,\mathrm{GeV}^2$, are shown by full points with statistical errors only, different histograms correspond to HERWIG with PRSOF = 0.0 and RAPGAP, as indicated. The quantity was studied for two regions of pseudorapidity of particles.

shape and the magnitude seen in the data.

Summarizing, the procedure described in Steps 1–3 gives for single MC event samples used for correction of the data to hadron level these optimal combinations:

$$\begin{array}{rcl} \text{Photopr.HERWIG} & : & \{c_{res} = 1.7, c_{dir} = 1.7, \mathrm{PRSOF} = 0.25\}, \\ \text{LowQ}^2\ \text{HERWIG} & : & \{c_{res} = 1.7, c_{dir} = 1.7, \mathrm{PRSOF} = 0.15, \omega = 0.2\ \mathrm{GeV}\}, \\ \text{LowQ}^2\ \text{RAPGAP} & : & \{c_{res} = 1.2, c_{dir} = 1.3, \quad \text{——} \quad, \omega = 0.2\ \mathrm{GeV}\}. \end{array} \tag{5.3}$$

The values of PRSOF are in a qualitative agreement with those obtained in [68]. It is worth stressing that we did not fit the values of mentioned parameters to match the data perfectly. The aim of this procedure was to reach a satisfactory description of the data by LO MC generators in selected distributions for the Q^2-range covered by this analysis.

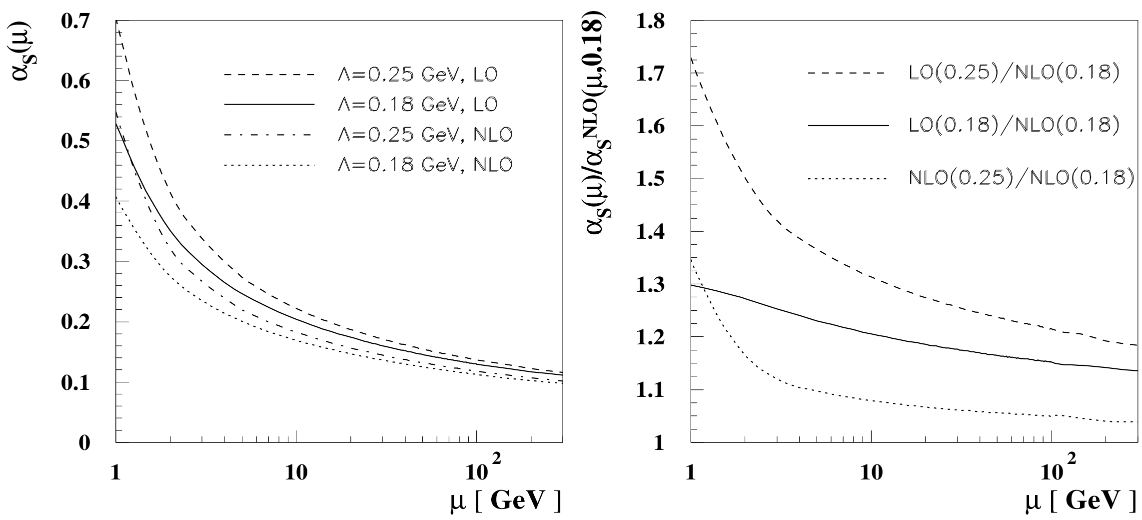


Figure 5.10: The comparison of the renormalization scale dependence of $\alpha_S^{NLO}(\mu)$ with $\alpha_S^{LO}(\mu)$ for two values of $\Lambda = 0.18$ and $0.25\,\text{GeV}$ (left) and their ratia (right).

5.2.3 Too low cross section in HERWIG for *ep* interactions: α_S^{LO} vs. α_S^{NLO}

We will attempt to shed light on the problem with too low cross sections in HERWIG for *ep* interactions. This observation was reported several times [28, 69, 70, 71]. It appears that the whole problem resides in the option taken for the order of α_s. By default, HERWIG uses the two-loop formula. The point is that for some quantities and in some kinematic regions, calculations with the two-loop formula (α_S^{NLO}) may differ significantly from those with the one-loop formula (α_S^{LO}) in the region of scales currently covered at HERA.

This is illustrated in Fig.5.10 where these two approximations are compared for two values of $\Lambda = 0.18$ and $0.25\,\text{GeV}$. In the region of scales covered by this analysis ($5 < E_T \lesssim 35\,\text{GeV}$) the ratio $\alpha_s^{LO}/\alpha_s^{NLO}$ for $\Lambda = 0.18\,\text{GeV}$ ranges between 1.23 and 1.13. Hence for scales $\mu \simeq 5$–$7\,\text{GeV}$, relevant for most of our dijet events and for direct photon processes which are proportional to α_s we get a factor around 1.2. Whereas for resolved photon processes that are proportional to α_s^2 one gets a factor 1.5, which can even increase up to 1.87 if we evaluate $\alpha_s^{LO}(\Lambda = 0.25)/\alpha_s^{NLO}(\Lambda = 0.18)$! The main effect of a change in the order of α_s or a change in the value of Λ within a fixed order of α_s is an overall change in the absolute normalization just like that observed. This fact should always be taken into account when HERWIG is compared with other generators, which usually have α_s^{LO} as a default.

If these facts can explain well the too low HERWIG cross sections for resolved photon processes, in direct ones this is not the case. See Fig.5.11 where the original HERWIG on the hadron level with α_s^{NLO} and parameter settings from (5.3) is compared to HERWIG with α_s^{LO} and without multiplicative factors. Relatively bigger proximity of these two HERWIG generations for photoproduction compared to the low Q^2 region is a result of the much smaller influence of direct processes and more frequent presence of the low generated hard scales in the photoproduction than in the low Q^2 region. In accordance with statements at the beginning of this chapter the unexplained high value of c_{dir} factor could be caused by possibly high NLO corrections and/or by insufficient influence of hadronization process (see discussion on model dependence and hadronization corrections in Sections 5.4.2, 5.5.2 and 6.7.6).

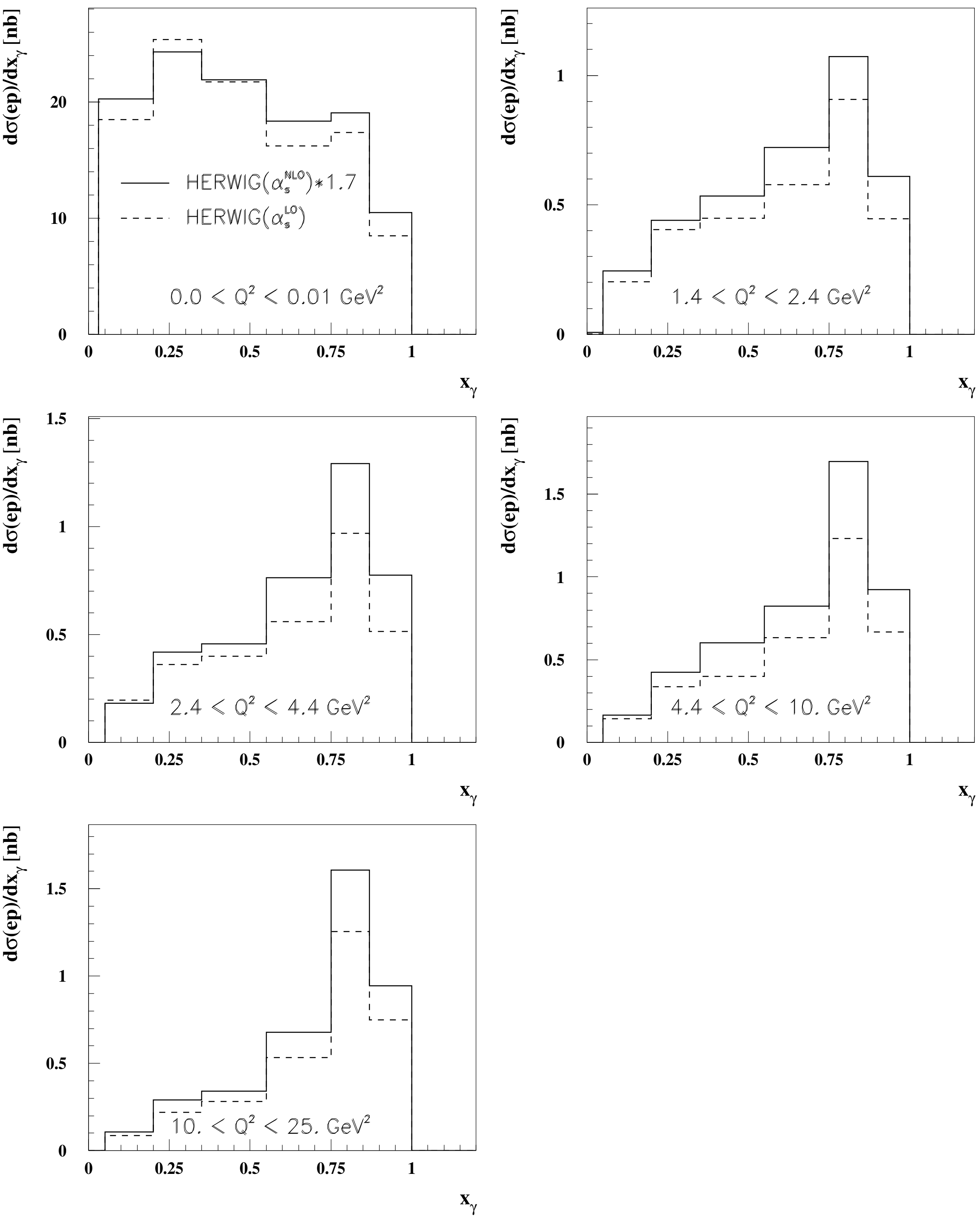


Figure 5.11: The inclusive dijet ep cross sections on hadron level as a function of x_γ shown in five Q^2 intervals. The full histogram represents HERWIG with $\alpha_S^{NLO}, c_{res} = c_{dir} = 1.7$ and GRV-HO parametrization for proton as well as photon PDF, the dashed histogram shows HERWIG with α_S^{LO}, without multiplicative factors and GRV-LO for proton and photon PDF. In photoproduction PRSOF $=$ 0.25, in the low Q^2 region PRSOF $= 0.15$ and $\omega = 0.2\,$GeV.

5.3 Comparison of the data with LO generators on detector level

In the following several control plots are shown concerning the comparison of the uncorrected data with the predictions of Monte Carlo event generator on the detector level. For the distributions shown all cuts summarized in Table 4.1 are imposed. The data will be shown by full points with statistical errors only and predictions of event generators by histograms. The full histogram represents HERWIG 5.9 and the dashed histogram RAPGAP 2.6, each with the parameter setting specified in (5.3).

5.3.1 Global characteristics

Vertex reconstruction

The position of the interaction vertex is measured using the forward and central drift chambers. The existence of the interaction vertex itself is a necessary condition since the reconstruction of the event kinematics is based on it. The condition not to take events with vertex outside the region $\pm 35\,\mathrm{cm}$ is important to suppress the cases when the colliding proton comes from an early or late proton satellite which is shifted in time by $\pm 4.8\,\mathrm{ns}$ with respect to the main proton bunch and collides with the main electron bunch around $\pm 70\,\mathrm{cm}$, respectively. The rejection of satellite bunches represents the main part of the total off-line correction of the integrated luminosity amounting to -3.6% $\pm$ 1.5% [44]. In Fig.5.12 and Fig.5.13 a slight shift between distributions of z-coordinate of the vertex in data and MC event generators is seen.

Hadronic centre of mass energy

Another important quantity of global character is the hadronic cms energy, W, calculated as $W = \sqrt{y_e \cdot s - Q_e^2}$ and shown in Fig.5.15. The spectrum of W for photoproduction is slightly different compared to the low Q^2 region due to the different y_e range but for meaningful comparisons of the results from these two Q^2 regions it is essential that the mean values of W are similar (they differ by 4%).

Balance

The distributions of such quantities like $p_{T,x}$, $p_{T,y}$ and $E - p_z$ presented in Fig.5.12 and 5.13 tell us how the electron and hadron system is balanced globally, i.e. for all selected events. The detailed studies related to this are widely commented in Chapter 4.

5.3.2 Electron Characteristics

In Fig.5.14 basic characteristics of the scattered electron are shown. Slightly different ends of spectra of the electron energy comparing the *photoproduction* and *low Q^2 samples* are due to different y_e ranges. Further, the polar angle, θ'_e, and the azimuthal angle, ϕ'_e, of the electron detected in the SpaCal are shown. The upper edge of the θ'_e-distribution is due to the fiducial cut on the SpaCal acceptance at a radius 8.7 cm (cut # 8). The distribution of ϕ'_e within statistical errors expresses a homogeneity of hits in the transverse plane of the SpaCal, as expected. Finally, the spectrum of virtualities of the exchanged photon is presented in Fig.5.14. In all quantities shown the event generators exhibit a satisfactory description.

5.3.3 Jet observables

In Fig.5.15 the number of jets satisfying all the cuts from Table 4.1 is shown. In the *photoproduction* (*low* Q^2) *sample* the number of three-jet events represents 12% (9%) of the total number of dijet events. No cut on angle in the azimuthal plane of the third jet was imposed.

In Fig.5.16, 5.17, 5.18 the spectra of the transverse energy, pseudorapidity and azimuthal angle for the highest and second highest E_T jet separately are shown in five Q^2 intervals. Both event generators give reasonable description. The distributions of azimuthal angle are flat within statistical errors, as required.

In Fig.5.19 and 5.21 the distributions of $|\Delta E_{T,jet}|$ and $|\Delta \phi_{jet}|$ are shown. In the LO picture, two outgoing partons in the diparton cms have the same transverse momentum and go back-to-back in the plane transverse to the direction of incoming partons. Largely due to initial state radiation and to a lesser extent to the intrinsic transverse momentum of the partons within photon and proton this diparton system may no longer be identical with the $\gamma^* p$ cms in which the jets are searched for. These facts together with the effect of hadronization account for the observed long tails in the two distributions.

The distribution of $|\Delta \eta_{jet}|$ (Fig.5.20) is directly related to the distribution of the scattering angle $\hat{\theta}$ of the outgoing parton in the diparton cms (cf. Eq.(2.5)). The distribution of this scattering angle is an important consequence of QCD. The description of the $\hat{\theta}$-distribution in the data is good and hence this basic prediction of QCD is well confirmed.

The fraction of the photon momentum carried by struck parton obtained using jets is plotted in Fig.5.22. The x_γ-variable was calculated according to formula (1.15). We see that the contribution of the resolved photon processes (coming from region $0 < x_\gamma \lesssim 0.75$) to the total jet cross section is a decreasing function of Q^2. Both event generators provide a satisfactory description of the data.

5.3.4 Energy flow

A good description of the jet profiles in the data by the MC generator is necessary for the unfolding of the data to the hard scattering level which is free of any soft underlying activity. In Fig.5.23, the transverse energy flow per event around the jet direction in a slice of $|\phi_{particle} - \phi_{jet}| < 1$ is shown versus the distance in η from the jet axis ("jet profile in η"). The jet profiles are given in different bins of E_T^{jet} and η_{jet}. Note that the jet profiles are asymmetric, showing a higher transverse energy flow in the direction of the proton remnant ($\Delta\eta > 0$) compared to the direction of the photon remnant ($\Delta\eta < 0$). The *photoproduction data* are reasonably described by HERWIG in the whole region of E_T and η, except perhaps for the proton remnant direction in events where the highest E_T jet goes most forward.

The jet profiles in ϕ are shown in Fig.5.24. The *low* Q^2 *data* are there compared to HERWIG and RAPGAP. While HERWIG provides a good description of the data over the whole kinematic region, RAPGAP does not reach this good level which confirms the observations found having studied the jet pedestals (Fig.5.6) and the E_t-flow outside the hard jets (Fig.5.9).

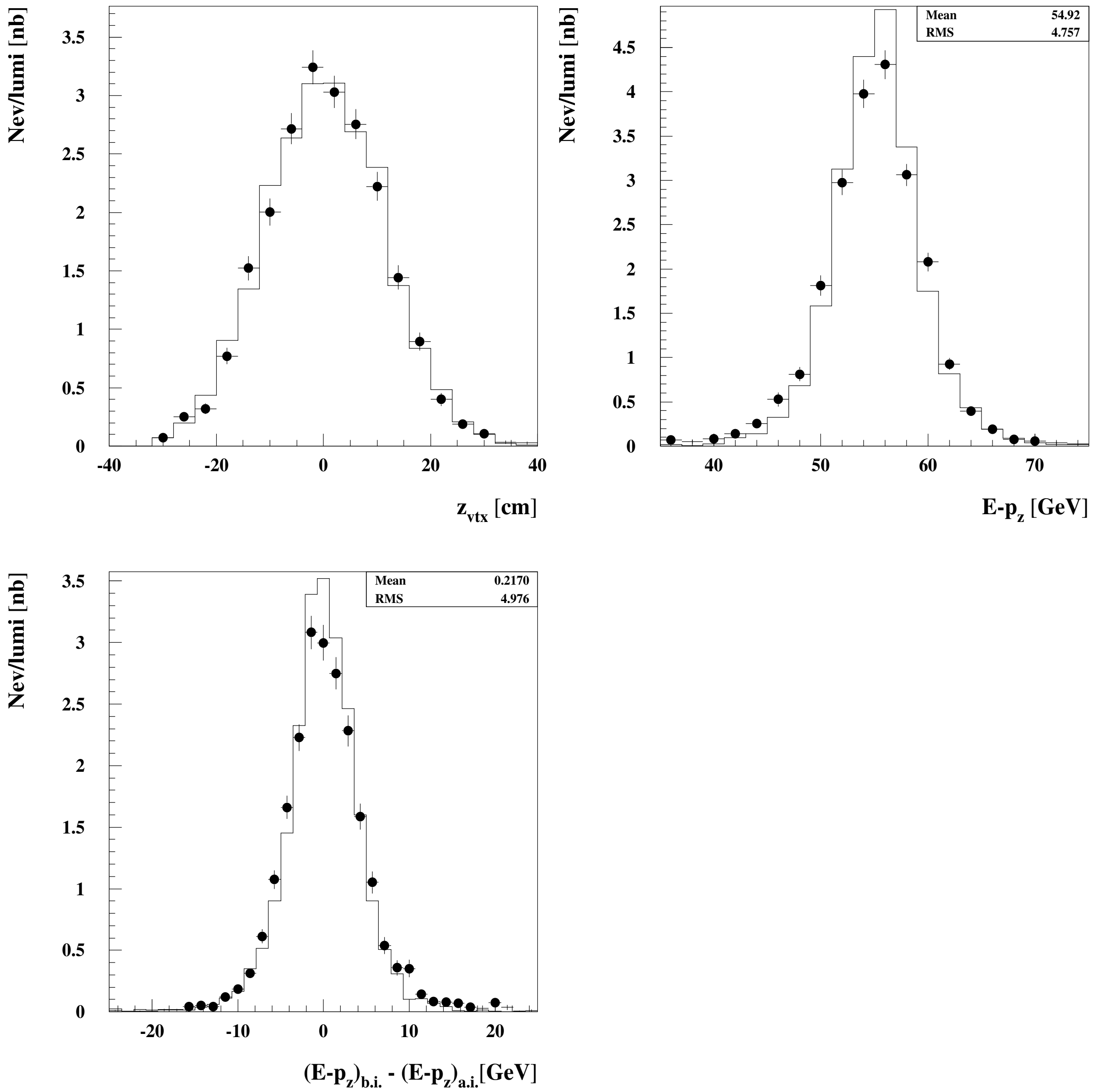


Figure 5.12: The comparison of the *photoproduction data* (full points) with a LO MC event generator on the detector level. The prediction of HERWIG 5.9 with the parameter setting $\{c_{res} = c_{dir} = 1.7$, PRSOF $= 0.25$ $\}$ is shown by the full histogram. The comparisons is made with the quantities: the z-coordinate of the interaction vertex, $E - p_z$ and difference between $E - p_z$ before and after interaction. The mean and RMS values of the distribution of the data are given.

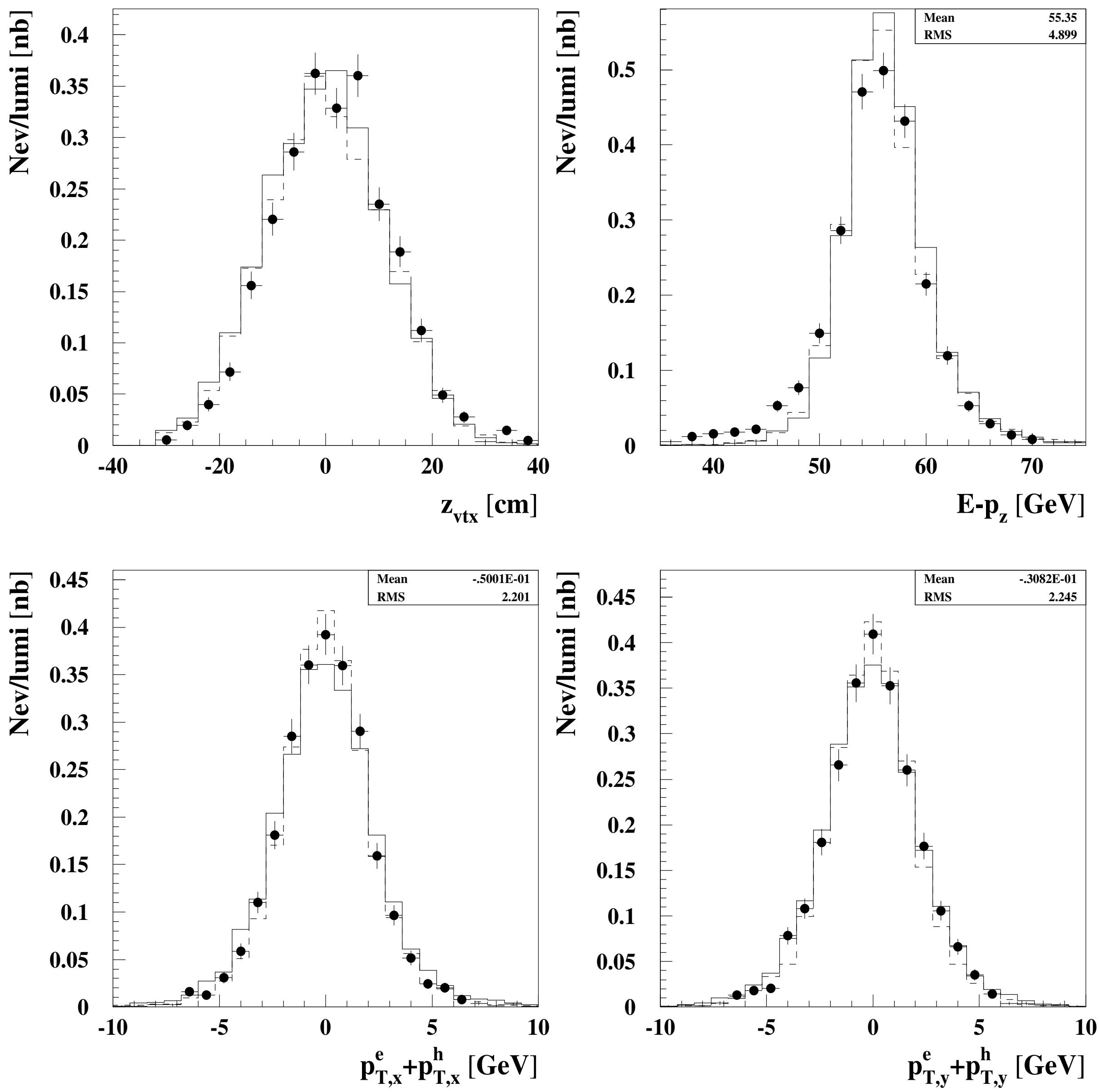


Figure 5.13: The comparison of the data (full points) with LO MC event generators on the detector level. The predictions of HERWIG 5.9 with the parameter setting $\{c_{res} = c_{dir} = 1.7$, PRSOF $= 0.15$, $\omega = 0.2\,\mathrm{GeV}\}$ are represented by full histogram, the predictions of RAPGAP 2.6 with the parameter setting $\{c_{res} = 1.2, c_{dir} = 1.3, \omega = 0.2\,\mathrm{GeV}\}$ by dashed histogram. The comparisons are made with the quantities: the z-coordinate of the interaction vertex, $E - p_z$, the x-component of the sum of the electron and hadron system momenta in transverse plane, the y-component of the latter. The mean and RMS values of the distribution of the data are given.

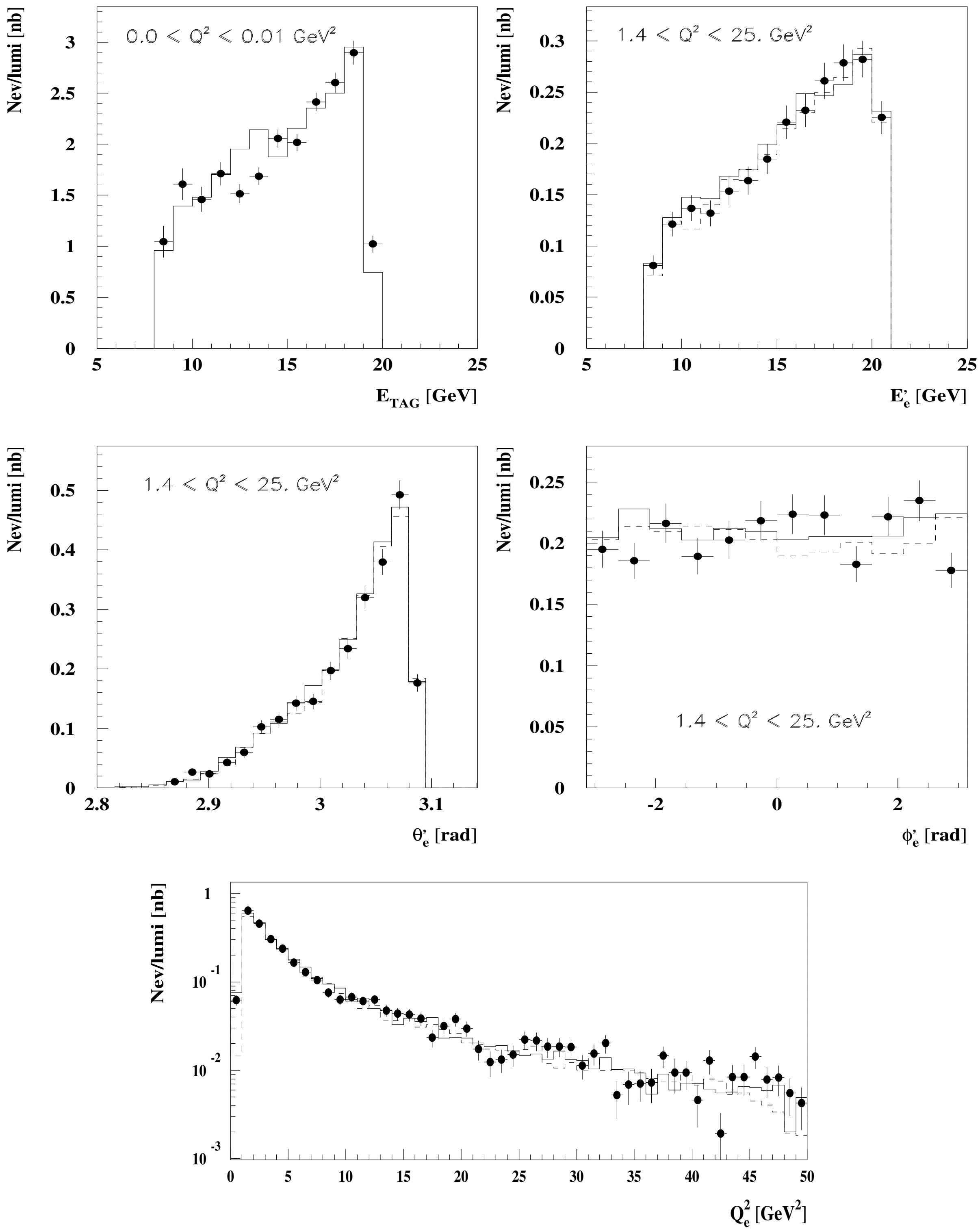


Figure 5.14: The comparison of the data (full points) with LO MC event generators on the detector level. The predictions of HERWIG 5.9 in the low Q^2 region with the parameter setting $\{c_{res} = c_{dir} = 1.7$, PRSOF = 0.15, $\omega = 0.2\,\text{GeV}\}$ are represented by full histogram, the predictions of RAPGAP 2.6 with the parameter setting $\{c_{res} = 1.2, c_{dir} = 1.3, \omega = 0.2\,\text{GeV}\}$ by dashed histogram. The parameter setting $\{c_{res} = c_{dir} = 1.7$, PRSOF = 0.25$\}$ was used in HERWIG 5.9 for photoproduction. The comparisons are made with the quantities: the energy, polar angle and azimuthal angle of the scattered electron and the virtuality of the exchanged photon.

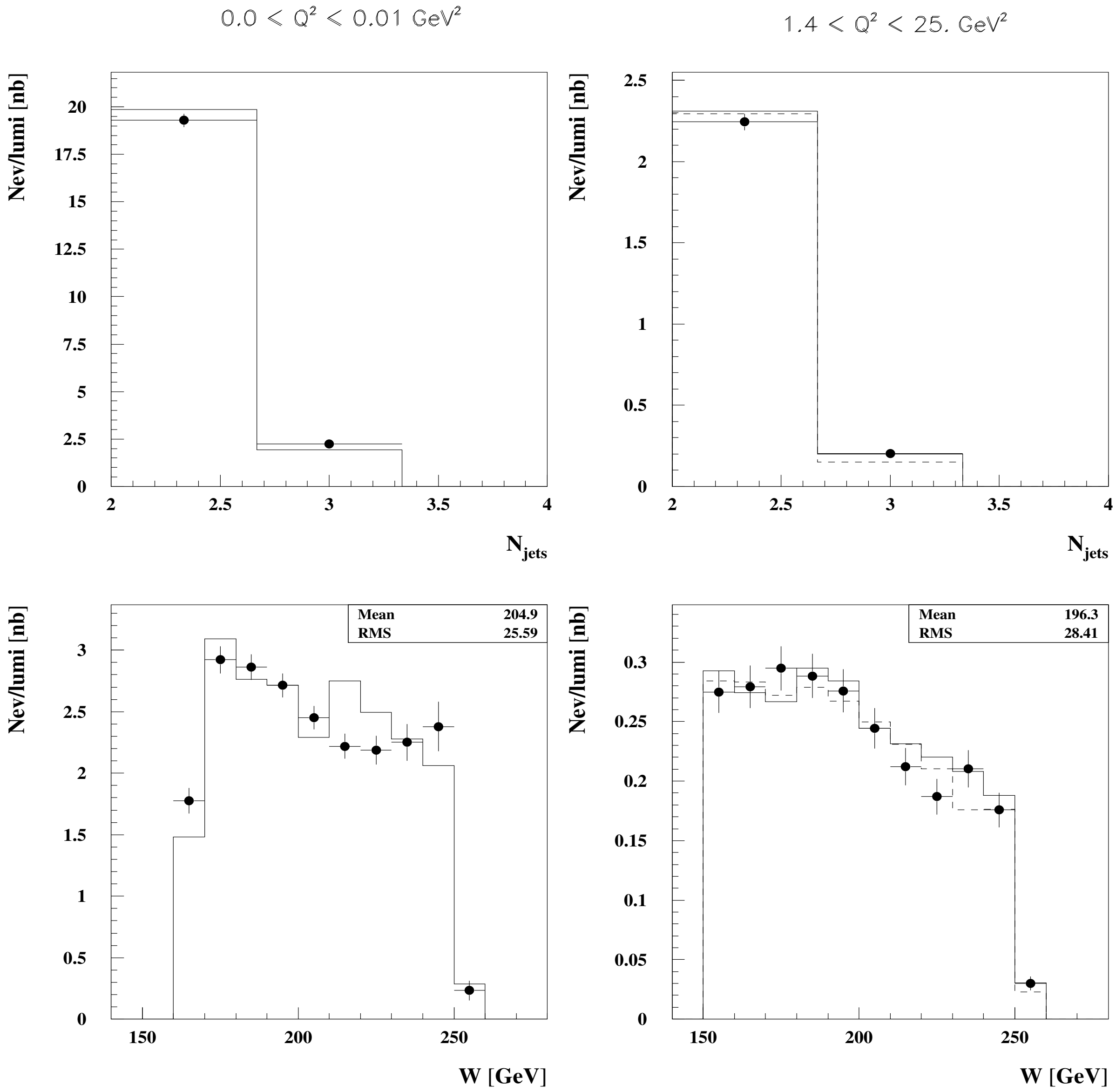


Figure 5.15: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantities: the number of jets and the hadronic centre-of-mass energy. The mean and RMS values of the distribution of the data are given. For description of event generators see the caption of Fig.5.14.

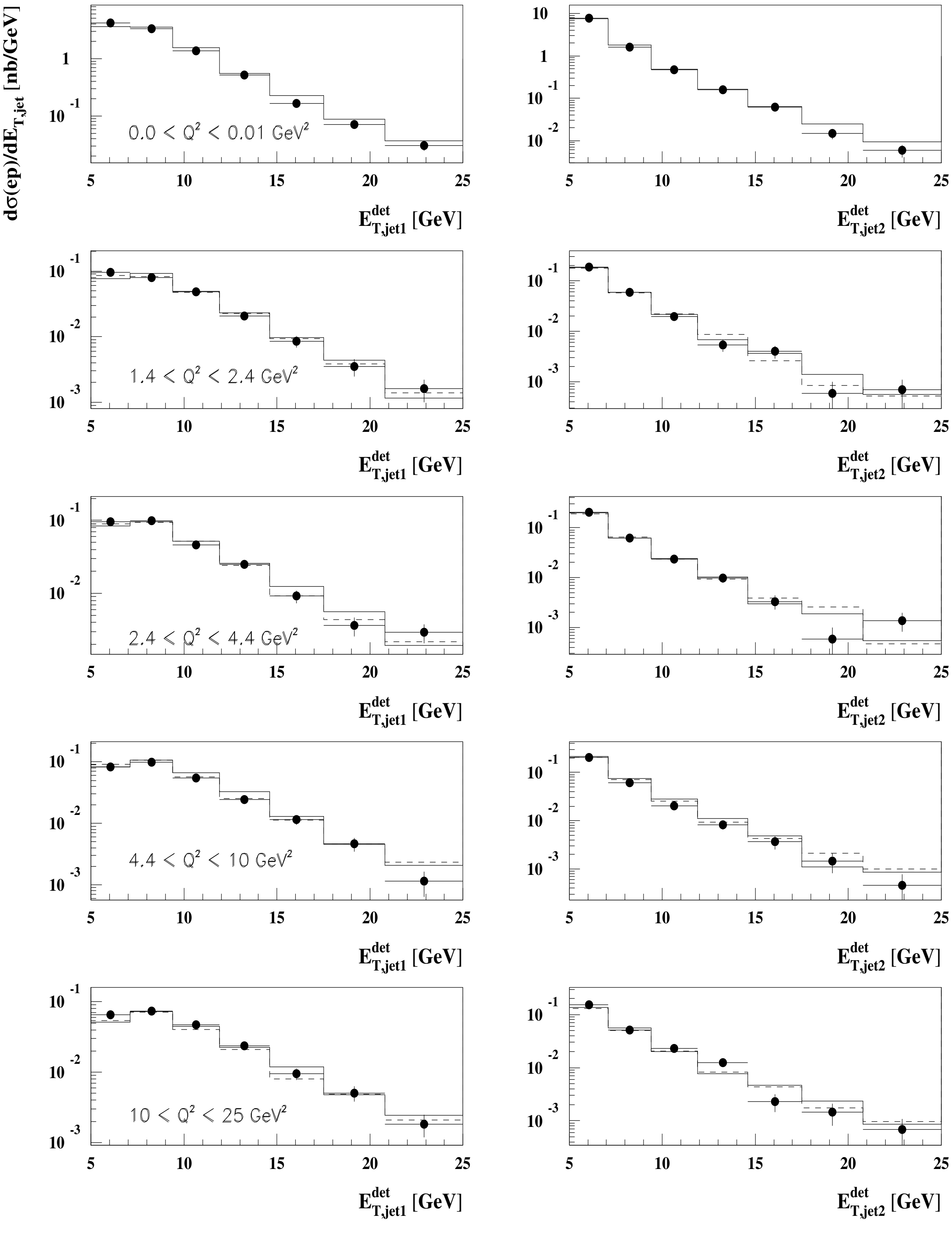


Figure 5.16: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantities: the transverse energy of the highest E_T jet (left column) and of the second highest E_T jet (right column). For description of event generators see the caption of Fig.5.14.

