Three Jet Events in Deep Inelastic Scattering

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Nils Krumnack
aus Bonn

Hamburg
2004

Gutachter der Dissertation:	K. Wick J. Terron
Gutachter der Disputation:	K. Wick W. Smith
Datum der Disputation:	22.4.2004
Vorsitzender des Prüfungsausschusses:	G. Heinzelmann
Vorsitzender des Promotionsausschusses:	R. Wiesendanger
Dekan des Fachbereichs Physik:	G. Huber

Abstract

Multijet production in neutral current deep inelastic positron-proton scattering has been studied for boson virtualities $25 < Q^2 < 5000\,\mathrm{GeV}^2$. The data were taken at the HERA collider with center-of-mass energy $\sqrt{s} = 318\,\mathrm{GeV}$ using the ZEUS detector and correspond to an integrated luminosity of $82.2\,\mathrm{pb}^{-1}$. Jets were identified using the longitudinally invariant k_T-cluster algorithm in the Breit Frame. Measurements of differential dijet and trijet cross sections are presented as functions of jet transverse energy in the Breit Frame (E_T^{breit}), the jet pseudorapidity in the Lab Frame (η^{lab}) and the boson virtuality (Q^2). The jets were selected in the range $-1 < \eta^{lab} < 2.5$ requiring $E_T^{Breit} > 5\,\mathrm{GeV}$ for an invariant mass of the jet system greater than 25 GeV. For the cross section in Q^2 the cross section ratio trijet to dijet has been calculated.

NLO calculations generated using the NLOJET++ program are compared to the data and the data are overall well described by these predictions. From the cross section ratio trijets to dijets a value of α_s has been extracted that is compatible with the world average:

$$\alpha_s(M_Z) = 0.1203 \pm 0.0010(stat)^{+0.0029}_{-0.0050}(sys) \pm 0.0047(theo)$$

Zusammenfassung

Die Produktion von Mehrjet Ereignissen in Neutral-Strom Tief-Inelastischer-Streuung wurde für Bosonenvirtualitäten von $25 < Q^2 < 5000\,\mathrm{GeV}^2$ untersucht. Die Daten wurden am HERA Beschleuniger mit einer Schwerpunktsenergie von $\sqrt{s} = 318\,\mathrm{GeV}$ mit dem ZEUS Detektor genommen und entsprechen einer integrierten Luminosität von $82.2\,\mathrm{pb}^{-1}$. Die Jets wurden im Breit-System mit dem longitudinal invarianten k_T Kluster Algorithmus rekonstruiert. Zweijet und Dreijet Wirkungsquerschnitte werden für die transversale Energie der Jets im Breit-System (E_T^{breit}), für die Pseudorapidität der Jets im Laborsystem (η^{lab}) and für die Bosonvirtualität (Q^2) gezeigt. Die Jets wurden im Bereich $-1 < \eta^{lab} < 2.5$ mit $E_T^{Breit} > 5\,\mathrm{GeV}$ und einer invarianten Masse des Jet-Systems groesser als $25\,\mathrm{GeV}$ ausgewählt. Für den Wirkungsquerschnitt in Q^2 wurde das Verhältnis von Dreijet- zu Zweijetwirkungsquerschnitt berechnet.

NLO Rechnungen wurden mit dem Program NLOJET++ generiert und liefern eine gute Beschreibung der Daten. Von dem Wirkungsquerschnittsverhältnis Dreijet- zu Zweijet wurde ein Wert für α_s bestimmt der kompatibel ist mit dem Weltmittelwert:

$$\alpha_s(M_Z) = 0.1203 \pm 0.0010(stat)^{+0.0029}_{-0.0050}(sys) \pm 0.0047(theo)$$

Contents

List of Figures

Chapter 1

Theoretical Introduction

1.1 The Standard Model

The standard model is the generally accepted theory of particle physics. It assumes that the elementary constituents of matter are quarks and leptons which are fermions of spin 1/2. There are six leptons (electron, muon, tau and their associated neutrinos) and six quarks (down, up, strange, charm, bottom and top). For each of them there exists an anti-particle with the same properties, but opposite quantum numbers. Each quark exists in three different colors.[1]

The forces between particles are described as the exchange of gauge bosons between these particles. The four known forces are:

1. Electromagnetic Force: At the macroscopic scale the interaction between charged bodies is described by the theory of Electrodynamics. At the microscopic scale quantum electrodynamics (QED) describes it as the exchange of photons between charged particles.

2. Weak Force: The weak force is mediated by the massive bosons Z^0 and $W^{\pm}$. Due to the mass of these bosons this force is only present at microscopic scale and of small effect compared to the electromagnetic force. Its presence can most clearly be seen in processes like the radioactive β decay which cannot be mediated by other forces. Together with the electromagnetic force it is described by the electroweak theory.

3. Strong Force: The force that keeps the colored quarks together to form the colorless hadrons. It is mediated by eight gluons that carry color charge themselves. This makes the gluons interact with each other, which allows the force to strengthen with increasing distances making colored objects only visible at the microscopic scale.

4. Gravitation: The weakest of all forces. At the macroscopic scale it has been successfully described by the general theory of relativity. For the microscopic scale the only confirmed prediction is that effects are much

[1] The colors here have nothing to do with actual colors; they are given in analogy to color mixing. They denote a kind of charge that has not only two (positive, negative), but six different manifestations (red, anti-red, green, anti-green, blue, anti-blue). A colorless (uncharged) combination consists either of a color and its anti-color or of all three (anti)colors.

	generations 1st	2nd	3rd	color	Q/e	chirality
leptons	ν_e $< 3\,\text{eV}$	ν_μ $< 0.19\,\text{MeV}$	ν_τ $< 18.2\,\text{MeV}$	no	0	L
	e $0.51\,\text{MeV}$	μ $106\,\text{MeV}$	τ $1.78\,\text{GeV}$		-1	L & R
quarks	u $6\,\text{MeV}$	c $1.2\,\text{GeV}$	t $174\,\text{GeV}$	yes	$+2/3$	
	d $3\,\text{MeV}$	s $120\,\text{MeV}$	b $4.25\,\text{GeV}$		$-1/3$	

Table 1.1: *Basic fermions of the standard model. The masses are taken from experimental data, all other quantities are predictions of the standard model.*

force	boson	types	mass	Q/e	color
strong	g (gluon)	8	0	0	yes
em	γ (photon)	1	0	0	no
weak	Z^0	1	$91.19\,\text{GeV}$	0	no
	$W^\pm$	2	$80.42\,\text{GeV}$	± 1	

Table 1.2: *The bosons mediating the forces as predicted by the standard model. The non-zero masses are taken from experimental data.*

weaker and completely shadowed by the other forces. There are theories of quantum gravity that predict the existence of a graviton (as force mediator), but such a particle has not been observed up to now.

1.1.1 Electroweak Theory

Electrodynamics is the mathematically simplest of the quantum gauge theories. It is invariant under local $U(1)$ gauge transformations. Due to the Abelian character of the gauge group the gauge bosons can't interact with each other. The theory has only one gauge boson, the photon.

For weak interactions left handed particles are grouped into doublets (d'–u, s'–c, b'–t and e, μ, τ with their neutrinos), right handed particles form singlets (there are no right handed neutrinos). The weak force can't be described on its own, but only combined with the electromagnetic force. Their combined gauge group is the $SU(2)_L \times U(1)$ group (the L indicates that the weak force only couples to left handed particles). This results in four gauge bosons two of which are charged ($W^\pm$) and change the flavor within a doublet. The photon is the superposition of the remaining two gauge bosons that has the same behavior as the photon from QED. The orthogonal superposition of these gauge bosons is called the Z^0.

The weak gauge bosons created this way have no mass. To impart mass to these (and all other massive particles), an interaction with a background field (called Higgs field) is assumed. The mass is then proportional to the coupling strength to the Higgs field.

1.1.2 Quantum Chromodynamics

The theory of Quantum Chromodynamics (QCD) describes the interaction of colored particles, i.e. quarks and the gluons introduced by the theory themselves. It is based on the $SU(3)$ color gauge group. The theory contains eight gluons carrying color and anti-color charge.

One unique property of QCD is that the coupling gets stronger as the distance increases. This leads to confinement, the fact that quarks are always bound together into colorless hadrons. Two quarks are bound together because if their color charge would separate, the gluons connecting the two would have increasingly more energy until a split of this gluonic connection and the creation of a new quark-antiquark pair becomes energetically favorable.

At high energies (short distances) the coupling strength decreases, leading to asymptotic freedom, which means the quarks in the hadrons interact independently of the surrounding hadrons.

1.1.3 Limits of the Standard Model

The standard model of the electroweak and strong interactions has proven itself very successful at describing all aspects of particle physics. No other theory has managed to make predictions that contradict the standard model and be confirmed by the data. Although the standard model is very successful, it has a number of shortcomings. It is incomplete and there are some experimental results that point at possible disagreement. These shortcomings are listed below:

1. It gives no predictions of the masses of the quarks and leptons.
2. It gives no explanation for the existence of exactly three families of quarks and leptons.
3. It does not contain a quantum level description of gravity or an explanation for why gravity is so weak.
4. For generating the particle masses it requires the existence of the Higgs particle, which has not been found yet, but its existence cannot be excluded.
5. It assumes neutrinos to be massless, but recent data shows the possibility of neutrino oscillations, which would be easiest to explain with a non-vanishing mass.

Regardless of all concerns, at the present no other theory can make predictions that disagree with the standard model and are confirmed by the data.

1.2 Deep Inelastic Scattering

Scattering experiments provide a good test of the standard model. In electron proton scattering the two particles interact by the exchange of an electroweak gauge boson. If the boson is highly virtual it can be seen as an point-like object probing the inner structure of the proton and the forces that bind the proton together. If the boson carries electrical charge ($W^{\pm}$) one speaks of charged current scattering and of neutral current otherwise (γ, Z^0).

1.2.1 Kinematic Variables

If one chooses to ignore the distribution of particles coming from the photon–proton interaction (hadronic final state), one can fully describe the scattering using the four momenta of the incoming and outgoing lepton and of the incoming proton (comp. figure 1.1). Using the information from the hadronic final state doesn't yield more information, due to four momentum conservation. The three four momenta would yield twelve degrees of freedom, but three of them are constrained by the fixed particle masses. Of the remaining nine degrees of freedom six are totally uninteresting because they correspond to the rotations and translations of the whole system, leaving three degrees of freedom that are actually physically relevant.

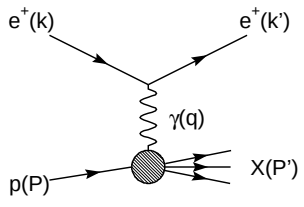

Figure 1.1: *The general kinematics of a DIS event*

These can be expressed by any set of three independent variables that seems to be useful. In most cases they are expressed using the Lorentz invariant variables:

- The virtuality of the exchanged boson:

$$Q^2 = -q^2 = -(k - k')^2 \tag{1.1}$$

- The fraction of the proton momentum taking part in the scatter, assuming an elastic scattering on a free constituent (quark parton model):

$$x = \frac{Q^2}{2P \cdot q} \tag{1.2}$$

- The inelasticity, which is the relative energy transfer from the positron to the proton in the protons rest frame:

$$y = \frac{P \cdot q}{P \cdot k} \tag{1.3}$$

- the center of mass energy of the lepton-proton system:

$$s = (k + P)^2 = 4E_e E_p = \frac{Q^2}{xy} \tag{1.4}$$

- center of mass energy of the boson-photon system:

$$W^2 = (q + p)^2 = sy - Q^2 \tag{1.5}$$

For the experimentalist who has one very special frame (the Lab Frame) it is also useful to define some variables in that frame, namely the scattering angle of the struck quark[2] and the scattering angle and energy of the electron.

[2]This assumes elastic scattering of a quark inside the proton. The events used in this analysis don't fall into that category, so the quark angle has no relevance except for the reconstruction of the Breit frame

1.2.2 Inclusive Cross Section

The general form of the inclusive DIS cross section is:

$$d\sigma \sim L_{\mu\nu} W^{\mu\nu} \tag{1.6}$$

The tensor $L_{\mu\nu}$ comes from the leptonic part of the interaction. The hadronic tensor $W^{\mu\nu}$ dependends on QCD, but from Lorentz invariance and the symmetry of $L_{\mu\nu}$ the most general form can be derived (M is the proton mass) [18]:

$$W^{\mu\nu} = -W_1 g^{\mu\nu} + W_2 \frac{p^\mu p^\nu}{M^2} + iW_3 \epsilon^{\mu\nu\rho\sigma} p_\rho q_\sigma + W_4 \frac{q^\mu q^\nu}{M^2} + W_5 \frac{p^\mu q^\nu + q^\mu p\nu}{M^2} \tag{1.7}$$

The four vector current conservation yields more constraints:

$$W_5 = -\frac{p \cdot q}{q^2} W_2 \tag{1.8}$$

$$W_4 = -\frac{p \cdot q}{q^2} W_5 + W_1 \frac{M^2}{q^2} \tag{1.9}$$

Using the more common definitions

$$F_1(Q^2, x) = M \cdot W_1(Q^2, x) \tag{1.10}$$

$$F_2(Q^2, x) = \frac{Q^2}{2Mx} \cdot W_2(Q^2, x) \tag{1.11}$$

$$F_3(Q^2, x) = \frac{Q^2}{2Mx} \cdot W_3(Q^2, x) \tag{1.12}$$

the neutral current ep cross section can be written:

$$\frac{d^2\sigma(e^\pm p)}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4} \left[y^2 x F_1(x, Q^2) + (1-y) F_2(x, Q^2) \mp y(1 - \frac{y}{2}) x F_3(x, Q^2) \right] \tag{1.13}$$

$$= \frac{4\pi\alpha^2}{xQ^4} \left[Y_+ F_2(x, Q^2) - y^2 F_L(x, Q^2) \mp Y_- x F_3(x, Q^2) \right] \tag{1.14}$$

with the definition $F_L = F_2 - 2xF_1$ and $Y_\pm = 1 \pm (1-y)^2$.