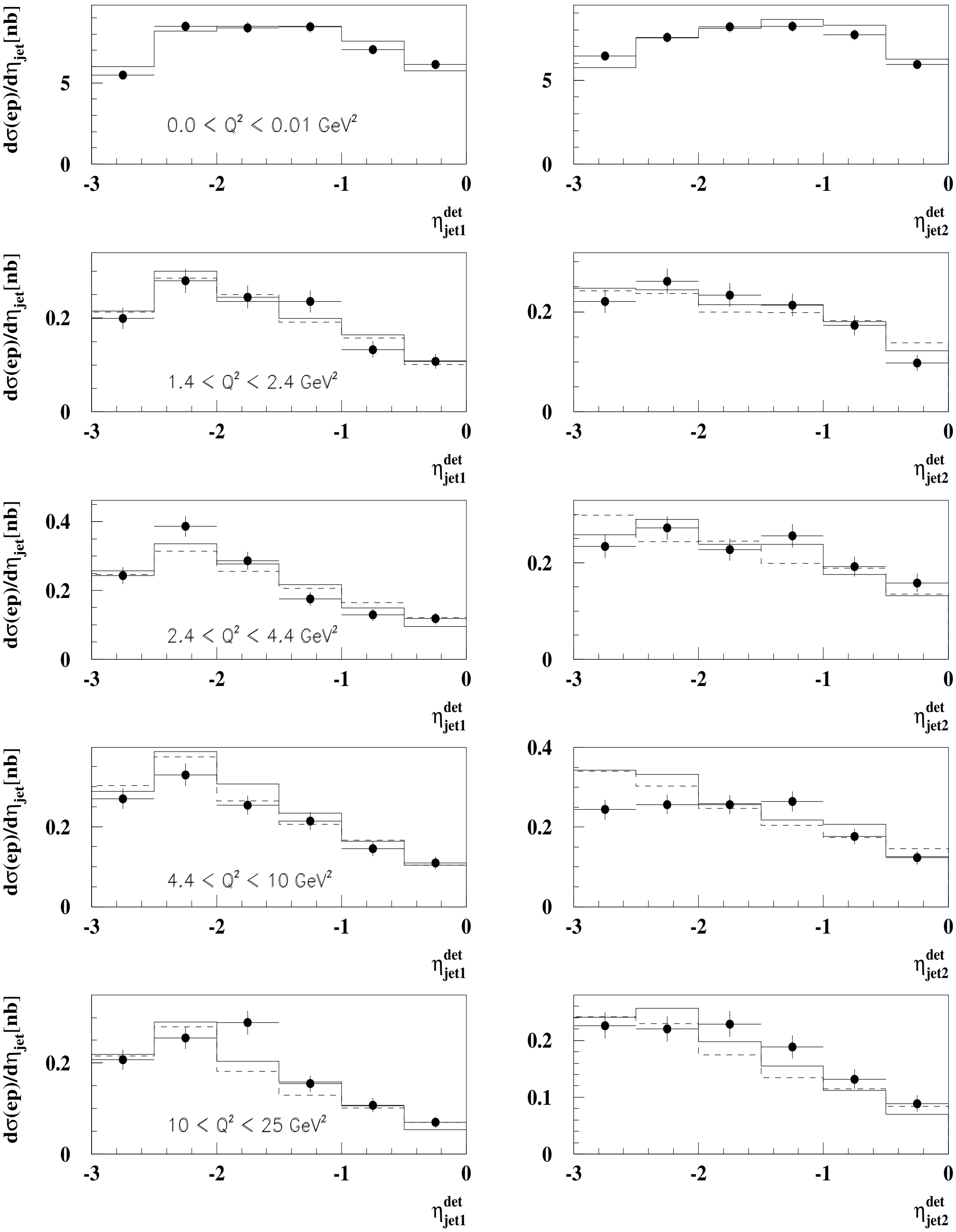


Figure 5.17: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantities: the pseudorapidity of the highest E_T jet (left column) and of the second highest E_T jet (right column). For description of event generators see the caption of Fig.5.14.

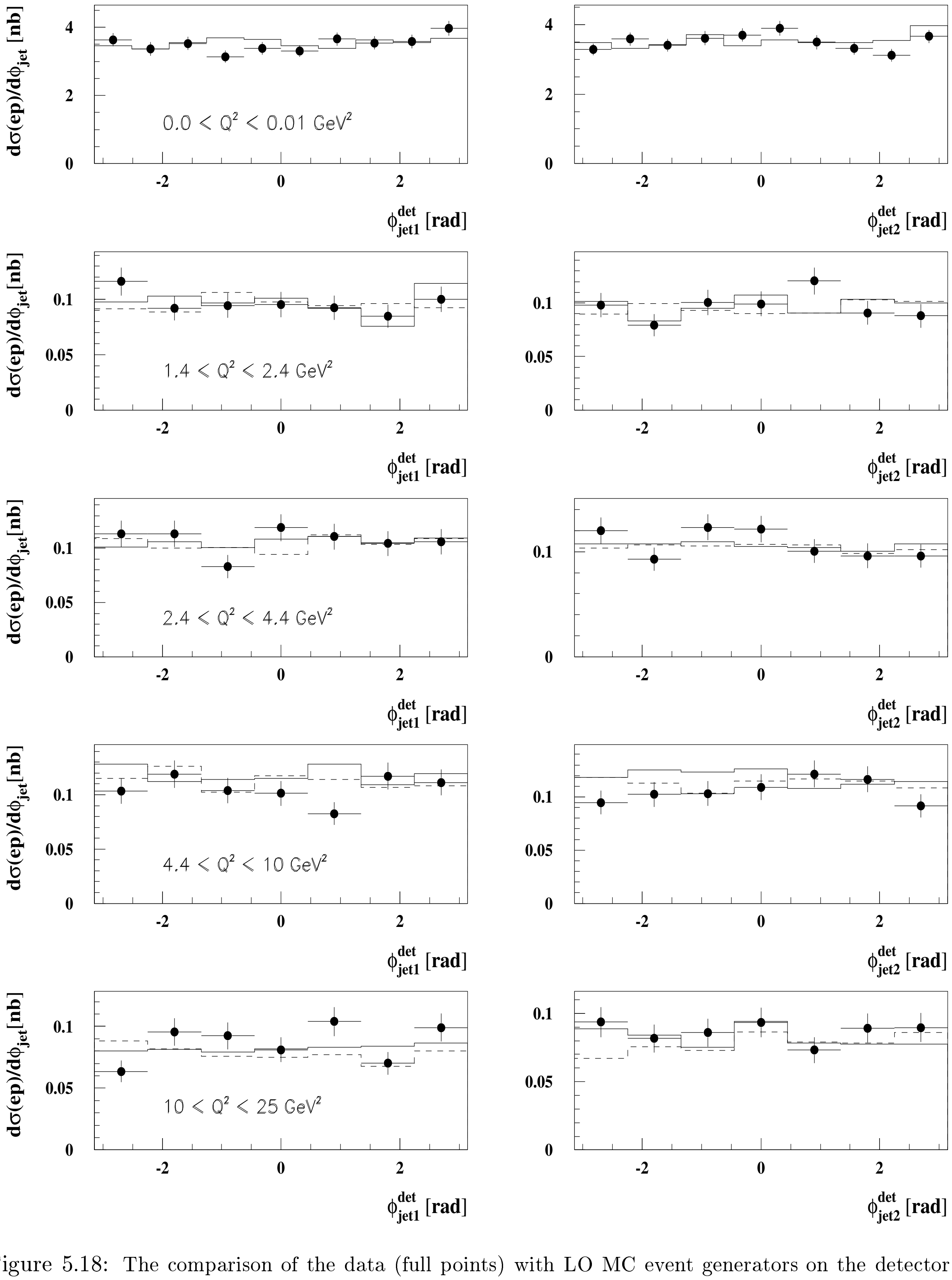


Figure 5.18: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantities: the azimuthal angle of the highest E_T jet (left column) and of the second highest E_T jet (right column). For description of event generators see the caption of Fig.5.14.

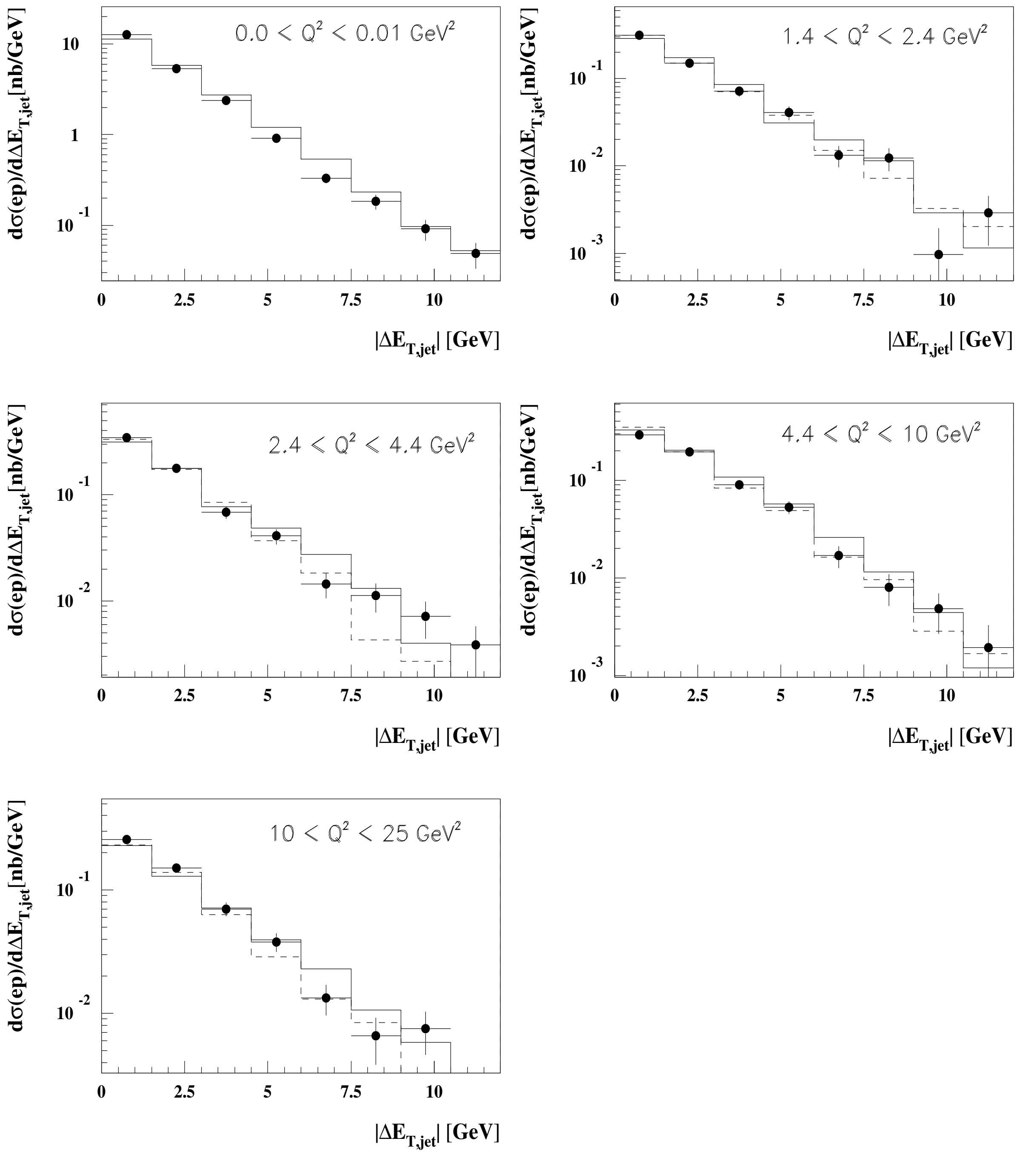


Figure 5.19: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantity: the absolute value of the difference between the transverse energy of the highest E_T and the second highest E_T jet. For description of event generators see the caption of Fig.5.14.

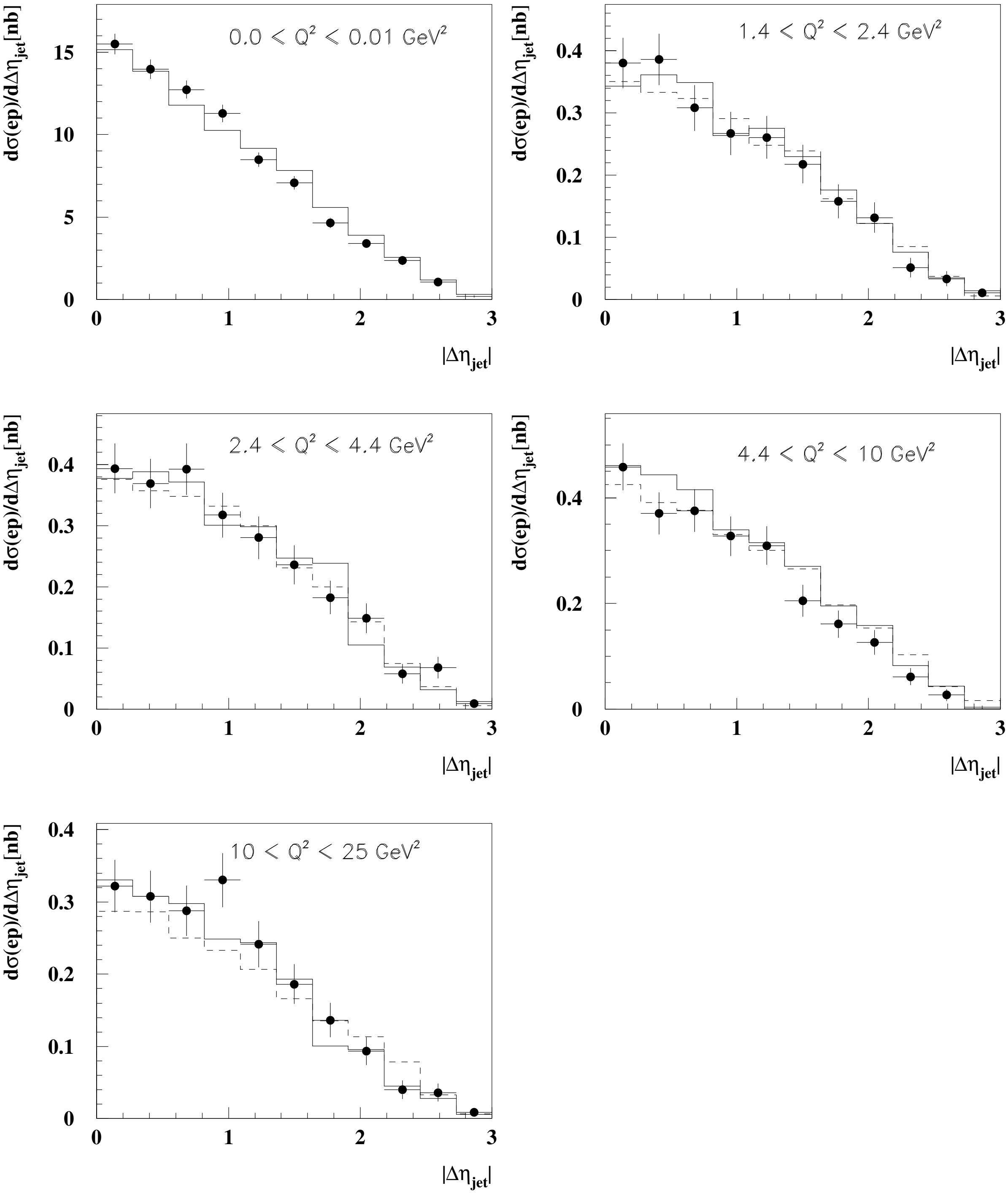


Figure 5.20: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantity: the absolute value of the difference between the pseudorapidity of the highest E_T and the second highest E_T jet. For description of event generators see the caption of Fig.5.14.

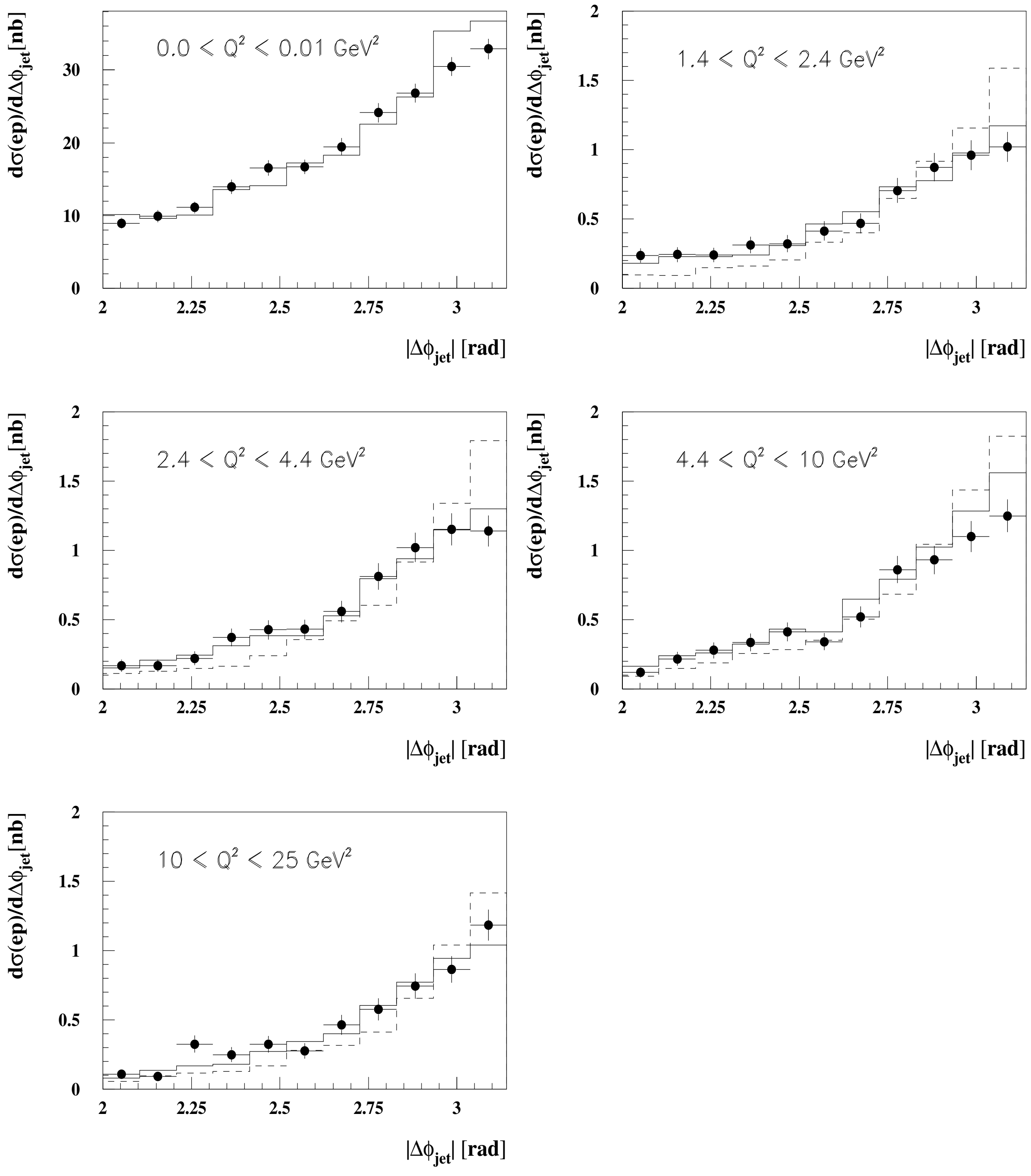


Figure 5.21: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantity: the absolute value of the difference between the azimuthal angle of the highest E_T and the second highest E_T jet. For description of event generators see the caption of Fig.5.14.

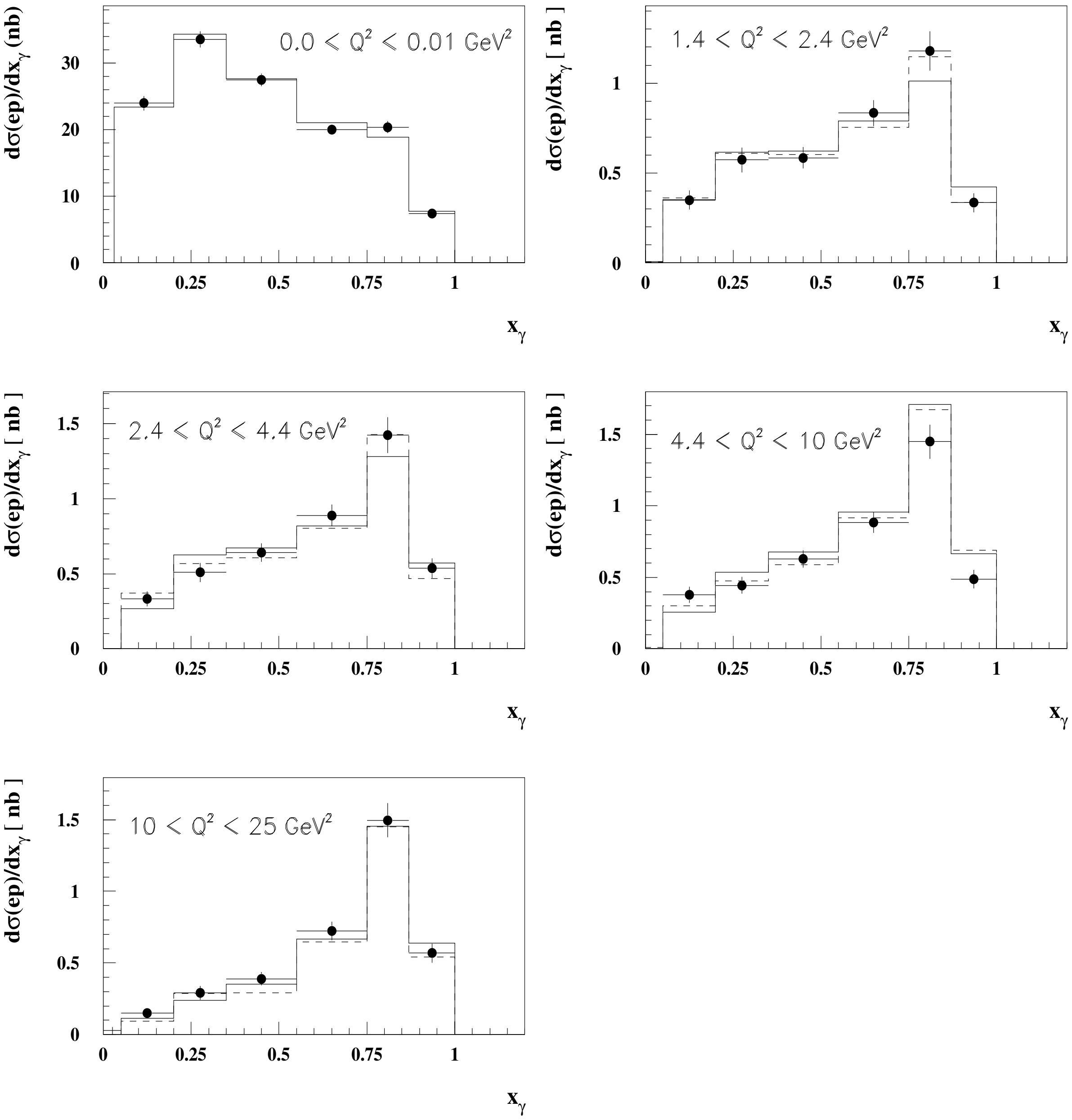


Figure 5.22: The comparison of the data (full points) with LO MC event generators on the detector level made with the quantity: x_γ using jets. For description of event generators see the caption of Fig.5.14.

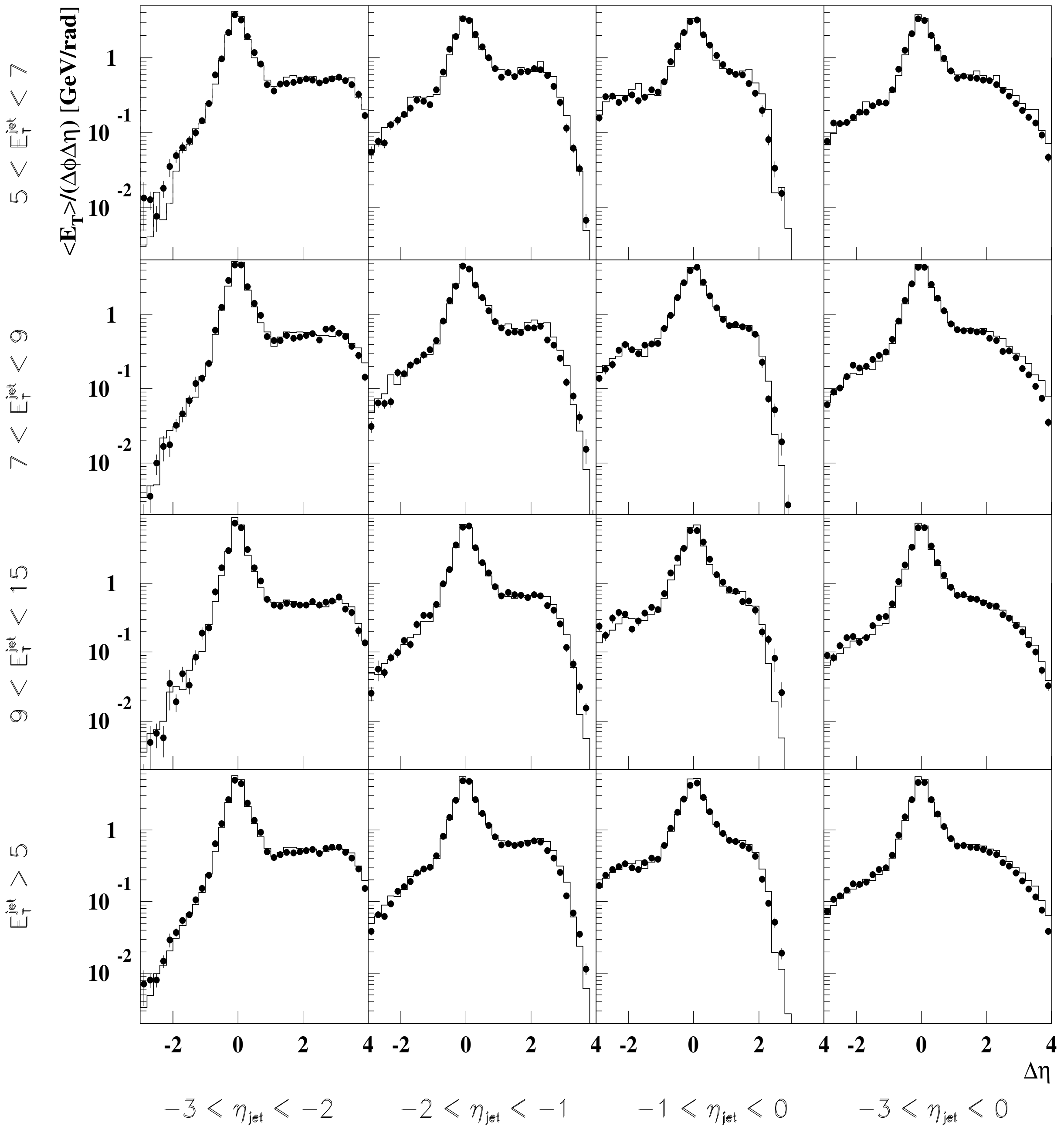


Figure 5.23: The comparison of the *photoproduction data* (full points) with LO MC event generators on the detector level for the jet profiles as a function of the distance of a particle from the jet axis, $\Delta\eta$. The particles were required to lie in the region $|\Delta\phi| \leq 1$. Performed for events where $x_\gamma < 0.75$. For description of the event generator see the caption of Fig.5.14.

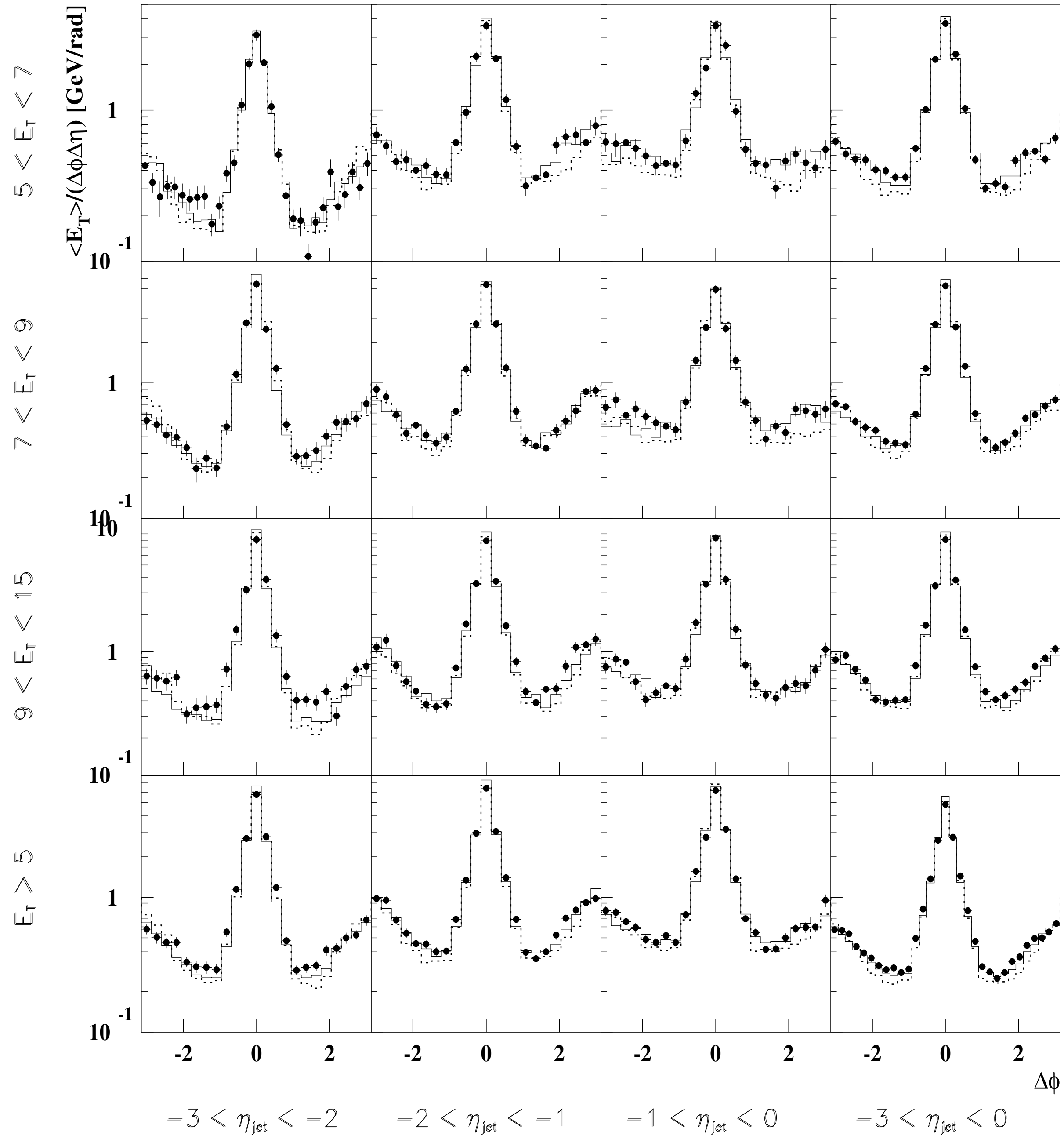


Figure 5.24: The comparison of the data from the region $1.4 < Q^2 < 25\,\mathrm{GeV}^2$ (full points) with LO MC event generators on the detector level for the jet profiles as a function of the distance of a particle from the jet axis, $\Delta\phi$. The particles were required to lie in the region $|\Delta\eta| \leq 1$. Performed for events where $x_\gamma < 0.75$. For description of event generators see the caption of Fig.5.14. The full histogram represents HERWIG 5.9, the dashed one RAPGAP 2.6.

5.4 Correction of the data to hadron level

The distribution of a measured observable always differs from that of the corresponding generated quantity, due to detector acceptance and resolution, hadronization, particle decays and QED and QCD radiative corrections.

The procedure mostly used to correct the data is called *bin-to-bin correction*. One evaluates with a Monte Carlo simulation the ratio between the number of events in a certain bin of the generated variable and the number of events in the same bin of the reconstructed variable. The number of corrected events is then estimated by multiplying the number of observed events in that bin by the corresponding ratio. It is clear that this method requires the same binning of the generated and experimental variables and hence it cannot take into account large migrations of events from one bin to another. Moreover it neglects the correlations between adjacent bins. This approximation is therefore usable only if the migration is negligible and if the standard deviation of the smearing is smaller than the bin size. Roughly speaking, the distributions suitable for this method are rather flat with sufficiently large bins.

The problem of the migrations is sometimes solved by constructing a matrix which connects the number of events generated in one bin with the number of events observed in the other bins (referred to as *smearing matrix*). This matrix is then inverted and applied to the measured distribution. Because it is not ensured that the matrix will always be regular we can get into trouble inverting it. The "regularized unfolding method" [72] has been proposed to overcome the troubles with matrix inversion. There also exist other unfolding methods [73, 74] from which that based on Bayes theorem [75] was chosen for this analysis.

Let us explain it in terms of *effects* and *causes*. In a measurement one has several possible effects E_j ($j = 1, 2, \ldots, n_E$) for a given cause C_i ($i = 1, 2, \ldots, n_c$). For each of them the Bayes formula holds:

$$P(C_i|E_j) = \frac{P(E_j|C_i)P_0(C_i)}{\sum_{l=1}^{n_c} P(E_j|C_l)P_0(C_l)}. \tag{5.4}$$

This can be read as follows: if we observe the jth effect (event) the probability that it was due to the ith cause is proportional to the probability of the cause $P_0(C_i)$ times the probability of the ith cause to produce the jth effect $P(E_j|C_i)$. The *initial* probability $P_0(C_i)$ as well as the *conditional* probability $P(E_j|C_i)$ (smearing matrix) can be evaluated from Monte Carlo simulation.

It is apparent that:

- $\sum_{i=1}^{n_c} P_0(C_i) = 1$
- $\sum_{i=1}^{n_c} P(C_i|E_j) = 1 \ldots$ each effect must come from one or more of the causes under examination
- $0 \leq \epsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|C_i) \leq 1 \ldots$ gives the efficiency of detecting the cause C_i in any of the possible effects.

A transition from the probabilities to the numbers of events is finally given as

$$n(C_i) = \sum_{j=1}^{n_E} n(E_j)P(C_i|E_j)/\epsilon_i, \tag{5.5}$$

where $n(E_j)$ is the number of events in the jth bin of an experimental distribution, $n(C_i)$ is the number of events for cause C_i - thus the *unfolded data*. The main advantages of the Bayesian approach with respect to the other unfolding methods are:

- it can be applied to multidimensional problems;
- it can use different bin sizes for the distribution of the generated and the experimental values;
- the domain of definition of the experimental values may differ from that of the generated values;
- it does not require matrix inversion;
- it provides the correlation matrix of the results

The weak point of the Bayes approach, namely the need for a knowledge of the initial distribution $P_0(C_i)$, is overcome by an iterative procedure.

In Fig.5.25 the bin correction factors from bin-to-bin method along with a graphical presentation of smearing matrices[2] are shown for each Q^2 interval as obtained in the MC event sample. The predictions of HERWIG as well as RAPGAP are presented in one figure as they do not differ significantly from each other.

The x_γ-dependence of the correction factors is neither flat nor monotonic but it does not change too fast. The smearing matrices do not show large migrations. The latter fact is reflected in a reasonable size of resolution in x_γ: it ranges between 23% for the first Q^2 window and 20% for the last one. Therefore we do not expect large differences from results from the bin-to-bin method.

[2] In the following, the term *smearing matrix* will be denoting both numerical and graphical presentation of the smearing matrix.

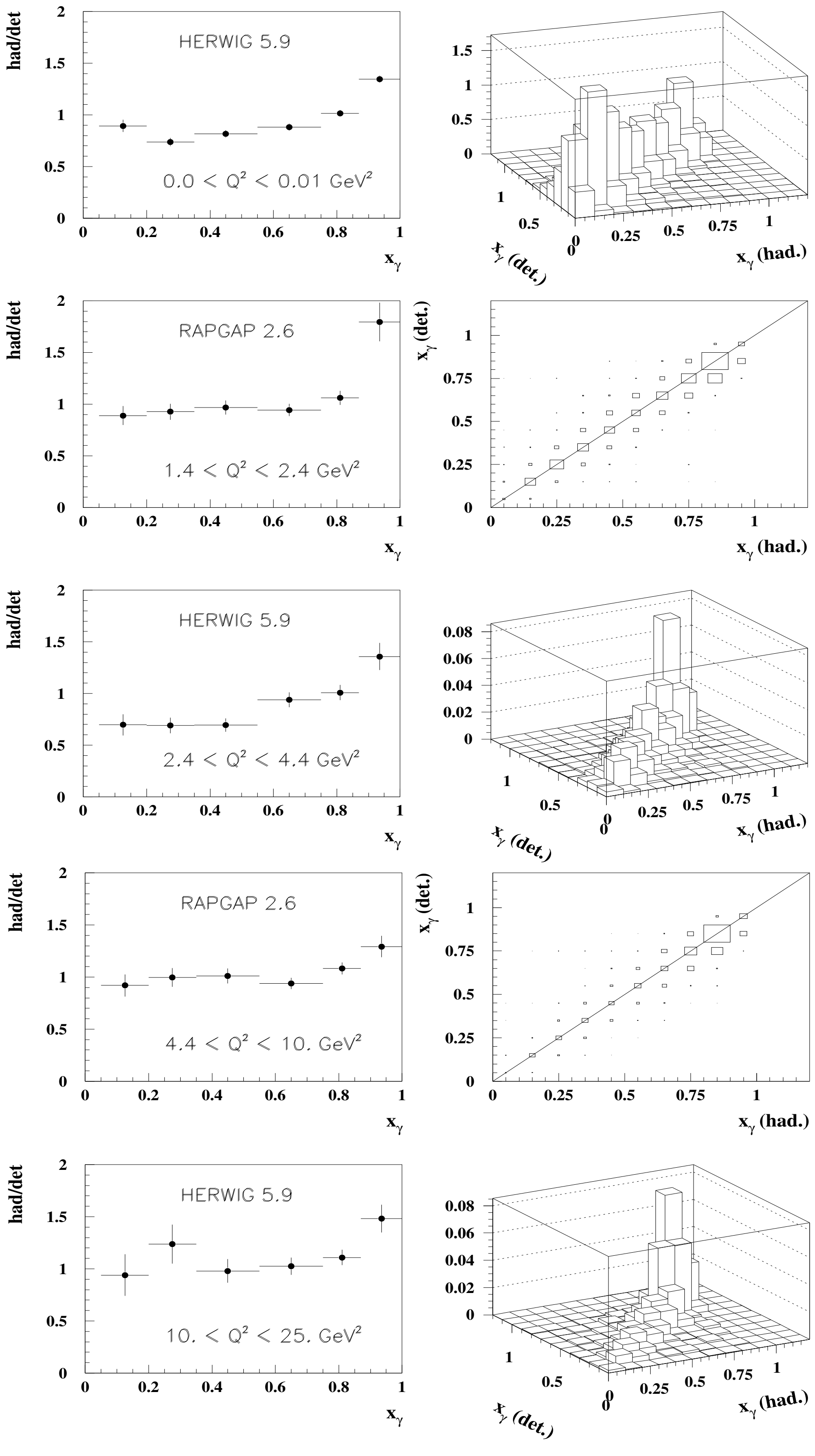


Figure 5.25: The correction factors from the bin-to-bin method and a graphical presentation of the smearing matrices for the unfolding program are shown. For each Q^2 interval the prediction of either HERWIG 5.9 or RAPGAP 2.6 is presented.

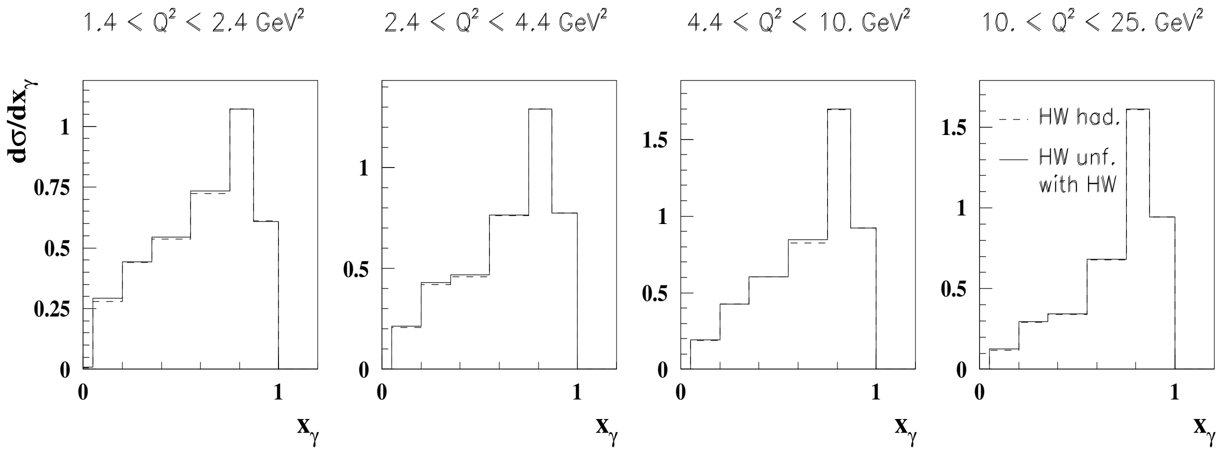


Figure 5.26: The test of the unfolding program: the full histogram represents the results of the unfolding of HERWIG using smearing matrices from HERWIG, the dashed histograms show the predictions of HERWIG on hadron level.

5.4.1 Tests of unfolding

In order to trust the results of the unfolding it is desirable to carry out several tests.

1. Instead of the data we put simulated Monte Carlo into the unfolding. If the unfolding program works properly we should get on the output what we put into, i.e. MC predictions on the hadron level. We unfolded HERWIG using smearing matrices from HERWIG. Fig.5.26 shows that the unfolded results almost coincide with the input.

2. Then we are obliged to convince ourselves that the unfolding program as an iterative routine converges. In Fig.5.27 we show the results of the unfolding for various number of iterations.

   As can be seen the differences between the results corresponding to two successive iterations get smaller as the number of iterations rises. It was found that the unfolding program ends after 4 or 5 iterations[3] in each Q^2 interval regardless of whether HERWIG or RAPGAP was used. We have chosen the fourth iteration as the reference iteration of the unfolded data. This number has been chosen also taking into account the results of the test following this one. Also shown are the results of the bin-to-bin method. The points are finally not very far from those corresponding to the fourth iteration as one might infer from Fig.5.25.

3. Another non-trivial test is a reweighting of the original MC event sample which served as the input of the unfolding program. The event falling to the ith bin in the distribution of the cause, $P(C_i)$, is assigned the (multiplicative) weight

$$w_i = \frac{n(C_i)}{P(C_i)} \tag{5.6}$$

   Ideally, if the data are unfolded precisely the distribution of the reweighted MC on detector level should match the distribution of uncorrected data. For brevity we show the effect

[3]The total number of iterations depends on the initial distribution $P_0(C_i)$ and the "termination parameter" ϵ. The iterative procedure stops if χ^2 of the results of two successive iterations is smaller than ϵ. We put $\epsilon = n_c/100$.

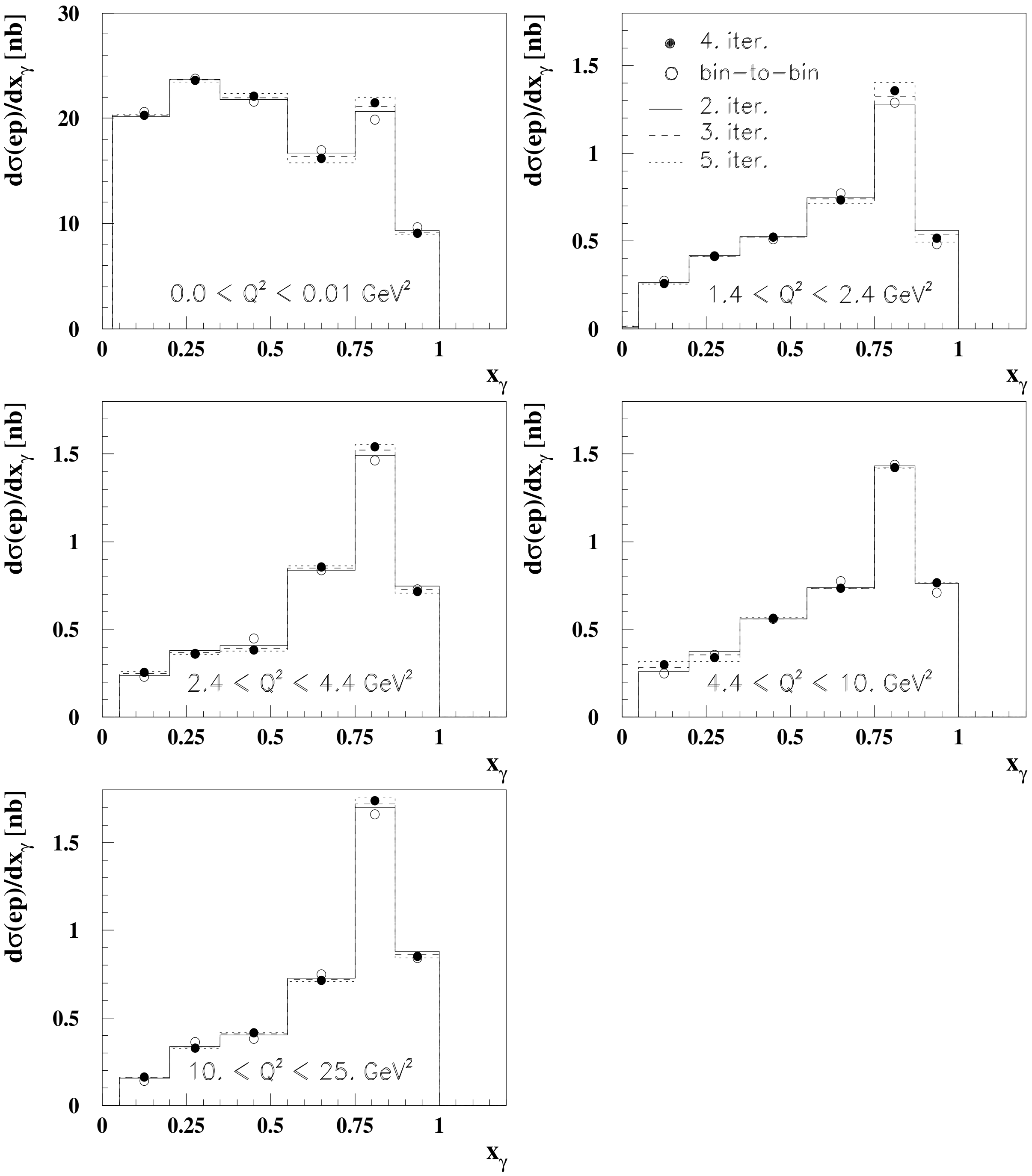


Figure 5.27: A dependence of the results of the unfolded data on the number of iterations using HERWIG. The full, dashed, dotted histograms correspond to the 2., 3. and 5. iteration, the full points correspond to the 4. iteration and open points to the results of the bin-to-bin method.

of the reweighting of both HERWIG as well as RAPGAP in one figure (Fig.5.28). The weights were given by the 4.iteration-results. We see that the reweighted HERWIG event sample describes the x_γ-distributions of the data almost perfectly! This fact however is not so astonishing because x_γ is the quantity in which the unfolding was done. A non-trivial consequence is a matching of the data and reweighted MC event sample in other quantities. For instance, we can see that reweighted RAPGAP gives a better description of the data in the η-distributions. A similar improvement is also reached in distributions of other quantities for both HERWIG and RAPGAP except for those in which the original RAPGAP fails to describe the data (see $|\Delta\phi|$-distributions in Fig.5.28). We cannot expect the reweighting itself to supply some missing non-perturbative effects etc. (see Section 5.3 for details on the description of the data by RAPGAP). Also shown are the distributions of global quantities, namely y and W, where the improvement is not so obvious as that reached in bins of the x_γ-distributions. Note that this test helps us to choose an optimal number of iterations. It should be that for which the description of the data by the reweighted MC sample is best.

4. As a stability test we should also check that the results of the unfolding program are independent of the input distribution ($P_0(C_i)$ from Eq.(5.4)). In principle, we may start from anything, even with a flat distribution. We tried out several input distributions where for each of them the optimal iteration had to be found according to the recipe from the point 3. The corresponding results were compared to the results of the fourth iteration with the original input distribution. The largest differences were considered to be the instability of the unfolding program and included into the total systematic error.

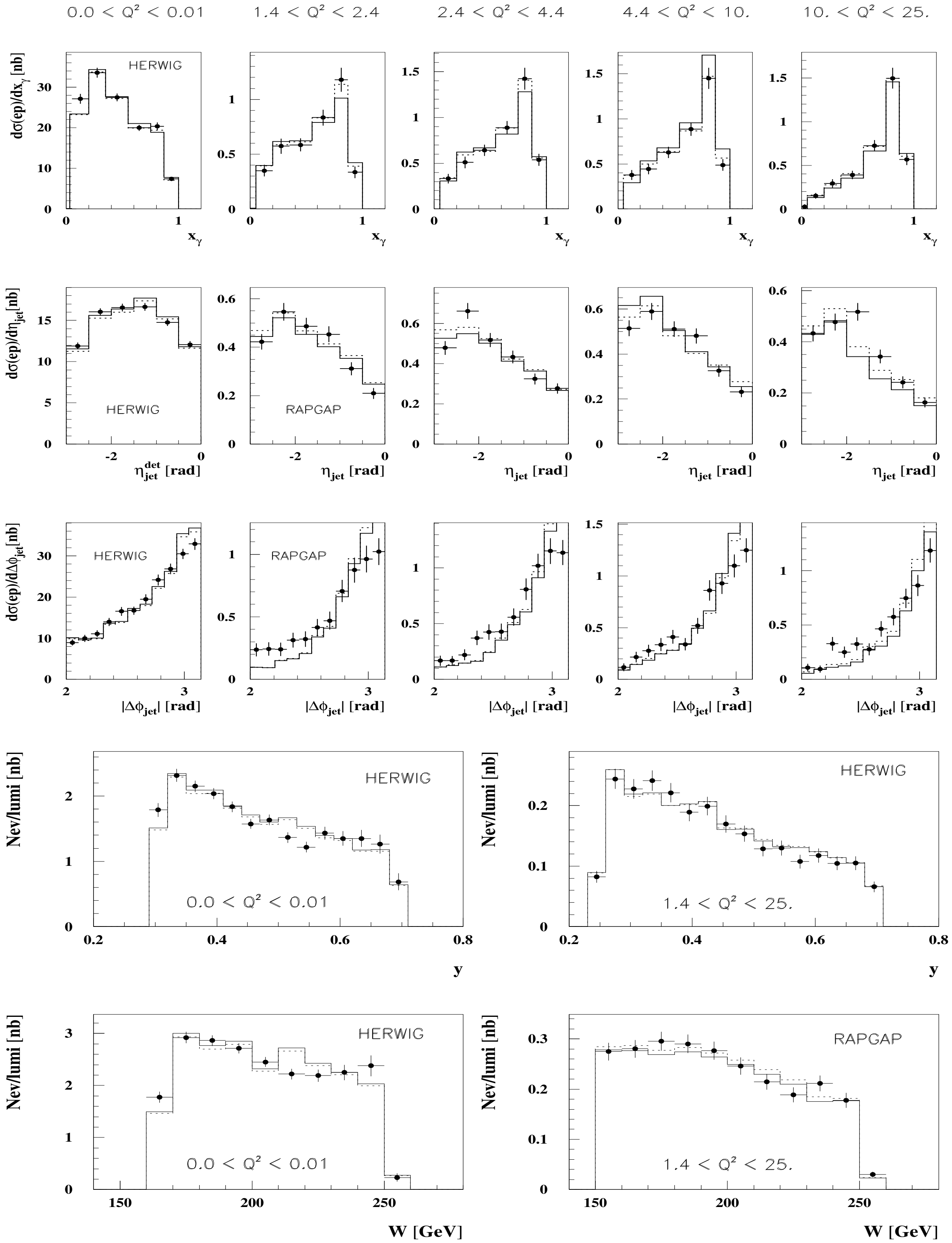


Figure 5.28: The effect of the reweighting of the detector level of MC event sample according to the results of the unfolding. The full histogram shows the original and the dashed histogram the reweighted predictions on the detector level. The full points represent the data.

The covariance matrix

The exact formulae for elements of covariance matrix are given in [75]. They take into account statistical errors of both the data and MC simulation. The statistical errors indicated in figures with unfolded data are given by square roots of diagonal elements of the covariance matrix in a given bin. From off-diagonal elements we can establish relative correlations, namely as

$$\text{Rel.corr.} = Cov(\sigma_i, \sigma_j)/\sqrt{Cov(\sigma_i, \sigma_i)Cov(\sigma_j, \sigma_j)} \tag{5.7}$$

which tells what part of the data that contributed to the ith bin contributed to the jth bin as well. In Table 5.2 all relative correlations larger than 40% are written out. As can be seen the largest bin correlations are concentrated between the 4. and 6. bin and do not exceed 50%. This results from the big change of the shape of the x_γ-spectrum in this region.

x_γ bins / Q^2 [GeV^2]	1-2	2-3	3-4	4-5	5-6
0.0 - 0.01	-	-	-	-	-
1.4 - 2.4	42.0	-	41.8	-	-
	-	-	-	-	48.7
2.4 - 4.4	-	44.5	40.3	-	-
	-	-	-	41.5	40.9
4.4 - 10.	-	-	-	42.1	-
	-	-	-	42.6	44.8
10. - 25.	-	43.2	-	42.1	44.9
	-	-	-	43.1	46.8

Table 5.2: The percentage for the relative correlations of unfolded numbers in x_γ-spectrum. The upper numbers in a row correspond to the use of HERWIG, the lower ones to RAPGAP. In photoproduction only HERWIG was used. Only correlations larger than 40% are shown.

5.4.2 Systematic errors

The total systematic error for the results on the hadron level was calculated as the quadratic sum of the following contributions.

PHENOMENOLOGICAL SOURCES:

- **model dependence**:
  This source turned out to give the largest contribution to the total systematic error. It was estimated by comparing the results of the unfolding using HERWIG with the results of the unfolding using RAPGAP. In Fig.5.29 we show the differences between the two mentioned procedures in a form of their ratio (full points). We see that almost all points obtained with RAPGAP are systematically above those obtained with HERWIG. In some bins the differences are sizable. We undertook several trials to trace back the cause of such model dependence. As noted in subsection 5.2.2 RAPGAP is not able to reproduce the behaviour of the data in such quantities as the jet pedestals (Fig.5.6), the E_t-flow outside the jets (Fig.5.9) and the $|\Delta\phi|$-distribution (Fig.5.21). Last but not least RAPGAP appeared to be insufficient also for photon remnant studies [9]. It was also observed that RAPGAP predictions in figures 5.6,5.9 resemble those of HERWIG with no SUE activity. Because there are no means of adding the missing SUE activity into RAPGAP we switched off this possibility in HERWIG in order to make the conditions as close as possible to those of RAPGAP. Then the comparison RAPGAP vs. modified HERWIG was performed again. There were no significant changes. In a second attempt to improve the agreement between RAPGAP without SUE activity and HERWIG the latter was reweighted according to $|\Delta\phi|$-distribution in RAPGAP. Although the unfolded data obtained with the reweighted version of HERWIG was indeed somewhat closer to those with RAPGAP, we cannot by far conclude that the insufficient amount of SUE activity in RAPGAP can explain the observed large model dependence.

  The systematic error stemming from this source was taken very conservatively: it was calculated as the difference between the results of the data unfolded with HERWIG and RAPGAP and then symmetrized.

  The problem of rather large model dependences surely deserves attention. It is worth emphasizing that there are not many event generators which would be able to describe the data in this very special kinematic region (resolved virtual photons, low E_T jets)[4].

  After a rather complicated and lengthy procedure for fixing the HERWIG and RAPGAP parameters their predictions on the detector level coincide (see Section 5.3). However, at the hadron level the HERWIG and RAPGAP results differ substantially (see open circles in Fig.5.29). This cannot simply be ascribed to the different hadronization models since it was found that the two generators also differ on the parton level because of the different default values for Λ in the formula for α_s, quark masses, treatment of initial and final state gluon radiation and the choice of the factorization and renormalization scales.

  Let us turn to Fig.5.29 once again. If the predictions of the two models on the detector level were very similar, the model dependence calculated as described above should essentially be seen in differences of the two models on the hadron level. Taking statistical errors of the MC event samples into account and ignoring two points with obviously low statistics

[4] See page 17.

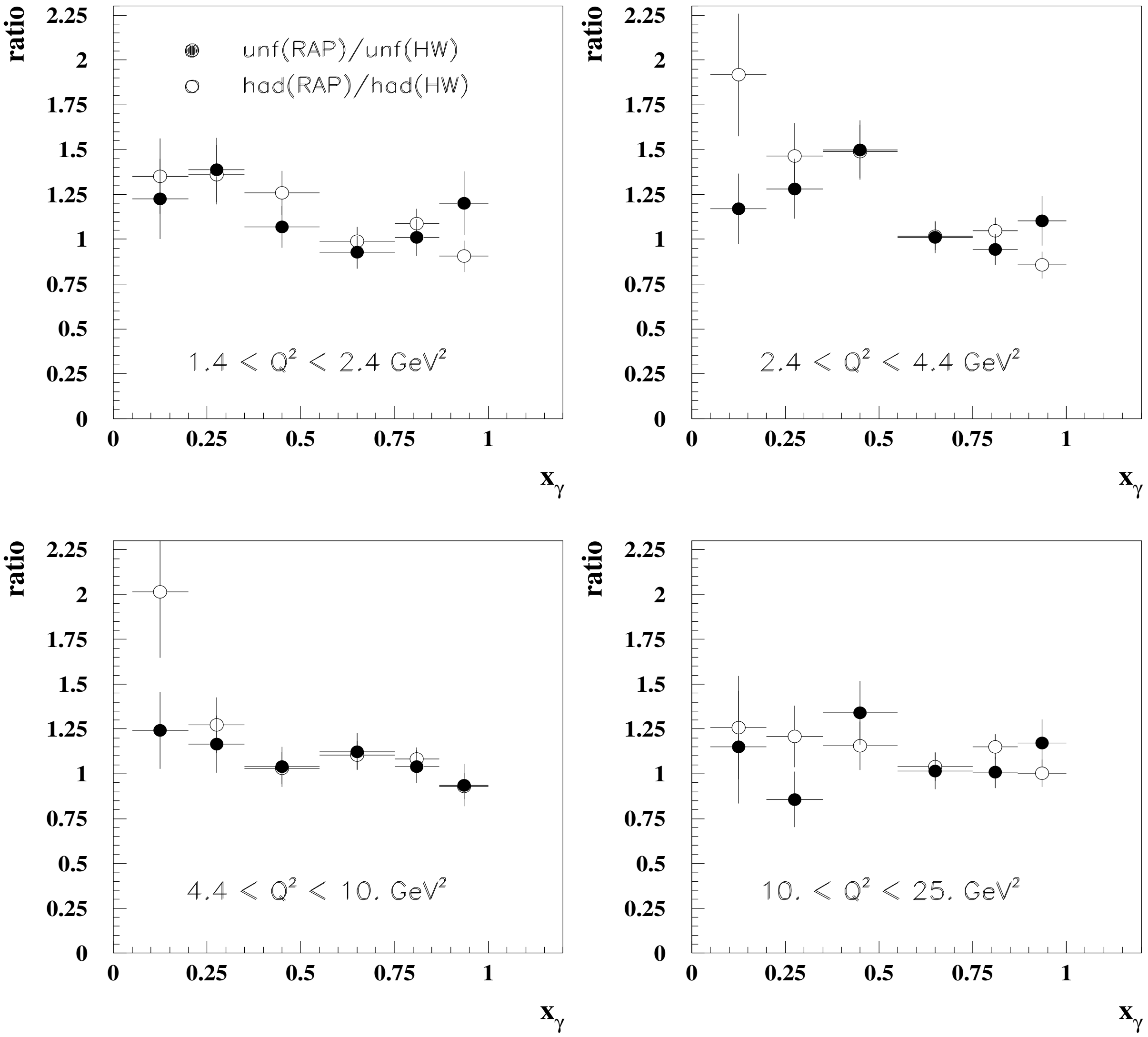


Figure 5.29: The model dependence for x_γ-distributions corrected for detector effects. The full circles show the ratio of the unfolded data using smearing matrices from RAPGAP to the unfolded data using smearing matrices from HERWIG. The open circles show the ratio of the prediction of RAPGAP to the prediction of HERWIG, both on hadron level. Specially for this study, HERWIG was used with PRSOF = 0.0 and reweighted according to $|\Delta\phi|$-distribution in RAPGAP.

we conclude this is the case. We realize that another procedure which could help in solving the problem with the large model dependence was multidimensional unfolding. However, the data statistics did not allow us to use it.
In the absence of another event generator in photoproduction the model dependence was estimated by comparing the predictions of HERWIG and RAPGAP on the hadron level. A typical error was 10–15%. This error was symmetrized.

- **uncertainty in determination of free parameters in HERWIG**
  A variation of all the free parameters around the optimal values noted in expressions 5.3 does not affect the cross section more than by 4%.

EXPERIMENTAL SOURCES:

- **4% - LAr energy scale uncertainty** (see Chapter 3)
  To check the dependence of the $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ cross sections on changes in the LAr energy scale, the effects of both ± 4% and ± 10% changes in the scale were studied. It is a natural question whether the systematic changes coming from ± 10% changes in LAr energy scale may be simply scaled by a factor 2.5 to obtain the systematic errors corresponding to ± 4% change. In the case of the x_γ-distribution it is not obvious. To solve this problem a special procedure was developed based on the fact that the only quantity depending on the LAr energy in the expression for x_γ (Eq.1.14) is E_T^{jet} which enters the expression in the first power. We generated a special HERWIG event sample weighted in $\hat{p}_T$ so as to reach higher cms energies with satisfactory statistics and then we fitted this $\hat{p}_T$-spectrum in each x_γ-bin. Since the change of the LAr energy scale by -10 (-4, $+4$, $+10$) % corresponds to a shift of the beginning of the E_T^{jet} spectrum from 5 GeV to 5.5 (5.2, 4.8, 4.5) GeV, the unknown scale factors are then figured out as the ratios

$$f_+ \stackrel{\text{def}}{=} f(\frac{+10\%}{+4\%}) = \frac{\int_{4.5}^{\infty} \frac{d\sigma}{d\hat{p}_T}\mathrm{d}\hat{p}_T}{\int_{4.8}^{\infty} \frac{d\sigma}{d\hat{p}_T}\mathrm{d}\hat{p}_T}, f_- \stackrel{\text{def}}{=} f(\frac{-10\%}{-4\%}) = \frac{\int_{5.5}^{\infty} \frac{d\sigma}{d\hat{p}_T}\mathrm{d}\hat{p}_T}{\int_{5.2}^{\infty} \frac{d\sigma}{d\hat{p}_T}\mathrm{d}\hat{p}_T} \tag{5.8}$$

for each x_γ-bin. It is obvious that due to a quite steep $\hat{p}_T$-dependence of the cross section the scale factors are foreseen to be not equal to just 2.5. In Table 5.3 the values of the scale factors calculated using Eq.(5.8) are summarized.

x_γ	0.05–0.20	0.20–0.35	0.35–0.55	0.55–0.75	0.75–0.87	0.87–1.00
f_-	3.8	3.6	3.5	3.3	3.1	3.0
f_+	1.9	1.9	2.0	2.0	2.1	2.1

Table 5.3: The transition factors from +10 (−10)% to +4 (−4)% change in the LAr energy scale for calculation of the systematic error in individual x_γ-bins.

By this procedure, the x_γ-dependence of sizes of the systematic error got smoothed and symmetrized. The typical error was 15–20%.

- **electron energy scale uncertainty in the SpaCal**
  As found in [51] it varies linearly from ± 0.7% at E'_e = 27.5 GeV to ± 2.5% at E'_e =

7 GeV. This source contributes typically about 8% to the total systematic error. It does not influence the photoproduction data.

- **θ'_e - measurement uncertainty in BDC**
  As found in [45] it amounts to $\pm$ 0.5 mrad. It does not contribute by more than 4–5% to the total systematic error. It does not influence the photoproduction data.

- **instability of the unfolding program**
  The error caused by the instability of the unfolding was estimated by comparing the results obtained by the unfolding program with several different input distributions $P_0(C_i)$. The differences were evaluated with regard to the results of four iterations of the unfolding with the original input distribution. The errors from this source range between 5 and 10%.

- **trigger inefficiency**
  As elaborated in Sections 4.2.6 and 4.3.5 the found the trigger inefficiency was estimated as a global systematic error +3% for the *low Q^2 sample* and +2% for the *photoproduction sample*.

- **uncertainty in luminosity measurement**
  As accounted for in [44] this source represents +1.4% error that was included into the total systematic error.

- **uncertainty of hadron energy in the SpaCal**
  As accounted for in [51] the energy of hadrons in the SpaCal is measured with 7% systematic uncertainty. This change has, however, a very small influence on the cross section, typically around 1%.

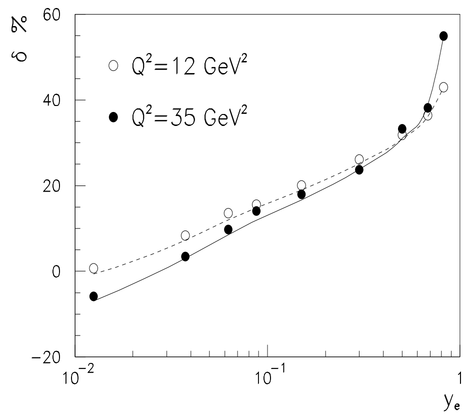
Electro-weak (radiative) corrections

The radiative corrections play a non-negligible role in measurements of the proton structure function. In the "leading log" approximation (LLA), two major sources of radiative corrections can be distinguished. The first one, called initial state radiation (**ISR**), indicates processes where the real (bremsstrahlung) photon is emitted almost collinearly with the beam electron. In processes, called final state radiation (**FSR**), the photon is emitted at a small angle from the scattered electron. A detailed description of radiative corrections can be found in [76].

Different methods for the kinematics reconstruction are affected differently by the radiative corrections. The electron method is most affected as it relies solely on the information about the scattered electron.

Ignoring all these processes leads to an overestimation of the energy of the scattered electron. Lower values of "true" Q^2 and y are produced by the lower electron beam energy in ISR events. In FSR processes, higher values of "true" Q^2 and lower values of "true" y are produced because of the reduced energy of the scattered electron. Due to the Q^2 and y-dependence of the DIS cross section the radiative corrections occur most often in events with low Q^2 and y. On the other hand, their size increases with increasing Q^2 and y. This can be seen in Fig.5.30 where the influence of the radiative corrections on the F_2^p measurement was studied.

The final state radiated photon typically appears in the vicinity of the scattered electron. To separate two clusters corresponding to the two particles requires a high granularity of a calorimeter and a sophisticated cluster algorithm. A study in [51] showed that it is possible to separate two electromagnetic clusters above a distance of 7 cm in the SpaCal plane. Still this



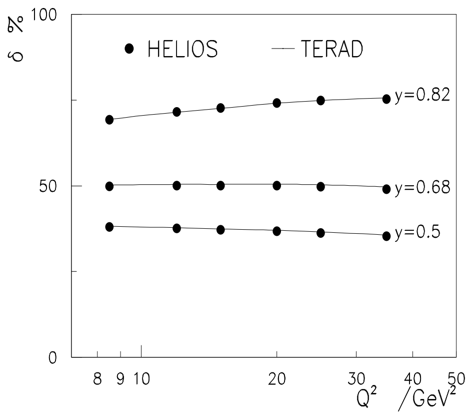


Figure 5.30: The radiative corrections as a function of y_e (left) and Q^2 (right), both for the electron method. The circles and lines correspond to various packages used for the calculations. The results are taken from [64] where the influence of the radiative corrections on F_2^p measurement was studied.

knowledge does not allow us to resolve more than about 1.5% of all FSR events generated by DJANGO [51]. Hence the background from FSR events is negligible.

In ISR processes the radiated photon typically escapes undetected along the backward beam pipe. These processes represent real background and in the following we will find out its influence in our kinematic region. Ideally, in order to correct the data for both detector and QED radiative effects one should evaluate the global correction factor for a bin as

$$c = \frac{\sigma^{\rm had}_{\rm MC,non-rad}}{\sigma^{\rm det}_{\rm MC,rad}} \tag{5.9}$$

by comparing predictions of the MC event sample on the detector level **with** QED radiative corrections to predictions of MC event sample on hadron level **without** radiative corrections. The Eq.(5.9) can be written also as

$$c = \frac{\sigma^{\rm had}_{\rm MC,non-rad}}{\sigma^{\rm had}_{\rm MC,rad}} \cdot \frac{\sigma^{\rm had}_{\rm MC,rad}}{\sigma^{\rm det}_{\rm MC,rad}} \equiv c_{\rm rad} \cdot c_{\rm det} \tag{5.10}$$

The problem however is that neither HERWIG nor RAPGAP (in resolved channel) contain radiative corrections. A model which allows the switching on of radiative corrections is DJANGO. However, this model contains only direct photon processes and consequently it fails to describe the data in the kinematic regions where the photon is really resolved [29]. Exploiting the fact that this event generator nevertheless gives effectively *some* prediction for the regions where the resolved processes dominate [29] we investigated the influence of radiative corrections using DJANGO on hadron level. To get closest to the experimental conditions we imposed these additional cuts on the hadron level: photons in an angle below 5° to the scattered lepton are merged to the lepton. All particles with $\theta_{lab} > 177°$ are ignored. The variables Q^2 and y are then calculated from the four-vector of the scattered lepton with merged photon. The results are written down in Table 5.4. It is apparent that up to $Q^2 = 10\,\mathrm{GeV}^2$ the corrections lie below 4% everywhere which is comparable with a statistical error. For higher Q^2 the corrections reach up to 8% in low x_γ-region. A similar result in a similar kinematic region was achieved in [77]. The radiative corrections were included into total systematic error where they may be considered negligible (compare with F_2^p measurement where $\delta_{RC} \sim 50\%$ at $y_e \sim 0.7$ [64]). This fact is given by the requirement on $E - p_z$ and presence of two jets.

x_γ / Q^2 [GeV^2]	0.05 – 0.35	0.35 – 0.75	0.75 – 1.0
1.4 - 2.4	+0.5	+0.4	+2.6
2.4 - 4.4	+4.0	-2.2	+0.1
4.4 - 10.	-4.0	-4.0	-3.3
10. - 25.	-7.9	-3.6	-4.0

Table 5.4: The percentage of the influence of radiative corrections calculated from the ratio of cross section with and without radiative corrections, both using DJANGO on the hadron level.

5.4.3 Results

In Fig.5.31 the virtuality dependence of the inclusive dijet cross section, $\mathrm{d}\sigma/\mathrm{d}x_\gamma$, corrected for detector effects is plotted [78]. The variable x_γ is calculated according to Eq.(1.15). Four Q^2-intervals in the low Q^2 region are considered, each approximately of the same statistical significance, together with photoproduction. The data are compared to HERWIG and RAPGAP, both with α_s^{LO} and GRV-LO parametrization of the photon PDF multiplied by the Drees-Godbole suppression factor L (Eq.1.34) for non-zero virtualities. We recall that no multiplicative factors were used in the generators. To estimate the sensitivity to ω within one MC generator and to check the model dependence, the data were compared with HERWIG predictions for $\omega =$ 0.2 and 0.3 GeV and with RAPGAP for $\omega = 0.2$ GeV. In HERWIG, the parameter PRSOF was set to 0.15 in the low Q^2 region and to 0.25 for photoproduction.

We observe that RAPGAP provides a satisfactory description of the data over the whole Q^2 range studied. HERWIG underestimates the data in the fifth bin where the direct photon processes dominate the cross section. The peak from direct photon processes is not in the last bin as expected which is due to smearing. In other bins all the predictions are in agreement with the data. As the value of ω is relevant only for the resolved photon processes we can conclude that both HERWIG and RAPGAP are consistent with the data for $\omega = 0.2$–0.3 GeV for $1.4 < Q^2 < 4.4\,\mathrm{GeV}^2$, while for $4.4 < Q^2 < 25\,\mathrm{GeV}^2$ the higher value is preferred, $\omega > 0.3$ GeV. The insufficient magnitude for the direct photon processes in HERWIG confirms the observations made in studies on the detector level (see Section 5.3) and remains not understood.

In Fig.5.32 the predictions of HERWIG for various parametrizations of the photon PDF are compared to the data. The differences are seen only in the region $x_\gamma \lesssim 0.75$ since the PDF of the photon affects only the resolved photon processes. Compared to GRV-LO (multiplied by L function for non-zero Q^2) both SaS parametrizations are fairly below the data. Whereas in the photoproduction SaS1D lies lower than SaS2D, for non-zero virtualities it is vice versa. This is due to different mixtures of the VDM and pointlike parts in the two parametrizations. Because the VDM component dies out rapidly with increasing Q^2 (as $1/Q^4$, see Chapter 1), the differences between the two parametrizations are much smaller for non-zero virtualities.

In Fig.5.33 the ratio $r \equiv \sigma^{\mathrm{res}}/\sigma^{\mathrm{dir}}$ of cross sections corresponding to resolved and direct contributions is plotted as a function of Q^2 [78]. On the basis of MC simulations and in agreement with [79] we defined $\sigma^{\mathrm{res}} \equiv \sigma(x_\gamma \leq 0.75)$ and $\sigma^{\mathrm{dir}} \equiv \sigma(x_\gamma > 0.75)$. The ratio $r(Q^2)$ decreases with Q^2 as measured by ZEUS in [28], but the results are not directly comparable due to different kinematic cuts. The observed decrease of $r(Q^2)$ is reasonably described by both MC generators for $\omega = 0.2$ GeV. Our measurement shows that the PDF of the virtual photon clearly requires virtuality suppression and qualitatively our result is in agreement with the conclusions of [29], where single jets were used.

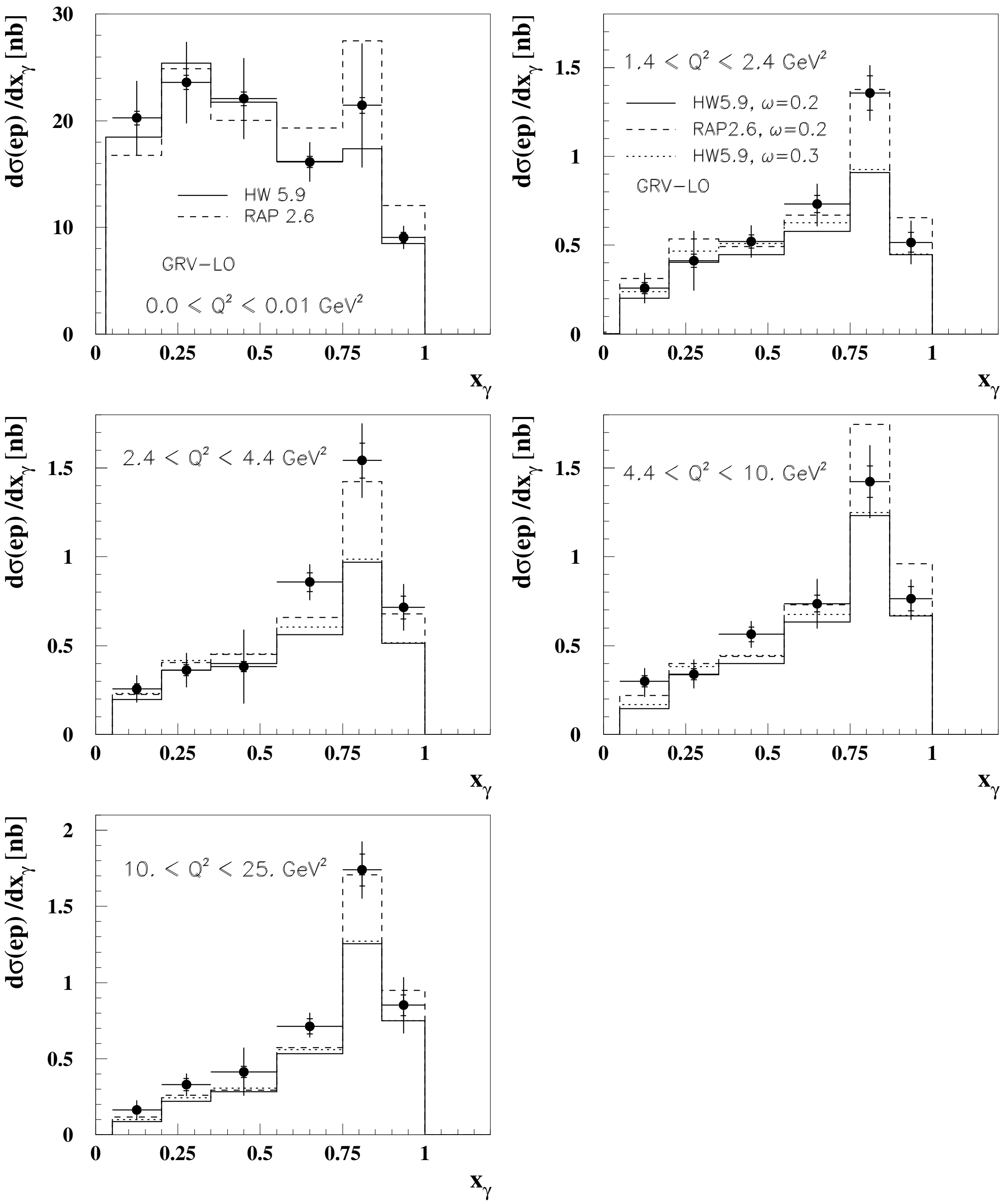


Figure 5.31: The inclusive dijet ep cross sections as a function of the x_γ in five intervals of the photon virtuality Q^2 (points) compared to HERWIG and RAPGAP. The full and dotted histograms correspond to the prediction of HERWIG with the photon structure function using GRV-LO parametrization multiplied by suppression function L (see Eq. (1.34)) with $\omega = 0.2$ and $0.3\,\mathrm{GeV}$. The dashed histogram corresponds to the prediction of RAPGAP with $\omega = 0.2\,\mathrm{GeV}$. The inner bars indicate the statistical errors and the total error bar shows the quadratic sum of statistical and systematic errors.