Writing it in this form allows associating the different functions with different contributions to the cross section. F_2 is the main contribution to the cross section. F_3 is the parity violating contribution coming from the Z^0 exchange and is small for $Q^2 \ll M_Z^2$. F_L is the contribution from the absorption of a longitudinally polarized photon (= 0 in the quark parton model).

1.2.3 Quark Parton Model

When DIS events were first sudied two models were developed to describe the structure of hadrons:

- The Quark Model by Gell-Mann and Zweig [8]. It uses combinatorics to deduce from the variety of observed hadrons that they are made up of two (mesons) or three (baryons) quarks.

- The Parton Model by Feynman [10]. It assumes that the proton is made up of quasi-free point-like constituents called partons. In the scattering process only one of the partons participates while the others stand by doing nothing (spectators).

The quarks and partons were then identified with each other, resulting in the name Quark Parton Model.

If the DIS interaction involves free partons, the total cross section is the incoherent sum of the cross sections for elastic scattering off the individual partons. This implies that the structure functions only depend on the likelihood to find a parton carrying the fraction x of the proton's momentum (first given by Bjorken [11]). Equivalently, the structure functions are independent of Q^2 (scaling):

$$F_i(x, Q^2) \longrightarrow F_i(x) \tag{1.15}$$

Assuming the partons are spin 1/2 particles their cross sections in lepton scattering can be calculated, summed and then compared to the general formula for the DIS cross section. This comparison gives the relation:

$$F_2(x) = 2xF_1(x) \Rightarrow F_L = 0 \tag{1.16}$$

Although the Quark Parton Model is rather simple it is quite successful in kinematic regions where the effects of gluons can be neglected. But in other kinematic regions the influence of gluons can be seen. Quarks radiate and absorb gluons. Although gluons can't directly couple to photons, they can split into quark-antiquark pairs (sea quarks). How much of this sea structure can be resolved depends on the photon virtuality. At high Q^2 more sea quarks are resolved while the valence quarks get shielded by the sea quarks around them. As the valence quarks carry a higher momentum fraction than the sea quarks F_2 rises with Q^2 at low x and falls with Q^2 at high x (comp. figure 1.2).

1.2.4 Perturbative QCD

The theory of QCD has one practical problem: it is given in the form of a local field equation. To get any predictions, an (approximate) solution to that equation has to be found. One attempt to find such solutions directly uses brute force numerical methods (lattice QCD). The more traditional method breaks the solution down into individual reactions that can be calculated analytically (perturbative QCD).

In this method the solution is calculated as a series expansion in the strong coupling constant. Each term of the expansion consists of one or more integrals. If one writes down the Feynman diagrams of the possible reactions of the involved particles, a one-to-one mapping between the integrals and the Feynman diagrams can be made. This is then used to associate each integral with the diagram and each part of the integrand with one interaction within the diagram. This leads to more complicated diagrams having a higher order in α_s, making their influence (hopefully) smaller and (presumably) safe to ignore.

One problem of this approach is that it has to deal with divergences as soon as loops are contained in the diagrams. One of these divergences comes from the gluons fluctuating into quark-antiquark pairs. Rather than explicitly calculating these diagrams, one can instead eliminate these divergencies by modifying the

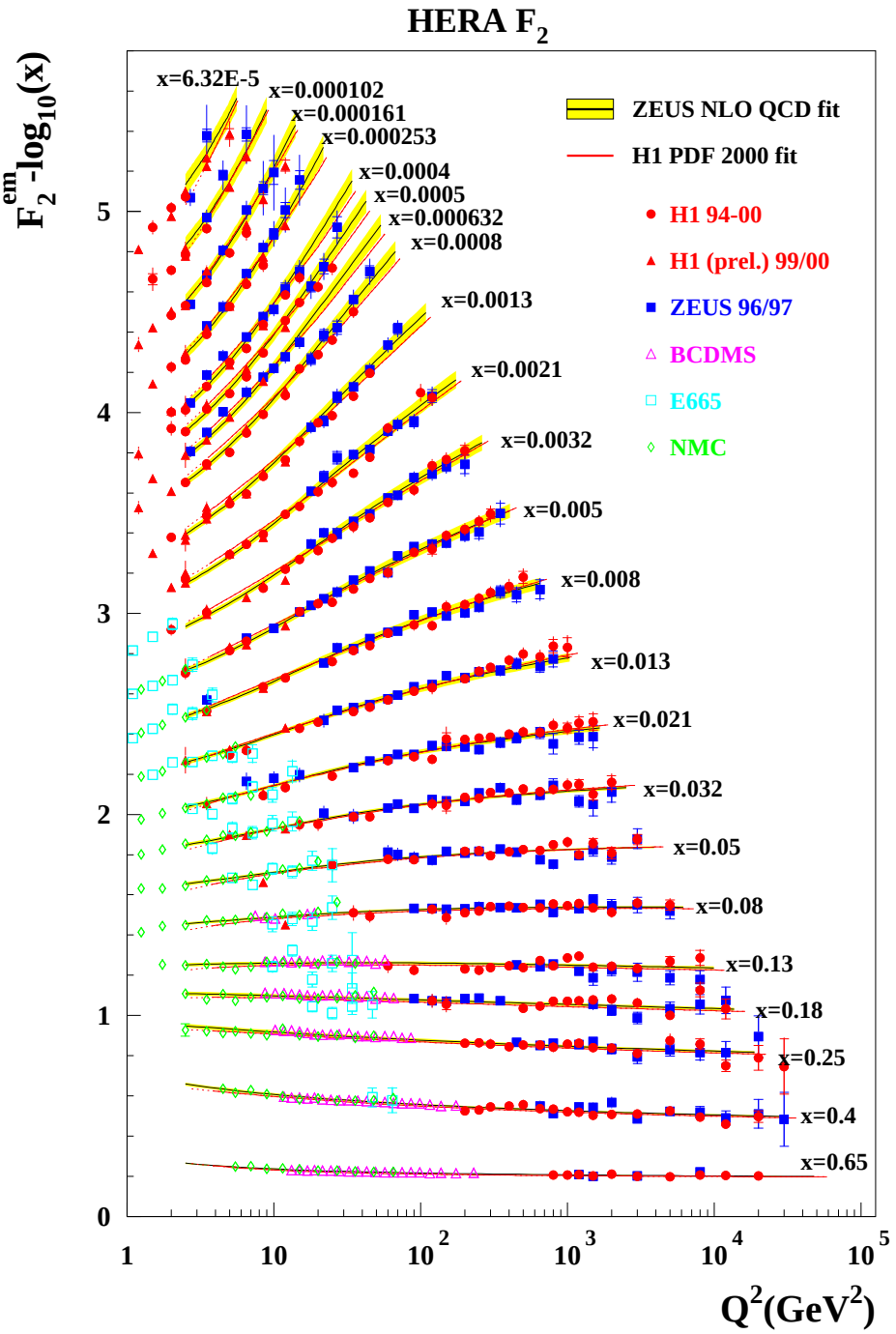

Figure 1.2: *The measured scaling behavior of F2. A scaling violation can be seen clearly at low x and not so clear at high x.*

definition of the coupling constant α_s to contain the effect of these diagrams. Unfortunately this leads to the introduction of an (arbitrary) renormalization scale μ_r. α_s decreases as the scale increases, reducing the influence of higher order effects. If the scale is large (hard), this means that the particles can be treated like free particles and the assumptions of perturbation theory are valid (asymptotic freedom).

On the other hand the absence of a hard scale means a large α_s and the end of the applicability of pQCD. One such case is the internal structure of the proton. Therefore all effects that are considered soft effects inside the proton are absorbed into a parton distribution function (PDF). To decide what is considered too soft a cut off parameter, the factorization scale μ_f, is introduced.

1.2.5 The order in α_s

When discussing theoretical predictions it is common to speak of LO and NLO predictions. LO means the prediction contains only the diagrams of the lowest order in α_s, whereas NLO predictions also contain the diagrams of the next order in α_s. It should be noted that for this analysis the lowest order in α_s means the lowest non-zero order in α_s. If the kinematic selection requires at least two partons in the final state all diagrams with one parton in the final state are cut out and the LO prediction is of $\mathcal{O}(\alpha_s^1)$. If it requires at least three partons LO is $\mathcal{O}(\alpha_s^2)$.

The leading order processes are rather simple: They are just all ways of splitting up the initial parton n times to form $n+1$ partons in the final state. These predictions are no longer considered accurate enough to compare with today's precision measurements, but are is still used in many simulations that then try to include the effects of higher order terms using phenomenological models.

The next to leading order processes are much more complicated to calculate, because the $\mathcal{O}(\alpha_s^n)$ processes with at least n partons in the final state show divergences:

- The processes with $n+1$ partons in the final state diverge if two of the partons become collinear.

- The processes with n partons in the final state need to have an internal loop to be of $\mathcal{O}(\alpha_s^n)$. This loop gives another free parameter that needs to be integrated over, giving a negative divergence.

As it turns out the divergences from both processes cancel out resulting in a finite cross section. These calculations are currently state of the art and give the most precise predictions.

Due to increasing experimental accuracy the experimental errors are in many cases smaller than the theoretical errors. To improve theoretical accuracy, NNLO contributions to the calculations are currently being computed. When these become available they will increase theoretical accuracy tremendously and will allow a very strong test of pQCD: They have to show agreement with all of today's precision measurements.

1.3 Jet Physics

1.3.1 Ideas and Principles

Perturbative QCD makes exact predictions about the likelihood of having a reaction with a certain number of partons scattered with a certain momentum. One might be inclined to computer the two parton cross section by just summing over all diagrams giving two partons in the final state, but this idea has several problems:

1. Due to virtual loops there is an infinite amount of diagrams with two partons in the final state and all except for the one with the lowest order in α_s diverge. However these divergences would cancel exactly with the divergences coming from the emittance of soft partons.

2. The partons don't exist as free physical particles, but are bound together by color forces and when they separate from each other they break apart in the process of hadronization.

3. Every detector has a finite resolution and therefore can't separate all particles and will give imprecise results if a fine separation is required.

All these problems are addressed by not looking at individual particles, but sprays of particles called jets. Jets with a sufficient (transverse) energy are supposed to originate from a parton emitted from the hard scattering process. Soft partonic radiation and particles produced during hadronization are supposed to be merged back to the jet from which they originate or form their own low energy jet (which is then ignored).

1.3.2 The Breit Frame

When studying scattering experiments one has the choice of which frame to study it in. Experimentally the easiest and most exact is the Lab frame, i.e. the frame in which the experiment is done. On the theoretical side the Lab frame has no advantages, so theorists devised a number of frames that have special properties and therefore advantages over the Lab frame.

During most of this analysis the Breit frame [19] is used. Its construction is easiest to understand assuming the photon scatters elastically off a quasi-free parton (quark parton model). In the Breit frame, the struck parton is simply reflected back (like hitting a brick wall). The z-axis is defined to point in the proton direction. This definition forces the exchanged photon to be space-like with all its momentum pointing into the negative z-direction.

The Breit frame has the advantage that everything belonging to the scattered parton is emitted in the negative z-direction whereas everything belonging to the unscattered partons flies in the positive z-direction, thereby giving maximum separation of the two.[3] In addition a one jet event would result in having no transverse momentum in the final state. By requiring transverse momentum higher order processes (the actual subject of this study) are more easily identified (comp. figure 1.3).

[3]For the multijet events studied in this analysis the momentum fraction ξ carried by the struck parton is actually higher than x, making the jets go more forward in the Breit frame. As the jet definition is independent of longitudinal boosts this doesn't make a difference in this analysis.