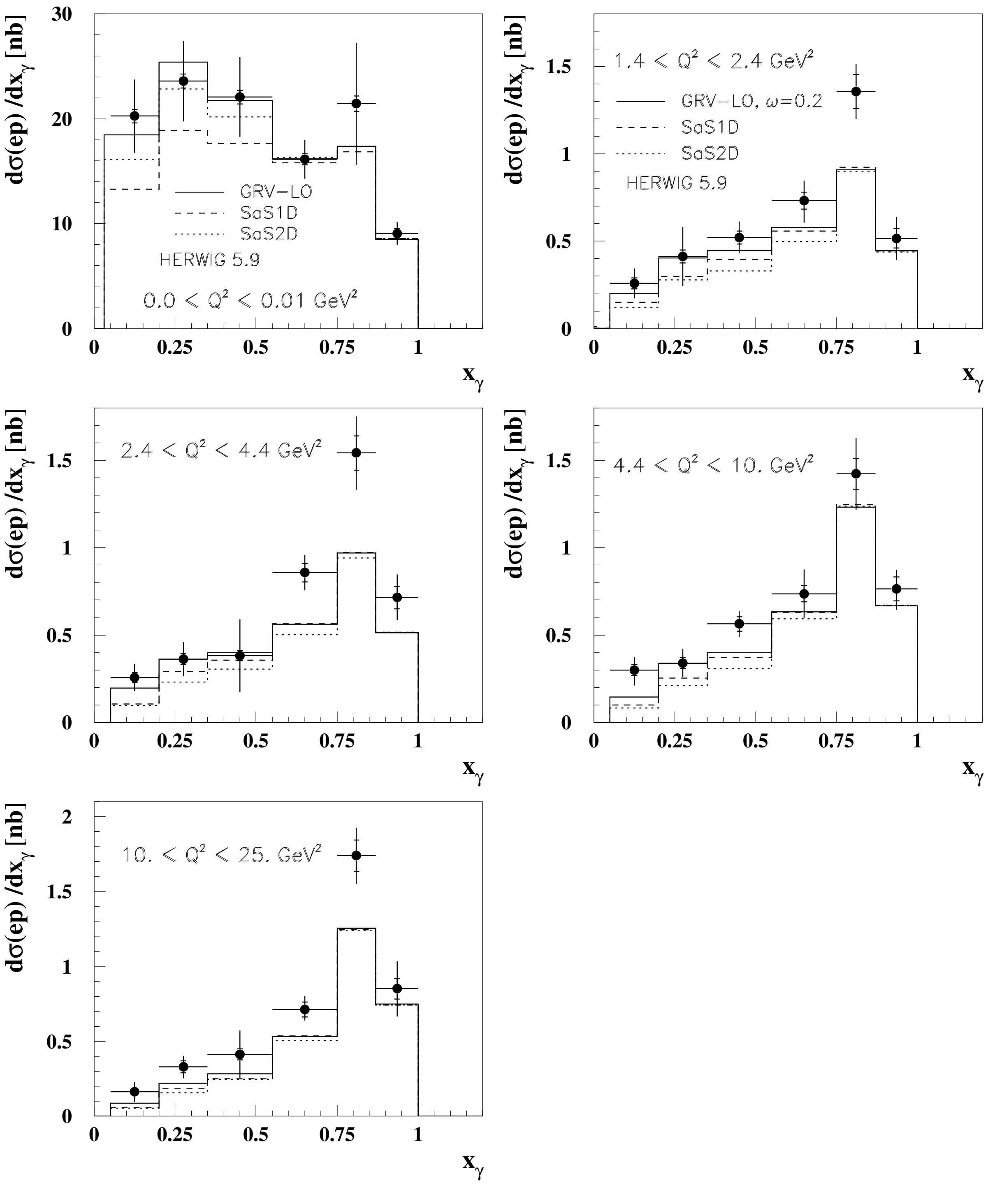


Figure 5.32: The inclusive dijet ep cross sections as a function of the x_γ in five intervals of the photon virtuality Q^2 (points) compared to HERWIG with different photon structure functions. The full histogram corresponds to GRV-LO parametrization multiplied by suppression function L (see Eq.(1.34)) with $\omega = 0.2\,\mathrm{GeV}$. The dashed and dotted histograms correspond to SaS1D and SaS2D parametrization. The inner bars indicate the statistical errors and the total error bar shows the quadratic sum of statistical and systematic errors.

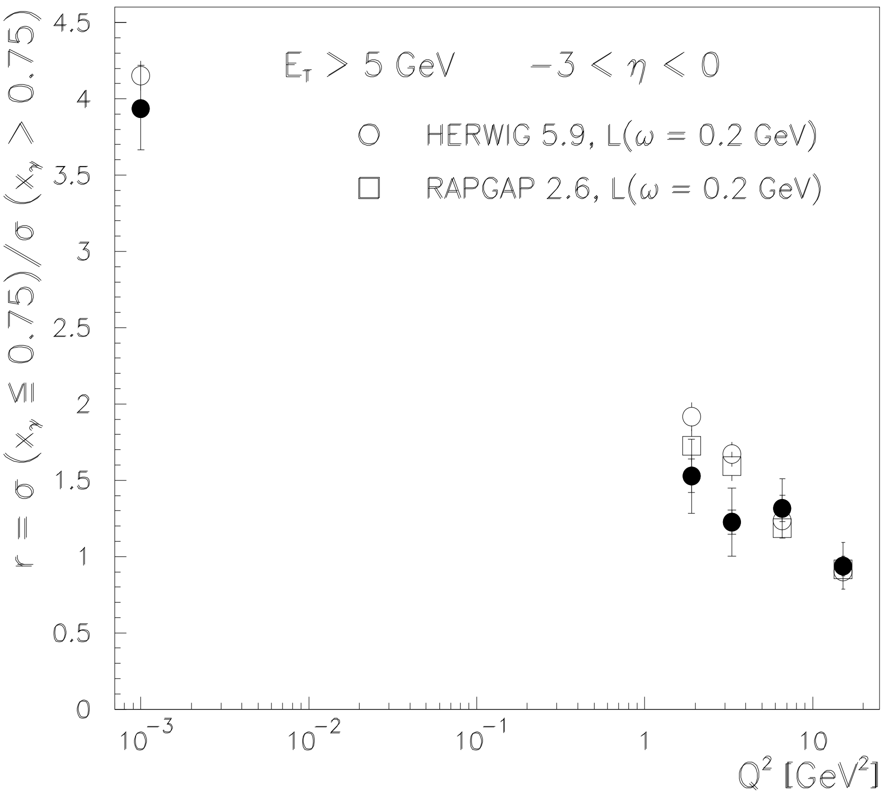


Figure 5.33: The ratio of the *resolved* to *direct* cross-section as a function of the photon virtuality Q^2. The resolved and direct contributions are defined by the cut on x_γ. The inner error bars represent the statistical errors, the outer error bars give the statistical and systematic errors, added in quadrature. The data (full circles) are compared to HERWIG (open circles) and RAPGAP (open squares) predictions with the structure function of the photon using the GRV-LO parametrization multiplied by suppression function L (see Eq.(1.34)) with $\omega = 0.2\,\mathrm{GeV}$.

5.5 Correction of the data to the hard scattering level

The main aim of this part of the analysis is the extraction of the effective parton distribution function of the virtual photon. As will be explained in section 5.5.4 this task makes sense only in the LO approximation. To make the analysis consistent we therefore generated, for photoproduction as well as for the low Q^2 region, HERWIG and RAPGAP event samples in which all structure functions as well as α_s were taken in LO. All parameters from these generations are listed in Table 5.5.

At the level of (partonic) hard scattering we study the processes (2.1) and (2.2)[5]. Unlike the parton level, defined in Chapter 2, the hard scattering level is free of the initial and final state parton showers. The unfolding of the data to the hard scattering level proceeds in two steps. In the first the data are corrected for detector effects to hadron level. The whole procedure with several tests of the unfolding program itself as well as systematic error assessment is described in Section 5.4. The second step is performed analogously to the first one: Bayesian unfolding is used by which the data corrected for detector effects are further unfolded to the level of the hard scattering. Again HERWIG and RAPGAP are used for correction of effects coming from hadronization, parton showers and soft underlying activity. The smearing matrices for the x_γ-quantity between the hard scattering level and hadron level is shown in Fig.5.34. The smearings in the region $x_\gamma < 0.75$ are not seen well in the *low Q^2 sample* due to a high peak in the bin $x_\gamma > 0.75$ which arises from the fact that all direct processes on the hard scattering level are assigned $x_\gamma = 1$ by definition. In the same figure the correction factors from the bin-to-bin method are added for illustration.

model	process	Q^2[GeV^2]	γ^* - PDF	N_{events}	$\mathcal{L}$ [pb^{-1}]
HERWIG 5.9	DIR	0.0 — 0.01	-	200 k	3.54
	RES	0.0 — 0.01	GRV-LO	200 k	0.32
	RES	0.0 — 0.01	SaS1D	200 k	0.49
	RES	0.0 — 0.01	SaS2D	200 k	0.42
	DIR	1.4 — 25.0	-	200 k	14.14
	RES	1.4 — 25.0	GRV-LO	200 k	15.26
	RES	1.4 — 25.0	SaS1D	200 k	22.26
	RES	1.4 — 25.0	SaS2D	200 k	27.18
RAPGAP 2.6	DIR	0.0 — 0.01	-	200 k	1.68
	RES	0.0 — 0.01	GRV-LO	200 k	0.27
	DIR	1.4 — 25.0	-	110 k	7.32
	RES	1.4 — 25.0	GRV-LO	200 k	8.48

Table 5.5: Characteristics of Monte Carlo event samples used for correction of the data to the hard scattering level and for making prediction on this level as well as on the hadron level. $p_T^{min} = 2.5$ GeV, GRV-LO parametrization for the proton PDF and α_s in LO approximation was set throughout. The luminosity was calculated from all generated events and cross section modified to describe the data.

[5]See footnote on page 5.

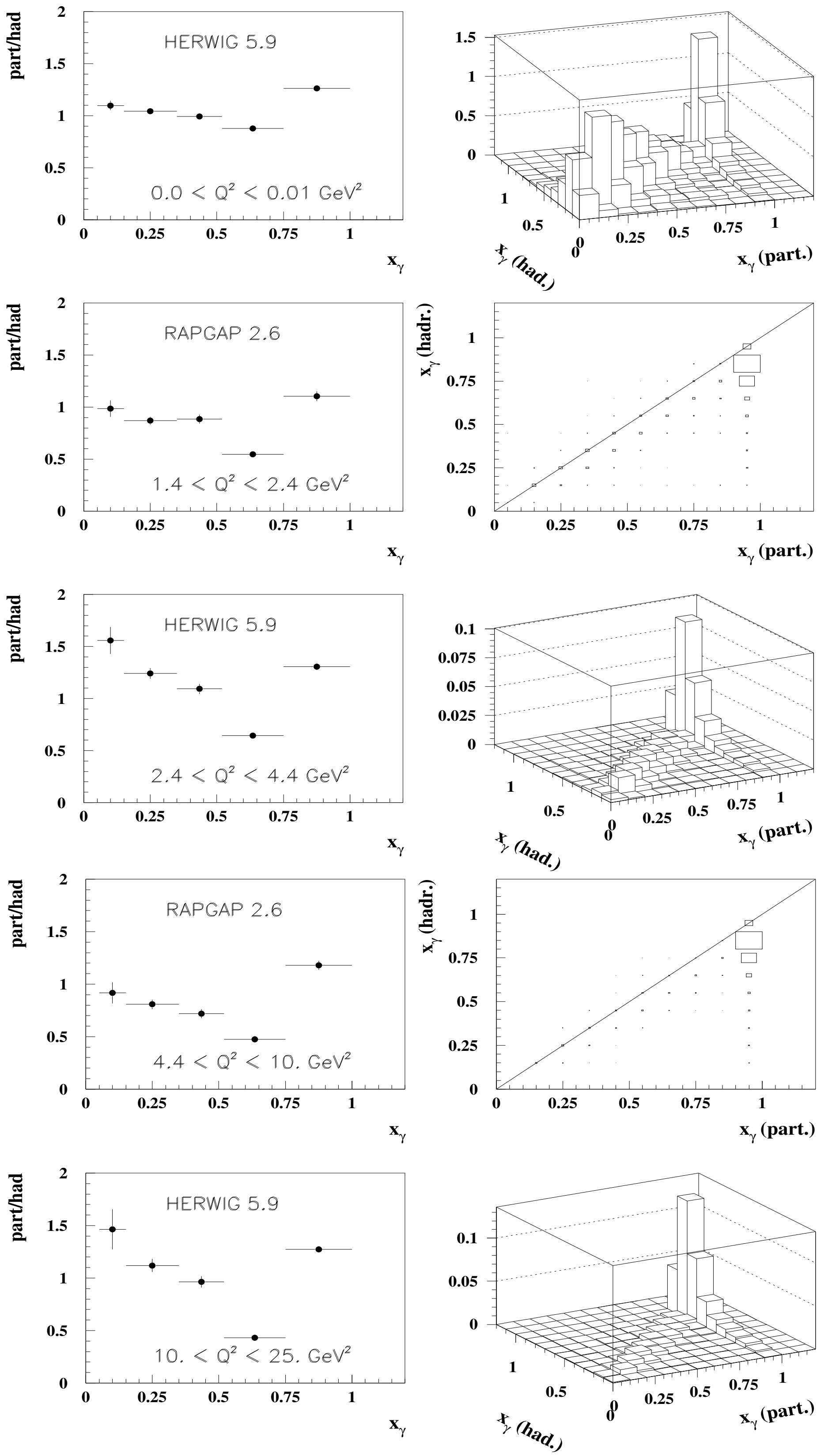


Figure 5.34: The correction factors from the bin-to-bin method and the smearing matrices for unfolding program are shown. For each Q^2 interval the prediction of either HERWIG 5.9 or RAPGAP 2.6 is presented.

5.5.1 Tests of unfolding

For this step of the unfolding two tests explained in detail in Section 5.4.1 were performed. In Fig.5.35 the results of the unfolding for various number of iterations are shown and a convergence of the results is observed as the number of iterations increases. The total number of iterations is 4 in case of the *low* Q^2 *sample* (for HERWIG as well as RAPGAP)[6]. In the *photoproduction sample* the number of iterations was higher which is given by much smaller statistical error of the data. On the basis of this and the following test the results of the 3. iteration for the *low* Q^2 *sample* and of the 5. iteration for the *photoproduction sample* were chosen as a reference. From Fig.5.35 we further see that in last two bins the results are essentially independent of number of iterations and moreover that the bin-to-bin method gives values relatively close to those from the reference iteration. This is due to a quite large size of the two bins whose borders intersect at the point $x_\gamma = 0.75$. A small region around this point is very special because it is assumed to be characterized by a well-balanced mixture of resolved and direct photon events. And as we aim at extracting the PDF of the photon from resolved photon processes only we have to assure their reliable separation. By choosing these two bins sufficiently large we will achieve a satisfactory stability (described above) of the unfolding procedure as well as a low contamination of the penultimate bin by direct photon processes and of the last bin by resolved photon ones.

In Fig.5.36 the effect of the reweighting of MC event sample according to formula 5.6 is plotted. The reweighting can only be done for x_γ-distribution because for other quantities the correction of the data for detector effects was not done. Both HERWIG and RAPGAP reweighted event samples describe the data on hadron level appreciably better than the original MC event sample.

The covariance matrix

For the exact formulae for elements of covariance matrix see [75]. See also the explanation and expression for the relative correlations of unfolded numbers (Eq.5.7) given in Section 5.4.1.

x_γ bins / Q^2 [GeV^2]	1-2	2-3	3-4	4-5
0.0 - 0.01	40.6	-	44.7	-
	-	-	49.6	41.2
1.4 - 2.4	63.8	61.4	64.0	42.7
	-	-	-	-
2.4 - 4.4	66.0	59.7	63.5	50.6
	-	-	-	41.5
4.4 - 10.	64.6	54.8	69.0	47.2
	-	-	-	42.6
10. - 25.	40.6	54.5	62.1	48.9
	-	-	-	43.1

Table 5.6: The percentage for the relative correlations of unfolded numbers in x_γ-spectrum. The upper numbers in a row correspond to the use of HERWIG, the lower ones to RAPGAP. Only correlations larger than 40% are shown.

[6]The termination parameter $\epsilon = n_c/100$, see also footnote on page 79.

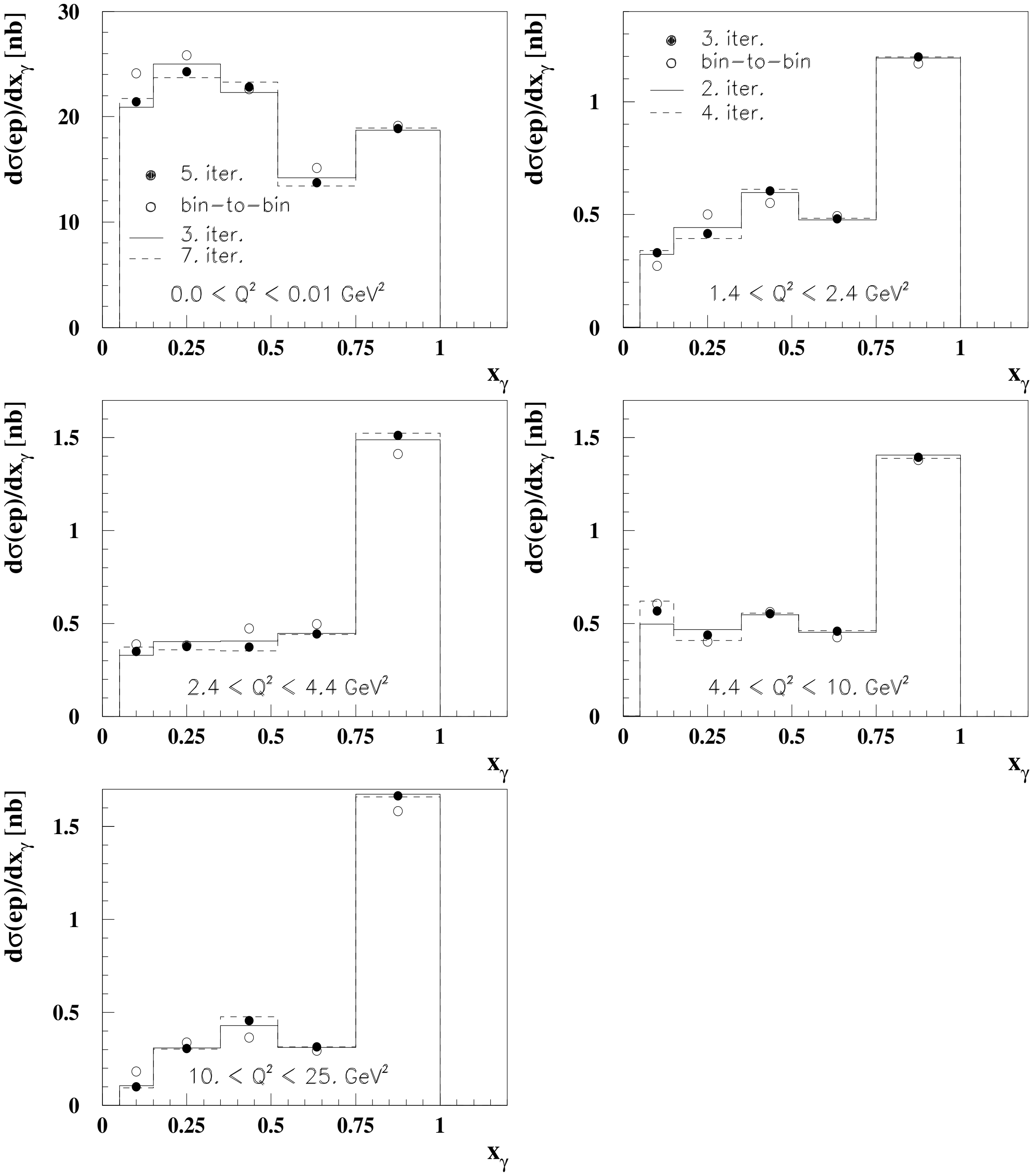


Figure 5.35: A dependence of the results of the unfolded data on the number of iterations using HERWIG in the low Q^2 region (photoproduction). The full and dashed histograms correspond to the 2. (3.) and 4. (7.) iteration, the full points to the 3. (5.) iteration. Open points correspond to the results of the bin-to-bin method.

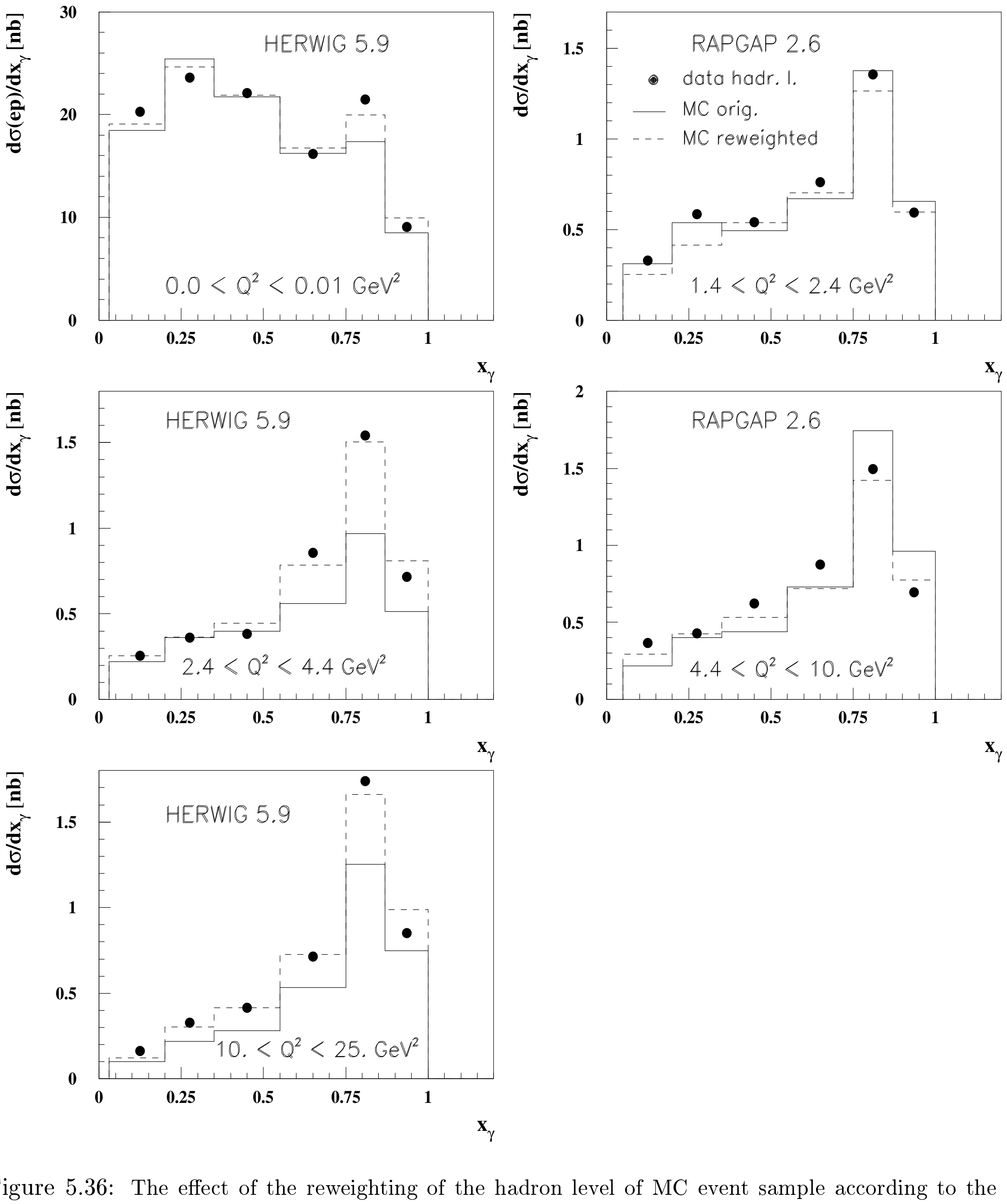


Figure 5.36: The effect of the reweighting of the hadron level of MC event sample according to the results of the data unfolding to the level of hard scattering. The full histogram shows the original and the dashed histogram the reweighted predictions on the hadron level. The full points represent the data unfolded to the hadron level using MC generator specified by the name in single figures.

5.5.2 Systematic errors

The total systematic error for the data unfolded to the hard scattering level was calculated as the quadratic sum of the following contributions:

PHENOMENOLOGICAL SOURCES:

- **model dependence**:
  the same procedure as in the first step of the unfolding was used: the systematic error of unfolded cross sections was taken as the difference between the results of the unfolding using HERWIG and RAPGAP. This error was then symmetrized in each x_γ bin. A typical difference was 15–25% and in some bins even higher. These quite large errors can be explained by non-negligible differences of the two models even on the parton level. We have to keep in mind that current MC generators may differ in a number of respects enumerated in Chapter 2, namely in the formula for α_s, the value of Λ, treatment of initial and final state radiation, the choice of the factorization and renormalization scale, the effect of quark masses, the treatment of α and the mean intrinsic p_T. By setting the corresponding parameters to the same values we can compare the hard scattering levels of all the event generators[7]. With this procedure the calculations of the pure LO matrix elements are checked but the parameter values used just for the comparisons may differ from those tuned to describe the data! For instance, the values of α_s obtained in RAPGAP can differ from those of HERWIG by as much as 13% at the scales explored in this analysis. The difference stems from different default values of Λ (0.18 GeV in HERWIG, 0.25 GeV in RAPGAP). See Fig.5.10 where the scale dependence of α_s calculated in RAPGAP coincides with that in HERWIG for $\Lambda = 0.25\,\text{GeV}$. It is worth emphasizing that due to different hadronization mechanisms a non-negligible model dependence of the results on the hard scattering level is expected. This is in contrast to the hadron level where the correction of the detector effects should not, in principle, depend on the MC generator used.

- **various parametrizations of photon PDF**:
  Beside the sample with GRV-LO parametrization combined in the low Q^2 region with the suppression function L with $\omega = 0.2$ GeV we generated resolved photon events with SaS1D or SaS2D parametrizations which have Q^2-suppression built in. The change of input photon PDF results in a change of the cross section for the unfolded data whose magnitude does not exceed 16%, a typical error is 5–7%. The reference cross section is that obtained with GRV-LO and $\omega = 0.2\,\text{GeV}$. This error was not symmetrized.

OTHER SOURCES:

- **propagation:**
  the total systematic error for hadron level was propagated further by the unfolding program to the hard scattering level.

- **instability of the unfolding program**:
  by using the procedure described in Section 5.4.2 we got a typical error whose magnitude is 10–15%.

[7]See e.g. *http://hepwww.rl.ac.uk/theory/seymour* for comparisons of HERWIG with LEPTO and DISENT.

5.5.3 Results

In Fig.5.37 the virtuality dependence of the diparton cross section, $\mathrm{d}\sigma/\mathrm{d}x_\gamma$, obtained by unfolding the dijet cross section on hadron level to the hard scattering level is plotted. The parton fractional momentum x_γ is calculated according to Eq.(1.15). The data are compared to HERWIG and RAPGAP with α_s^{LO} and GRV-LO parametrization of the photon structure function which is scaled by Drees-Godbole suppression factor L. The MC results are free of any multiplicative factors. Analogously to the Fig.5.31, the data were compared with HERWIG predictions for $\omega = 0.2$ and $0.3\,\mathrm{GeV}$ and with RAPGAP for $\omega = 0.2\,\mathrm{GeV}$. In HERWIG, the parameter PRSOF was set to 0.15 in the low Q^2 region and 0.25 for photoproduction.

Due to large systematic errors of the data we cannot make a clear preference for any of the models used. In photoproduction both generators give a reasonable description of the data. In low Q^2 and $x_\gamma < 0.75$ region HERWIG describes the data well for both values of ω, while in bin $x_\gamma > 0.75$ its prediction underestimates the data. RAPGAP gives the lowest prediction in region $x_\gamma < 0.75$ over the Q^2 region studied, in some bins even beyond the error of the data. On the other hand, RAPGAP gives a very good description of the data in the bin $x_\gamma > 0.75$. The data tend to prefer $\omega \sim 0.2\,\mathrm{GeV}$ for lower Q^2 and $\omega \sim 0.3$ or $0.4\,\mathrm{GeV}$ in higher Q^2.

In Fig.5.38 the data from Fig.5.37 are compared to HERWIG with various input PDFs of the photon. The GRV-LO parametrization in the low Q^2 region is scaled by suppression function with $\omega = 0.2\,\mathrm{GeV}$.

As the change of the photon PDF affects only the resolved photon processes, the statement made above about the poor description of the data in the bin $x_\gamma > 0.75$ is unchanged. Both SaS parametrizations give lower cross sections compared to that for GRV-LO with $\omega = 0.2\,\mathrm{GeV}$. While in photoproduction SaS1D gives the lowest cross section, in low x_γ region even beyond the errors of the data, the cross section with SaS2D is of a similar shape and magnitude as the data. On the other hand, in the low Q^2 region the cross section with SaS2D is below that with SaS1D. For $Q^2 < 4.4\,\mathrm{GeV}^2$ predictions with SaS1D are still able to describe the data, for $Q^2 > 4.4\,\mathrm{GeV}^2$ the predictions are too low. Integrally, the cross sections obtained by HERWIG with SaS2D in region of low Q^2 and $x_\gamma < 0.75$ differ from those with GRV-LO and $\omega = 0.2$ GeV by 31–38%.

5.5.4 The effective parton distribution function for virtual photons

In LO, the expression for the resolved photon contribution to the diparton cross section in ep collisions reads

$$\frac{\mathrm{d}\sigma^{\mathrm{ep}}}{\mathrm{d}y\mathrm{d}x_\gamma\mathrm{d}x_p\mathrm{dcos}\hat{\theta}\mathrm{d}Q^2} = \frac{1}{32\pi s_{ep}}\sum_{r=T,L}\frac{f^r_{\gamma/e}(y,Q^2)}{y}\sum_{i,j,k,l}\frac{f^r_{i/\gamma}(x_\gamma,M^2,Q^2)}{x_\gamma}\frac{f_{j/p}(x_p,M^2)}{x_p}|M^{kl}_{ij}(\mathrm{cos}\hat{\theta})|^2 \tag{5.11}$$

where $f^T_{\gamma/e}, f^L_{\gamma/e}$ denote the fluxes of transverse and longitudinal virtual photons inside the electron, $f^T_{i/\gamma}, f^L_{i/\gamma}, f_{i/p}$ stand for PDF of species i in transverse and longitudinal photons and proton respectively, taken at the factorisation scale M, and M^{kl}_{ij} are appropriately normalized LO matrix elements of the binary hard processes $ij \to kl$, proportional to $\alpha_s^2(\mu)$. These matrix elements are functions of a single variable, for which we took in (5.11) the scattering angle in the diparton rest frame. The quantity s_{ep} is the square of the cms energy in the ep collisions. Although in general the renormalization and factorisation scales μ and M are independent free parameters labelling two different ambiguities of the perturbative QCD calculations, we followed the con-

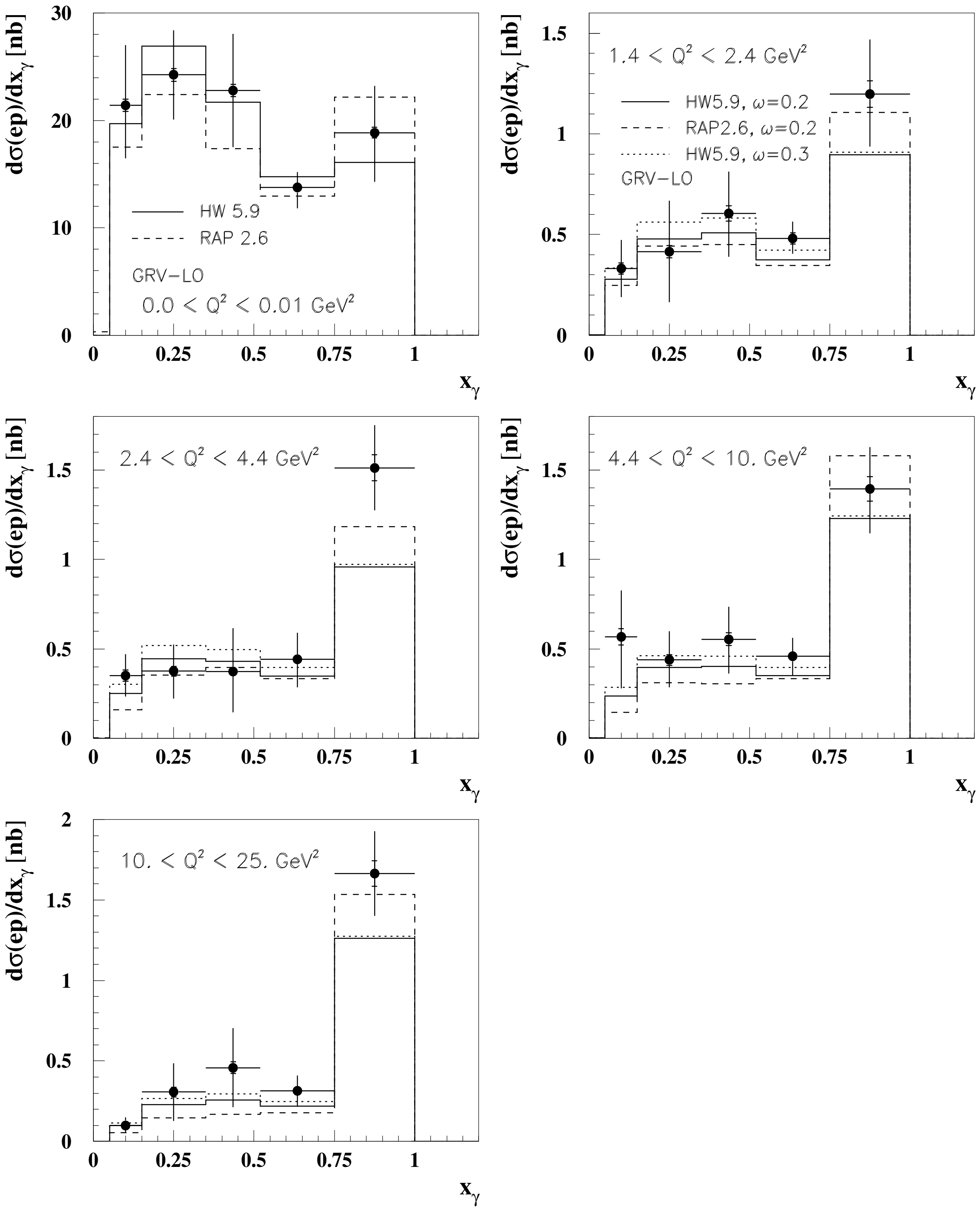


Figure 5.37: The ep diparton cross section as a function of the parton fractional energy x_γ in five intervals of the photon virtuality Q^2 (points) compared to HERWIG and RAPGAP. The full, dotted line resp. correspond to the prediction of HERWIG with the photon structure function using GRV-LO parametrization multiplied by suppression function L with $\omega = 0.2, 0.3\,\text{GeV}$. The dashed line corresponds to the prediction of RAPGAP with $\omega = 0.2\,\text{GeV}$. The inner bars indicate the statistical errors and the total error bar shows the quadratic sum of statistical and systematic errors.

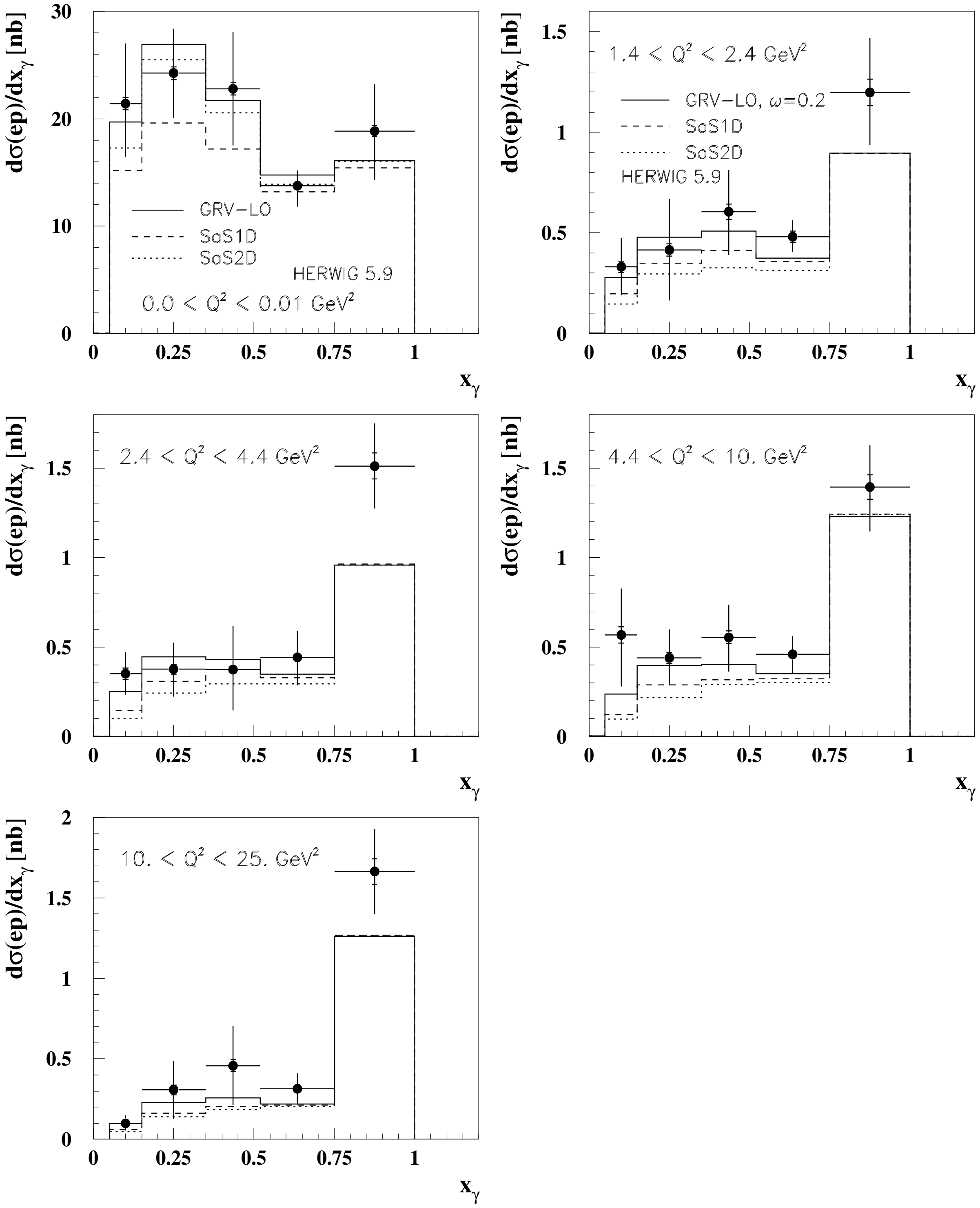


Figure 5.38: The data from Fig.5.37 (points) are compared with predictions of HERWIG for various parametrizations of the photon PDF. The full line corresponds to GRV-LO, the dashed line to SaS1D and dotted line to SaS2D.

ventional practice and assumed $M = \mu$. Furthermore, in all our consideration we set $\mu = E_T$. The fluxes of transverse and longitudinal photons are given by Eq.(1.35) and (1.36). The data do not allow us to separate various diparton sub-processes or photon polarization states. To obtain a factorizable form of the cross section, we first approximate equation (5.11) by:

$$\frac{\mathrm{d}\sigma^{\mathrm{ep}}}{\mathrm{d}y\mathrm{d}x_\gamma\mathrm{d}x_p\mathrm{d}\cos\hat{\theta}\mathrm{d}Q^2} \approx \frac{1}{32\pi s_{ep}}\frac{f^T_{\gamma/e}(y,Q^2)}{y}\sum_{i,j,k,l}\frac{f_{i/\gamma}(x_\gamma,E_T^2,Q^2)}{x_\gamma}\frac{f_{j/p}(x_p,E_T^2)}{x_p}|M^{kl}_{ij}(\cos\hat{\theta})|^2. \tag{5.12}$$

We have defined a set of photon polarisation-averaged parton densities:

$$f_{i/\gamma}(x_\gamma,E_T^2,Q^2) \equiv f^T_{i/\gamma}(x_\gamma,E_T^2,Q^2) + \bar{\epsilon} f^L_{i/\gamma}(x_\gamma,E_T^2,Q^2), \tag{5.13}$$

where $\bar{\epsilon} \sim 0.8$ is the ratio of longitudinal to transverse photon fluxes, averaged over the y-range of the data. At the moment there are no parametrizations available for the PDF of the longitudinal virtual photon and no attempts to extract this information from the data were made. Therefore neither in HERWIG nor in RAPGAP the contribution of the longitudinal virtual photon was generated. Recent crude estimates of this contribution to the effective PDF of the photon [11, 12] suggest that this contribution may not be negligible and that this question deserves further experimental as well as theoretical investigations.