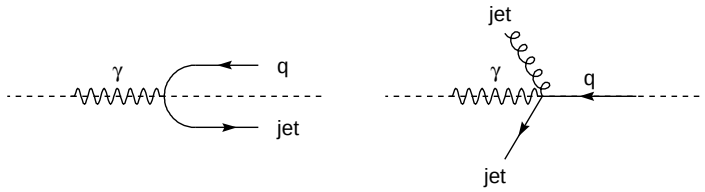

Figure 1.3: *The diagrams for one-jet (left) and dijet (right) events in the Breit frame. The z-axis points from right to left.*

1.3.3 Jet Algorithms

The purpose of a jet algorithm is to identify a spray of particles that is supposed to originate from one hard parton of the hard scattering process. In addition it is supposed to combine soft emissions of partons with the original hard partons, thereby making the positive and negative divergencies in NLO calculations cancel out.

several requirements are needed to achieve these aims:

- For typical events the hard jets should be found with comparable properties at the parton, hadron and detector level.
- The definition of the jets should be stable towards losses of small amounts of energy. There will always be some jets that change drastically when they gain or lose a little energy, but the number of affected jets should drastically decrease as the amount of energy added or removed decreases.
- Having two collinear particles should result in the same jet as with one particle with their combined energy.

A geometrical definition of jets is most convenient for satisfying these requirements.Particles which are close to each other are combined into a jet. There are two possibilities for defining what's close. In e^+e^- collisions where the lab frame is also the center of mass frame, it's custom to use the angle between the particles. In hadronic collisions the fraction of momentum taking part in the interaction is different between different events, and the center of mass frame gets boosted along the z-axis differently for each event. Therefore one chooses a coordinate system that is invariant under longitudinal boosts: the η-ϕ-space in combination with the transverse energy. ϕ is the well known azimuthal angle and η is the pseudorapidity which is derived from the polar angle in a way that longitudinal boosts don't change differences in pseudorapidity:

$$\eta = -\ln \tan \theta/2 \tag{1.17}$$

The Cone Algorithm

The basic idea of the cone algorithm is to define circles in η-ϕ space that define the jets and every deposit[4] within the circle is assigned to the jet. In detail it works like this:

[4]I'm speaking of energy deposits instead of particles here, as there is no necessity to assign the energy in the detector to individual particles.

1. Find a number of starting (seed) positions at points of high energy deposition.
2. Calculate the jet position using an energy weighted sum of the deposits within a circle with a given radius around the seed.
3. Move the jets to the new positions and determine their new content.
4. If the composition of a jet changed go back to 2.
5. If two jets overlap either join them or split up the overlap.

The k_T Algorithm

The k_T algorithm [17] merges energy deposits that are near to each other until all energy deposits are well separated. As the order of merging is important the deposits that seem to be most likely of the same jet are merged first:

1. Calculate for all deposits
$$d_i = E_{T,i}^2 \tag{1.18}$$
2. Calculate for all pairs of deposits
$$d_{i,j} = \min(E_{T,i}^2, E_{T,j}^2) \left[\Delta\eta^2 + \Delta\phi^2\right] \tag{1.19}$$
3. Find the lowest d_i or $d_{i,j}$.
4. If a d_i is the lowest set the corresponding deposit aside as an already finished jet.
5. If a $d_{i,j}$ is the lowest merge the two corresponding deposits using an energy weighted sum for the new position.
6. Repeat until all deposits are categorized as jets.

Chapter 2

Event Simulation

2.1 Ideas and Methods

An important part of experimental science is to compare the experimental results with the theoretical predictions. This allows one to judge between good and bad predictions and to tune parameters of the predictions. If one assumes the processes in the detector are in general correctly described one can use this to describe the experimental results in terms of the underlying theory and not only in terms of the experimental setup.

In particle physics the theoretical predictions are normally given separately for each possible reaction (diagram) that can happen. For each diagram one has to solve an integral over the complete phase space accessible by the final state. For many diagrams this integration is extremely difficult or impossible to do analytically. The requirement to allow more or less arbitrary cuts on the phase space makes this integration even more difficult.

A common technique for solving these problems numerically is the Monte Carlo integration. For Monte Carlo integration individual points of phase space are picked at random and for each point a weight is calculated which is basically the probability of picking that point compared to the probability of the process ending up at that point in phase space. The normalized sum of these weights (for the points in the accessible phase space) is a good approximation of the integral to be calculated.

Now it is possible to turn the interpretation around and interpret each generated point of phase space as one (simulated) event being generated. In this interpretation it is easy to see that one can take the simulated event and add processes afterwards (parton showers, hadronization and measurements in the detector). These are normally done step by step and at each step there is a probability specified for each possible reaction to happen. This way it is rather easy to specify processes like hadronization, particle decays and particle showers.

By adding a complete simulation of the detector and by playing some mathematical tricks to make all weights equal to 1, it is possible to generate events that are very similar to data. These events can be analyzed the same way as the data while at the same time giving information about the (supposed) true processes underneath.

2.2 Event Generators

There are a number of programs that can be used to generate events for ep interactions at ZEUS. These generate the events in several steps:

1. Generate an event at LO using pQCD matrix elements. At the end of this stage there are the final state partons and the proton remnant.
2. Apply a model for parton cascades that generates additional partons. The purpose of the parton cascade is to compensate for higher order effects. At the end of this stage there are a rather large number of partons and the proton remnant.
3. Apply a model for hadronization. The partons from the last stage carry color charge that connects them. In the process of hadronization the forces between the colored partons (and the remnant) create colorless hadrons.
4. These hadrons are then fed into a simulation of the detector where they decay further or interact with the detector.

It's important to note that the processes of parton cascades and hadronization are only described by using models that are tuned to match the data. These models are likely to fail in describing data they have not been tuned to.

2.3 Models for Parton Cascades

There are several models for parton cascades. The programs used in this thesis use the parton shower (LEPTO [13]) and color dipole model (ARIADNE [12]). These models try to generate additional partons one at a time by either splitting a gluon in a $q\bar{q}$ pair or by generating an additional gluon radiation. By repeatedly doing this a cascade of partons is created.

The parton shower model starts from the DGLAP description of the branching processes $q \to qg$, $g \to gg$ and $g \to q\bar{q}$. The showers are controlled by looking at the virtuality of the partons, separating between on-shell ($m^2 \approx 0$) partons and off-shell partons which are either space-like ($m^2 < 0$) or time-like ($m^2 > 0$). If a parton splits up, the sum of virtualities of the two split products is less than the virtuality of the original parton. Therefore one can start a time-like shower with a time-like parton and generate a cascade of partons with decreasing virtualities until all partons are on-shell. To allow for initial state showers there are also space-like showers in which one parton becomes increasingly space-like by splitting off time-like partons (which in turn make time-like showers).

The color dipole model is inspired by classical dipole radiation. The struck quark and the remnant carry color and anti-color charge. Therefore they can be interpreted as a dipole and any gluon created as dipole radiation. The gluon itself carries color and anti-color (and either the quark or gluon changes color). Thereby two new color dipoles are created, one between quark and gluon and one between the remnant and gluon. These dipoles are then treated as independent and the procedure is repeated. To allow for creation of quarks a competing process is introduced that splits up a gluon in a $q\bar{q}$ pair. To allow for the creation of BGF[1] events the first dipole is also allowed to emit the anti-partner of the struck quark thereby reinterpreting the event as a BGF event.

[1] In one jet events the photon can only couple to quarks but not to gluons (at LO). For

2.4 Hadronization Models

There are several models for hadronization. The programs used in this thesis (LEPTO,ARIADNE) use the Lund string fragmentation model.

The Lund string model assumes that the color connection between the partons is in form of a color flux tube. This tube or string is considered to be one dimensional and contain energy proportional to its length. The string consists of gluons and has a quark or anti-quark at each end. As the quarks move apart the string gains length and energy which leads to breaking the string into two by creating a quark-antiquark pair. Sufficiently short pieces of string can then be interpreted as mesons. To create baryons diquarks instead of quarks are created.

2.5 Detector simulation

The hadrons that emerge from the hadronization process are then fed into a simulation of the complete ZEUS detector. This simulation is based on GEANT 3.13 [20] and reasonably close to the real detector. The resulting signals are then passed through a complete simulation of the ZEUS trigger. The resulting data is then stored in the same format as the ZEUS data, which allows one to analyze it in the same way. This data is refered to as the detector level for both data and simulation. Then data and Monte Carlo events pass through the same reconstruction program.

2.6 Definition of the Parton and Hadron Level

For the simulation, different particles are created beginning from the hard interaction passing thorugh the different stages discussed in sections 2.3–2.5, resulting in a tree-like structure, with particles ordered in space and time. From this tree of particles individual stages of particles can be selected that correspond to the event at different times. Selecting the stage at the very end corresponds to selecting the detector level.

As NLO calculations (see chapter 3) are only available at the parton level, a selection of the parton level in the LO calculation is required (comp. fig. 2.1). The parton level is meant to consist of all particles that come out of the hard interaction and then enter into the soft process of hadronization. Therefore all particles that are present after the parton shower are considered partons. This includes not only colored particles which then enter the hadronization process, but also uncolored particles (mainly the proton remnant turning into an excited baryon in interactions with a sea quark).

Although this selection of the parton level seems reasonable (after the hard showers, but before the soft hadronization) it has a major problem: It is rather arbitrary. Hard and soft aren't exact definitions, but just terms to express which kind of calculation is more likely to give correct results. Correspondingly there is a tunable parameter in the Monte Carlo indicating where to cut off hard processes and start with soft processes, which at the same time changes

dijet events the photon can couple to gluons if these split into quark-antiquark pairs. These events are called Boson-Gluon Fusion.

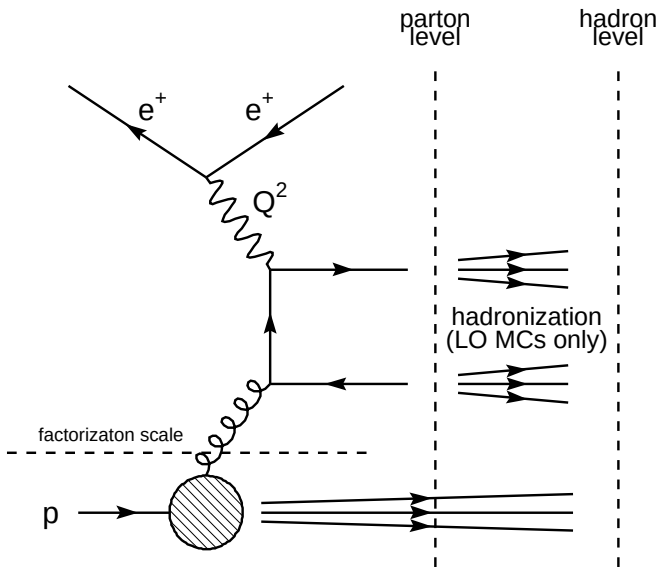

Figure 2.1: *Illustration of the meaning of parton and hadron level. The hadronization process and the hadron level are only implemented in the LO Monte Carlo. The factorization scale is the scale that separates the hard interaction from the soft processes in the proton.*

the parton level. By comparing jets at parton level the situation is less severe, as the jet algorithm is designed to recombine particles that are split in soft processes.

For comparison with other experiments and theoretical predictions the hadron level is much better suited. It describes the free and colorless particles coming out of the hadronization process before they interact with the detector.[2] As most of the particles seen in the detector are unstable and would decay if they did not hit the detector, it is important to specify where in the decay line the hadrons for the hadron level selection (stable hadrons) are localized. For this analysis the stable hadrons from the standard version of the AMADEUS generator (i.e.: the particles that enter the detector simulation) are used as stable hadrons.

[2]Not applying the hadronization corrections to the data cross sections also allows for later improvements. If a new model gives more precise predictions on the hadronisation corrections these can be immediatly applied and tested. Whereas the QED corrections described in section 3.8 are not expected to improve and can therefore be safely applied to the data.

Chapter 3

NLO Calculations

3.1 Introduction

Although the NLO programs described in this chapter may seem very similar to the LO programs described in chapter 2, they are rather different:

- The LO matrix elements are by far not exact as the NLO matrix elements. The precision of NLO calculations are considered to be exact enough to compare directly with the data.
- The LO calculations try to offset this disadvantage by emulating the effect of higher order contributions using parton shower models. The parton showers not only emulate the effects of higher orders, but they also provide soft partonic radiation. For NLO calculations it is much more difficult to incorporate parton showers as one has to avoid double counting. Therefore there is presently no program available doing this.
- The LO calculations also include models for the hadronization process, giving hadrons in the final state. These hadrons are the prerequisite for interfacing the calculations with the detector simulation, making the LO calculations available at the parton, hadron and detector level. The NLO calculations are only available at the parton level.
- As the LO calculations employ phenomenological models, they have a large number of interdependent parameters that have been tuned to the data. On the other hand the NLO calculations are only based on QCD. The only tunable parameters are the ones which (at present) can not be calculated (α_s and PDF) or which are free parameters necessary to perform the calculation (μ_r and μ_f).
- Due to the low number of parameters in NLO calculations it is easy to estimate the uncertainty introduced by the choice of these parameters. The large number of interdependent parameters makes such an estimation unfeasable for the LO calculations. This also gives NLO calculations a distinct advantage in the extraction of pQCD parameters.
- The NLO calculations start at $\mathcal{O}(\alpha_s^n)$ where n can be chosen to match the cross section calculated (thereby ignoring the orders that are zero).