The concept of the *effective parton distribution functions* had been developed [80] to simplify jet analyses at the CERN $\overline{p}p$ collider and recently was used to investigate real photon structure [61]. The Single Effective Subprocess (SES) approximation exploits the fact that the dominant contribution to the resolved photon cross section (Eq.(5.11)) comes from processes $GG \to GG, qG \to qG$ and $qq' \to qq'$ whose matrix elements have similar shapes and differ mainly by their associated colour factors, namely

$$|M_{qq'\to qq'}|^2 : |M_{qG\to qG}|^2 : |M_{GG\to GG}|^2 \approx 1 : \frac{9}{4} : \left(\frac{9}{4}\right)^2. \tag{5.14}$$

Hence the sum in Eq.(5.12) can be approximated by a simple product

$$\frac{\mathrm{d}\sigma^{ep}}{\mathrm{d}y\mathrm{d}x_\gamma\mathrm{d}x_p\mathrm{d}\cos\hat{\theta}\mathrm{d}Q^2} \doteq \frac{f^T_{\gamma/e}(y,Q^2)}{y}\frac{f^\gamma_{\mathrm{eff}}(x_\gamma,E_T^2,Q^2)}{x_\gamma}\frac{f^p_{\mathrm{eff}}(x_p,E_T^2)}{x_p}|M_{\mathrm{SES}}|^2 \tag{5.15}$$

where the *single effective subprocess* matrix element M_{SES} was defined in [80] as

$$|M_{\mathrm{SES}}(\cos\hat{\theta})|^2 = |M_{qq'\to qq'}(\cos\hat{\theta})|^2$$

and the *effective PDF* of the photon and proton are defined as

$$f^\gamma_{\mathrm{eff}}(x,E_T^2,Q^2) = \sum_i^{n_f}[q_i^\gamma(x,E_T^2,Q^2) + \overline{q}_i^\gamma(x,E_T^2,Q^2)] + \frac{9}{4}g^\gamma(x,E_T^2,Q^2). \tag{5.16}$$

$$f^p_{\mathrm{eff}}(x,E_T^2) = \sum_i^{n_f}[q_i^p(x,E_T^2) + \overline{q}_i^p(x,E_T^2)] + \frac{9}{4}g^p(x,E_T^2). \tag{5.17}$$

We see that the extraction of the effective PDF has a sense only in LO due to the fact that SES approximation itself is defined in LO. The accuracy of this approximation for the real photon is discussed in detail in [81]. Our studies showed that the upper bound of the inaccuracy of SES approximation in our kinematic region is 30%. But the accuracy is a function of dijet invariant mass and E_T of jets and in most of the phase space covered in this analysis it is even better than 10%. This conclusion is in agreement with that made in [81].

The effective PDF of the virtual photon, $f^{\gamma}_{\rm eff}$, was extracted from the data by comparing the unfolded x_γ-dependence of the diparton cross section with that calculated by MC event generator with some photon PDF and correcting the input photon PDF correspondingly. For example, the correction of GRV-LO parametrization in the low Q^2 region proceeded this way:

$$f_{\rm eff}^{\gamma,{\rm DATA}}(x_\gamma, E_T^2, Q^2) = f_{\rm eff}^{\gamma,{\rm GRV-LO}}(x_\gamma, E_T^2, Q^2) L(E_T^2, Q^2, \omega^2) \frac{{\rm d}\sigma^{\rm DATA}/{\rm d}x_\gamma}{{\rm d}\sigma^{\rm MC,GRV-LO}/{\rm d}x_\gamma} \tag{5.18}$$

where ${\rm d}\sigma^{\rm DATA}/{\rm d}x_\gamma$ are the data unfolded to the hard scattering level by HERWIG with GRV-LO photon parametrization. The same procedure was used with RAPGAP and GRV-LO parametrization. By comparing the two results in each x_γ bin we obtain an estimate of the model dependence of this procedure. A typical error was 15–25%, in some bins higher. This error was symmetrized. In spite of the fact that in the ratio in Eq.(5.18) both terms may differ when using different generators, globally this model dependence is of the same size as that estimated for measured cross sections (Section 5.5.2). A similar procedure was then used to estimate the dependence of the result (5.18) on the input photon PDF within one generator. The effective PDF extracted from the data by using HERWIG with SaS1D differed from that obtained with HERWIG with GRV-LO typically by 6–8% in magnitude. A similar effect was observed when SaS2D was used. This error was not symmetrized.

In Fig.5.39 the leading order effective PDF of the virtual photon is presented in five intervals of Q^2 [78]. The data are compared with predictions of GRV-LO parametrization including the suppression factor L for $\omega = 0.2$ and $0.3\,{\rm GeV}$ and also with SAS1D and SAS2D sets. For each interval of Q^2 we also indicate the average value of Q^2 and E_T^2. These numbers show that the condition $Q^2 \ll E_T^2$ is satisfied in the whole kinematic region studied. The virtual photon has time to develop its non-trivial structure and the measurement of its structure is, therefore, meaningful.

The large systematic errors of the data points in Fig.5.39 do not allow us to establish too precisely the value of ω for an optimal description of the data. We conclude that it lies in the range 0.1–0.4 GeV, with slight preference of $\omega = 0.1$–$0.2\,{\rm GeV}$ up to $Q^2 = 4.4\,{\rm GeV}^2$ and $\omega > 0.3\,{\rm GeV}$ for higher Q^2. It may indicate that the functional form of the Drees–Godbole Q^2-suppression (1.34) in the data is slightly different—it falls slowly with Q^2 at low Q^2 and faster at higher Q^2 than the formula (1.34). Nevertheless, despite of the large errors the data suggest that SaS1D in photoproduction and SaS2D in the low Q^2 region lie too low compared to the data. This fact also explains those rather low cross sections of HERWIG with SaS parametrizations seen in Fig.5.38. Overall, both SaS parametrizations decrease with Q^2 at low x_γ faster than the data.

In Fig.5.40 the difference between x_γ-independent and x_γ-dependent Q^2-suppression of the effective PDF is demonstrated, as used by Drees-Godbole approach and in SaS parametrizations. The data from Fig.5.39 are plotted for each x_γ bin as a function of Q^2. Within large errors the Drees-Godbole suppression of GRV-LO with $\omega \sim 0.2$–$0.3\,{\rm GeV}$ describes the data satisfactorily, while neither SaS1D nor SaS2D give a satisfactory description, especially for low x_γ. The data are in a qualitative agreement with those in [31].

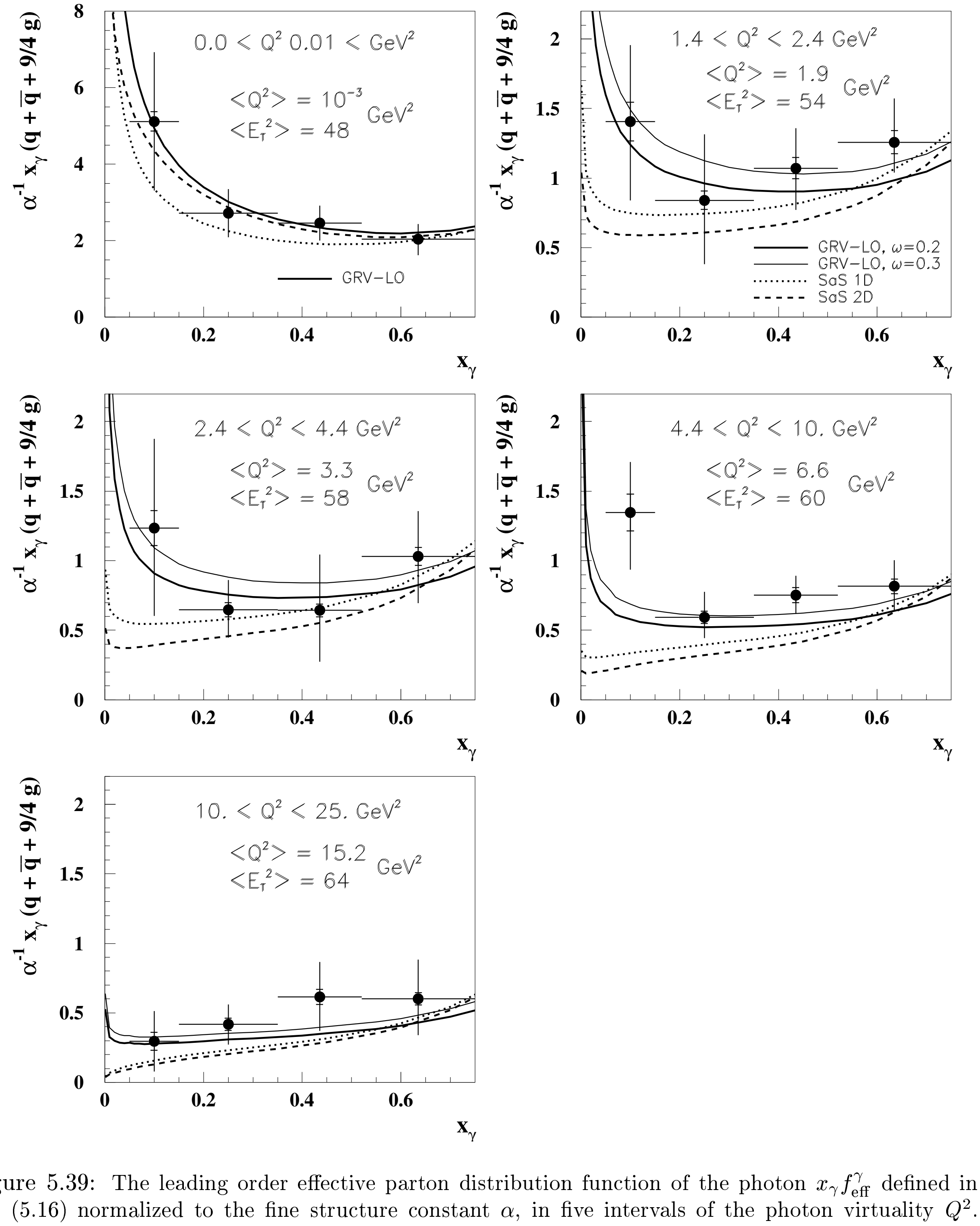


Figure 5.39: The leading order effective parton distribution function of the photon $x_\gamma f_{\rm eff}^\gamma$ defined in Eq. (5.16) normalized to the fine structure constant α, in five intervals of the photon virtuality Q^2. The data (points) are compared to the effective parton distribution of the photon using the GRV-LO parametrization multiplied by suppression function L with ω =0.2, 0.3 GeV (full thick, full thin curve resp.) and SaS1D (dotted curve), SaS2D (dashed curve) parametrizations. Also shown are average values of Q^2 and E_T^2 in single Q^2 interval. The inner error bars represent the statistical errors, the outer error bars give the statistical and systematic errors, added in quadrature.

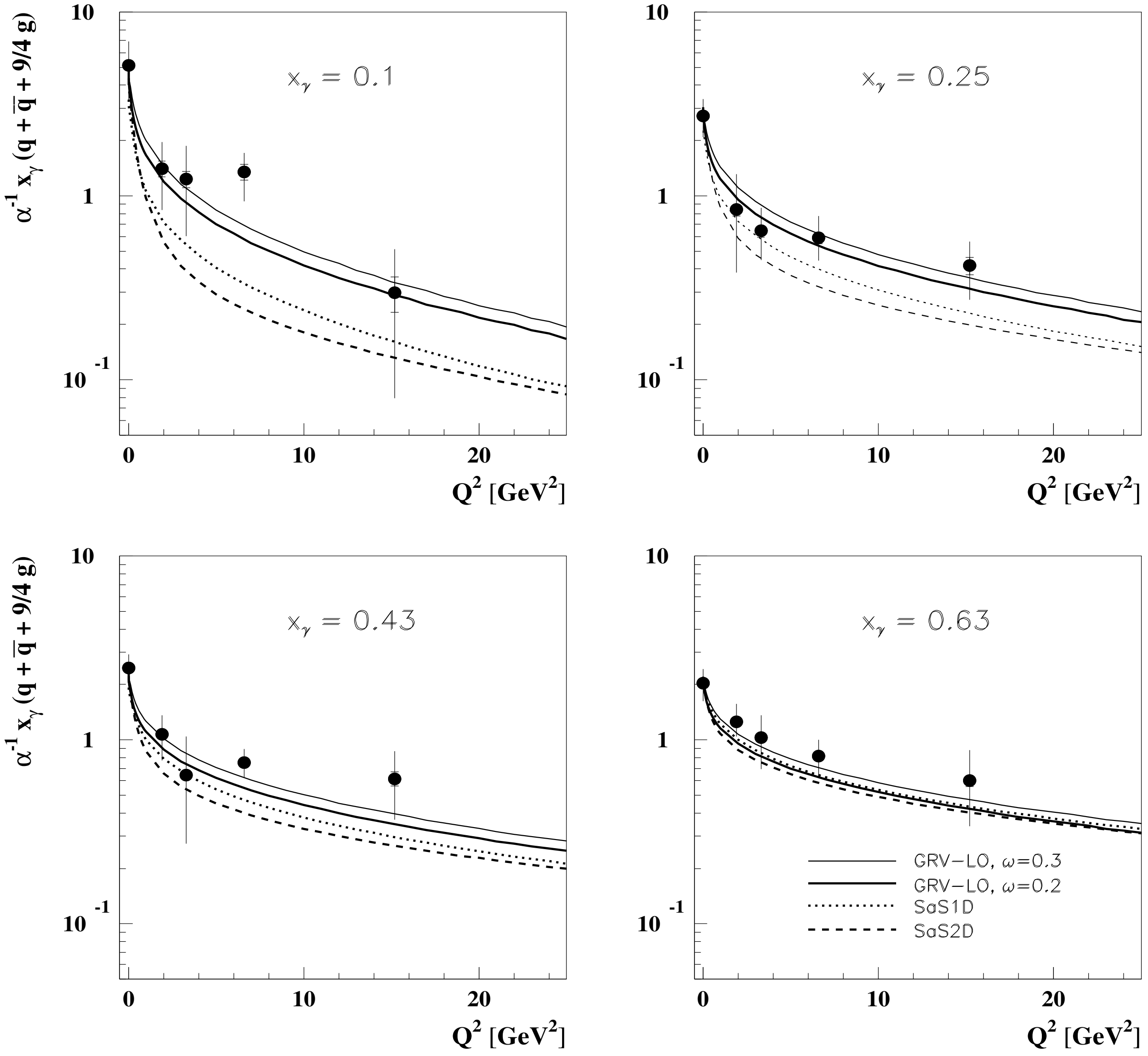


Figure 5.40: The same data points as in Fig.5.39 are compared to Q^2-dependence of various parametrizations of the effective PDF for the virtual photon. The full thick, the full thin, dashed and dotted lines represent GRV-LO multiplied by L with $\omega = 0.2\,\mathrm{GeV}$, GRV-LO multiplied by L with $\omega = 0.3\,\mathrm{GeV}$, SaS1D and SaS2D, resp.

5.6 Conclusions

In this chapter we performed comparisons of the data with Monte Carlo event generators. We obtained several results on different levels of data correction.

On the detector level we described the procedure of fixing the parameters in HERWIG and RAPGAP event samples. With configurations (5.3) the description of the data by HERWIG is very good in all the distributions studied over the whole Q^2 range explored. The description of the data by RAPGAP is good as well except for distributions connected with the soft underlying activity (the jet pedestals and the E_t-flow outside the jets) and with the initial state radiation ($|\Delta\phi|$).

On the hadron level the virtuality dependence of the inclusive dijet cross section $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ was measured in the range $0.0 < Q^2 < 0.01\,\mathrm{GeV}^2$ and $1.4 < Q^2 < 25.\,\mathrm{GeV}^2$. The data were compared to HERWIG and RAPGAP for two values of ω in the suppression function L. In photoproduction a reasonable description of the data by both models is observed. For low x_γ and for $1.4 < Q^2 < 4.4$ GeV^2 both models provide the satisfactory description of the data for $\omega = 0.2$–$0.3\,\mathrm{GeV}$. For $4.4 < Q^2 < 25\,\mathrm{GeV}^2$ the preferred value is $\omega > 0.3\,\mathrm{GeV}$. In the region where the direct photon processes dominate (specifically in the bin $0.75 < x_\gamma < 0.87$) HERWIG underestimates the data. This confirms the observation made on the detector level.

The data were also compared to HERWIG with various parametrizations for the virtual photon PDF, namely GRV-LO suppressed by L with $\omega = 0.2\,\mathrm{GeV}$, SaS1D and SaS2D. The cross sections obtained with both SaS parametrizations lie below the data. The cross sections with SaS1D are lower than those with SaS2D in photoproduction, while in the low Q^2 region it is vice versa.

The ratio $r = \sigma^{\mathrm{res}}/\sigma^{\mathrm{dir}}$ was measured as a function of Q^2. The decrease of r with increasing Q^2 is reasonably described by HERWIG and RAPGAP for $\omega = 0.2\,\mathrm{GeV}$.

On the hard scattering level the virtuality dependence of the diparton cross section $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ was measured in the virtuality range specified above. The same comparisons as on the hadron level were performed and also observations are similar. In photoproduction HERWIG and RAPGAP describe the data satisfactorily. Comparisons have shown that the preferred value of ω is $\sim 0.2\,\mathrm{GeV}$ for $Q^2 < 4.4\,\mathrm{GeV}^2$ and $\omega = 0.2$–$0.3\,\mathrm{GeV}$ for higher Q^2. HERWIG underestimates the data in the bin $x_\gamma > 0.75$. Globally the cross sections calculated using HERWIG with both SaS parametrizations are below those with GRV-LO and $\omega = 0.2\,\mathrm{GeV}$. The cross sections corresponding to SaS1D are lower than those from SaS2D in photoproduction and vice versa in the the low Q^2 region.

On the basis of the comparisons in cross sections we recommend using α_s in LO approximation for ep interactions instead of α_s in NLO which is the default in HERWIG.

The effective PDF of the photon was extracted and its virtuality dependence was studied. The Q^2-dependence as seen in the data is different from that obtained with SaS1D as well as SaS2D and with the suppression of GRV-LO using $\omega = 0.2\,\mathrm{GeV}$. Again it was confirmed that the data prefer $\omega = 0.1$–$0.2\,\mathrm{GeV}$ for $Q^2 < 4.4\,\mathrm{GeV}$ and $\omega > 0.3\,\mathrm{GeV}$ for higher Q^2. In agreement with observations on the hadron level SaS1D in photoproduction and SaS2D at low Q^2 tend to underestimate the data but in many points they are still consistent with the data because of large systematic errors.

Chapter 6

Comparison of the data with NLO calculation

The comparisons of the data with LO MC event generators performed in the previous chapter have convincingly shown that the contribution of resolved photon processes is substantial for a satisfactory description of the data. The studies made in previous chapter, however, revealed several issues. In this chapter we will turn to the comparison of the data with NLO parton level Monte Carlo calculations. As already stated the region of interest, $E_T^2 \gg Q^2 > \langle Q^2 \rangle_{\mathrm{phot}}$, finds itself at a transition between the jet production in photoproduction ($E_T^2 \gg \langle Q^2 \rangle_{\mathrm{phot}}$[1]) and jet production in DIS (predominantly $E_T^2 < Q^2 \gg 0$)

For photoproduction the NLO calculations have been performed by several groups [82, 83, 84]. In these calculations, the addition of resolved photon processes introducing the concept of PDF of the real photon turns out to be indispensable for two reasons: for a successful description of the data and for a theoretically consistent treatment because the singularity coming from the real photon coupling to a collinear pair of massless quarks has to be included in the definition of the photon PDF.

There exist several programs for jet production in DIS, for instance **DISENT**, **MEPJET** or **DISASTER** which do not introduce the concept of PDF of virtual photons. As will be explained later, in the approach where resolved photon processes are taken into account the cross sections given by the above programs are often called DIRECT UNSUBTRACTED.

For a virtual photon the situation is different compared to the real photon as its virtuality shields off the mentioned singularity of the unsubtracted NLO direct contribution and the results are therefore finite even for massless quarks. Consequently, for sufficiently large Q^2, where perturbative QCD can be applied, concept of a PDF of virtual photons does not have to be introduced and the calculations reduce to the direct unsubtracted contribution. Nevertheless, so long as $Q^2 \ll E_T^2$ (as it is in photoproduction) the concept of PDF of the virtual photon makes sense and is actually quite useful phenomenologically. To illustrate this point we shall now discuss dijet cross sections calculated by means of **JETVIP** [85, 86], currently the only NLO parton level program that includes both the direct and resolved photon contributions. The comparisons made in the following using dijet cross sections - for the first time in the kinematic range specified above - fill the gap between comparisons for photoproduction and DIS processes. The kinematic region for the comparisons has been modified compared to that used in previous

[1] In photoproduction processes with tagged electron in H1 the maximal value of virtuality is about 10^{-2} GeV^2 and average value $\langle Q^2 \rangle \simeq 10^{-3}$ GeV^2.

chapter (see Table 4.1) in three aspects: we left out the cut on $\Delta\phi$ that makes the data sample more "LO-like" (the anticorrelation in ϕ-angle of outgoing partons is what one expects in LO picture for the basic hard subprocess with equalized p_T of the partons), and imposed more restrictive cuts on jet E_T and η. The reasons for these additional requirements are explained in the following sections.

The comparisons are carried out for dijet cross sections $\mathrm{d}\sigma/\mathrm{d}E_T$ and $\mathrm{d}\sigma/\mathrm{d}\eta$, defined as the sums of the corresponding distributions of jets with the highest and second highest E_T:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_T} = \frac{\mathrm{d}\sigma}{\mathrm{d}E_{T1}} + \frac{\mathrm{d}\sigma}{\mathrm{d}E_{T2}}, \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}\eta} = \frac{\mathrm{d}\sigma}{\mathrm{d}\eta_1} + \frac{\mathrm{d}\sigma}{\mathrm{d}\eta_2} \tag{6.1}$$

It must be stressed that the comparison is made with the data unfolded to the hadron level while the NLO calculations are performed on the parton level. We therefore need to know the behaviour of hadronization corrections. The problems of hadronization corrections are elaborated in Section 6.7.6.

6.1 Structure of JETVIP

All the above mentioned parton level NLO MC programs contain the same full set of partonic cross sections for the direct photon contribution up to the order $\alpha\alpha_s^2$. Moreover all these programs also contain one-loop corrections to the LO tree diagrams. They differ mainly in the technique used to regularize mass singularities: MEPJET and JETVIP employ the slicing method while DISENT and DISASTER use the subtraction method. As far as the direct processes are concerned, they should provide identical predictions. A nice agreement between JETVIP and DISENT is documented in [85]. The complete set of NNLO (i.e. $\mathcal{O}(\alpha\alpha_s^3)$) matrix elements of the direct photon contribution would require the evaluation of all tree diagrams of order $\alpha\alpha_s^3$ as well as one-loop corrections to $\alpha\alpha_s^2$ diagrams and two-loop corrections to $\alpha\alpha_s$ diagrams. So far, such calculations are not available.

In addition to complete NLO direct photon contribution JETVIP also includes the resolved photon one. Once the concept of virtual photon structure is introduced, part of the direct photon contribution, namely the splitting term (which for the virtual photon is nonsingular, see Eq.1.24) has to be subtracted because it is already included in PDF appearing in the resolved photon contribution. So far only structure functions of transversely polarized photons are known so that the contribution of longitudinally polarized ones are not considered. To avoid confusion we shall use the term "direct unsubtracted" ($\mathrm{DIR_{uns}}$) to denote the NLO direct photon contributions before this subtraction and the term "direct subtracted" ($\mathrm{DIR_s}$) for the results after this subtraction. In this terminology the complete calculation is then given by the sum of direct subtracted and resolved parts and denoted as NLO TOTAL = $\mathrm{DIR_s}$+RES.

The addition of the LO resolved photon contribution means including diagrams which involve convolutions of the PDFs from both proton and photon sides with α_s^2 tree partonic cross sections. For a complete $\mathcal{O}(\alpha_s^2)$ calculation this is all that *has* to be added to the $\mathcal{O}(\alpha\alpha_s^2)$ direct part. However, for reasons discussed in the next subsection, JETVIP also includes NLO resolved contributions, which involve convolutions of PDF with complete α_s^3 partonic cross sections. This might seem inconsistent as no corresponding $\alpha\alpha_s^3$ direct photon terms are included. Nevertheless, this procedure makes sense precisely because of a clear physical meaning of PDF of the virtual photon! Numerically, the inclusion of the NLO resolved terms turns out to be very important and in certain parts of the phase space leads to substantial increase of JETVIP results compared to those of DISENT, MEPJET or DISASTER.

6.2 Factorization mechanism in γp interactions

The main argument for adding the α_s^3 resolved photon terms to $\mathcal{O}(\alpha\alpha_s^2)$ direct and $\mathcal{O}(\alpha_s^2)$ resolved photon contributions is based on the specific way the factorization mechanism works for processes involving initial photons. This point is crucial but subtler and we therefore merely summarize the conclusions and refer to [12, 16] for details. In the absence of $\alpha\alpha_s^3$ direct photon calculations we have two options:

- To stay within the framework of complete $\mathcal{O}(\alpha\alpha_s^2)$ calculations including the LO resolved photon and NLO direct photon contributions, but with no mechanism for the cancellation of the dependence of PDF of the virtual photon on the factorization scale M.
- To add to the previous framework the α_s^3 resolved photon contributions which provide the necessary cancellation mechanism with respect to that part of factorization scale dependence of PDF "generated" by the homogeneous part of the evolution equations (1.17-1.19). The drawback of this procedure is the fact that the NLO resolved photon terms do not represent a complete set of $\mathcal{O}(\alpha_s^3)$ contributions.

In our view the second strategy, adopted in JETVIP, is more appropriate. In fact one can look at the α_s^3 resolved photon terms as the results of an approximate evaluation of the so far uncalculated $\alpha\alpha_s^3$ direct photon diagrams in the collinear kinematics. There are of course $\alpha\alpha_s^3$ direct photon contributions that cannot be obtained in this way, but still it is better than waiting for complete $\mathcal{O}(\alpha\alpha_s^3)$ calculations - particularly in a kinematic region where the NLO direct photon calculations clearly do not suffice for a good description of the data.

Note that the as yet unknown $\alpha\alpha_s^3$ direct photon contribution (as a sum of the collinear and non-collinear gluon emissions) provides the first chance of generating the pointlike part of the gluon distribution function inside the photon. In summary, although the pointlike parts of quark and gluon distribution functions of the virtual photon are in a sense included in higher order perturbative corrections and can therefore be considered as expressions of "interactions" rather than "structure", their uniqueness and phenomenological usefulness definitely warrant their introduction as well as their name.

6.3 E_T cuts

At the NLO one encounters a complication concerning the definition of cuts on jet E_T discussed for instance in [84]. We consider three classes of E_T cuts:

- Symmetric (Fig. 6.1a): $E_{T1} \geq E_T^c, \quad E_{T2} \geq E_T^c$,
- Asymmetric (Fig 6.1b): $E_{T1} \geq E_T^c + \Delta \quad, E_{T2} \geq E_T^c$,
- Sum (Fig. 6.1c): $E_{T1} + E_{T2} \geq 2E_T^c, \quad E_{T2} \geq E_T^c - \Delta$,

where E_T^c and Δ are two parameters specifying the selected region. The symmetric cuts option has been used in most earlier experimental analyses and their comparisons with LO parton level calculations or MC event generators. At the NLO this definition of E_T cuts leads, however, to problems [84]. They come from the region $E_{T1} \approx E_{T2}$, where the proximity of transverse energies of the two highest E_T jets prevents the full cancellation of negative NLO contribution of

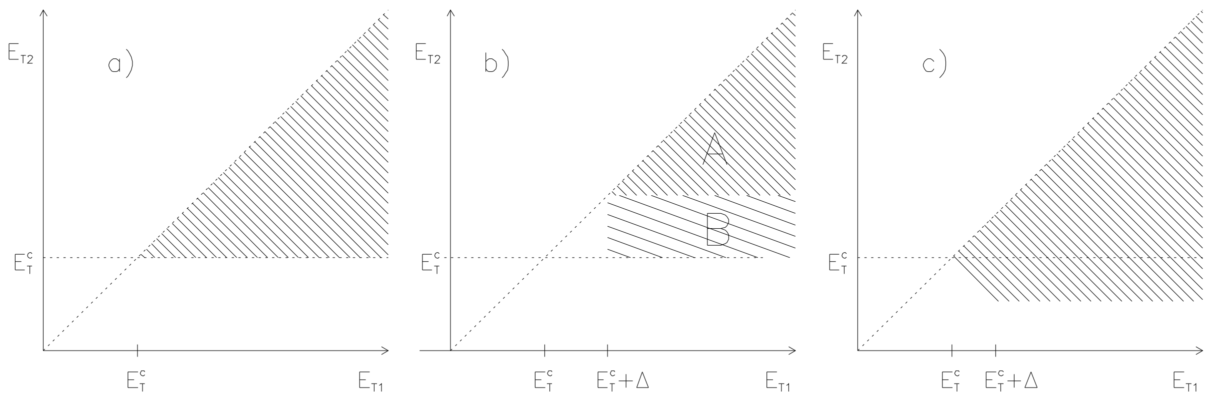


Figure 6.1: Regions of E_{T1}, E_{T2} corresponding to the three options defined in the text.

virtual corrections to LO diagrams by positive NLO contribution of real parton emissions. This leads to the unphysical behaviour of NLO calculations as a function of Δ, displayed in Fig.1 of [84], as $\Delta \to 0$: while the phase space increases, the NLO cross section for the asymmetric cut option nevertheless decreases. At fixed (next-to-leading) order there is no way how to avoid this problem entirely [2]. Cutting out the whole region $E_{T1} \approx E_{T2}$ is no remedy as then the $O(\alpha_s^2)$ calculations become of LO only. Keeping part of the phase space which contains events with $E_{T1} \approx E_{T2}$ is therefore a necessity. In these conditions the proper definition of E_T cuts depends on the quantity measured. The asymmetric cut option is appropriate if one plots separately the E_{T1} and E_{T2}-distributions or their sum as in Eq.(6.1). On the other hand, if, as in [61], the average E_T of first two jets, $E_T \equiv (E_{T1}+E_{T2})/2$, is measured then the sum cut option of Fig.6.1c should be chosen. For calculating the cross sections defined in Eq.(6.1) we therefore choose the asymmetric cut option. So as not to lose much statistics in the data and simultaneously to get reliable cross sections we have to investigate the behaviour of the integral over the selected region (labelled A+B in Fig.6.1b) in dependence on Δ. The results are shown in Fig.6.2. Note that for $\Delta \to 0$ (the phase space increases) the TOTAL result given by NLO calculations decreases! This unphysical behaviour which is due to the limited phase space for real parton emission can never occur in the data and one has to therefore choose Δ somewhere behind the bending point. Without doubt, the option $\Delta = 2$ GeV is very safe.

6.4 Correction procedure

The correction procedure is the same as in the case of x_γ-distribution described in Section 5.4, namely making use of the unfolding program based on Bayes theorem. The correction factors from the bin-to-bin method together with the smearing matrices are shown in Fig.6.3. While the correlations between detector and hadron level for η-distributions are almost diagonal the correlations for E_T-distributions behave differently. These facts are reflected in the resolution obtained which amounts to 13% (20% in photoproduction) for η and 25% (30% in photopro-

[2]The problem can be cured only by exponentiating the leading soft and collinear emissions.

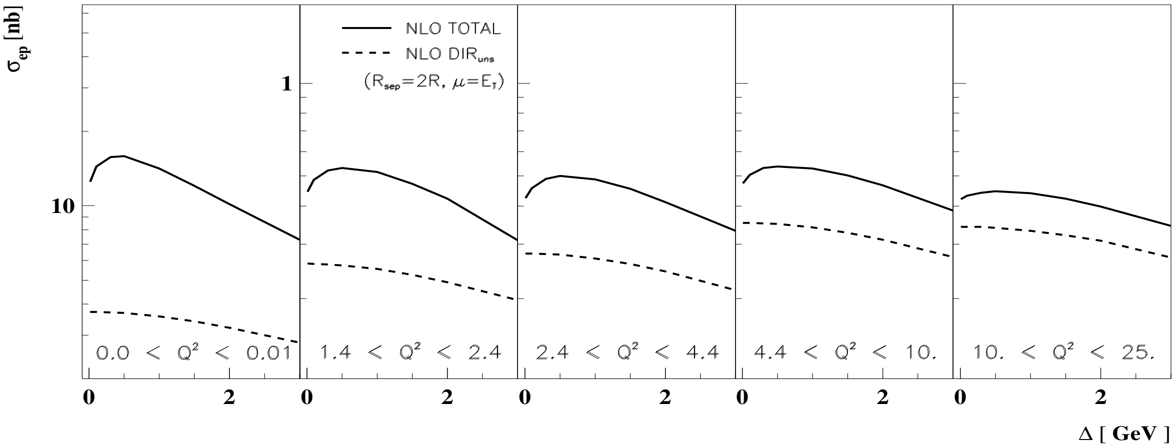


Figure 6.2: The dependence of integral $\sigma(\Delta)$ over the cut area A+B in Fig.6.1b on the asymmetry parameter Δ. The full line corresponds to the TOTAL result, the dashed one to the $\mathrm{DIR_{uns}}$ component, both of the NLO calculations by JETVIP.

duction) for E_T-distributions. For convergent running of the unfolding program however it is important that some correlation exists. Studying the convergence of the unfolding program and the description of the data by the reweighted MC event sample for different numbers of iterations we chose the third iteration for η and the fourth iteration for E_T-distributions as the reference iterations.

The effect of reweighting both HERWIG and RAPGAP MC event samples according to the E_T and η-distributions is shown in Fig.6.4. The only distributions on which the comparison can be made are $\mathrm{d}\sigma/\mathrm{d}E_T$, $\mathrm{d}\sigma/\mathrm{d}\eta$ and $\mathrm{d}\sigma/\mathrm{d}\phi$ since they are constructed as sums of distributions of the two highest E_T jets. The unfoldings of $\mathrm{d}\sigma/\mathrm{d}E_{T1}$, $\mathrm{d}\sigma/\mathrm{d}E_{T2}$, $\mathrm{d}\sigma/\mathrm{d}\eta_1$ and $\mathrm{d}\sigma/\mathrm{d}\eta_2$ distributions each provide their own set of weights (see Eq.(5.6)) so it is straightforward to reweight these distributions in contrast to finding some weight for reweighting the distributions of a quantity characterizing the whole event (e.g. $\mathrm{d}\sigma/\mathrm{d}Q^2$). In Fig.6.4 the results represent predictions of the original and the reweighted MC samples with regard to the data on detector level. In the upper row the $\mathrm{d}\sigma/\mathrm{d}E_{T1}$ distribution was reweighted according to the results of the unfolding in $\mathrm{d}\sigma/\mathrm{d}\eta_2$. In the following rows $\mathrm{d}\sigma/\mathrm{d}E_{T2}$ distribution was reweighted according to $\mathrm{d}\sigma/\mathrm{d}\eta_1$ distribution, $\mathrm{d}\sigma/\mathrm{d}\eta_1$ according to $\mathrm{d}\sigma/\mathrm{d}E_{T2}$ and $\mathrm{d}\sigma/\mathrm{d}\eta_2$ according to $\mathrm{d}\sigma/\mathrm{d}E_{T1}$ distributions. We see that some improvement was achieved everywhere. The last row is to show that a sizable improvement is reached whenever the reweighting of distribution of a quantity is driven by the unfolding of this quantity ($\mathrm{d}\sigma/\mathrm{d}\eta_2$ according to $\mathrm{d}\sigma/\mathrm{d}\eta_2$ in this case).