The LO calculations are always at $\mathcal{O}(\alpha_s^0)$ (sometimes incorporating $\mathcal{O}(\alpha_s^1)$ effects). This makes them far less suited for multijet studies where the lower orders are zero and all higher orders have to be estimated using parton shower models.

- The NLO calculations are numerically less stable, as they have to achieve a cancellation of divergencies of positive and negative weight diagrams.

- The LO calculations act as event generators, i.e. they randomly discard events to achieve equal weights for all events. For NLO there is no upper limit on the weight of the events, making such a technique impossible.

This leads to different use of the calculations in this analysis. The LO calculations are only used for correcting from the detector or parton level to the hadron level. The NLO calculations can't be used for that, as they are only available at the parton level. The NLO calculations give the theory predictions the data is compared to. In addition they are used for the extraction of α_s.

3.2 Parton Density Functions (PDF) and Factorization Scale

One problem in calculating the cross section in ep scattering is that the proton isn't a point-like particle and doesn't take part in the interaction as a whole. Instead one of its constituents will take part in the scatter and this one constituent can interact with the other constituents. Luckily the calculation can be factorised into a parton density function (PDF) and a matrix element (ME). The parton density function contains the soft effects and is independent of the hard process. The matrix element depends only on the hard process and is calculable in pQCD. This factorisation requires the introduction of an arbitrary parameter called factorization scale μ_f, which is a parameter to the calculation of the PDF and the ME.

The PDF takes two arguments: the momentum fraction ξ carried by the struck parton and the factorization scale μ_f which characterizes how hard the processes are that are absorbed into the PDF. It should be noted that ξ can only take one value due to momentum conservation, while μ_f is an arbitrary parameter and can take in principle any value. The value assigned to μ_f is one of the free parameters of the calculation, and most programs allow to set it on a per event basis. For one-jet events it turns out that $\xi = x$ while at the same time $\mu_f = Q$ seems to be the most reasonable choice. As this analysis tests multijet events it was chosen differently as $\mu_f^2 = (Q^2 + E_T^2)/4$. To test the uncertainty coming from the choice of the factorization scale, the calculation is repeated with the factorization scale varied up or down by a factor of 2.

To calculate the PDF a parameterization of the PDF is chosen which is then fit to the experimental data [21]. This fit leaves a certain uncertainty on the PDF and therefore on the theoretical prediction. The size of this uncertainty can be estimated by not only using the mean value of the fit, but repeating the calculation with variations of it. In principle the PDF is a QCD prediction and should be calculable within the framework of QCD (not pQCD), but such a prediction is not available.

3.3 α_s and Renormalization Scale

One important parameter in the calculation of the cross section is the strong coupling constant α_s. Constant is quite a misnomer as it is dependent on the renormalization scale μ_r. The renormalization scale can in principle be set to arbitrary values (that have the dimension of energy). It is normally set to the same value as μ_f ($\mu_r^2 = \mu_f^2 = (Q^2 + E_T^2)/4$) and varied together with μ_f in the calculations to estimate the uncertainty introduced by the choice of μ_r.

The dependence of the result on the renormalization scale decreases as number of orders included increases. If all orders were to be included the renormalization scale would have no effect on the result. This leads to using the dependence on the renormalization scale as an estimate for the effect of omitting the higher order terms. By convention the renormalization scale is varied up and down by a factor of 2 and the resulting variation in the cross section is taken as the theoretical error. This is however just a convention bare of any real justification.

The renormalization process (see section 1.2.4) gives only the scaling behavior of $\alpha_s(\mu_r)$ but not its absolute scale. To fix that its value has to be determined for one value of the renormalization scale, traditionally chosen to be $\alpha_s(M_Z)$. Although the value of α_s is not a fundamental parameter of QCD, it is a free parameter of pQCD and determined from experiments (like the one described in this thesis). An alternative would be to use lattice QCD which allows to calculate α_s from first principles. However the precision achieved at the moment is not competitive with the experiments.

3.4 The Program NLOJET

The program NLOJET [15] is a rather new program, that offers the calculation of DIS di- and trijet cross sections at leading and next to leading order and of four-jet cross sections at leading order. In doing so it is the first program to offer an NLO calculation for trijet events in DIS, which is one essential requirement for this measurement. What this program does is the numeric phase space integration of all diagrams relevant to the cross section at NLO.

The divergences in real and virtual corrections at NLO present a challenge to this calculation. Although they are expected to cancel each other and give a finite result they make a direct numeric integration impossible. The program NLOJET solves that problem using the subtraction method. For this method a term is added to the virtual corrections which makes the divergences disappear. The same term is subtracted from the real corrections making also those divergences disappear. As the same term is added and subtracted it makes no difference to the final result.

3.5 The Program DISENT

The program DISENT [14] is already well established and offers the calculation of one-jet and dijet cross sections at LO and NLO. Like NLOJET it employs the subtraction method for calculating the cross section. It has been used to test the correctness of the dijet calculations in NLOJET. It showed good agreement

if α_{em} was set to the fixed value 1/137 (in contrast to the NLOJET default of having it run) (comp. fig. 3.1).

3.6 Asymmetric Jet/Invariant Mass Cut

In contrast to calculating cross sections at LO the NLO calculation depends largely on the cancellation of positive and negative weight events (each class of diagrams separately diverges, but in the sum the divergences cancel).

When defining the phase space for the measurement it is necessary to make sure not to cut just one class of diagrams and leave the other class of diagrams still in. The typical problem is that the loop diagrams result in the same kinematics as the LO diagrams while the emission of soft particles will reduce the energy of one of the jets. Putting an energy cut on the jets will then cut out only the radiation diagrams, but not the loop diagrams, reducing the calculated cross section.

Although this can not be avoided (partly it's even desired), it is normally negligeble and only results in a big effect if a lot of events are affected. That's what happens when placing the same transverse energy cut on the first and second jet for dijet events in the Breit frame (where both jets have equal E_T at LO): Most jets have very low energy and if one of them emits a soft parton that ends up in a separate jet, the emitting jet will drop below the threshold and the complete event will be discarded.

This problem does not arise for trijet events which are the main point of study in this analysis, but since the trijet cross section is compared to the dijet cross section it must still be addressed. There are several possible solutions available (see also [6]):

- Just look at the first jet, if it has transverse momentum in the Breit frame it must be a at least a dijet event (see section 1.3.2). As it only requires finding one jet this method will provide a higher efficiency and purity sample than selecting on both jets. However any kind of observable that requires the knowledge of both jets can not be studied. Another drawback is that this method can't be naturally extended to trijet events.

- Make an asymmetric cut on the E_T, i.e. make the E_T requirement higher for the first than for the second jet. This allows for enough kinematic space to allow soft radiation.

- Put a cut on the invariant mass of the dijet system. This cut depopulates the low energetic jet configurations thereby removing the problematic configurations. As the goal is to put comparable cuts on the trijet and the dijet sample this cut is prefered, as it doesn't deplete the statistic of the trijet sample.

3.7 Hadronization Corrections

The output of the NLO calculation is a cross section at parton level. To give a result that is comparable with the data it is necessary to correct the NLO calculation to the hadron level.

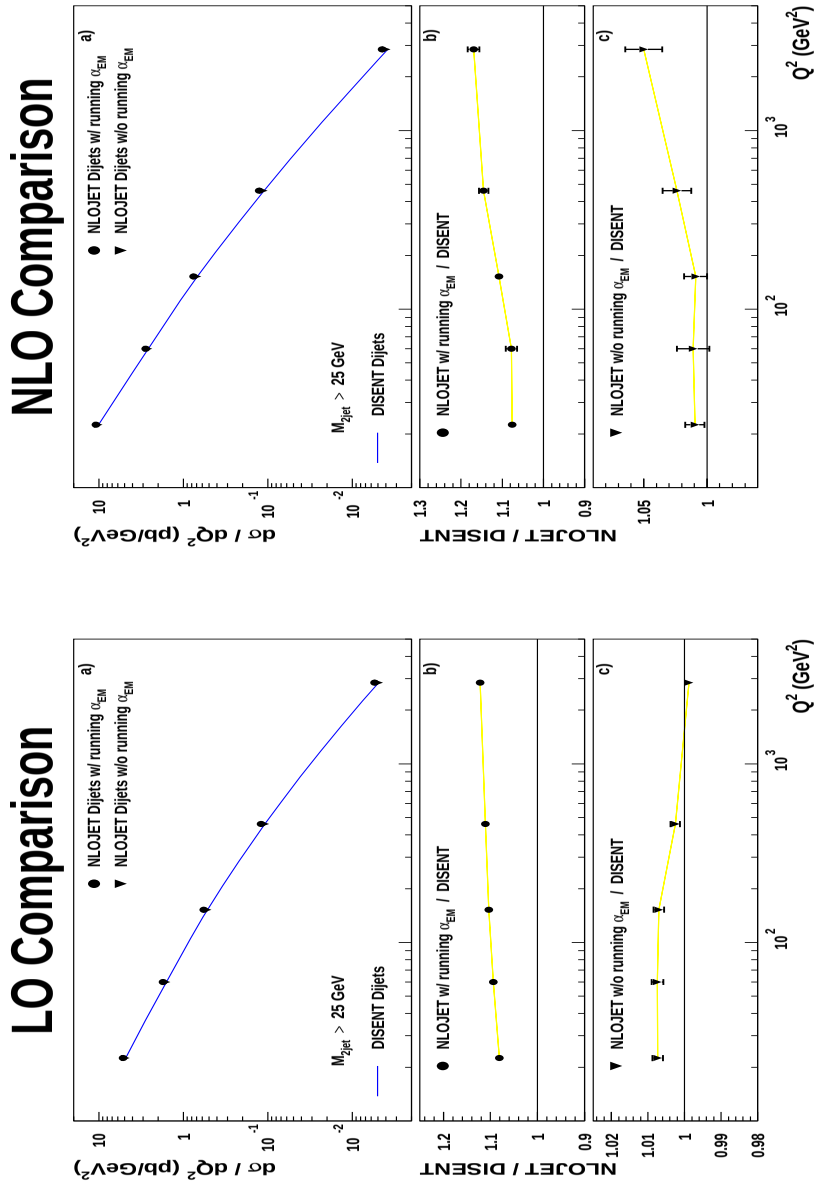

Figure 3.1: *The predicted dijet cross section at LO and NLO as a function of Q^2 for both DISENT and NLOJET. The LO prediction is a good test to see that the general parameters of the program are set up right, whereas the NLO comparison tests if the NLO contributions are modeled the same way for both programs. For NLOJET the prediction has been generated once with running and once with constant α_{em}. It can be seen that there is good agreement for constant, but not for running α_{em}.*

To do this the hadronization from the LO Monte Carlo Models is used. The LO Monte Carlo is used to predict the hadron and the parton level cross section. From this a correction factor is calculated:

$$C^{had} = \frac{\sigma^{LO}_{had}}{\sigma^{LO}_{par}} \tag{3.1}$$

This factor is then used to correct the NLO cross section to hadron level:

$$\sigma^{NLO}_{had} = C^{had} \cdot \sigma^{NLO}_{par} \tag{3.2}$$

For this method to work reliably, two conditions have to be fulfilled. Obviously the soft partonic radiation and the hadronization process have to be modelled reasonably close to reality. In addition the parton level of the LO Monte Carlo has to describe the NLO cross section reasonably close to model migrations properly.

The alternative of correcting the data to the parton level has the disadvantage that it is not so easy to define what exactly the parton level is. In addition it will add the uncertainty coming from the hadronization process to the experimental result making future improvements on the hadronization process impossible.

3.8 QED Corrections

The goal of this analysis is a comparison of measurements and NLO calculations. The measurements are at detector level and the NLO is at parton level and both of them have to be corrected to the hadron level. The NLO contains only a simplified version of the electroweak theory that is easier to implement. To allow for easy comparisons with many different programs the data is corrected to include only the QED effects the NLO implements.

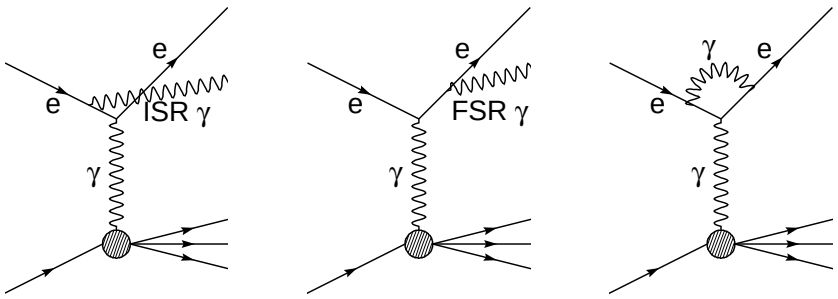

Figure 3.2: *NLO QED diagrams: Initial State Radiation (ISR), Final State Radiation (FSR) and Virtual Loop. The contribution from these diagrams are taken out from the measured cross section to compare with the theory predictions.*

The NLO calculation[1] only uses processes that are LO in QED effects; higher order contributions like ISR/FSR are not included (comp. Fig. 3.2). The NLO calculation only includes contributions from processes with γ exchange, but not

[1] as implemented in most of the current programs, e.g. NLOJET, DISENT, DISASTER, MEPJET, ...

with Z_0 exchange. The NLO calculation keeps α_{em} fixed at the value 1/137 instead of letting it increase with Q^2.