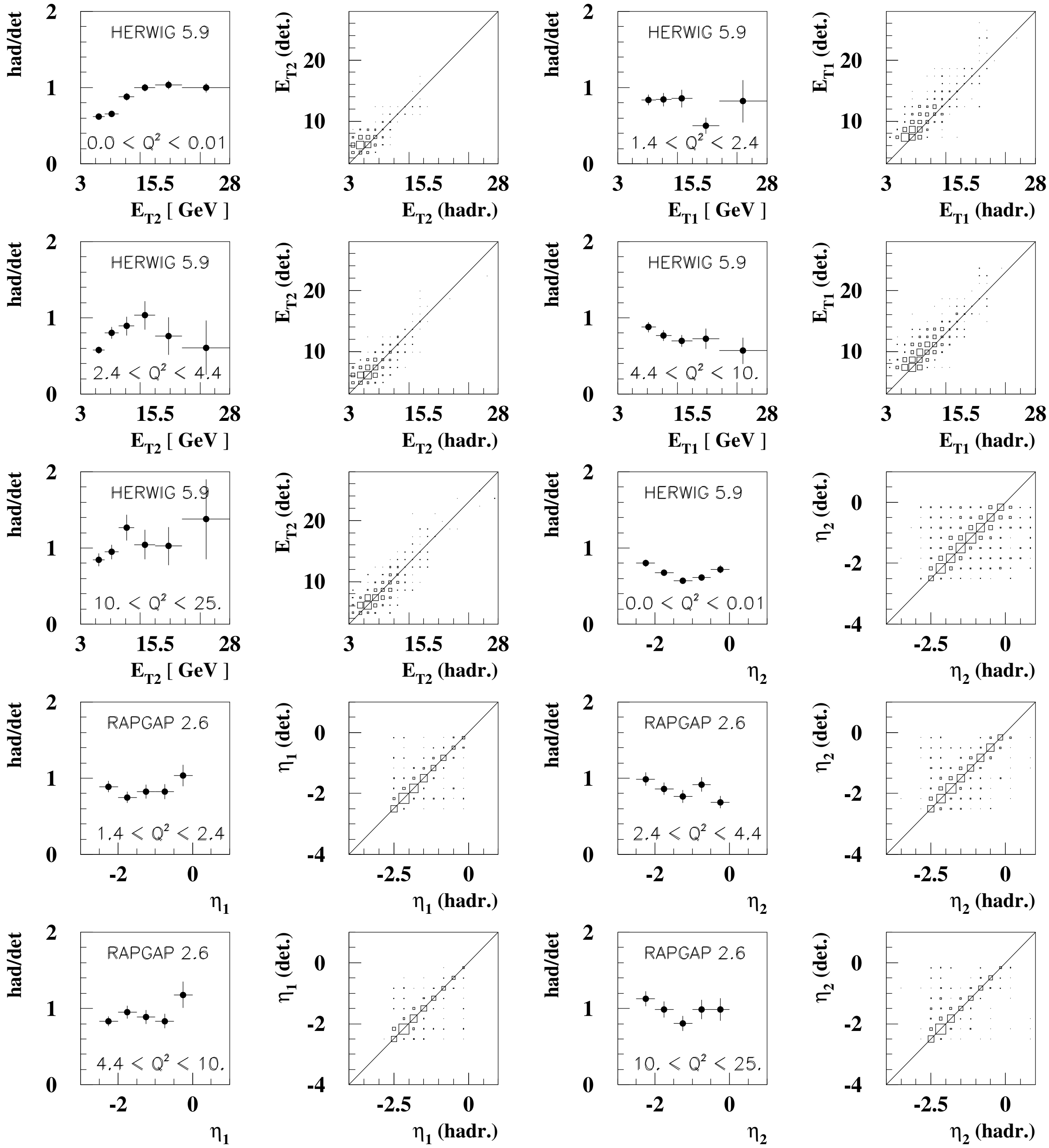


Figure 6.3: The correction factors from the bin-to-bin method and the smearing matrices of detector vs. hadron level for the unfolding program are shown. For E_T-distributions predictions of HERWIG and for η-distributions of RAPGAP are shown.

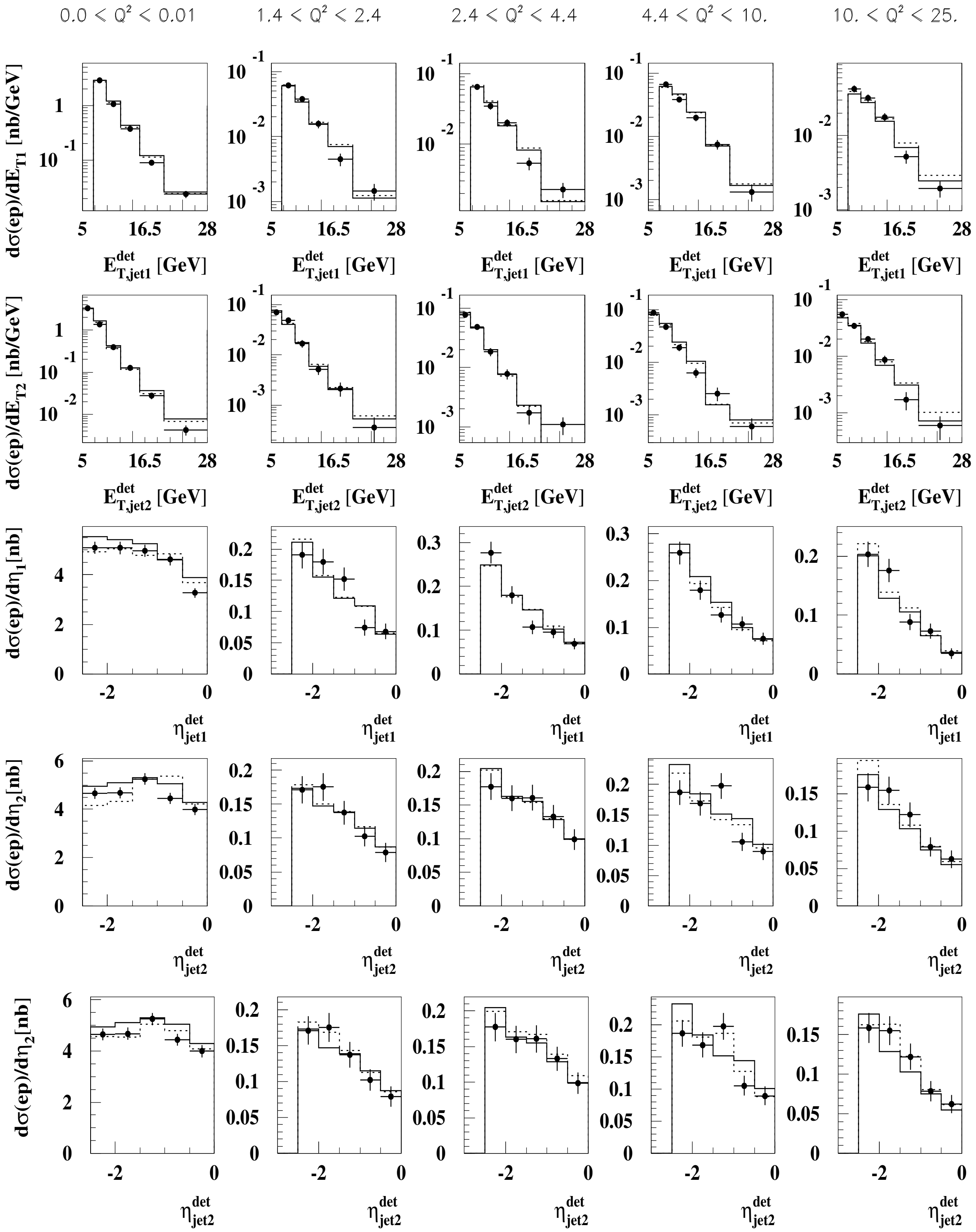


Figure 6.4: The effect of the reweighting of detector level of MC event sample according to the results of the unfolding. The full histogram shows the original and the dashed histogram the reweighted predictions of HERWIG on the detector level. The full points represent the uncorrected data. The reweighting in the 1. (2., 3., 4., 5.) row was done according to the results of the unfolding of η_2 (η_1, E_{T2}, E_{T1}, η_2)-distribution.

The covariance matrix

The definition of the covariance matrix is given in Section 5.4.1. The correlations of the resulting numbers in bins of η-distributions are below 45% everywhere and they are not shown. The correlations of bins in E_T-distributions show rather higher values because of the steeply falling cross section. The numbers are summarized in Tables 6.1 and 6.2.

E_{T1} bins / Q^2 [GeV^2]	1-2	2-3	3-4	4-5	5-6
0.0 - 0.01	-	44.1	53.2	55.8	49.7
	-				
1.4 - 2.4	-	42.1	47.4	58.9	-
	-	48.1	49.3	-	59.3
2.4 - 4.4	-	46.6	59.4	53.0	47.8
	-	52.2	-	-	46.5
4.4 - 10.	-	53.9	51.2	45.6	-
	-	54.7	-	55.9	-
10. - 25.	-	47.8	55.0	54.9	46.9
	-	53.6	45.3	44.0	-

Table 6.1: The percentage for the off-diagonal elements of the covariance matrix of the unfolded numbers in E_{T1}-spectrum. The upper numbers in a row correspond to the use of HERWIG, the lower ones to RAPGAP. In photoproduction only HERWIG was used. Only correlations larger than 40% are shown.

E_{T2} bins / Q^2 [GeV^2]	1-2	2-3	3-4	4-5	5-6
0.0 - 0.01	49.4	53.7	56.8	50.3	42.3
1.4 - 2.4	63.5	50.7	49.4	44.9	41.1
	-	-	53.0	48.5	48.7
2.4 - 4.4	69.6	52.8	46.6	-	-
	-	-	-	41.5	40.9
4.4 - 10.	57.2	50.0	44.9	61.2	49.2
	50.3	71.4	63.4	-	-
10. - 25.	52.8	52.9	47.4	51.1	44.6
	56.5	51.1	-	48.0	46.7

Table 6.2: The percentage for the off-diagonal elements of the covariance matrix of the unfolded numbers in E_{T2}-spectrum. The upper numbers in a row correspond to the use of HERWIG, the lower ones to RAPGAP. In photoproduction only HERWIG was used. Only correlations larger than 40% are shown.

6.5 Systematic errors

When evaluating the systematic errors we considered the same set of sources as specified in Section 5.4. The largest contribution comes from the model dependence and it ranges typically between 15 and 30%, the LAr-energy scale uncertainty amounts typically to 10 - 15%. The uncertainties in measuring the electron energy in the SpaCal and polar angle in the BDC do not contribute more than 10%. The instability of the unfolding does not exceed 10%.

6.6 Technical parameters of NLO calculations

- The number of points chosen for each integration (100k–120k) gives a statistical error on the NLO calculations smaller than 1% everywhere. As a nontrivial check on the theoretical calculations we have checked that the integrals over the inclusive spectra in η and E_T coincide to better than 3%.
- The default value of the phase space slicing parameter was $y_c = 3.10^{-3}$. The independence of the NLO calculations on y_c was checked for $y_c \in\ < 10^{-2}, 10^{-4} >$.
- The default values were taken for the number of iterations (5) and the precision of the integration (10^{-5}).

6.7 Ambiguities in NLO calculations

Finite order perturbation theory suffers from ambiguities resulting from the truncation of perturbation expansions. Moreover, there are other ambiguities or corrections that must be taken into account before drawing conclusions from the comparison of theoretical calculations with experimental data. The most important are:

6.7.1 Jet algorithm ambiguities

The jets in the data are found by the cone algorithm. The point is that at the NLO parton level all jet algorithms are equivalent to the cone algorithm when supplemented with the parameter $R_{\rm sep}$ introduced in [87] in order to bridge the gap between the application of the cone algorithm to hadronic systems (from data or MC) and to NLO parton level calculations where one encounters ambiguities concerning the seed selection and jet merging. In a general cone algorithm two objects (partons, hadrons or calorimetric cells) belong to a jet if they are within the distance R from the jet centre. Their relative distance may, however, be bigger as it satisfies a weaker condition

$$\Delta R_{ij} = \sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2} \leq \frac{E_{T_i} + E_{T_j}}{\max(E_{T_i}, E_{T_j})} R. \tag{6.2}$$

In an attempt to model different seed selection procedures at the hadron level, the parameter $R_{\rm sep}$ has been introduced into NLO parton level calculations by demanding that two partons form a jet only if their relative distance ΔR_{ij} satisfies the condition

$$\Delta R_{ij} \leq \min\left[\frac{E_{T_i} + E_{T_j}}{\max(E_{T_i}, E_{T_j})} R, R_{\rm sep}\right]. \tag{6.3}$$

The region of meaningful $R_{\rm sep}$ is obviously $R \leq R_{\rm sep} \leq 2R$. The parameter $R_{\rm sep}$ therefore governs the maximal distance between two partons inside a single jet. For two partons of equal E_T, $R_{\rm sep} = 2R$ means that they may be $2R$ apart, but still within the same jet. The question of which value of $R_{\rm sep}$ to choose for the comparison of NLO parton level calculations with the results of applying the cone algorithm at the hadron level is nontrivial and we shall therefore present NLO results for both extreme choices $R_{\rm sep} = 1R$ and $R_{\rm sep} = 2R$. To define momenta of jets JETVIP uses the standard E_T-weighting recombination procedure which leads to massless jets. The NLO calculations were performed for the same value of the cone size $R = 1$ as the data.

The Figs. 6.5 and 6.6 demonstrate the dependence of NLO results on $R_{\rm sep}$ parameter. The influence on the $\rm DIR_{uns}$ is very small, it reaches at most -6%, namely at $\eta = -2.5$ for the last Q^2 interval. The influence on the TOTAL result is rather bigger, being largest at the same bin (-11% independently of Q^2), and globally even enlarging the difference between the data and calculations.

6.7.2 Factorization scheme dependence

We shall disregard this ambiguity although its numerical importance may not be negligible and is closely related to the factorization scale dependence.

6.7.3 Scale dependence

Factorization scale dependence

In principle proton and photon (in resolved channel) are associated with different factorization scales M_p and M_γ, but we shall follow the standard practice and set $M \equiv M_p = M_\gamma$. In JETVIP one has $M = E_{T1}$.

Renormalization scale dependence

The dependence of the calculations on the renormalization scale μ is in principle a separate ambiguity, unrelated to that of the factorization scale, but we again follow the common practice and identify these two scales, i.e. set $\mu = M$. To reflect this identification, in the following we shall simply use the term "scale" dependence to describe the dependence on this common scale.

The scale dependence was estimated by relating the calculations for two values of μ: $\mu = E_{T1}/2$ and $2E_{T1}$ to those with $\mu = E_{T1}$, in all three cases taking $R_{\rm sep} = 2R$. As evident from Figs.6.7 and 6.8 in general the $\rm DIR_{uns}$ term is considerably affected. The scale change shifts the $\rm DIR_{uns}$ contribution globally upwards for $\mu = E_{T1}/2$ (from $+15\%$ at $\eta = -2.5$ to $+37\%$ at $\eta = 0$) and downwards for $\mu = 2E_{T1}$ (from -12% at $\eta = -2.5$ to -25% at $\eta = 0$), independently of Q^2. The scale dependence of the TOTAL result is a little more complicated. Looking at η-spectra we observe an increasing influence of the scale change with increasing Q^2 and increasing η whereas in the lowest Q^2 interval a small but perceptible change is seen only in the most backward region (upper bounds are given by -12% and $+17\%$ both for $\eta = 0$ and the last Q^2 interval). These dependences are transformed to E_T-spectra in the following way: the beginnings of the spectra become more and more affected as Q^2 increases, the other bins are changed equally independently of Q^2. These findings may be understood by realizing three well-known facts: in the direct channels the dominant effect of decreasing μ is an increase of $\alpha_s(\mu)$, while the corresponding variation of proton PDF is much weaker. In resolved photon processes, the scale dependence of the virtual photon PDF also comes into play. Since it is driven by the ratio

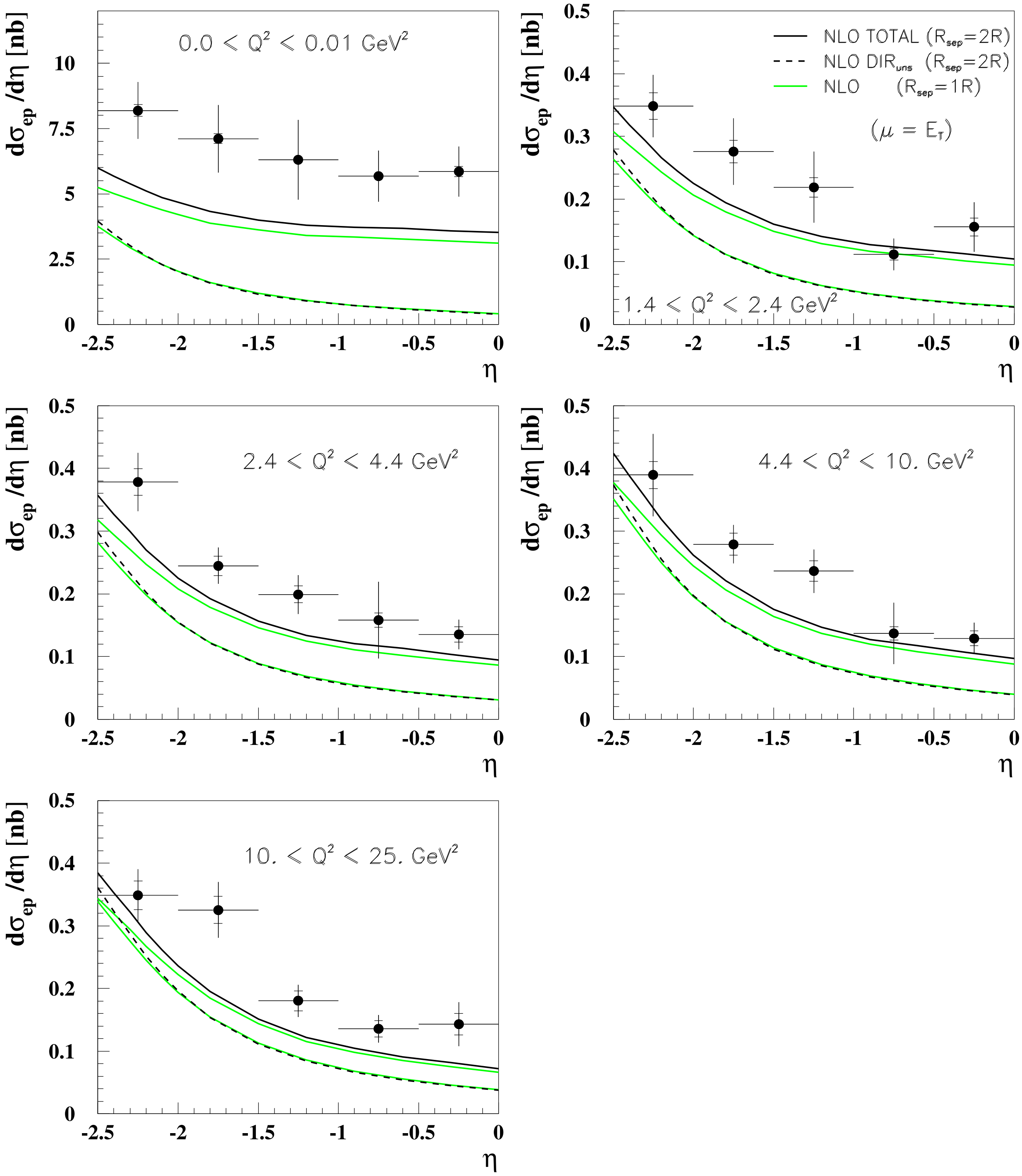


Figure 6.5: The inclusive dijet ep cross sections as a function of the jet pseudorapidity η in five intervals of the photon virtuality Q^2 (points) compared to NLO calculations of JETVIP using the CTEQ4M parametrization for proton and SaS1D for photon PDF. The full curve represents the TOTAL result, the dashed curve represents the direct unsubtracted term, both calculated for $R_{\rm sep} = 2R$, the dotted curves correspond to $R_{\rm sep} = 1R$, all for the scale $\mu = E_{T1}$. The inner bars indicate the statistical errors and the total error bar shows the quadratic sum of statistical and systematic errors.

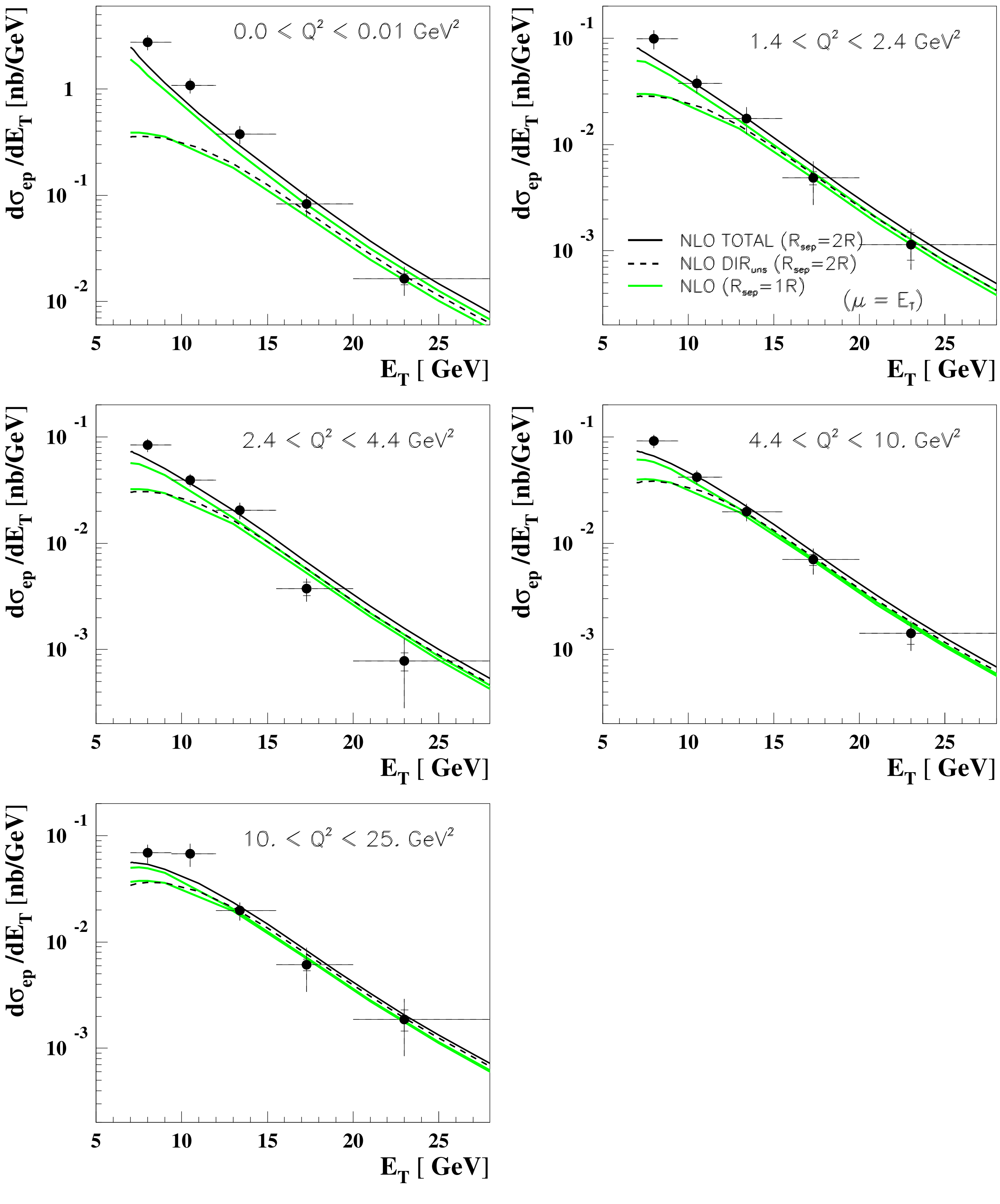


Figure 6.6: The inclusive dijet *ep* cross sections as a function of the jet transverse energy E_T in five intervals of the photon virtuality Q^2 (points) compared to NLO calculations of JETVIP using the CTEQ4M parametrization for proton and SaS1D for photon PDF. The full curve represents the total cross-section, the dashed curve represents the direct unsubtracted term, both calculated for $R_{\mathrm{sep}} = 2R$, the dotted curves correspond to $R_{\mathrm{sep}} = 1R$, all for the scale $\mu = E_{T1}$. The inner bars indicate the statistical errors and the total error bar shows the quadratic sum of statistical and systematic errors.

μ^2/Q^2 a decrease with decreasing scale can be large (as documented in [11, 12]). Replacing $\mu^2 = (E_{T1})^2$ with $\mu^2 = (E_{T1})^2 + Q^2$ changes the $\mathrm{DIR}_{\mathrm{uns}}$ contribution very little, the variation being perceptible only in the highest Q^2 interval.

6.7.4 Choice of PDF

As recommended by the authors of JETVIP we have taken CTEQ4M [41] and SAS1D [22, 23] sets of PDF of the proton and photon respectively. Both of these sets treat quarks, including c and b ones, as massless above their respective mass thresholds, as required by JETVIP, which uses LO and NLO matrix elements of massless partons. CTEQ4M version is a NLO parametrization with $\overline{\mathrm{MS}}$ factorization scheme and $\Lambda^{(5)}_{\overline{\mathrm{MS}}} = 204\,\mathrm{MeV}$. This Λ value is also used to calculate α_s from the two-loop formula at the scale μ. Contrary to CTEQ4M set all SaS sets are LO parametrizations only. This implies incomplete cancellation of the M_γ-dependence within the set of NLO resolved photon contributions. Despite this drawback, we prefer SaS parametrizations to the GRS one [27], the only existing parametrization of virtuality dependent PDF of the photon that is claimed to be of the NLO. Our reasons are threefold. First, as argued in [16], none of the currently available parametrizations of photonic PDF that claim to be of the NLO incorporates all of the NLO effects. Secondly the PDF of GRS correspond to $N_f = 3$ flavours, so that the charm and bottom contributions must be added as an extra contribution which is not done so in JETVIP. Thirdly, SaS parametrizations provide separate parametrizations of VDM and pointlike components of all PDF, which is crucial for the physically transparent interpretation of JETVIP results.

The PDF of the proton is fairly well determined from global analyses of the CTEQ and MRS groups and we have therefore estimated the residual uncertainty related to the choice of proton PDF by comparing the results obtained with CTEQ4M to those obtained with MRS(2R) set. The differences for both $\mathrm{DIR}_{\mathrm{uns}}$ and TOTAL are very small, both between 1% at $\eta = -2.5$ and 3.5% at $\eta = 0$ and between 1% at $E_T = 7$ GeV and 6% at $E_T = 28$ GeV, independently of Q^2.

6.7.5 Choice of number of flavours

As said in the previous subsection JETVIP calculates matrix elements for massless partons. However the b quark may be considered massless only in processes where its mass (≈ 4.6 GeV) is negligible with regard to the hard scale. In our processes $\mu \geq 6\,\mathrm{GeV}$ and average value $\langle\mu\rangle \simeq 10\,\mathrm{GeV}$ which implies a number of flavours somewhere between 4 and 5. We set $N_f = 5$ everywhere which means that a contribution from $b\bar{b}$-pair production was added[3]. By setting $N_f = 5$ we get an upper bound for JETVIP predictions as far as the choice of N_f is concerned. By decreasing N_f from 5 to 4 the results of $\mathrm{DIR}_{\mathrm{uns}}$ contribution lower uniformly by 8–9% while the TOTAL result comes down by 3–7%, independently of Q^2.

6.7.6 Hadronization corrections

In the NLO QCD calculations jet is formed out of one or two partons, whereas the jets at hadron level to which the data are unfolded are formed by many hadrons. The question of estimating hadronization corrections is by itself a rather complicated task. In order to compare these two different levels we have to know what the hadronization process makes from the outgoing

[3]Contribution of incoming b quarks from the proton is tiny according to CTEQ4M parametrization

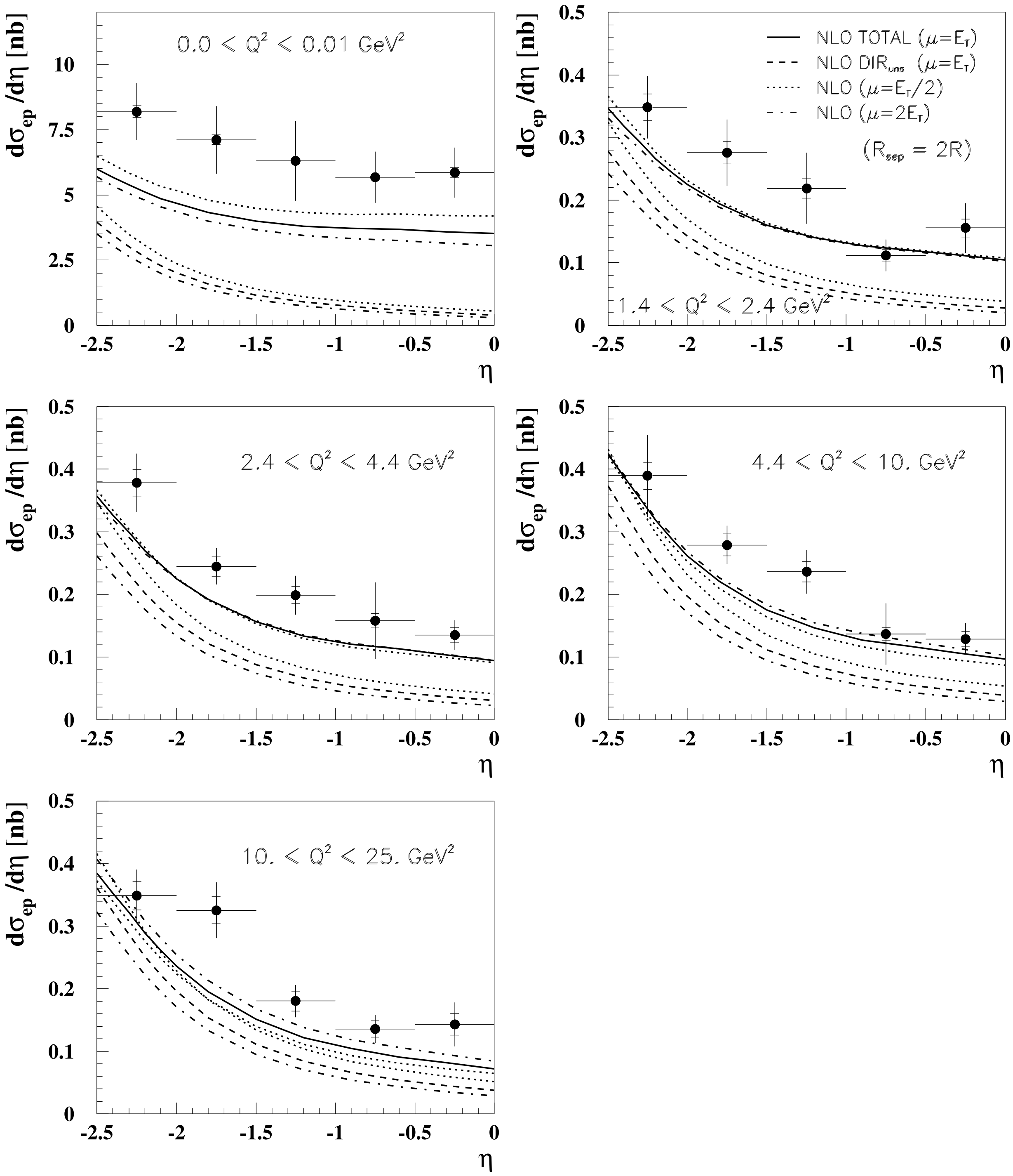


Figure 6.7: The scale dependence of the NLO calculations. The data points, solid and dashed lines are taken over from Fig.6.5, the dotted lines correspond to the scale $\mu = E_{T1}/2$, the dashed-dotted to the scale $\mu = 2E_{T1}$. $R_{\rm sep} = 2R$ for all the curves.

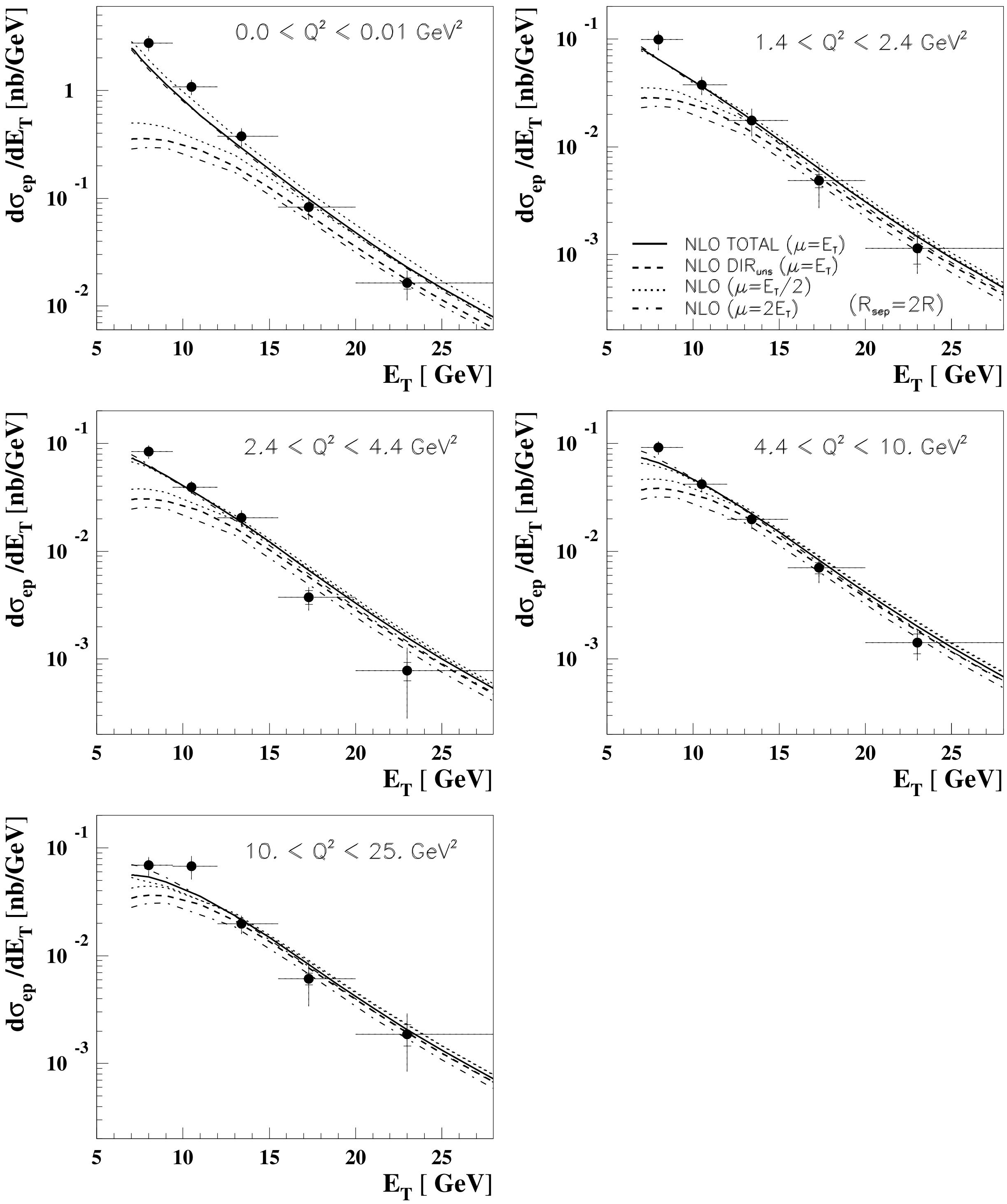


Figure 6.8: The scale dependence of the NLO calculations. The data points, solid and dashed lines are taken over from Fig.6.6, the dotted lines correspond to the scale $\mu = E_{T1}/2$, the dashed-dotted to the scale $\mu = 2E_{T1}$. $R_{\rm sep} = 2R$ for all the curves.

partons. The hadronization corrections are therefore quite naturally defined as ratio of cross sections corresponding to these two levels, namely

$$hc(E_T) \equiv \frac{\mathrm{d}\sigma^{\mathrm{parton}}}{\mathrm{d}E_T} \Bigg/ \frac{\mathrm{d}\sigma^{\mathrm{hadron}}}{\mathrm{d}E_T} \tag{6.4}$$

and analogously for η-distributions.

However, the problem is that all MC event generators use LO matrix element only supplemented with the parton showers in LLA, whereas in NLO calculations the integrations are being performed over the full angular range. The approximation currently used for the estimating of the hadronization corrections is to look for jets found on partons from the hard scattering and QCD parton showers. The parton showers are the only chance of getting unbalanced transverse energies of the parton jets.

A further problem arises from the question of selection of kinematic ranges for comparison of cross sections corresponding to the two levels. For our discussion the following three possibilities are relevant:

1. to apply the same cuts on quantities on both levels independently of each other
2. in events selected by cuts on parton level quantities to apply cuts on hadron level quantities
3. in events selected by cuts on hadron level quantities to leave parton level quantities without restriction

Compared to the first one the second as well as the third possibility certainly offers more information. The second possibility tells us what would happen with results of NLO calculations after application of some hadronization mechanism. The selected hadron jets observed in experiments, however, do not only originate from such parton jets but also from those which did not pass through the selection criteria. We have chosen the third possibility as it provides information about all parton jets giving rise the selected hadron jets, thus it tells us about event migration. It is obtainable from LO event generators because they allow the hard parton subprocesses to start with much smaller $\hat{p}_T$ than the required minimum E_T of hadron jets.

In Fig.6.9 the hadronization corrections as obtained from HERWIG and RAPGAP as functions of E_T and η are presented. The results are shown for both direct photon and the TOTAL results. We see that the generators give very similar predictions for direct photon processes. Systematically higher hadronization corrections from RAPGAP compared to those from HERWIG for the total jet cross sections are expected due to the absence of the model for additional transverse activity in RAPGAP. The difference is most evident in photoproduction where in HERWIG PRSOF= 0.25 (0.15 in the low Q^2 region). The higher value of PRSOF, the higher number of hadron jets that pass the selection criteria. We also compare the hadronization corrections for two different cut scenarios. The corrections corresponding to scenario 3. range between -30% and $+15\%$ in E_T and between -20% and 0% in η-distributions for both direct photon and the TOTAL results. The application of scenario 1. gives slightly different values compared to the latter, with the largest deviations of the order of 10–15 % residing in regions of small E_T and large η where the difference between the data and $\mathrm{DIR}_{\mathrm{uns}}$ is anyhow largest.

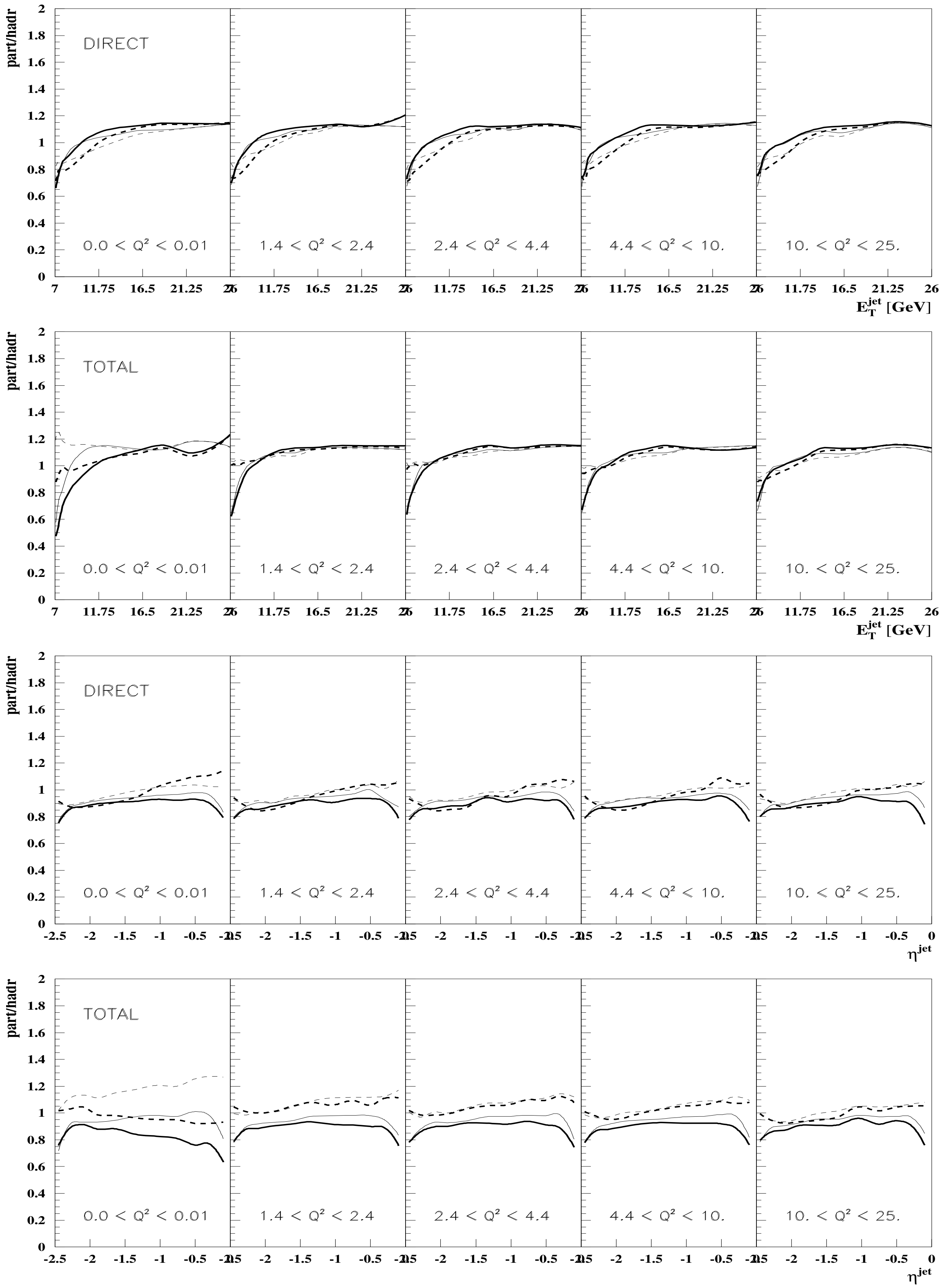


Figure 6.9: The hadronization corrections from HERWIG (thick lines) and RAPGAP (thin lines). The dashed line corresponds to the application of the same cuts independently on both levels, the full line corresponds to the application of cuts on hadron level leaving parton level without restriction.

6.8 η-range or historical remark

We now return to the question of the η-range used for comparisons with NLO calculations. Originally we performed a similar study for jets in the region $-3 < \eta < 0$ which was also used for the x_γ-distributions and effective PDF (Chapter 5). A detailed experimental procedure, discussion of results and conclusions are given in [88]. The results discussed in this section contributed to [89, 90, 91] as well. A summary of main facts from [88] follows:

$\mathcal{A}$: $\mathrm{DIR_{uns}}$ contribution lies significantly below the data over the whole η-range except for the bin $-3 < \eta < -2.5$ bin.

$\mathcal{B}$: $\mathrm{DIR_{uns}}$ contribution lies significantly below the data only in the range $5 \le E_T \lesssim 7\,\mathrm{GeV}$, otherwise it describes the data. In this first bin the hadronization corrections are also largest.

$\mathcal{C}$: The TOTAL result describes the data in the η-distributions if we take systematic errors into account.

$\mathcal{D}$: The bin $-3 < \eta < -2.5$ is "insubordinate": the scale dependence as well as R_{sep} dependence is largest there. In particular, the hadronization corrections are sizable there (40%) while in other bins they range around 10%.

In [88] we concluded that $\mathrm{DIR_{uns}}$ contribution is not able to describe the data. The predictions of the NLO calculations moved closer to the data when the resolved photon contribution was included. However, the facts summarized in the point $\mathcal{D}$, and especially the large hadronization correction, did not allow us to draw clear conclusion about comparisons in the bin $-3 < \eta < -2.5$ where the data were below the $\mathrm{DIR_{uns}}$ results. Therefore we decided to cut back the η-region from $-3 < \eta < 0$ to $-2.5 < \eta < 0$. We decided as well not to show the very first bin in E_T-spectra, i.e. the bin $5 \le E_T \le 7$ GeV, because of large hadronization corrections and the fact that in this bin the NLO calculations are of LO nature only.

6.9 Discussion

Let us come back to the region $-2.5 < \eta < 0$. Though the data points are accompanied by large systematic errors Fig.6.5 and 6.7 convincingly show that $\mathrm{DIR_{uns}}$ contribution is systematically below the data over the whole region of η [4]. The inclusion of the resolved photon contribution shifts the NLO calculations closer to the data; the shape is in a very good agreement with that of the data but the TOTAL result does not suffice to describe the magnitude. The Fig.6.5 suggests that the nontrivial structure of the virtual photon PDF (i.e. the difference between what $\mathrm{DIR_{uns}}$ predicts and what the data require) is clearly required up to 25 GeV^2.

In E_T-spectra (Fig.6.6 and 6.8) we observe that $\mathrm{DIR_{uns}}$ contribution is not able to describe the data in the range $7 \le E_T \lesssim 15$ GeV [5]. In the first bin this discrepancy is very large and shrinks with rising E_T [6]. It is also the case that in this range the inclusion of the resolved photon contribution brings the TOTAL result considerably closer to the data but in the first bin it is still insufficient.

[4]Compare with point $\mathcal{A}$ in Section 6.8

[5]Compare with point $\mathcal{B}$ in Section 6.8

[6]The largest difference, however, still resides in the bin $5 \le E_T \le 7$ which we decided not to show.

A study devoted to the ambiguities of the NLO calculations clearly showed that the largest changes of $\mathrm{DIR_{uns}}$ results are caused by the scale ambiguity while the largest changes of TOTAL results are caused by freedom of R_{sep}. None the less these facts do not invalidate the main conclusions drawn from this comparison. The Fig.6.5 and 6.6 prove that the $\mathrm{DIR_{uns}}$ results depend only weakly on R_{sep}. As stated in Section 6.7.6 the hadronization corrections are also kept under control. We can therefore conclude that by omitting the bin $-3 < \eta < -2.5$ from the analysis the region with largest theoretical ambiguities was removed. We thereby made the predictions of NLO calculations more solid.

We have also investigated the question of from which regions of E_{T1} and E_{T2} the large discrepancy between the data and $\mathrm{DIR_{uns}}$ contribution actually originates. From Fig.6.10 it is apparent that the regions responsible for the discrepancy are $7 \leq E_{T1} \lesssim 15$–$20\,\mathrm{GeV}$ and $5 \leq E_{T2} \lesssim 12.5$ GeV. The $\mathrm{d}\sigma/\mathrm{d}\eta_1$ and $\mathrm{d}\sigma/\mathrm{d}\eta_2$ distributions are not shown since they look almost identical.

The need for a resolved photon contribution in NLO calculations is visualized in Fig.6.11 where the ratios of the data over the $\mathrm{DIR_{uns}}$ and the TOTAL result are plotted as functions of Q^2. These ratia are quite large as one might infer already from Fig.6.5. The ratio diminishes with increasing Q^2 which is of course expected due to the fact that a lifetime of the virtual photon is inversely proportional to its virtuality. The difference between the data and the TOTAL result is much less Q^2-dependent, it amounts to about 60% in photoproduction and 40% in the low Q^2 region. There could be two explanations for this observation:

1. SaS1D needs to be increased
2. the contribution of the longitudinally polarized virtual photon is non-negligible in the kinematic region studied here

or, most likely, a combination of both.

The first idea was checked by scaling the virtual photon PDF by a factor 1.4. The result is shown in Fig.6.12 for the first Q^2 window. We can see that a mere multiplication of SaS1D parametrization makes the NLO calculations a good approximation to the data over the whole region of η. The slope of the E_T-spectrum in the data seems to be steeper than that in the NLO calculations. A similar effect was observed in the remaining Q^2 intervals. From this one can conclude that in the kinematic range studied in this analysis the SaS1D parametrization needs a global multiplicative factor which is independent of η and Q^2 but which is likely to be dependent on E_T. As stated earlier crude estimates of the contribution of the longitudinal photons in our kinematic range were made in [11, 12] which suggest that it may not be negligible. The high statistics data accumulated this year offers the chance to extract this contribution from the data.

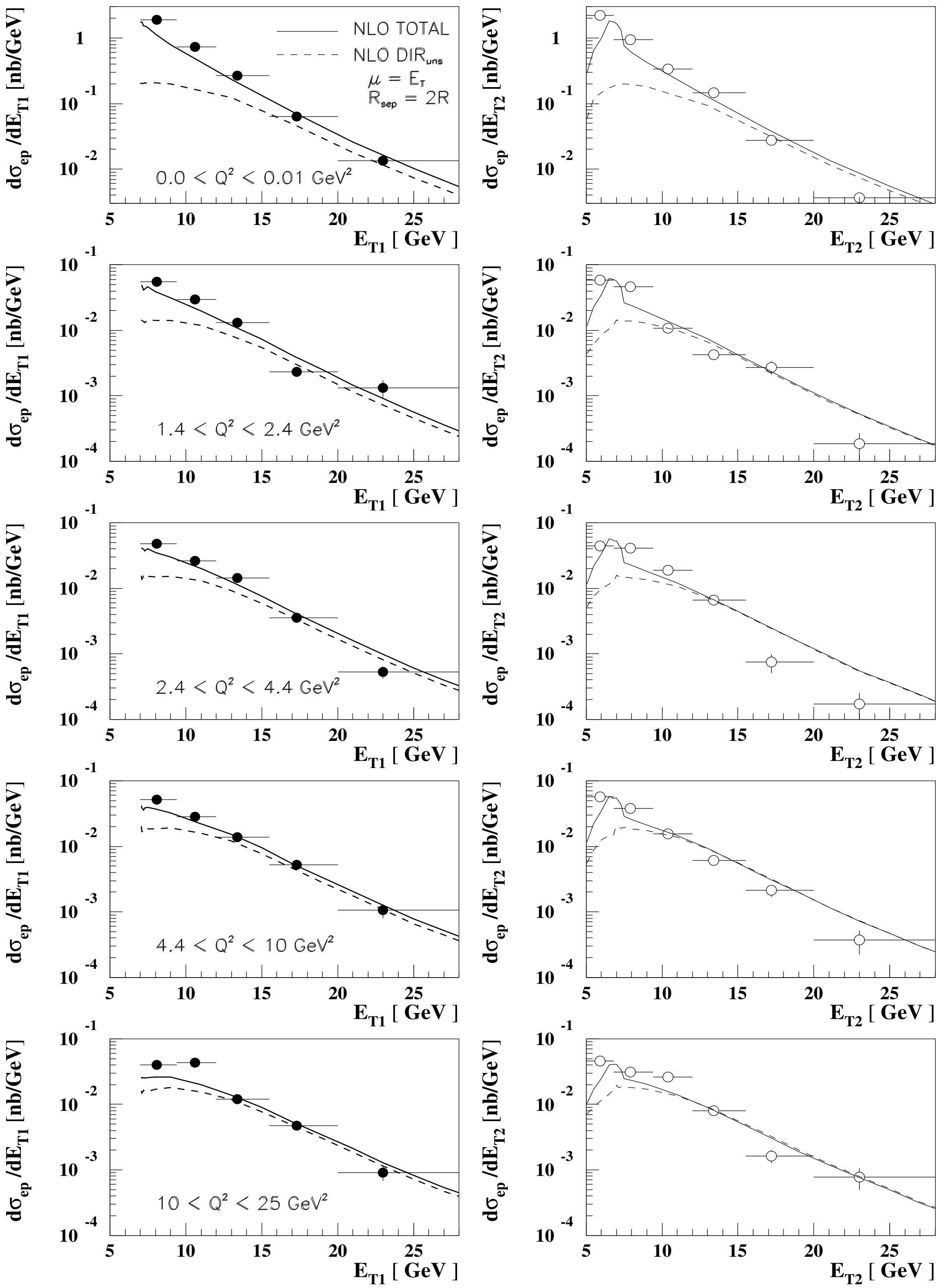


Figure 6.10: The inclusive dijet ep cross sections as a function of the jet transverse energy of the highest E_T jet (E_{T1}) and the second highest E_T jet (E_{T2}) in five intervals of the photon virtuality Q^2 (points) compared to NLO calculations (curves). The data points have only statistical errors and the NLO calculations were performed with CTEQ4M and SaS1D parametrizations for proton, resp. photon PDF, $R_{\rm sep} = 2R$ and $\mu = E_{T1}$.

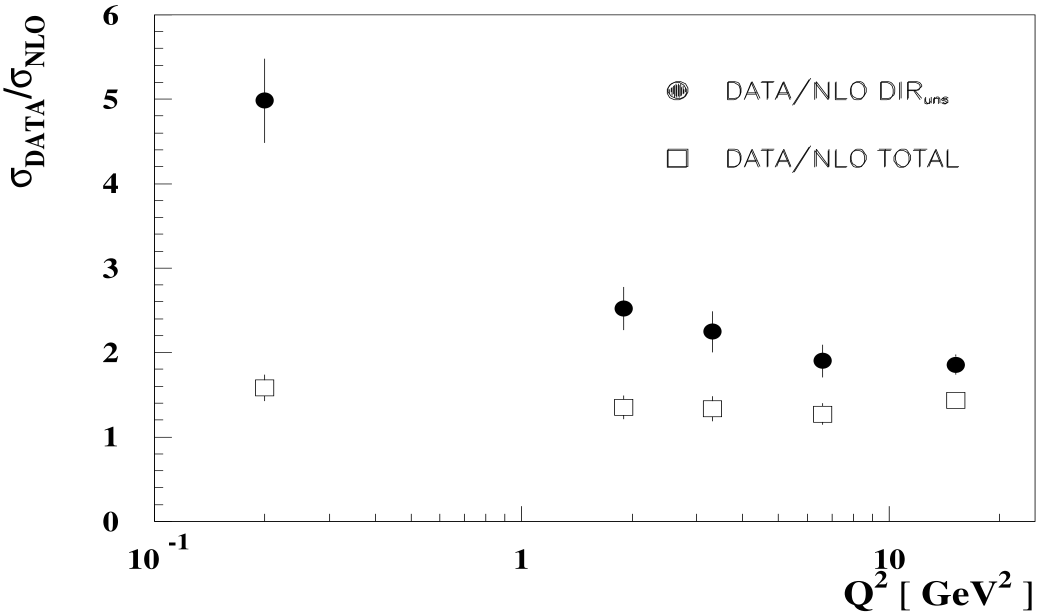


Figure 6.11: The ratio of cross sections from the data and NLO calculations integrated over the Q^2-range specified by one of five Q^2-intervals defined above. The error bars reflect the statistical and systematic errors of the data points from Fig.6.5.

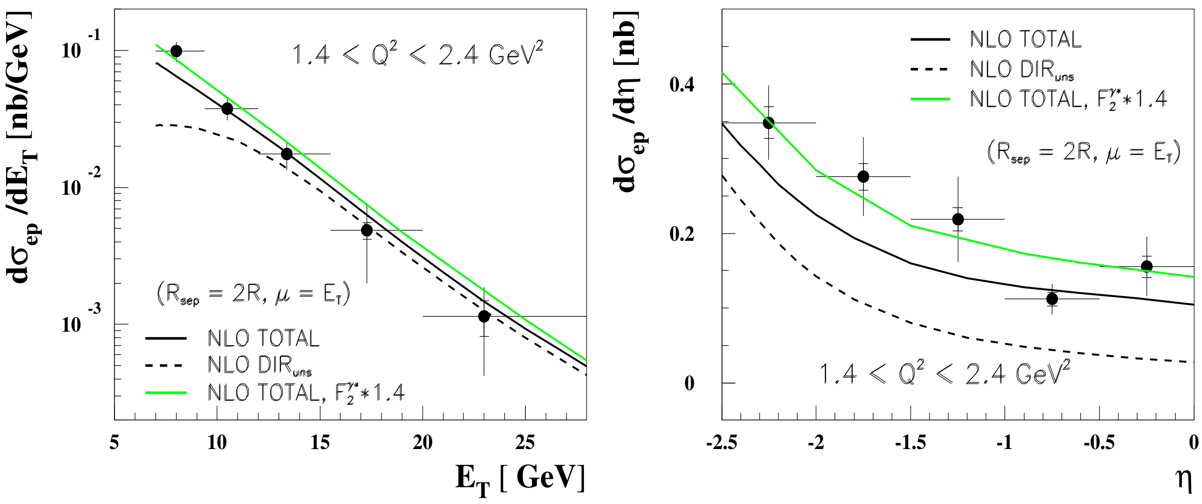
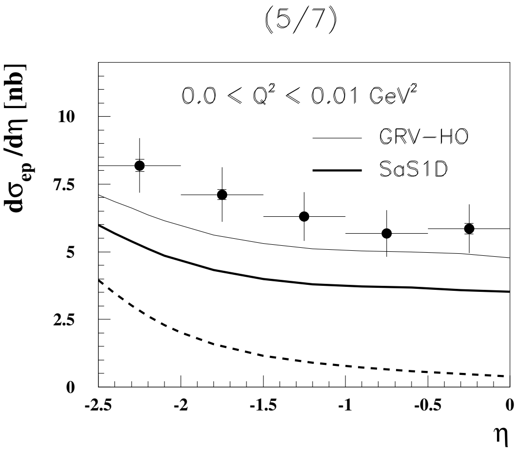


Figure 6.12: The same as in Figs. 6.5 and 6.6. The grey full line represents the TOTAL result of NLO calculations for which the original SaS1D parametrization of PDF was scaled by the factor 1.4.



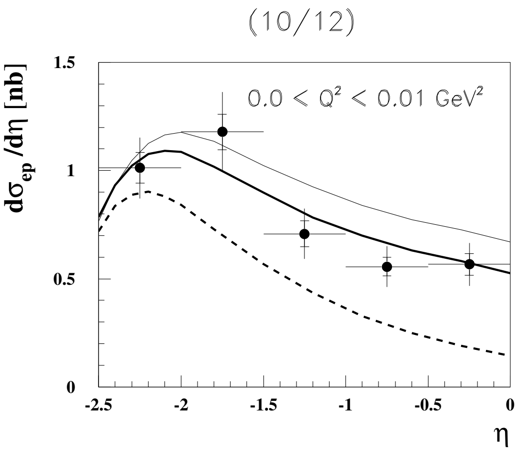


Figure 6.13: The inclusive dijet ep cross sections for photoproduction as a function of η compared with JETVIP. On the left the cut scenario ($E_{T1} > 7\,\mathrm{GeV}$, $E_{T2} > 5\,\mathrm{GeV}$), on the right the cut scenario ($E_{T1} > 12\,\mathrm{GeV}$, $E_{T2} > 10\,\mathrm{GeV}$) was used. The thick full lines represent the TOTAL results for SaS1D, the thin full lines represent the TOTAL results for GRV-HO parametrization for the photon PDF.

We have seen a large discrepancy between the data and the TOTAL result in photoproduction. This observation is in agreement with the results of studies in the previous chapter where the diparton and dijet inclusive cross sections $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ were compared to HERWIG with SaS1D. It was also noted that the use of SaS2D brought HERWIG results significantly closer to the data. In this chapter we compared the application of SaS1D and GRV-HO parametrizations for the photon PDF in the NLO calculations. We find GRV-HO more successful in the description of the data (see Fig.6.13 on the left). The remaining difference between the data and the TOTAL result with GRV-HO can be filled by application of the hadronization corrections (see Fig.6.9). The differences between existing parametrizations of the real photon PDF in the low E_T region stem from a small amount of experimental data in this region. We have therefore checked whether in the region of higher E_T where more data on the real photon are available the TOTAL results with the two parametrizations get closer. In Fig.6.13 on the right we see that this is the case and moreover both are close to the data. We account for the latter fact by a smaller influence from non-perturbative phenomena, like soft underlying activity, for the high E_T region. The relation of the data to the NLO calculations using GRV-HO in photoproduction is in agreement with that observed in [92].

Consistency with other analyses

Recently, the NLO calculations have also been compared to the data in [30], but mostly in the complementary region where $Q^2 \simeq E_T^2$. The kinematic range was specified by the following conditions:

$$5 \le Q^2 \le 100\,\mathrm{GeV}^2,\ \ 0.05 < y \lesssim 0.6,\ \ -3.5 \le \eta_{jet} \le 1,\ \ |\Delta\eta_{jet}| < 2.$$

The jets were searched for in the $\gamma^* p$ cms using the same jet algorithm as in this analysis and requiring $R = 1$ as well. The data were compared to the NLO calculations of JETVIP, specifically for two options of E_T cut: asymmetric ($E_{T1} > 7\,\mathrm{GeV}$, $E_{T2} > 5$ GeV) and sum ($E_{T1} + E_{T2} \ge 13\,\mathrm{GeV}$, $E_{T1}, E_{T2} > 5\,\mathrm{GeV}$). The scale used in the calculations was $\mu^2 = Q^2 + 50\,\mathrm{GeV}^2$.

The message of the comparison in [30] for asymmetric E_T scenario is consistent with that of the previous paragraphs: the $\mathrm{DIR_{uns}}$ result falls below the data for $Q^2 \lesssim 20\,\mathrm{GeV}^2$, while at higher Q^2 it is able to describe the data and the resolved photon result essentially vanishes. The smaller difference between the data and $\mathrm{DIR_{uns}}$ in [30] compared to observations in this thesis is caused by the different cut on y thanks to which our data sample is much more enriched by the resolved photon processes.

In the case of the sum E_T scenario the message from the comparison made in [30] is slightly different. Here, the $\mathrm{DIR_{uns}}$ contribution is able to describe the data except in the region $Q^2 \lesssim 11\,\mathrm{GeV}^2$. Therefore we also carried out the comparisons in this (sum) E_T scenario. The results shown in Fig.6.14 fully confirm the message from the comparisons made for asymmetric E_T scenario.

6.10 Components of the resolved photon part in JETVIP

To trace back the origins of the large differences observed between the TOTAL and $\mathrm{DIR_{uns}}$ results we did the following exercises with JETVIP under the conditions specified in Section 6.7 (see also [11, 12]). In the left column of Fig.6.15 the TOTAL and $\mathrm{DIR_{uns}}$ results are compared to the $\mathrm{DIR_s}$ ones. The difference between TOTAL and $\mathrm{DIR_s}$ curves, giving the RES contribution, $\mathrm{d}\sigma^{\mathrm{res}}/\mathrm{d}\eta$, is further dissected into the following contributions:

- VDM part of the photon PDF convoluted with complete NLO (i.e. up to the order α_s^3) parton level cross sections (denoted *NLO VDM*).
- Pointlike quarks[7] and pointlike gluons convoluted with
  - α_s^2 parton level cross sections (denoted $\sigma_q^{\mathrm{res}}(\alpha_s^2) \otimes q^{\mathrm{PL}}$ and $\sigma_G^{\mathrm{res}}(\alpha_s^2) \otimes G^{\mathrm{PL}}$),
  - α_s^3 parton level cross sections (denoted $\sigma_q^{\mathrm{res}}(\alpha_s^3) \otimes q^{\mathrm{PL}}$ and $\sigma_G^{\mathrm{res}}(\alpha_s^3) \otimes G^{\mathrm{PL}}$),

and plotted to the right column of Fig.6.15. The full NLO RES contribution is then obviously:

$$\text{NLO RES} = \text{NLO VDM} + \sigma_q^{\mathrm{res}}(\alpha_s^2)\otimes q^{\mathrm{PL}} + \sigma_G^{\mathrm{res}}(\alpha_s^2)\otimes G^{\mathrm{PL}} + \sigma_q^{\mathrm{res}}(\alpha_s^3)\otimes q^{\mathrm{PL}} + \sigma_G^{\mathrm{res}}(\alpha_s^3)\otimes G^{\mathrm{PL}}. \quad (6.5)$$

Fractional contributions of LO (i.e. α_s^2) and NLO (i.e. α_s^2 and α_s^3) terms to σ^{res} are plotted in Fig.6.16 as functions of η. Several important conclusions can be drawn from Figs. 6.15 and 6.16:

- The difference between the TOTAL and $\mathrm{DIR_{uns}}$ is significant in the whole range of η. It gradually rises with increasing η up to the upper edge $\eta = 0$ where the TOTAL results exceed the $\mathrm{DIR_{uns}}$ ones by a factor of more than 3 in the second Q^2 window. Furthermore, as Q^2 increases the $\mathrm{DIR_{uns}}$ contribution represents an increasing fraction of the TOTAL result reflecting the decrease in the relative importance of the nontrivial part of photon PDF (given by JETVIP with SaS1D) as its virtuality increases.
- The contribution of the photon splitting term (the difference $\mathrm{DIR_{uns}} - \mathrm{DIR_s}$) is a decreasing function of η and looks almost the same in each Q^2 window.
- The inclusion of α_s^3 resolved photon contributions is numerically very important in the whole considered range where the results proportional to α_s^2 are close to the α_s^3 ones.

[7] Defined as pointlike part of the quark distribution function (likewise with gluon), see Chapter 1.

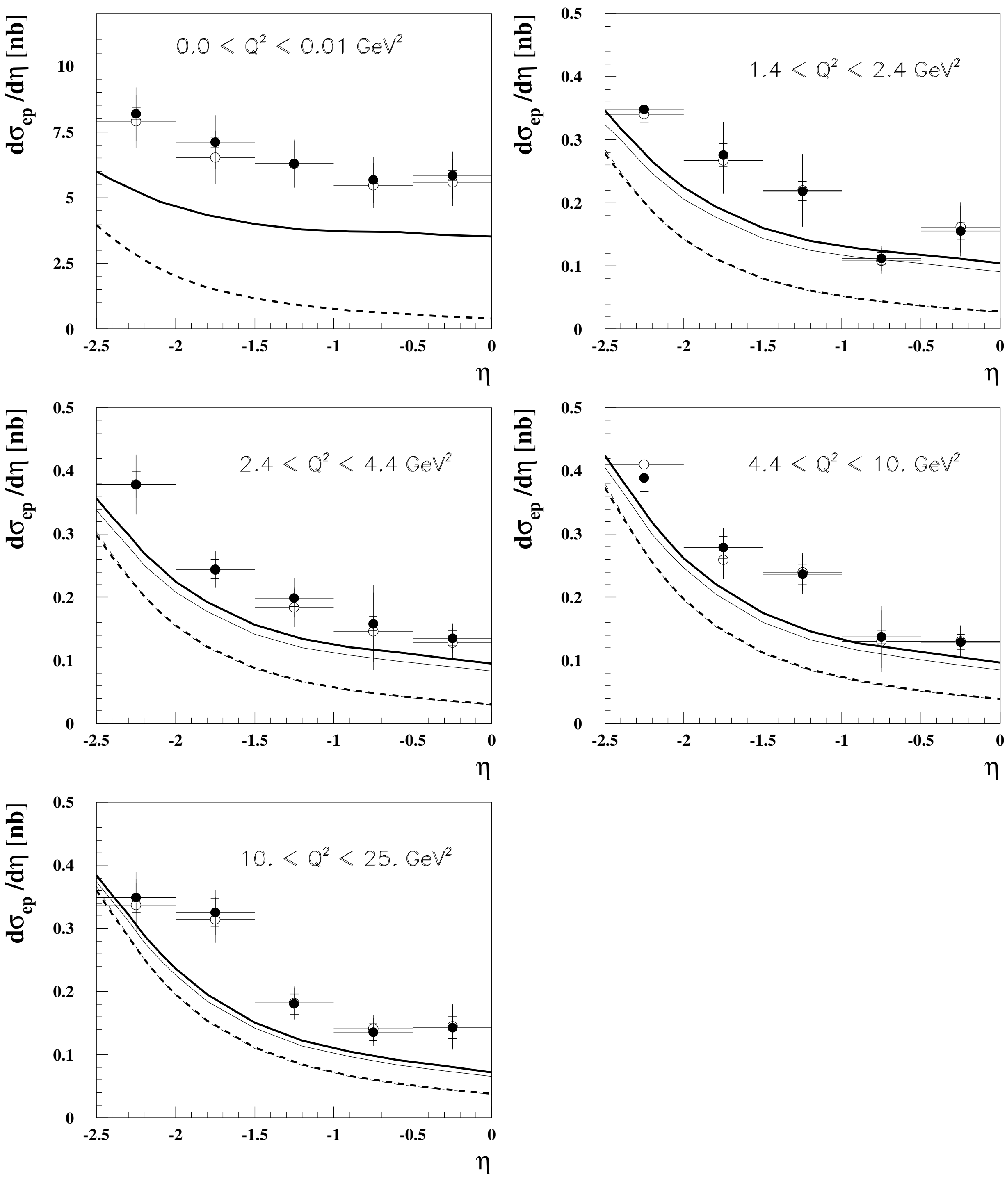


Figure 6.14: The comparison of the results of two E_T cut scenarios. The full points, thick full and thick dashed lines correspond to the asymmetric E_T scenario ($E_{T1} > 7\,\mathrm{GeV}$, $E_{T2} > 5\,\mathrm{GeV}$), the open points, thin full and thin dashed lines correspond to the sum E_T scenario ($E_{T1} + E_{T2} \geq 13\,\mathrm{GeV}$, $E_{T1}, E_{T2} > 5\,\mathrm{GeV}$). CTEQ4M parametrization was used for the proton PDF, SaS1D for the photon PDF. The scale $\mu = E_{T1}$ and $R_{\mathrm{sep}} = 2R$ was used everywhere.

- At both α_s^2 and α_s^3 orders pointlike quarks dominate $\mathrm{d}\sigma^{res}/\mathrm{d}\eta$ at large negative η, whereas as $\eta \to 0$ the fraction of $\mathrm{d}\sigma^{res}/\mathrm{d}\eta$ coming from pointlike gluons increases (up to about 40% at $\eta = 0$ in the second Q^2 interval). Moreover, as Q^2 increases the contribution of the pointlike gluons fades and it is superseded by the contribution of the pointlike quarks.

- The contribution of NLO VDM is a rising function of η. In photoproduction it represents a significant part reaching about 30%. It however vanishes rapidly with increasing Q^2 since already at $Q^2 \simeq 2\,\mathrm{GeV}^2$ it falls down to 3%. Using SaS2D parametrization for non-zero virtualities would roughly double this number.

We emphasize that pointlike gluons already carry nontrivial information in convolution with α_s^2 partonic cross sections because in the $\mathrm{DIR_{uns}}$ part such contributions would appear first at the order α_s^3. The results of pointlike quarks convoluted with α_s^3 partonic cross sections would be included in $\mathrm{DIR_{uns}}$ calculations starting at the order $\alpha\alpha_s^3$, whereas for pointlike gluons this would require evaluation of $\mathrm{DIR_{uns}}$ terms of even higher order $\alpha\alpha_s^4$!

We have seen that the VDM part of the resolved photon contribution in JETVIP is tiny in the region of photon virtualities, $1.4 < Q^2 < 25\,\mathrm{GeV}^2$. Therefore, in this region, it makes sense to characterize the nontrivial aspects of taking into account $\mathcal{O}(\alpha_s^3)$ resolved photon contributions in JETVIP with help of pointlike distribution functions only. We define "nontriviality fractions" R_3 and R_4 as

$$R_3 \equiv \frac{\sigma_q^{\mathrm{res}}(\alpha_s^3) \otimes q^{\mathrm{PL}} + \sigma_G^{\mathrm{res}}(\alpha_s^2) \otimes G^{\mathrm{PL}}}{\sigma^{\mathrm{res}}}, \quad R_4 \equiv \frac{\sigma_G^{\mathrm{res}}(\alpha_s^3) \otimes G^{\mathrm{PL}}}{\sigma^{\mathrm{res}}}, \tag{6.6}$$

which quantify the fractions of σ^{res} that are not included in $\mathrm{DIR_{uns}}$ calculations. These fractions are plotted as functions of η and Q^2 in Fig.6.17. Note that at $\eta = 0$ and interval $1.4 < Q^2 < 2.4\,\mathrm{GeV}^2$ almost 70% of σ^{res} comes from these origins. This fraction rises even further in the region $\eta > 0$. As Q^2 increases the nontriviality factor R_4 (which comes entirely from pointlike gluons) decreases, whereas R_3, which is dominated by pointlike quarks and flat in η, is almost independent of Q^2.

All these features of the JETVIP results reflect the fundamental fact that as Q^2 rises towards the factorization scale $M^2 \approx E_T^2$ the higher order effects incorporated in pointlike parts of photonic PDF vanish and consequently the $\mathrm{DIR_{uns}}$ results approach the $\mathrm{DIR_s}$+RES ones. The crucial point is that for pointlike quarks and gluons this approach is governed by the ratio Q^2/M^2 (roughly like $(1 - Q^2/M^2)$). The nontrivial effects included in PDF of the virtual photon will thus be large for arbitrarily large Q^2, provided we stay in the region where $Q^2 \ll M^2$.

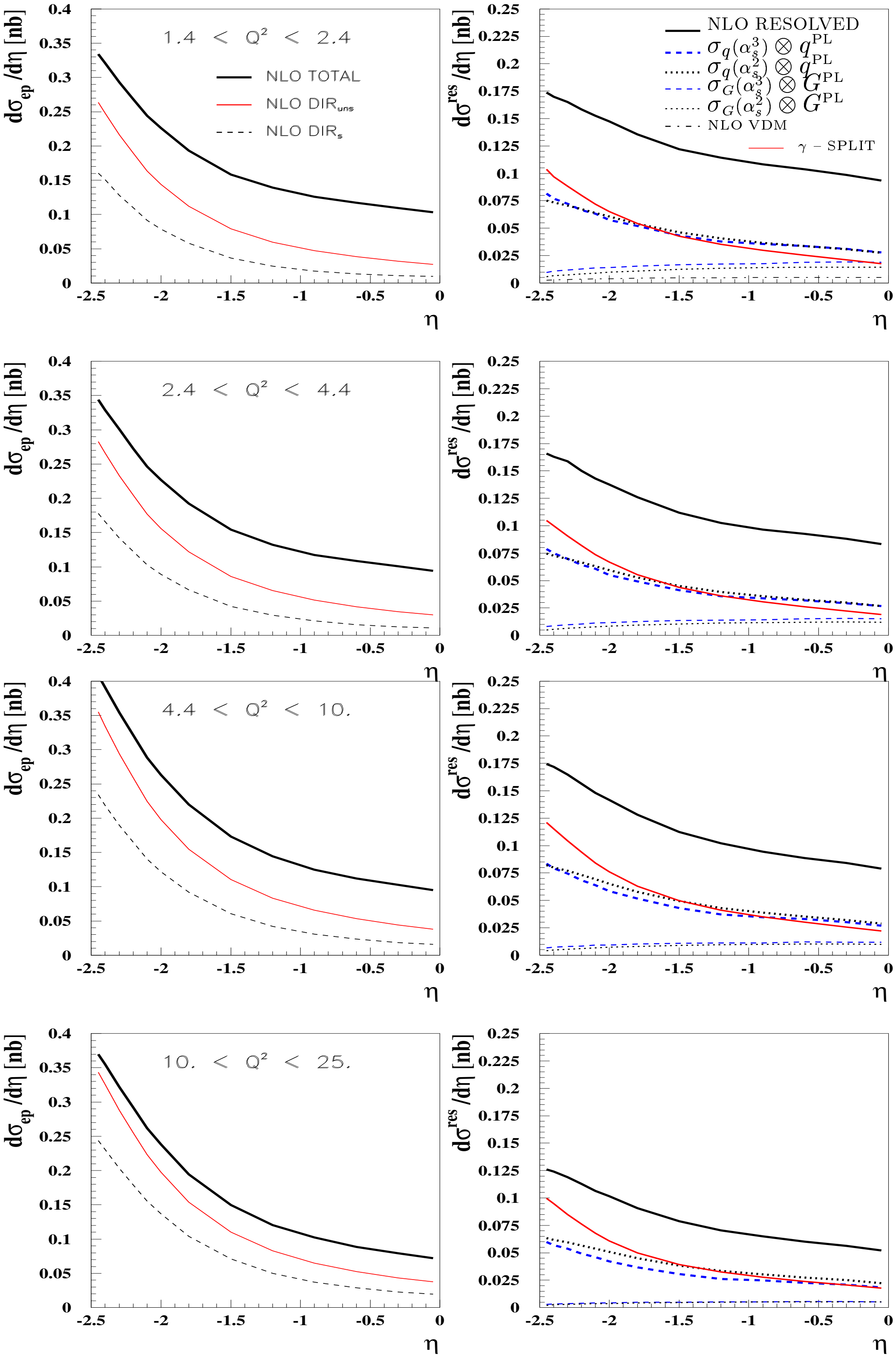


Figure 6.15: A separation of single contributions to the cross section calculated by JETVIP with CTEQ4M parametrization for the proton PDF and SaS1D for the virtual photon PDF. The scale chosen is $\mu = E_{T1}$ and $R_{sep} = 2R$. In the left column the contributions to the TOTAL result are shown: DIR$_s$, DIR$_{uns}$ and TOTAL = DIR$_s$+RES. In the right column the resolved photon cross section is further dissected to single components.