To correct the data cross section LEPTO is run twice: once with all QED effects activated and once with the QED effects simplified as for the NLO calculations. The calculated hadron level cross sections are then used to get a correction factor that can be multiplied with the data (see section 7.1):

$$c_{QED} = \frac{\sigma_{noQED}}{\sigma_{QED}} \tag{3.3}$$

Chapter 4

Experimental Setup

4.1 DESY

In 1959 the Deutsche Elektronen Synchrotron (DESY) laboratory was founded in Hamburg, Germany as a national research center with the goal to develop, build and run accelerators for German Universities. By now it has become an international research center for high energy physics and synchrotron radiation.

The two most prominent accelerators at DESY are the Hadron Electron Ring Anlage (HERA) and the Doppel Ring Speicher (DORIS). HERA started operation in 1991 and is used exclusively for high energy physics experiments. DORIS started operation in 1974 and was turned into a pure synchrotron radiation source in 1992. The ring PETRA which has become famous for the discovery of the gluon in 1979 is now used as a synchrotron source as well. The original DESY ring which started operation in 1965 can be used for testbeam experiments with new detectors (comp. chapter C).

The accelerators are connected together to allow the use of the smaller accelerators as pre-accelerators for the bigger ones (comp. Fig. 4.1). The LINAC injects particles into the DESY ring. DESY accelerates them and fills them into either DORIS or PETRA. PETRA further accelerates the particles before filling them into HERA.

4.2 HERA

The Hadron Elektron Ring Anlage (HERA) is the world's first lepton proton collider. It was built to reach much higher center of mass energies than can be achieved with fixed target experiments. For a collider the center of mass energy squared is $s = 4E_eE_p$ whereas for the fixed target experiment it is $s = E_em_p$.

At HERA there are four experiments: the two colliding beam experiments ZEUS and H1 and the fixed target experiments HERMES and HERA-B (see figure 4.1). Both H1 (north hall) and ZEUS (south hall) are general purpose experiments which are designed to allow a wide range of physics to be done. They are meant to be complementary by having independent detector designs which were optimized differently, and also meant to validate each other's result.

In 1995 the HERMES experiment was added in the east hall. It is a fixed target experiment that's used to study the spin structure of the nucleon (proton/

neutron). For this the partly polarized electron beam hits a polarized gas target (the proton beam is unused). HERMES studies the spin contribution of the constituents of the nucleon (quarks and gluons) to its total spin.

In 1999 the HERA-B experiment was added in the west hall. It is a fixed target experiment which was built for studying the CP violation in the $b\bar{b}$ system. For this the proton beam halo is brought into collision with wire target (the electron beam is unused). The HERA-B experiment presented a great technical challenge as it is searching for very rare events in a very high rate environment. In solving them HERA-B helped to improve the detectors to be built for the LHC.

The electron beam at HERA runs at an energy of 27.52 GeV. Originally the proton beam had an energy of 820 GeV, but that was increased to 920 GeV from 1998 on. This results in a center-of-mass energy of ~ 318 GeV (originally ~ 300 GeV). The luminosity taken can be seen in Figure 4.2.

The beam itself is organized into bunches of electrons and protons with a bunch separation of 96 ns. In total there can be 210 bunches of each kind. They are set up in a way that each bunch of electrons has a specific bunch of protons it will collide with. But not all bunches are actually filled. Some bunches have an empty partner to allow background studies. There is also an additional series of 10 empty bunches to allow to change the magnet settings for dumping the beam.

4.3 ZEUS

ZEUS has been planned, built and is maintained by a large international collaboration of about 300 physicist from over 50 institutes in 12 different nations.

The ZEUS detector is situated ~ 30 m underground in the South Hall of the HERA ring outside of the DESY site. It has a size of $12 \times 11 \times 20\,\mathrm{m}^3$ and weighs 3600 tons. It is a general purpose detector designed to study a wide range of physics.

I will provide a general overview of the detector in this section and of individual components important to this thesis in the following sections. The readers interested in a more detailed description should have a look at [1] which gives a complete description for each component.

Cross sections of the ZEUS detector can be seen in figures 4.3 and 4.4. ZEUS uses a right handed coordinate system with the origin at the nominal interaction point.[1] The positive z-axis points in the direction of the proton beam, the y-axis upwards, and the positive x-axis towards the center of HERA. In polar coordinates the polar (azimuthal) angle θ (ϕ) is measured with respect to the z-axis (x-axis). Instead of θ one often uses the pseudorapidity $\eta = -\ln\tan(\theta/2)$. The positive z-direction (proton direction) is called forward and the negative z-direction (electron direction) rear.

Directly around the interaction region is the tracking system that detects charged particles. Around the tracking system is the solenoid that creates a magnetic field that bends the tracks of charged particles and thereby allows measurement of particle momenta. The tracking system is surrounded by the high precision uranium calorimeter (UCAL). Outside of the calorimeter are the

[1] Actually this is the nominal interaction point as defined by the central tracker. The interaction point as defined by HERA might have an offset to this one.

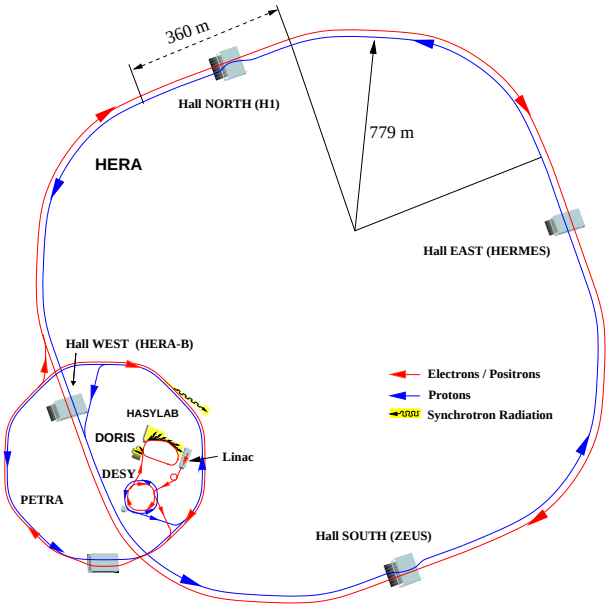

Figure 4.1: *The DESY accelerator structure. The LINAC injects particles into the DESY ring. DESY accelerates them and fills them into either DORIS or PETRA. PETRA further accelerates the particles before filling them into HERA.*

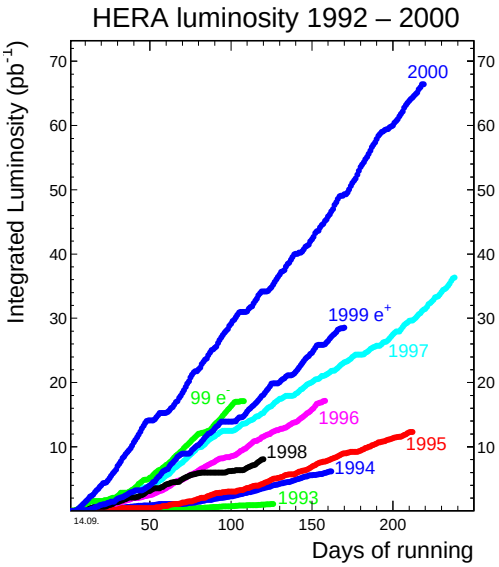

Figure 4.2: *Integrated luminosity delivered by HERA.*

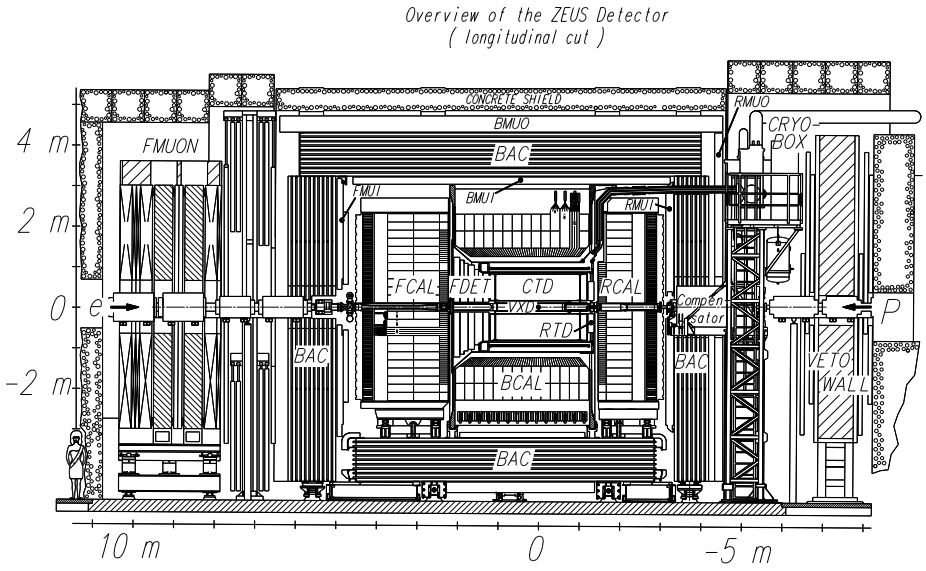

Figure 4.3: *Cut through the ZEUS detector in beam direction.*

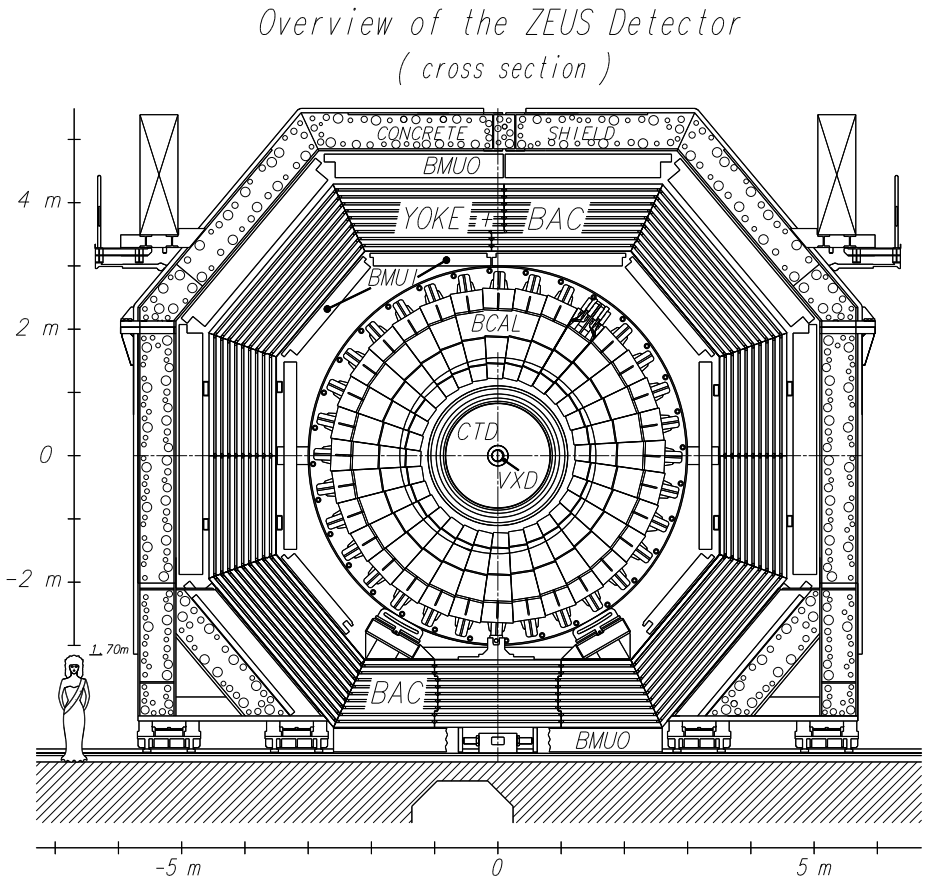

Figure 4.4: *Cut through the ZEUS detector perpendicular to the beam direction.*

inner muon chambers, the backing calorimeter and the outer muon chambers. The backing calorimeter is made of scintillators and steel and also serves as a return yoke for the solenoid.

There are also a number of specialty detectors. The forward plug calorimeter (FPC) and the beam pipe calorimeter (BPC) increase the angular coverage even further. A number of detectors in the forward (LPS, FNC, PRT) are set up for measuring the proton or its remnant. In the rear at $z = -7.5m$ an iron and scintillator veto wall is installed to detect upstream proton beam-gas interactions. The C5 counter at $z = -3.14\,\mathrm{m}$ can measure the interaction timing and thereby separate *ep* from beam gas interactions. A presampler is installed in front of the calorimeter to help correct the energy of particles that started showering before reaching it. After three radiation lengths inside the forward and rear part of the calorimeter[2] the hadron electron separator (HES) is installed for distinguishing hadronic from electromagnetic showers. The luminosity system consists of a photon calorimeter at $z = -104\,\mathrm{m}$ and an electron tagger at $-34\,\mathrm{m}$.

4.4 Interaction of particles with matter

When a charged particle traverses through matter it ionizes some of the atoms in the matter and thereby loses some energy. If this ionization is measured the path of the particle can be reconstructed. If the charged particle moves through a magnetic field it will be bent depending on the particle's charge and momentum, thereby allowing a momentum and charge measurement.

The amount of ionization done/energy lost depends on the particle's charge and velocity. Most particles can be assumed to be minimum ionizing, i.e. that their energy loss depends only on the charge. For these particles the energy loss effectively counts the number of particles going through the medium.

If a particle travels through matter for a sufficient distance it will have a hard interaction with one of the nuclei or electrons usually ending up with the particle being absorbed or deflected and several new particles being created that carry part of the particle's energy and momentum. Thereby a shower of particles is created that becomes wider and denser as it penetrates into the material. At the same time the particles become less and less energetic, at some point failing to generate any new particles on interaction, causing the shower to get smaller and disappear.

The kind of interactions that actually happen depend largely on the kind of particle involved. The simplest cases are electrons and photons. Both of them result in a cascade of electrons and photons called an electromagnetic shower. At high energy the photons will mainly pair produce and the electron will emit Bremsstrahlung photons. At lower particle energies Compton scattering for the photon and ionization loses for the electron become predominant.

If the shower starts from a hadron it is dominated by inelastic hadronic interactions. At high energy these mainly include multi particle production and particle emission from nuclear decay and excited nuclei. Since many π^0 particles are created that decay into two photons and thereby decay electromagnetically, the hadronic showers also have an electromagnetic component.

From the other leptons there are essentially two particles of special interest.

[2] FCAL and RCAL, see section 4.6

Muons move right through matter depositing minimum ionizing energy. Neutrinos normally don't interact with matter at all, therefore leaving no traces.

4.5 Basics of Calorimetry

Calorimetry involves measuring particle energies and the devices that do so are called calorimeters. Many calorimeters are segmented, which allows the reconstruction of the particle direction, distinguishing between single/multiple particles, assists in particle identification, etc.

The usual way to build a calorimeter is to have a massive chunk of high Z matter. When it is hit by a particle, the particle turns into a shower that then deposits all its energy in the calorimeter. Measuring the energy of the particle is done by measuring the energy deposited in the calorimeter.

One way to measure that energy is to use scintillators. A scintillator is a material that turns a certain fraction of the energy deposited in it into light. That light can then be measured using photodiodes or photomultipliers. As most scintillators have a long radiation length the resulting calorimeter would be rather big. To reduce that size one mixes the scintillator with a material with short radiation length, called absorber.

Calorimeters that use scintillators and absorbers are called sampling calorimeters, because they measure only the fraction of the shower's energy that's deposited in the scintillator. (For the 6m tagger described in appendix C this fraction is as low as 1–2%.) Most properties these calorimeters have come from the way that scintillators and absorbers are combined. Therefore several different designs have been tested. Two popular designs are the sandwich calorimeter which has alternating layers of scintillator and absorber and the spaghetti calorimeter that uses scintillating fibers going lengthwise through the calorimeter.

As mentioned in the last section electromagnetic and hadronic showers develop differently. The signal from hadronic showers is usually less as a fraction of the energy is lost due to nuclear processes that don't result in measurable signal or experience leakage due to muons, neutrinos and slow neutrons. As a varying fraction of the shower is electromagnetic one has to guess the electromagnetic component afterwards to make a correction for these effects. Alternatively one can try to compensate for these losses with effects that lower the ratio of electron to hadron signal like using U^{238} as an absorber. The spallation neutrons from the uranium can be detected using plastic scintillators in which the neutrons create recoil protons when scattering on hydrogen nuclei. By the right tuning of the uranium to scintillator ratio compensation can be achieved.

The most important property of a calorimeter is its energy resolution. For sampling calorimeters the most important contribution are the sampling fluctuations. However the resolution is influenced by several effects that are differentiated by how they scale with the energy:

- statistical effects ($\Delta E_{stat}/E \sim 1/\sqrt{E}$): These are mostly sampling fluctuations, but to a certain degree also other things like photo statistics.
- noise ($\Delta E_{noise}/E \sim 1/E$): This can be electronic noise or anything else that does not depend on a signal being present.
- constant ($\Delta E_{const}/E \sim 1$): Contributions that scale with the energy are mainly due to miscalibrations of individual channels

- leakage ($\Delta E_{leak}/E \sim \ln E$): If there is longitudinal leakage of the shower out of the detector, one experiences this kind of contribution.

Apart from the energy resolution there are a number of other properties that characterize a calorimeter:

- the calibration uncertainty: If the whole calibration is off by a certain factor this will lead to a systematic shift in the energy reconstruction that's very hard to detect using the calorimeter signals alone.
- uniformity of the energy reconstruction: due to manufacturing and calibration tolerances there might be position-dependent variations in the signal response. Depending on the analysis this can create a serious bias or might just degrade the resolution.
- the linearity of the energy reconstruction: there can be several reasons for nonlinearity, such as nonlinear photomultiplier responses or leakage.
- the angular coverage
- the angular/position resolution
- the linearity of the angular/position reconstruction

4.6 Uranium Scintillator Calorimeter (UCAL)

When designing the ZEUS calorimeter the goal was to get the best possible calorimeter for measuring jets [1]. For this the following goals were set (priority ordered):

1. (almost) full solid angle coverage
2. good resolution of the jet energy measurement
3. good calibration of the absolute energy scale and of the individual calorimeter sections with respect to each other
4. a good angular resolution
5. hadron-electron separation for both isolated electron and electrons in jets

The design chosen was a depleted uranium-plastic scintillator sandwich calorimeter. Due to the use of uranium and the right ratio of scintillator and absorber thicknesses an equal response to electrons and hadrons was achieved (compensation). The radioactivity of the uranium provides for a very stable calibration of the detector. The calorimeter is longitudinally segmented to allow for easy separation of electromagnetic and hadronic showers.

The calorimeter consists of three parts the cylindrical, hollow Barrel Calorimeter (BCAL) that contains all the tracking systems and the flat Forward and Rear Calorimeter (FCAL and RCAL) at both ends (see Fig. 4.5). These parts are subdivided into towers consisting of stacks of absorber and scintilator. These towers are grouped into modules (see Fig. 4.6). FCAL and RCAL consist of 23 calorimeter modules which consist of up to 23 towers near the middle and less towers on the side so that an approximately circular shape is achieved. For the

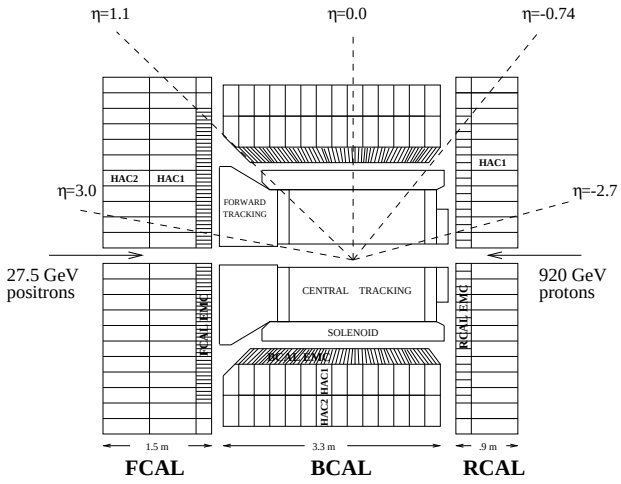

Figure 4.5: *A shematic x-z view of the ZEUS Calorimeter. For several directions the pseudorapidity is shown.*

middle module the middle tower was omitted to give space for the beam pipe. In FCAL each tower consists of four electromagnetic (EMC: $5 \times 20\,\mathrm{cm}^2$) and two hadronic (HAC: $20 \times 20\,\mathrm{cm}^2$) sections. The EMC sections are stacked up on top of each other while the HAC sections are lined up behind them. Away from the beam pipe where the BCAL covers the front of the towers the four EMC sections are combined into one section (HAC0). For RCAL only two EMC sections and one HAC section are used, otherwise the design is very similar. BCAL again uses four EMC sections and two HAC sections, but the towers are wedge shaped due to the cylindrical shape and to provide projectivity. The outermost towers have fewer EMC cells that are slightly tilted toward the interaction point.

Each section (also called cell) consists of several layers of uranium and scintillator. When a shower hits the section it generates light in the scintillators. This light is then transported to both sides where it is picked up by wavelength shifters. These convert the light to a longer wavelength and transport it to the back of the tower where it is read out by photomultipliers. Using two photomultipliers which read out the signal from both sides has several advantages:

- hardware failures normally only affect one side and therefore don't leave a dead spot in the detector.
- photomultipliers and their infrastructure can generate a noise pulse from time to time. By checking that both sides have a signal these can be mostly cut out.
- averaging the signal from both sides gives a more uniform detector response then just taking the signal from one side.
- from the left/right ratio one can get an approximation of the transverse location of the shower in the cell.

The calorimeter has a very good timing resolution of order of a 1–2 nanoseconds [1]. This avoids any pileup effects, i.e. no energy from previous events is

left in the detector. It can also be used to reject non ep background. If a time of 0 labels when the signal whould arrive at the calorimeter, a large negative time in RCAL indicates proton beamgas interaction upstream in the beam pipe. A time difference between the upper and the lower part of the detector indicates a cosmic ray passing through the detector.

4.7 Central Tracking Detector (CTD)

The central tracking detector is a cylindrical wire chamber operating in the magnetic field of the solenoid covering the angular range of $15° < \theta < 164°$. It consists of 9 superlayers (SL) (see Fig. 4.7). Five of them have the wires going paralell to the CTD axis (axial superlayers) and four of them have the wires at a small angle ($\sim 5°$) to allow for reconstruction of the z position. The angular acceptance can be extended in the forward using the FTD and in the rear using the RTD.

The CTD is filled with a gas mixture of argon, carbon dioxide and ethane. When a charged particle passes, it creates a track of ionised particles. To detect these particles a high voltage is applied. This causes the electrons to move to the positive sense wires and the ions to move to the negative field wires. In the field of the sense wire the electrons ionize more atoms producing a multiplication of the electrons by about a factor of 10^4. From the time it takes the electrons to reach the sense wire it can be estimated how far they travelled (but not from which side they came).

Each of the superlayers consists of eight layers of sense wires and is divided into cells. Each cell contains one sense wire from each layer of sense wires. The sense wires which belong to one cell are placed at an angleof 45° to radial. Thereby a particle moving radially outward has to cross at least two cells allowing a determination on which side of the sense wire the hit was.

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

4.8 Luminosity System

The definition of a cross section is number of events per luminosity, corrected for accaptance. To measure the luminosity one can in principle measure the event rate for any type of event. To get an accurate measurement one has to have exact knowledge of the cross section and also a good understanding of the acceptance. Of course one also needs reasonable statistics, where reasonable not only means that the statistical errors are negligible, but also that one can get a fairly good measurement of the instantaneous luminosity by sampling over a short period of time.