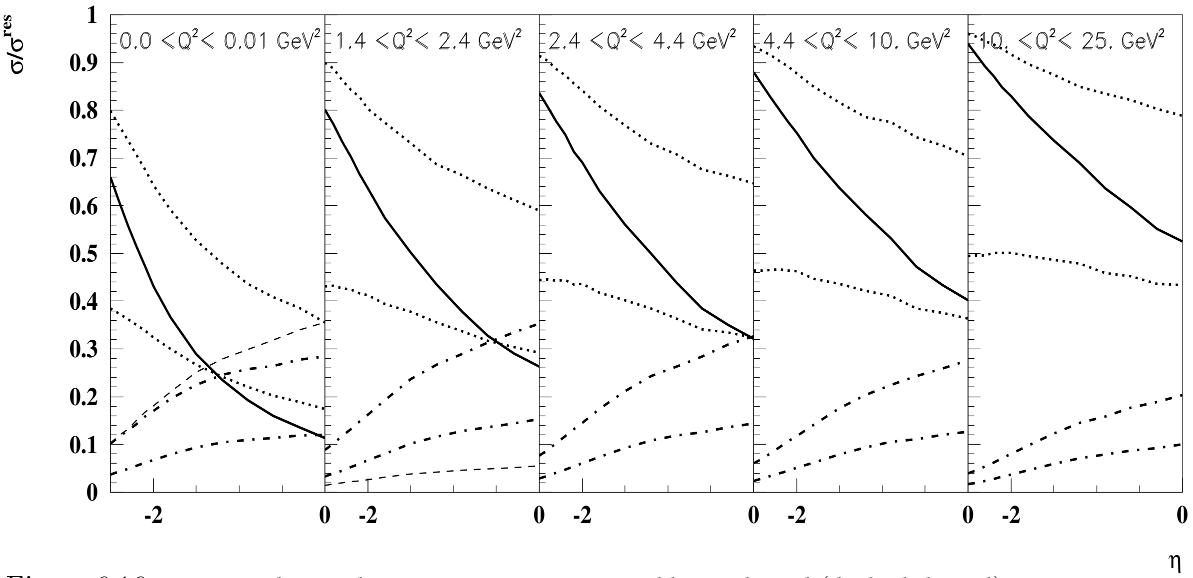


Figure 6.16: Fractional contributions to σ^{res}. Upper and lower dotted (dashed-dotted) curves correspond to pointlike quarks (gluons) convoluted with $\mathcal{O}(\alpha_s^3)$ and $\mathcal{O}(\alpha_s^2)$ parton cross sections. The dashed curves in the first two Q^2 windows give the NLO VDM contributions. The solid curves show the ratia of direct unsubtracted ($\mathrm{DIR_{uns}}$) cross sections to the NLO $\mathrm{DIR_s}$+RES ones.

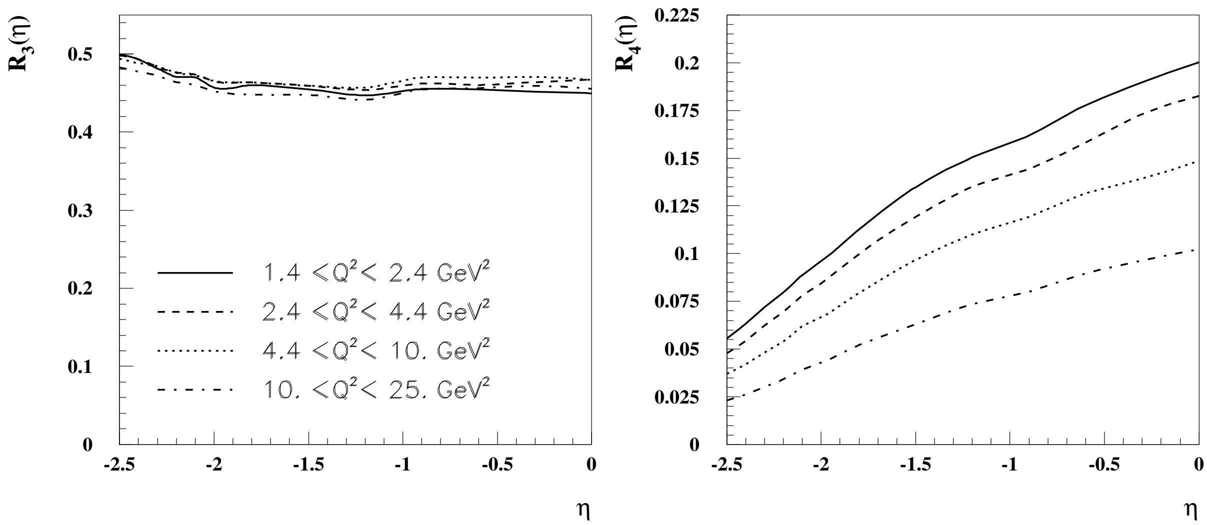


Figure 6.17: Nontriviality fractions R_3 and R_4, defined in Eq.(6.6), as functions of η and Q^2.

6.11 Conclusions

In this chapter the data were compared with the NLO parton level calculations of the JETVIP program. A comprehensive discussion of the ambiguities in NLO calculations was performed. It was found that the biggest ambiguity in calculations concerns the scale dependence of the $\mathrm{DIR}_{\mathrm{uns}}$ contributions which amounts roughly to +15% for $\eta = -2.5$ and +38% for $\eta = 0$ estimated as the ratio of calculations with $\mu = E_{T1}$ to those with $\mu = E_{T1}/2$. The biggest ambiguity for the calculation of the TOTAL result stems from the uncertainty in R_{sep} parameter. The ratio of these calculations for $R_{\mathrm{sep}} = 2R$ to those for $R_{\mathrm{sep}} = 1R$ is almost uniform over the η-range and amounts to −11% at most. We made estimates of the hadronization corrections which range between −20% and +10% in E_T-distributions (−30% for $E_T \approx 7\,\mathrm{GeV}$) and between −20% and +0% in η-distributions.

The main conclusion is that the data cannot be described by the $\mathrm{DIR}_{\mathrm{uns}}$ contribution only. This conclusion remains valid even when all the ambiguities of NLO calculations and systematic errors of the data are taken into account. The largest difference between the data and $\mathrm{DIR}_{\mathrm{uns}}$ contribution comes from regions of small E_T and large η where the resolved photon contribution is dominant part of the total jet cross section.

The addition of the NLO resolved photon contribution in JETVIP significantly shifts the prediction of the NLO calculations towards the data. The comparisons, however, clearly show that the TOTAL result of JETVIP with SaS1D parametrization for the photon PDF underestimated the data by roughly 40%. We have seen that the discrepancy between the data and the TOTAL result of JETVIP can be removed by multiplying the SaS1D parametrization by a factor 1.4. But there is also a possibility of including the contribution from the longitudinal photons which should be investigated in detail.

The study of the resolved photon contribution in JETVIP in the framework of SaS parametrizations showed that the inherent ambiguity of separation of the pointlike and the hadronic part of the photon PDF, being numerically large for the real photon, is almost irrelevant for the virtual photons above $Q^2 \simeq 3\,\mathrm{GeV}^2$. In this region the photon PDF is dominated by its pointlike part which has a clear physical origin.

Summary

We have studied the manifestations of the structure of the virtual photon in dijet production in electron proton scattering ($\sqrt{s_{ep}} \simeq 300\,\mathrm{GeV}$) using the H1 detector at HERA. We have measured the virtuality dependence of the inclusive dijet and diparton ep cross sections $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ and the effective parton distribution function of the virtual photon in the kinematic range

$$0.0 < Q^2 < 0.01\,\mathrm{GeV}^2\ (0.3 < y < 0.7), \quad 1.4 < Q^2 < 25\,\mathrm{GeV}^2\ (0.25 < y < 0.7),$$

$$\text{for } E_T > 5\,\mathrm{GeV}, \quad -3 < \eta < 0.$$

From comparisons of the cross sections of the data to HERWIG and RAPGAP, the leading order Monte Carlo generators, with various parametrizations of the parton distribution function of the photon we conclude that the photoproduction data are well described by both generators using GRV-LO and SAS2D parametrizations. For non-zero virtualities the virtuality dependence of $\mathrm{d}\sigma/\mathrm{d}x_\gamma$ in the data is different from that predicted by SaS1D and SaS2D as well as by GRV-LO suppressed by Drees-Godbole function. These conclusions were confirmed in comparisons of the effective parton distribution function of the virtual photon extracted from the data to SaS1D, SaS2D parametrizations as well as to GRV-LO suppressed by Drees-Godbole function.

We have measured the virtuality dependence of the inclusive dijet ep cross sections $\mathrm{d}\sigma/\mathrm{d}E_T$ and $\mathrm{d}\sigma/\mathrm{d}\eta$ in the kinematic range

$$0.0 < Q^2 < 0.01\,\mathrm{GeV}^2\ (0.3 < y < 0.7), \quad 1.4 < Q^2 < 25\,\mathrm{GeV}^2\ (0.25 < y < 0.7),$$

$$E_{T1} > 7\,\mathrm{GeV},\ E_{T2} > 5\,\mathrm{GeV}, \quad -2.5 < \eta < 0.$$

From comparisons of the data to the NLO calculations of JETVIP with SaS1D parametrization of the photon structure function we conclude that calculations which do not introduce the concept of virtual photon structure are not able to describe the data. The largest discrepancies between the calculations and the data are observed in the region of small E_T and large η where the contribution of resolved photon processes is the dominant part of the total jet cross section. The inclusion of the resolved photon processes moves the predictions of the NLO calculations significantly towards the data but it is still insufficient.

The measurements show that the structure function of the virtual photon decreases with increasing virtuality as expected due to the fact that the lifetime of the virtual photon decreases with increasing virtuality.

All the comparisons suggest that SaS1D in photoproduction as well as SaS1D and SaS2D parametrizations for non-zero virtualities underestimate the data. A better description of the data can be achieved by a modification of these parametrizations and/or by taking into account the contribution of the longitudinal photons.

Appendix A

Tables of experimental results

Q^2 [GeV^2] min	Q^2 [GeV^2] max	x_γ min	x_γ max	cross section $\frac{d\sigma^{ep}}{dx_\gamma}$ [nb]	error [nb] stat.	syst. model	syst. exper.	total
0.0	0.01	0.05	0.20	20.26	± 0.65	± 1.42	± 3.14	± 3.50
		0.20	0.35	23.59	± 0.67	± 1.89	± 3.27	± 3.83
		0.35	0.55	22.07	± 0.64	± 2.43	± 2.86	± 3.81
		0.55	0.75	16.16	± 0.52	± 0.32	±1.75	± 1.85
		0.75	0.87	21.45	± 0.76	± 5.36	± 2.19	± 5.84
		0.87	1.00	9.07	± 0.46	± 0.73	± 0.68	± 1.10
1.4	2.4	0.05	0.20	0.258	± 0.031	± 0.057	± 0.056	± 0.086
		0.20	0.35	0.412	± 0.036	± 0.161	± 0.037	± 0.169
		0.35	0.55	0.521	± 0.038	± 0.036	± 0.075	± 0.092
		0.55	0.75	0.733	± 0.048	± 0.051	± 0.097	± 0.120
		0.75	0.87	1.357	± 0.097	± 0.014	± 0.121	± 0.156
		0.87	1.00	0.516	± 0.056	± 0.103	± 0.033	± 0.122
2.4	4.4	0.05	0.20	0.256	± 0.028	± 0.044	± 0.057	± 0.077
		0.20	0.35	0.362	± 0.032	± 0.083	± 0.037	± 0.096
		0.35	0.55	0.383	± 0.029	± 0.191	± 0.078	± 0.209
		0.55	0.75	0.857	± 0.054	± 0.009	± 0.085	± 0.101
		0.75	0.87	1.542	± 0.098	± 0.093	± 0.160	± 0.210
		0.87	1.00	0.715	± 0.064	± 0.071	± 0.089	± 0.131
4.4	10.0	0.05	0.20	0.299	± 0.032	± 0.072	± 0.030	± 0.084
		0.20	0.35	0.339	± 0.031	± 0.058	± 0.053	± 0.084
		0.35	0.55	0.563	± 0.042	± 0.023	± 0.058	± 0.075
		0.55	0.75	0.736	± 0.047	± 0.088	± 0.099	± 0.141
		0.75	0.87	1.423	± 0.089	± 0.057	± 0.176	± 0.205
		0.87	1.00	0.765	± 0.068	± 0.046	± 0.079	± 0.114
10.0	25.0	0.05	0.20	0.163	± 0.029	± 0.024	± 0.051	± 0.064
		0.20	0.35	0.329	± 0.040	± 0.046	± 0.042	± 0.074
		0.35	0.55	0.415	± 0.037	± 0.141	± 0.061	± 0.158
		0.55	0.75	0.714	± 0.050	± 0.007	± 0.071	± 0.087
		0.75	0.87	1.739	± 0.104	± 0.017	± 0.156	± 0.188
		0.87	1.00	0.851	± 0.068	± 0.144	± 0.090	± 0.183

Table A.1: The inclusive dijet ep cross section for jets with $E_T > 5$ GeV and $-3 < \eta < 0$ measured in the range $0.25 < y < 0.7$ in bins of x_γ and Q^2. For each bin a cross section number is given together with the statistical and systematic uncertainties. The systematic errors are separated into contributions from model dependences and experimental uncertainties.

Q^2 [GeV^2] min	Q^2 [GeV^2] max	x_γ min	x_γ max	$\frac{1}{\alpha}x_\gamma f_{\rm eff}^\gamma(x_\gamma, E_{\rm T}^2, Q^2)$	error stat.	error syst. model	error syst. exper.	error total
0.0	0.01	0.05	0.15	5.12	$\pm$ 0.25	$\pm$ 1.64	$\pm$ 0.72	$\pm$ 1.81
		0.15	0.35	2.72	$\pm$ 0.09	$\pm$ 0.53	$\pm$ 0.33	$\pm$ 0.63
		0.35	0.52	2.46	$\pm$ 0.10	$\pm$ 0.35	$\pm$ 0.27	$\pm$ 0.45
		0.52	0.75	2.04	$\pm$ 0.08	$\pm$ 0.35	$\pm$ 0.16	$\pm$ 0.40
1.4	2.4	0.05	0.15	1.40	$\pm$ 0.14	$^{+0.46}_{-0.48}$	$\pm$ 0.27	$^{+0.55}_{-0.57}$
		0.15	0.35	0.84	$\pm$ 0.07	$^{+0.46}_{-0.44}$	$\pm$ 0.12	$^{+0.48}_{-0.46}$
		0.35	0.52	1.07	$\pm$ 0.08	$^{+0.25}_{-0.26}$	$\pm$ 0.13	$^{+0.29}_{-0.30}$
		0.52	0.75	1.26	$\pm$ 0.08	$^{+0.25}_{-0.12}$	$\pm$ 0.16	$^{+0.31}_{-0.22}$
2.4	4.4	0.05	0.15	1.23	$\pm$ 0.13	$^{+0.55}_{-0.53}$	$\pm$ 0.31	$^{+0.64}_{-0.63}$
		0.15	0.35	0.65	$\pm$ 0.05	$^{+0.16}_{-0.14}$	$\pm$ 0.13	$^{+0.21}_{-0.20}$
		0.35	0.52	0.64	$\pm$ 0.05	$^{+0.39}_{-0.36}$	$\pm$ 0.08	$^{+0.40}_{-0.37}$
		0.52	0.75	1.03	$\pm$ 0.06	$^{+0.28}_{-0.29}$	$\pm$ 0.15	$^{+0.33}_{-0.34}$
4.4	10.0	0.05	0.15	1.35	$\pm$ 0.13	$^{+0.23}_{-0.31}$	$\pm$ 0.24	$^{+0.36}_{-0.41}$
		0.15	0.35	0.59	$\pm$ 0.04	$^{+0.15}_{-0.11}$	$\pm$ 0.09	$^{+0.18}_{-0.15}$
		0.35	0.52	0.75	$\pm$ 0.05	$^{+0.08}_{-0.07}$	$\pm$ 0.10	$^{+0.14}_{-0.13}$
		0.52	0.75	0.82	$\pm$ 0.05	$^{+0.16}_{-0.15}$	$\pm$ 0.09	$^{+0.19}_{-0.18}$
10.0	25.0	0.05	0.15	0.27	$\pm$ 0.06	$\pm$ 0.19	$\pm$ 0.09	$\pm$ 0.22
		0.15	0.35	0.32	$\pm$ 0.04	$\pm$ 0.11	$\pm$ 0.08	$\pm$ 0.14
		0.35	0.52	0.34	$\pm$ 0.06	$\pm$ 0.22	$\pm$ 0.12	$\pm$ 0.25
		0.52	0.75	0.36	$\pm$ 0.04	$^{+0.26}_{-0.24}$	$\pm$ 0.08	$^{+0.28}_{-0.26}$

Table A.2: The effective parton distribution function of the photon $\frac{1}{\alpha}x_\gamma f_{\rm eff}^\gamma$ measured in the bins of x_γ and Q^2. For each bin a value is given together with the statistical and systematic uncertainties. The systematic errors are separated into contributions from model dependences and experimental uncertainties.

Q^2 [GeV^2] min	Q^2 [GeV^2] max	η min	η max	cross section $\frac{d\sigma^{ep}}{d\eta}$ [nb]	error [nb] stat.	syst. model	syst. exper.	total
0.0	0.01	-2.5	-2.0	8.18	± 0.23	± 0.41	± 0.98	± 1.09
		-2.0	-1.5	7.11	± 0.19	± 0.78	± 1.01	± 1.29
		-1.5	-1.0	6.31	± 0.17	± 1.26	± 0.84	± 1.52
		-1.0	-0.5	5.68	± 0.16	± 0.45	± 0.85	± 0.98
		-0.5	0.0	5.85	± 0.19	± 0.35	± 0.87	± 0.96
1.4	2.4	-2.5	-2.0	0.348	± 0.021	± 0.017	± 0.042	± 0.050
		-2.0	-1.5	0.276	± 0.018	± 0.041	± 0.027	± 0.053
		-1.5	-1.0	0.219	± 0.016	± 0.048	± 0.026	± 0.057
		-1.0	-0.5	0.112	± 0.009	± 0.016	± 0.018	± 0.025
		-0.5	0.0	0.155	± 0.014	± 0.028	± 0.025	± 0.040
2.4	4.4	-2.5	-2.0	0.378	± 0.021	± 0.007	± 0.041	± 0.047
		-2.0	-1.5	0.245	± 0.015	± 0.007	± 0.023	± 0.029
		-1.5	-1.0	0.200	± 0.014	± 0.020	± 0.020	± 0.031
		-1.0	-0.5	0.158	± 0.011	± 0.057	± 0.019	± 0.061
		-0.5	0.0	0.135	± 0.012	± 0.004	± 0.020	± 0.023
4.4	10.0	-2.5	-2.0	0.389	± 0.022	± 0.043	± 0.045	± 0.066
		-2.0	-1.5	0.279	± 0.018	± 0.008	± 0.024	± 0.031
		-1.5	-1.0	0.236	± 0.016	± 0.019	± 0.024	± 0.035
		-1.0	-0.5	0.137	± 0.010	± 0.045	± 0.016	± 0.049
		-0.5	0.0	0.129	± 0.012	± 0.012	± 0.019	± 0.025
10.0	25.0	-2.5	-2.0	0.349	± 0.023	± 0.004	± 0.035	± 0.042
		-2.0	-1.5	0.325	± 0.022	± 0.029	± 0.026	± 0.045
		-1.5	-1.0	0.180	± 0.016	± 0.009	± 0.018	± 0.026
		-1.0	-0.5	0.136	± 0.013	± 0.011	± 0.014	± 0.022
		-0.5	0.0	0.143	± 0.017	± 0.024	± 0.018	± 0.035

Table A.3: The inclusive dijet ep cross section for jets with asymmetric E_T cut ($E_{T1} > 7\,\mathrm{GeV}$, $E_{T2} > 5\,\mathrm{GeV}$) measured in the range $0.25 < y < 0.7$ in bins of η and Q^2. For each bin a cross section number is given together with the statistical and systematic uncertainties. The systematic errors are separated into contributions from model dependences and experimental uncertainties.

Q^2 [GeV^2] min	Q^2 [GeV^2] max	E_T [GeV] min	E_T [GeV] max	cross section $\frac{d\sigma^{ep}}{dE_T}$ [nb/GeV]	error [nb/GeV] stat.	syst. model	syst. exper.	total
0.0	0.01	7.0	9.4	2.76	± 0.06	± 0.19	± 0.40	± 0.45
		9.4	12.0	1.08	± 0.03	± 0.11	± 0.14	± 0.18
		12.0	15.5	0.376	± 0.016	± 0.045	± 0.055	± 0.073
		15.5	20.0	0.083	± 0.005	± 0.012	± 0.014	± 0.020
		20.0	28.0	0.016	± 0.002	± 0.004	± 0.003	± 0.005
1.4	2.4	7.0	9.4	0.099	± 0.005	± 0.016	± 0.011	± 0.020
		9.4	12.0	0.038	± 0.003	± 0.005	± 0.004	± 0.007
		12.0	15.5	0.017	± 0.002	± 0.004	± 0.002	± 0.005
		15.5	20.0	4.8×10^{-3}	$\pm0.7\times10^{-3}$	$\pm1.8\times10^{-3}$	$\pm0.8\times10^{-3}$	$\pm2.1\times10^{-3}$
		20.0	28.0	1.1×10^{-3}	$\pm0.3\times10^{-3}$	$\pm0.2\times10^{-3}$	$\pm0.1\times10^{-3}$	$\pm0.4\times10^{-3}$
2.4	4.4	7.0	9.4	0.084	± 0.004	± 0.001	± 0.011	± 0.012
		9.4	12.0	0.039	± 0.002	± 0.001	± 0.005	± 0.005
		12.0	15.5	0.020	± 0.002	± 0.001	± 0.003	± 0.004
		15.5	20.0	3.7×10^{-3}	$\pm0.6\times10^{-3}$	$\pm0.6\times10^{-3}$	$\pm0.5\times10^{-3}$	$\pm0.9\times10^{-3}$
		20.0	28.0	0.8×10^{-3}	$\pm0.2\times10^{-3}$	$\pm0.5\times10^{-3}$	$\pm0.1\times10^{-3}$	$\pm0.5\times10^{-3}$
4.4	10.0	7.0	9.4	0.092	± 0.005	± 0.005	± 0.011	± 0.013
		9.4	12.0	0.042	± 0.003	± 0.001	± 0.006	± 0.006
		12.0	15.5	0.020	± 0.002	± 0.002	± 0.003	± 0.004
		15.5	20.0	7.0×10^{-3}	$\pm0.8\times10^{-3}$	$\pm1.3\times10^{-3}$	$\pm1.1\times10^{-3}$	$\pm1.9\times10^{-3}$
		20.0	28.0	1.4×10^{-3}	$\pm0.3\times10^{-3}$	$\pm0.2\times10^{-3}$	$\pm0.3\times10^{-3}$	$\pm0.4\times10^{-3}$
10.0	25.0	7.0	9.4	0.069	± 0.004	± 0.010	± 0.008	± 0.014
		9.4	12.0	0.067	± 0.004	± 0.014	± 0.008	± 0.017
		12.0	15.5	0.020	± 0.002	± 0.002	± 0.003	± 0.004
		15.5	20.0	6.1×10^{-3}	$\pm0.7\times10^{-3}$	$\pm2.6\times10^{-3}$	$\pm0.6\times10^{-3}$	$\pm2.7\times10^{-3}$
		20.0	28.0	1.9×10^{-3}	$\pm0.4\times10^{-3}$	$\pm0.9\times10^{-3}$	$\pm0.3\times10^{-3}$	$\pm1.0\times10^{-3}$

Table A.4: The inclusive dijet ep cross section for jets with asymmetric E_T cut ($E_{T1} > 7\,\mathrm{GeV}$, $E_{T2} > 5\,\mathrm{GeV}$) measured in the range $0.25 < y < 0.7$ and $-2.5 < \eta < 0$ in bins of E_T and Q^2. For each bin a cross section number is given together with the statistical and systematic uncertainties. The systematic errors are separated into contributions from model dependences and experimental uncertainties.

Bibliography

[1] H1 Collab., S. Aid et al., Z. Phys. **C69** (1995) 27.

[2] J. J. Sakurai, Ann. Phys. **11** (1960) 1.

[3] S. Söldner–Rembold, *Proceedings of the International Symposium on Lepton-Photon Interactions*, eds. A. De Roeck, A. Wagner, Hamburg 1997, p. 97.

[4] R. E. Taylor, J. Friedman, H. W. Kendall, Rev. Mod. Phys. **63** (1991) 573, 596, 615.

[5] Max Klein, *Proceedings of the International Symposium on Lepton and Photon Interactions at High Energies*, Stanford 1999, to be published.

[6] PLUTO Collab., Ch. Berger et al., Phys. Lett. **B142** (1984) 111;
PLUTO Collab., Ch. Berger et al., Phys. Lett. **B281** (1987) 365;
PLUTO Collab., Ch. Berger et al., Phys. Lett. **B107** (1981) 168;
JADE Collab., W. Bartel et al., Z. Phys. **C24** (1984) 231;
TASSO Collab., M. Althoff et al., Z. Phys. **C31** (1986) 527;
TPC/2γ Collab., H. Aihara et al., Phys. Rev. Lett. **58** (1987) 97;
TPC/2γ Collab., H. Aihara et al., Z. Phys. **C34** (1987) 1;
CELLO Collab., H.-J. Behrend et al., contributed to the XXVth Int. Con. on HEP, Singapore 1990, unpublished;
AMY Collab., T. Sasaki et al., Phys. Lett. **B252** (1990) 491;
AMY Collab., B. J. Kim et al., Phys. Lett. **B325** (1994) 248;
TOPAZ Collab., H. Hayashii et al., Phys. Lett. **B314** (1993) 149;
PLUTO Collab., Ch. Berger et al., Phys. Lett. **B142** (1984) 119.

[7] OPAL Collab., R. Ackers et al., Z. Phys. **C61** (1994) 199;
ALEPH Collab., D. Buskulic et al., Phys. Lett. **B313** (1993) 509;
DELPHI Collab., P. Abreu et al., Phys. Lett **B342** (1995) 402;
ZEUS Collab., M. Derrick et al., Phys. Lett. **B322** (1994) 287;
H1 Collab., T. Ahmed et al., Nucl. Phys. **B445** (1995) 195;
H1 Collab., S. Aid et al., Z. Phys. **C70** (1996) 17;
ZEUS Collab., M. Derrick et al., Phys. Lett. **B348** (1995) 665;
ZEUS Collab., M. Derrick et al., Phys. Lett **B384** (1996) 401;
ZEUS Collab., J. Breitweg et al., Eur. Phys. J. **C1** (1998) 109;
H1 Collab., C. Adloff et al., Eur. Phys. J. **C1** (1998) 97.

[8] OPAL Collab., G. Abbiendi et al., CERN-EP-99-010, submitted to Eur. Phys. J. C.

[9] Alice Valkárová, *Proceedings of the International Conference on the Structure and Interactions of the Photon*, eds. S. Söldner-Rembold, Freiburg 1999, to be published.

[10] M. Wobisch, *Proceedings of the International Workshop on Deep Inelastic Scattering and QCD*, Zeuthen 1999, eds. J. Blümlein, T. Riemann, to be published, PITHA-99-17.

[11] J. Chýla and M. Taševský, *Proceedings of the International Conference on the Structure and Interactions of the Photon*, eds. S. Söldner-Rembold, Freiburg 1999, to be published.

[12] J. Chýla and M. Taševský, Contribution to the Workshop Monte Carlo Generators for HERA Physics 1999, hep-ph/9905444.

[13] E. Witten: Nucl. Phys. **120** (1977), 189.

[14] A. Vogt, in *Proceedings of the International Conference on the Structure and Interactions of the Photon*, Egmond aan Zee 1997, eds. A. Buijs, F. C. Erné, p. 3.

[15] J. Chýla, Phys. Lett. **B 320** (1994), 186.

[16] J. Chýla, hep–ph/9811455

[17] H. D. Politzer, Nucl. Phys. **B 194** (1982) 493.

[18] A. Levy, *Proceedings of the International Workshop on Deep Inelastic Scattering and QCD*, eds. A. Nigro, Rome 1996, p. 114.

[19] Yu. L. Dokshitzer, *Proceedings of the International Conference on High Energy Physics*, eds. D.Lellouch, G. Mikenberg and E. Rabovici, Jerusalem 1997, p. 47.

[20] S. J. Brodsky, *Proceedings of the International Conference on the Structure and Interactions of the Photon*, Egmond aan Zee 1997, eds. A. Buijs, F. C. Erné, p. 197.

[21] M. Glück, E. Reya, A. Vogt, Phys. Rev. **D 45** (1992) 3986.

[22] G. Schuler, T. Sjöstrand, Z. Phys. **C 68** (1995) 607.

[23] G. Schuler, T. Sjöstrand, Phys. Lett. **B 376** (1996) 193.

[24] J. Chýla, Lecture notes, http://www-hep.fzu.cz/Theory/notes/text.ps.gz.

[25] M. Drees, R. Godbole, Phys. Rev. **D 50** (1994) 3124.

[26] J. D. Bjorken, SLAC-PUB-5013, 1989 (unpublished).

[27] M. Glück, E. Reya, M. Stratman, Phys. Rev. **D 51** (1995) 3220.

[28] L. E. Sinclair, *Proceedings of the International Conference on High Energy Physics*, eds. A. Astbury, D. Axen and J. Robinson, Vancouver 1998, p. 765.

[29] H1 Collab., C. Adloff et al., Phys. Lett. **B 415** (1997) 418.

[30] H1 Collab., C. Adloff et al., DESY 98-076, hep-ex/9806029, subm. to Eur. Phys. J. C.

[31] H1 Collab., C. Adloff et al., DESY-98-205, subm to Eur. Phys. J. C, hep-ex/9812024

[32] See for example, V. M. Budnev et al., Phys. Rep. **C 15** (1974) 181.

[33] H1 Collab., I. Abt et al., Nucl. Instr. and Meth. **A 386** (1997) 310 and 348.

[34] R. Brun et al., "GEANT User's guide", CERN-DD/EE-84-1 Geneva (1987).

[35] G. Marchesini et al., Comput. Phys. Commun. **67** (1992) 1.

[36] H. Jung, Comput. Phys. Commun. **86** (1995) 147.

[37] T. Sjöstrand, CERN TH-7112/93.

[38] R. Engel, Z. Phys. **C 66** (1995) 203.

[39] J. Chýla, J. Cvach, *Proceedings of the International Conference on the Structure and Interactions of the Photon*, Egmond aan Zee 1997, eds. A. Buijs, F. C. Erné, p. 156.

[40] A. D. Martin, R. G. Roberts and W. J. Stirling, Phys. Lett. **B 387** (1996) 419.

[41] H.L. Lai, J. Huston, S. Kuhlmann, F. Olness, J. Owens, D. Soper, W.K. Tung, H. Weerts, Phys. Rev. **D 55** (1997) 1280.

[42] H1 Collab., S. Aid et al., Z. Phys. **C 70** (1996) 17.

[43] H. Bethe, and W. Heitler, Proc. Roy. Soc. **A 146** (1934) 83.

[44] N. Goguitidze and S. Levonian, Internal Note H1-IN-471 (02/1996).

[45] J. Katzy, Ph.D. Thesis (in German), Heidelberg 1997.

[46] H1 SpaCal group, Nucl. Instr. Meth. **A 386** (1997) 397.

[47] H1 SpaCal group, Nucl. Instr. Meth. **A 382** (1996) 395.

[48] H1 SpaCal group, Nucl. Instr. Meth. **A 374** (1996) 149.

[49] F. Sefkow, E. Elsen, H. Krehbiel, U. Straumann, Internal Note H1-IN-407 (11/1994)

[50] H1 Collab., C. Adloff et al., Nucl. Phys. **B 497** (1997) 3.

[51] A. Meyer, Ph.D. Thesis, Hamburg 1997.

[52] F. Lehner, Ph.D. Thesis, Hamburg 1998.

[53] H1 Collab., C. Adloff et al., Nucl. Phys. **B 545** (1999) 3.

[54] D.E.Soper, *Proceedings of the 32nd Rencontres de Moriond: QCD and High Energy Hadronic Interactions*, Les Arcs 1997, p. 177.

[55] M.H.Seymour, *Proceedings of the 11th Les Rencontres de Physique de la Vallee d'Aoste: Results and Perspectives in Particle Physics*, La Thuile 1997, p. 475 and Nucl.Phys. **B 513** (1998) 269.

[56] L.A. del Pozo, Ph.D. Thesis, University of Cambridge 1993, RALT-002.

[57] J.Huth et al., *Proceedings of the Summer Study on High Energy Physics*, Snowmass, Colorado 1990, p. 134.

[58] CDF Collab., F. Abe et al., Phys. Rev. **D 45** (1992) 1448.

[59] OPAL Collab., R. Akers et al., Z. Phys. **C 63** (1994) 197 and Z. Phys. **C 68** (1995) 179.

[60] H1 Collab., C. Adloff et al., DESY-97-210, Eur. Phys. J. **C1** (1998) 495.

[61] H1 Collab., C. Adloff et al., Eur. Phys. J. **C1** (1998) 97.

[62] S. Egli et al., Internal Note H1-IN-517 (04/1997).

[63] S. Spielmann, Ph.D. Thesis (in French), Paris 1996.

[64] A. Glazov, Ph.D. Thesis, Berlin 1998.

[65] T. Krämerkämper, Ph.D. Thesis (in German), Dortmund 1998.

[66] M. Arpagaus, Ph.D. Thesis, Zurich 1995.

[67] A. Bouniatian, Ph.D. Thesis, Hamburg 1995.

[68] K. Sedlák, Diploma Thesis, Prague 1998.

[69] ZEUS Collab., J. Breitweg et al., Eur. Phys. J. **C1** (1998) 109.

[70] D. Milstead, *Proceedings of the International Conference on High Energy Physics*, eds. A. Astbury, D. Axen and J. Robinson, Vancouver 1998, p. 760.

[71] M. Hayes, *Proceedings of the International Conference on the Structure and Interactions of the Photon*, Egmond aan Zee 1997, eds. A. Buijs, F. C. Erné, p. 134.

[72] V. Blobel, *Proceedings of CERN School of Computing*, Aiguablava 1984, p. 88.

[73] G. Zech, DESY-95-113.

[74] Höcker and Kartvelishvilli, MC-TH-95/15, LAL-95/55, hep-ph/950937.

[75] G. D'Agostini, Nucl. Instr. Meth. **A 362** (1995) 487.

[76] H. Spiesberger et al., *Proceedings of the HERA Workshop*, Hamburg 1991, p. 798

[77] J. Spiekermann, Ph.D. Thesis, Dortmund 1997.

[78] M. Taševský, *Proceedings of the International Workshop on Deep Inelastic Scattering and QCD*, Brussels 1998, eds. Gh. Coremans, R. Roosen, p. 183.

[79] ZEUS Collab., M. Derrick et al., Phys. Lett. **B 322** (1994) 287.

[80] B. V. Combridge, C. J. Maxwell, Nucl. Phys. **B 239** (1984) 429.

[81] H. Rick, Ph.D. Thesis (in German), Dortmund 1997.

[82] M. Klasen, G. Kramer, Z. Phys. **C 72** (1996) 107.

[83] B. W. Harris and J. F. Owens, Phys. Rev. **D 57** (1998) 5555.

[84] S. Frixione, G. Ridolfi, Nucl. Phys. **B 507** (1997) 295.

[85] B. Pötter, hep-ph/9804373;
G. Kramer and B. Pötter, Eur. Phys. J. **C5** (1998) 665.

[86] B. Pötter, Comp. Phys. Comm. **119** (1999) 45 for the JETVIP code.

[87] S.D. Ellis, Z. Kunszt, D.E. Soper, Phys. Rev. Lett. **69** (1992) 3615.

[88] J. Chýla, J. Cvach, M. Taševský, A. Valkárová, Internal Note H1-IN-551 (08/1998).

[89] J. Chýla, J. Cvach, M. Taševský, A. Valkárová, *Sborník 13. konference slovenských a českých fyziků* (in Czech), Zvolen 1999, to be published.

[90] J. Cvach, Nucl. Phys. B (Proc. Suppl.) 79 (1999) 501.

[91] B. Pötter, *Proceedings of the Workshop on the Photon Interactions and the Photon Structure*, Lund 1998, p. 29.

[92] ZEUS Collab., J.Breitweg et al., Eur. Phys. J. **C1** (1998) 109.

Acknowledgements

The biggest and warmest thanks go to my wife Ivana and son Štěpán. Their immense patience allowed me to devote a maximum of time to this work. They often had to understand that there were some "events" for which I could not be with them. I am indebted to them very much for that. This work is dedicated just to my family.

I would like to express special thanks to Jiří Chýla. How many long discussions we had! It was he who opened up for me new horizons in physics, new views of it. In many cases he afforded me very detailed ideas about physics problems we were talking about. I am convinced it is worth developing these ideas further. I thank him for initiating me into the secrets of QED and QCD, for great assistance with the interpretation of the achieved results, for continual support and careful reading of the manuscript.

A big word of thanks goes to Alice Valkárová, my supervisor. I am grateful to her for constant support and patient help in overcoming pit-falls of the analysis and for help in interpretation of the results. I thank her for guiding my first steps into the world of physics, i.e. into the environment of the big experiment H1. I especially appreciate her humane approach.

I would like to thank Jaroslav Cvach for much valuable advise concerning the data processing in which his sense for experimental apparatus appeared, for help in interpretation of results and for support. I am thankful also for his material support and for his creation of splendid working conditions at my home institute.

I would like to express my gratitude to these three physicists for the experience they gave me. I have the honour that I can cooperate with them.

I would like to take this opportunity to thank my colleagues from H1 experiment. The collaboration has been really great. I met many of them on the way to the final results and all of them endeavoured to help me.

Particularly I want to thank Hartmut Rick for his incredible patience. He was ready to spend hours with me explaining details of his analysis. He helped me significantly by giving me his experience. It was a pleasure and enrichment to collaborate with him as well as with Steve Maxfield. I had the opportunity to work together with him and Tania McMahon on the topic of the structure of the virtual photon. I consider this cooperation exemplary. I thank them for help and advise.

I had good luck that I had been working out my analysis at the time when the spokesman of H1 was John Dainton and the physics coordinator Martin Erdmann. The support and interest of these wise men has been stimulating.

I would also like to thank Andrey Rostovtsev and Sergey Levonian for many helpful discussions.

I want to thank Andreas Meyer for writing his thesis.

I wish to thank those people in green who enabled me not to lose contact with my analysis.

I am grateful to the DESY directorate for the generous hospitality extended to me while working at DESY.

This research was supported by the GA AV ČR under grant no. A1010821, by the GA ČR under grant no. 202/96/0214 and by the GA CU under grant number 177.