At ZEUS the luminosity is measured using the Bethe-Heitler process which can be calculated with high precision:

$$ep \rightarrow ep\gamma \tag{4.1}$$

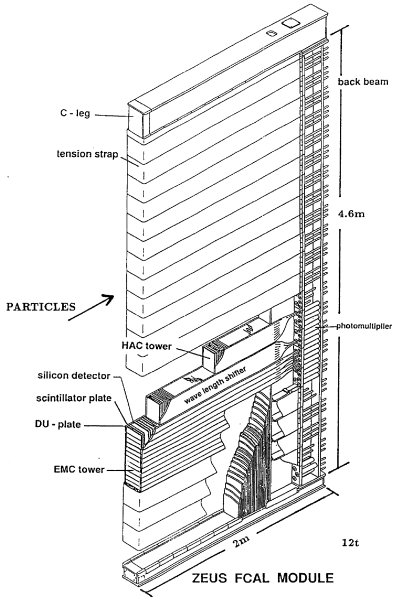

Figure 4.6: *One FCAL module*

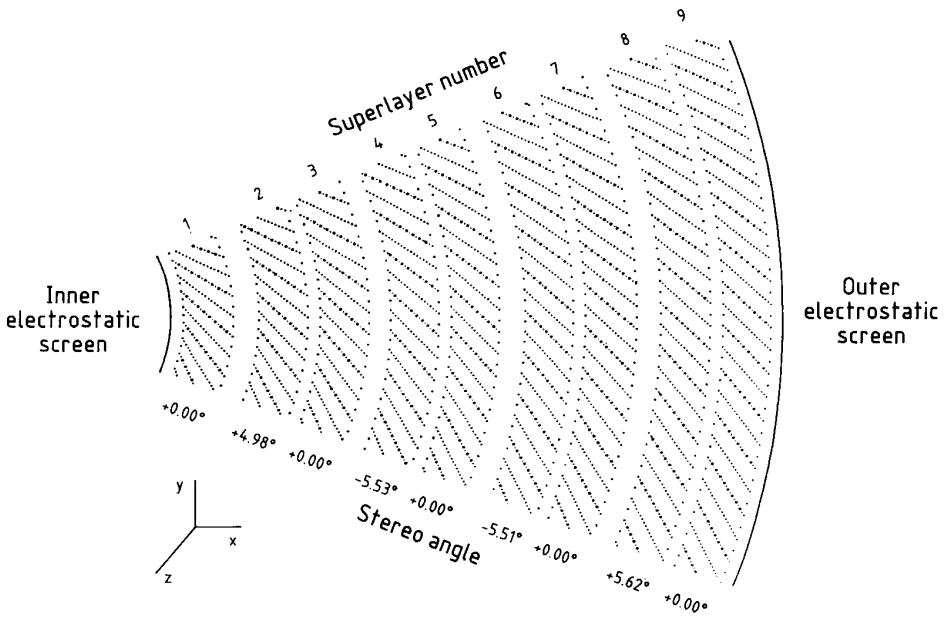

Figure 4.7: *One octant of the CTD.*

4.8.1 Old Luminosity System

For measuring the event rate both the final state electron and photon are measured. After the interaction both electron and photon move collinearly down the beam pipe. The photon isn't affected by the magnets that bent the electron beam on its path, so the photon moves straight on leaving through an exit window 92.5 m from the interaction point where it is detected by the photon calorimeter. The final state electron has lost a fraction of the initial energy and can therefore be separated from the beam electrons with a dipole magnet. It leaves through an exit window at -27.5 m and is then detected by the electron calorimeter. The electron calorimeter can also be used to tag the electrons from photoproduction events.

4.8.2 New Luminosity System

For the Lumi Upgrade that was part of the HERA II program the whole Lumi system has been redesigned. The photons are now detected by two detectors. One is the upgraded photon calorimeters that had to be shielded from the synchrotron radiation by extra filters that decreased its resolution and efficiency. The other is the pair spectrometer that measures the fraction of photons that pair produce in the exit window. The resulting electron-positron pairs are separated by a dipole magnet and detected by two calorimeters (the former BPC). The final state electrons now leave the beam pipe at an exit window at 5.85 m and are detected by a very compact electromagnetic calorimeter (the 6m tagger). Having two photon detectors allows for two independent luminosity measurements (each photon can only be detected by one of the two detectors). Tagging the electron allows to directly measure certain parameters like acceptance and energy scale thereby allowing to reduce the systematic uncertainties coming from them.

Due to a water leakage in 2002 the 6m tagger had been destroyed and had to be completely rebuilt under the responsibility of the author. For this the detector had to be deassembled and the tungsten absorber plates chemically cleaned. The detector was then reassembled and tested in the DESY testbeam. The testbeam results were compared to Monte Carlo simulations and the Monte Carlo model was improved until an agreement was reached. This is documented in appendix C starting on page 117.

4.9 Trigger System

At HERA we have 96 ns between the individual bunch crossings, resulting in an interaction rate of more than 10 MHz. This creates several problems: It is virtually impossible to read out the entire detector at that rate and even if that were possible it would be totally impossible to store all that data. To solve these problems the detector is equipped with a three level trigger system which step by step reduces the event rate down to a rate of a few Hz (see Fig. 4.8).

The purpose of the First Level Trigger (FLT) is to reduce the event rate to the level that makes it possible to read out the detector. As the FLT decision takes some time, the components store their data in a pipeline and read it out if the FLT decides to keep an event. For making the FLT decision the components are equipped with a second readout system, that takes a reduced

amount of data at the full rate, once every 96 ns. This data is then analyzed in dedicated hardware and sent to the GFLT. The GFLT then decides within 2 μs whether the event was interesting. If it decides to take the event it sends out a signal to read out the detector.

The purpose of the Second Level Trigger (SLT) is to reduce the event rate to the level that makes it possible to analyze the data in more detail. The SLT forms it decision after all data has been read in. It is software based running on a network of transputers distributed over the components. As it has only limited processing power available it can do only simple checks. If the SLT decides to keep an event, the event builder collects the data from all the components and puts it into a format that the TLT can process.

The purpose of the Third Level Trigger (TLT) is to reduce the event rate to the level that allows writing the events to tape. The TLT is software-based and runs on a cluster of commodity computers. It runs several simplified physics analysis on each event and then decides whether one might actually use the event for at least one of these analysis. All events accepted by the TLT are written to tape and are available for the offline analysis.

Internally each trigger level works by checking the event for a number of conditions (called trigger slots) and if it fulfills at least one of them it accepts the event. For the event written to tape the decision for each trigger slot is stored as well as other trigger variables. This information can then be used to understand the trigger decision.

After setting up the trigger slots in a way that selects all events that are possibly interesting the usual result is that one exceeds the rate requirements of the system. So one has several possibilities to cut down on the rate of some slots. For this there are two easy solutions:

- Raise the thresholds for the slot and thereby lose the less interesting events.
- Take only a fraction of the events in the slot (prescale the slot). This cuts down the statistics, but avoids any bias for a certain kind of event.

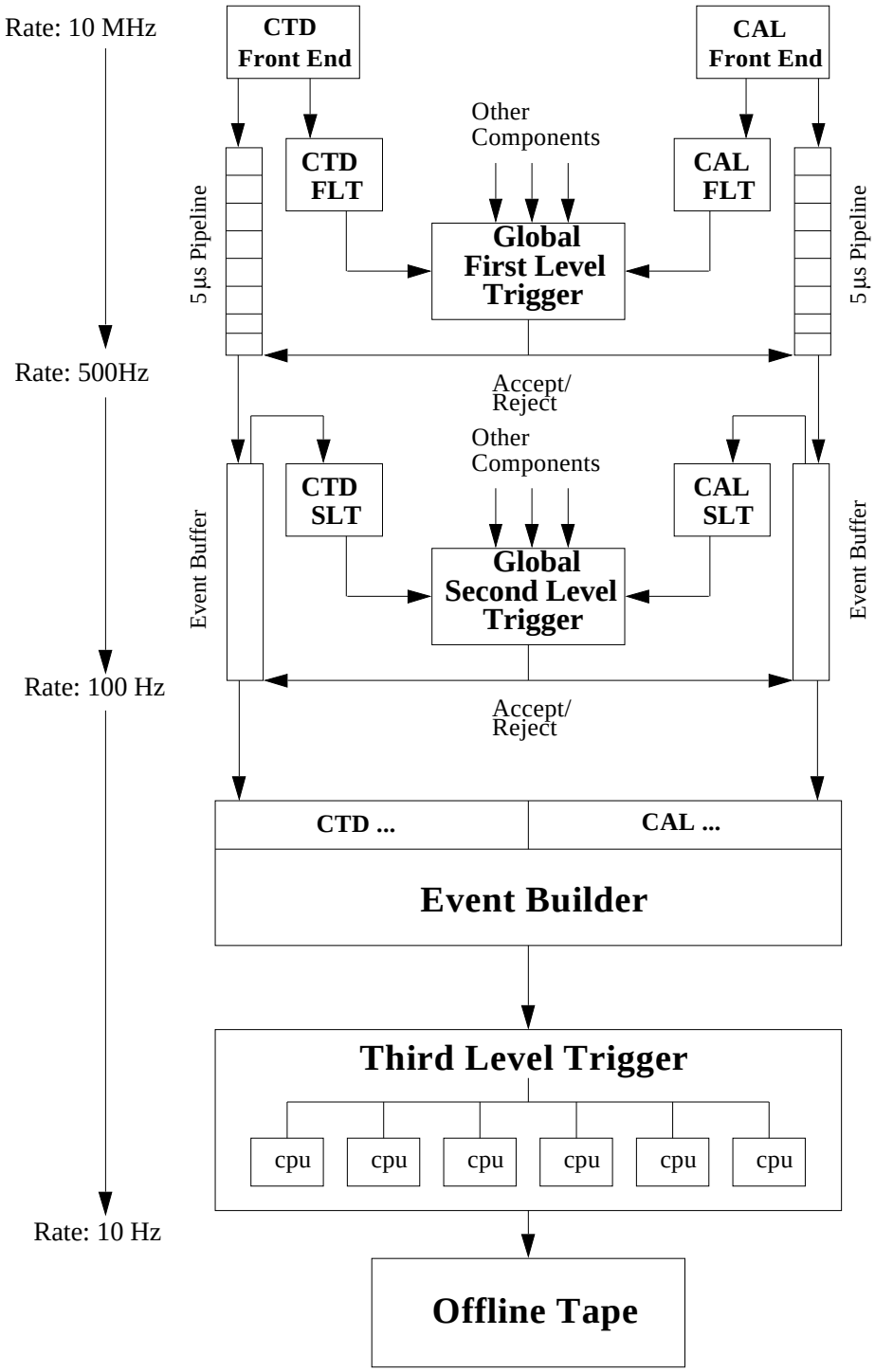

Figure 4.8: *Diagram of the ZEUS trigger and data acquisition system.*

Chapter 5

Event Reconstruction

5.1 Calorimeter noise surpression

The main detector component used for this analysis is the Uranium Calorimeter. The calorimeter measurement consists of a (zero surpressed) list of energies in the different calorimeter cells. To improve the calorimeter performance the cells that are affected by noise are excluded from the further reconstruction [22]. Cells are assumed to contain noise if they:

- have one photomultiplier with much higher signal than the other. These are likely to come from noise pulses of one of the photomultipliers.
- are known to be rather noisy. For each running period a list of these noisy cells is maintained.
- are isolated and of low energy. ($E_{emc} < 0.1\,\mathrm{GeV}$ or $E_{hac} < 0.15\,\mathrm{GeV}$)

All these cuts have little impact when looking at high E_T jets, which consist of several cells with rather high energies.

5.2 Track Reconstruction

Although the primary object of study in this analysis are energy depositions in the calorimeter, track information is used for background rejection and for more precise position and momentum information.

As described in section 4.7 charged particles create hits along their way through the CTD. To these hits the tracks of the particles are fit [7]. A pattern recognition software groups the individual hits into tracks which are then fit with a five parameter curve. The tracks can be connected to the primary vertex (the interaction vertex) or a secondary vertex (a particle decay vertex). The connection to a vertex adds additional constraints on the track and at the same time constrains the position of the vertex.

The vertex position is used in combination with the impact position on the calorimeter to get a precise measurement of the direction of the outgoing particles. The tracks are also used as additional input for the reconstruction of energy flow objects and of the electron. The fit result is also used for a number

of cuts applied in the trigger logic and the offline analysis to improve the quality of reconstruction and to remove backgrounds.

5.3 Electron Reconstruction

The prime signature of neutral current DIS events is the scattered electron found in the calorimeter. To find it a special program (Sinistra95) has been developed [23]. It takes all calorimeter cells after noise suppression (comp. 5.1) and groups them into islands. Islands within CTD acceptance, but without a track are not considered.

The lateral and longitudinal distribution of energy within the islands is then expressed as a series expansion. The first terms of this expansion are then fed into a neural network trained to find electrons. From the neural network a probability for each candidate is obtained. If the candidate with the highest probability has a probability greater than 0.9 it is assumed to be the kinematic electron. If there is more than one candidate with a probability greater than 0.9, only the one with the highest probability is considered.

If an electron candidate has been selected its energy is calculated as the sum of cell energies in the island. This energy is then corrected for different shower developments using the preshower detectors (if available). The energy is corrected with dead material maps for energy loses in dead material in front of the calorimeter. These maps are different for data and simulation to compensate for incorrect description of the dead material in the simulation.

5.4 Kinematic Reconstruction

To reconstruct the kinematic variables and the Breit frame an exact reconstruction of the four momentum of the exchanged photon is necessary. For the reconstruction one can use any combination of the electron beam energy $E_{e,beam}$, the angle θ_{el} and energy E_{el} of the scattered electron and the transverse momentum and the difference of energy and longitudinal momentum of the hadronic system.

For the reconstruction of the kinematic variables any combination of three of these variables can be used. For the reconstruction of the exchanged photon one needs to reconstruct the azimuthal angle. For this the azimuthal angle of the electron is used, as it has higher precision than the one obtained from the hadronic system.

One usual choice for an input variable is the electron beam energy which is set to the nominal beam energy.[1] This gives a fixed center of mass energy and reduces the number of degrees of freedom to two (normally chosen to be y and Q^2). These can be taken solely from the electron or the hadronic system. It is also common just to use the angles from the electron and the hadronic system.

[1] This is not correct for events with initial state radiation (ISR). These make up only a small fraction of the events and are normally ignored, but for special applications it is possible to measure the energy lost by ISR indirectly from other variables or directly from the photon calorimeter of the luminosity system.

5.4.1 Electron Method (el)

The electron method is the conceptually simplest method to reconstruct the kinematics as it only uses the scattered electron for the kinematic reconstruction. The difference between four momentum of outgoing and incoming electron is equal to the four momentum of the exchanged photon. This gives the following formulas:

$$y_{el} = 1 - \frac{E_{el}}{2E_{e,beam}}(1 - \cos\theta) \tag{5.1}$$

$$Q^2_{el} = 2E_{e,beam}E_{el}(1 + \cos\theta) \tag{5.2}$$

5.4.2 Jaquet-Blondel Method (jb)

Independent from the electron one can reconstruct the kinematics using only the hadronic system. The hadronic final state consists of many particles of which some belong to the remnant and escape through the beam pipe. Jaquet and Blondel [9] proposed a kinematic reconstruction for which the four momentum is summed up for all these particles. To avoid biases from the particles lost through the beam pipe one only looks at the transverse momentum $p_{T,had}$ and the difference between energy and longitudinal momentum $(E - p_z)_{had}$. Both of these variables are close to zero for remnant particles going down the beam pipe. The resulting formulas are:

$$y_{jb} = \frac{1}{2E_{e,beam}}(E - p_z)_{had} = \frac{1}{2E_{e,beam}}\sum(E_i - p_{z,i}) \tag{5.3}$$

$$Q^2_{jb} = \frac{1}{1 - y_{jb}}p^2_{T,had} = \frac{1}{1 - y_{jb}}\left[\left(\sum p_{x,i}\right)^2 + \left(\sum p_{y,i}\right)^2\right] \tag{5.4}$$

5.4.3 Double Angle Method (da)

The basic idea of the double angle method is to use the scattering angles of the electron and the struck quark.[2] This results in the formulas for the kinematics:

$$y_{da} = \frac{\sin\theta(1 - cos\gamma_h)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h + \theta)} \tag{5.5}$$

$$Q^2_{da} = 4E^2_{e,beam}\frac{\sin\gamma_h(1 + \cos\theta)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h + \theta)} \tag{5.6}$$

The angle of the struck quark can't be measured directly, but it can be calculated solely from hadronic variables using the Jaquet-Blondel Method:

$$\cos\gamma_h = \frac{Q^2_{jb}(1 - y_{jb}) - 4y^2_{jb}E^2_{e,beam}}{Q^2_{jb}(1 - y_{jb}) + 4y^2_{jb}E^2_{e,beam}} \tag{5.7}$$

$$= \frac{(\sum p_x)^2 + (\sum p_y)^2 - (\sum E - p_z)^2}{(\sum p_x)^2 + (\sum p_y)^2 + (\sum E - p_z)^2} \tag{5.8}$$

The advantage of using angles is that angular resolution is normally better than energy resolution, giving a more precise measurement.

[2] This derivation assumes an elastic scatter on a quasi-free quark (quark parton model), which is incorrect for the events studied. However the expressions derived are still valid.

5.4.4 Choosing the Kinematic Reconstruction

Given the various possibilities for the kinematic reconstruction one can select the kinematic reconstruction which achieves the best results. The quality of the kinematic reconstruction depends largely on the kinematic region examined, several methods are normally used with the choice depending on the kinematic region.

It is possible to use separate reconstructions for the different kinematics cuts depending on which reconstruction is the best at the cut boundary. For the reconstruction of the Breit frame one has to choose a reconstruction (on a per event basis). The double angle method has a good resolution at high Q^2 whereas the electron method has a better resolution at low Q^2. To determine the transition point a study has been done in [6] for dijet events. It showed that a good choice is to use the double angle method if the electron has a track associated with it and $\gamma_h < \pi/2$. Otherwise the electron method gives a better resolution.

5.5 Energy Flow Objects (ZUFOS)

The jet measurement depends on the measurement of energy deposits. For that measurement the calorimeter information would be sufficient. However the tracking system has a much better angular resolution and a better energy resolution at low energies. Therefore an improved measurement of the energy deposits can be made by combining calorimeter and tracking information into energy flow objects. At ZEUS the energy flow objects are called ZUFOs (for details see [5]).

For this all cells (except the ones belonging to the scattered electron) and all good primary vertex tracks are used. To match tracks to calorimeter information it is necessary to combine cells belonging to one shower into one island. For this the next neighbor algorithm is used (which doesn't connect cells diagonally). However this grouping decreases the granularity of the calorimeter (especially in the forward).

The matches of tracks and islands are the ZUFOs. A ZUFO can have more than one track or island and is not required to have both a track and an island. The decision on how to assign position and energy to the ZUFO is based on the following procedure:

- Unmatched tracks or islands are counted directly. For calculating the energy of tracks they are assumed to be ions.

- Islands with more than three associated tracks are measured with the calorimeter.

- If the calorimeter measures a much higher energy than the track it is assumed that neutral particles hit at the same position as the charged particles and the calorimeter measurement is used.

- If the tracking system has a better energy resolution (at low energies) than the calorimeter the tracking system is used, otherwise the calorimeter (at high energies).

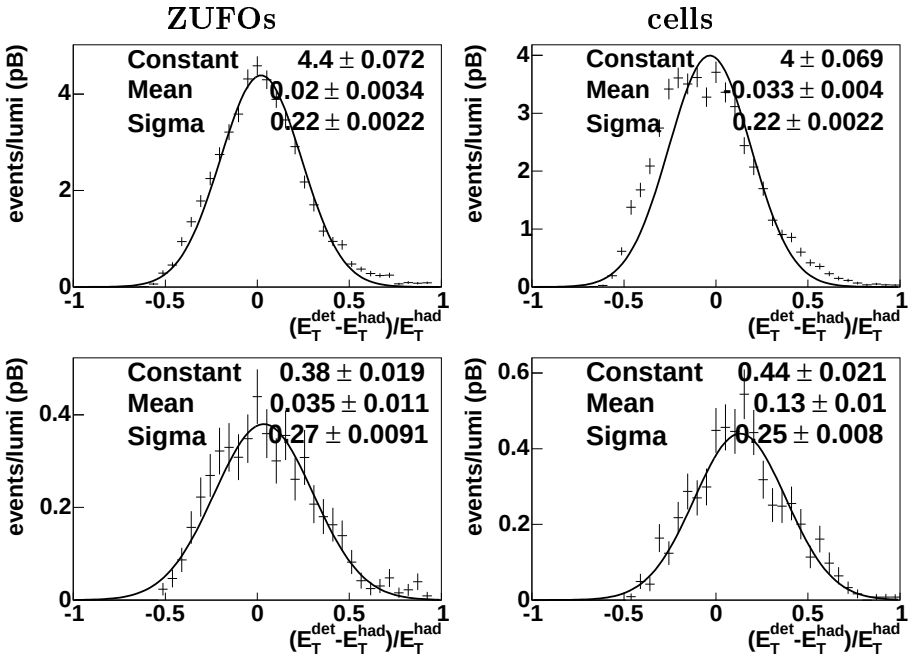

Figure 5.1: *Resolution in E_T in the range $5\,GeV < E_T < 7\,GeV$ using ZUFOs (left) and cells (right) for $-1 < \eta < 2.5$ (top) and $2.3 < \eta < 2.5$ (bottom). The resolutions have been determined as described in section 7.5.*

- If a match consists of more than one island (track) the information from the islands (tracks) is first combined and then used for this algorithm.
- If a single track matches one or two islands the positional information is taken from the tracking system.

The result of all this work is an improved overall resolution in E_T and η and m_{jets} (compared to using simple cells) although the resolution is worse in the forward region (fig. 5.1, 5.2 and 5.3).

5.6 Jet Reconstruction and Energy Correction

The ZUFOs are boosted into the Breit frame and then grouped into jets using the longitudinally invariant k_T algorithm. To improve the accuracy of the energy scale of the calorimeter the jets are then corrected to match the energy scale of the simulation. After that the jet energy is corrected for energy losses in the material in front of the calorimeter.

For measuring the energy scale a sample of pure one-jet events is studied in the lab frame. These events are assumed to be elastic scattering of the photon on a quasi-free quark (quark parton model) for which the energy can be calculated using the double angle method. The ratio of measured energy from the calorimeter and reconstructed energy from the double angle method is directly sensitive to the energy scale. From the ratio of data and simulation the energy scale can be determined more precisely. This has been done in [4]

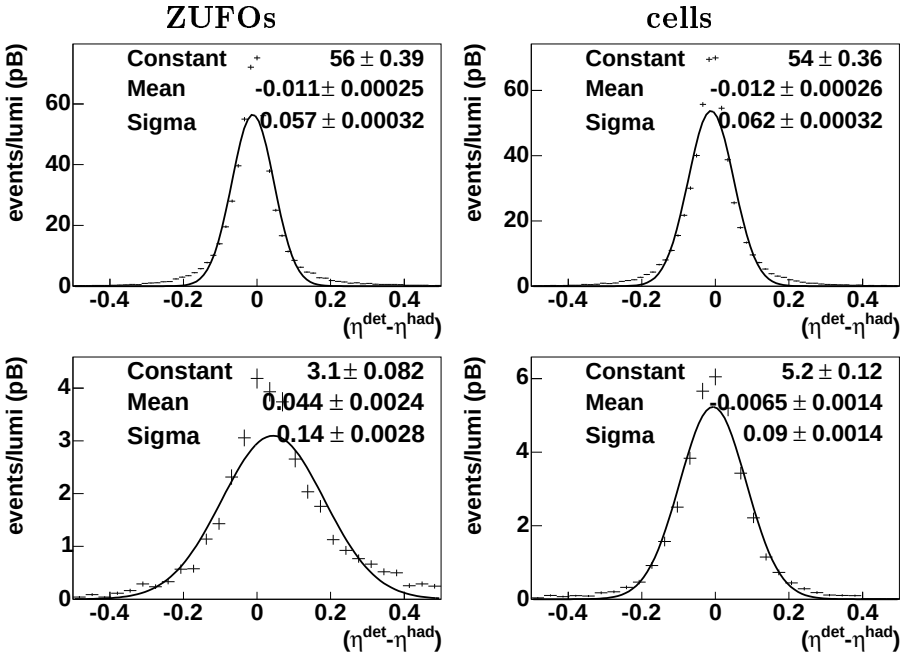

Figure 5.2: *Resolution in η in the range $-1 < \eta < 2.5$ (top) and $2.3 < \eta < 2.5$ (bottom) using ZUFOS (left) and cells (right). The resolutions have been determined as described in section 7.5.*

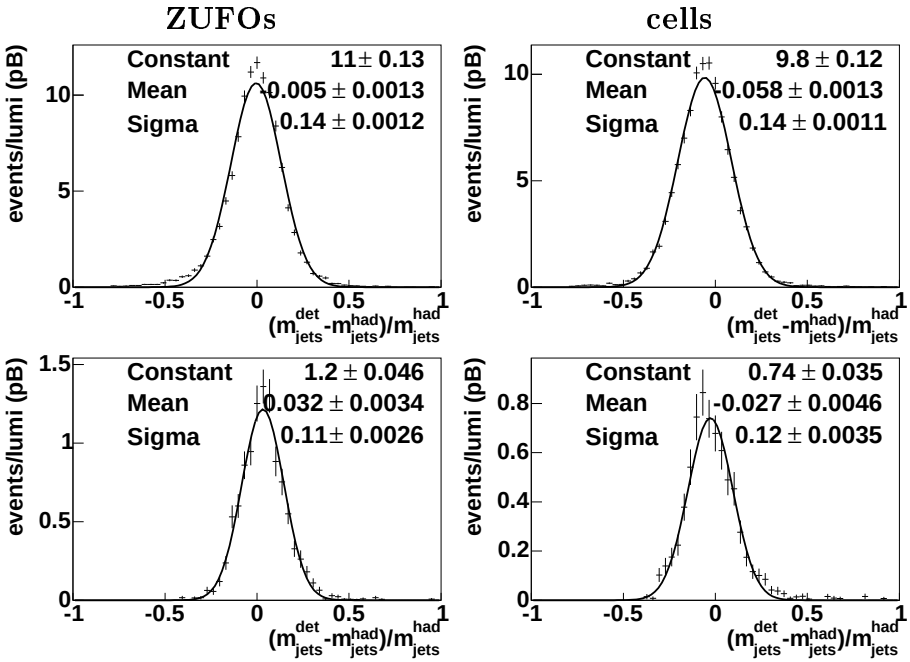

Figure 5.3: *Resolution in m_{jets} in the range $25\,GeV < m_{jets} < 30\,GeV$ using ZUFOS (left) and cells (right) for dijets (top) and trijets (bottom). The resolutions have been determined as described in section 7.5.*

for jets above 10 GeV transverse momentum. It was shown that applying these correction factors improves the energy scale uncertainty from 3% to 1%.

After that the jets are corrected for energy loses. For this the simulation is used. For events that pass the detector level kinematic and cleaning cuts the jets are then matched at hadron level by requiring hadron and detector level jets to be closer than 1 in η-ϕ-space. The dependence of detector level E_T^{lab} on hadron level E_T^{lab} is then calculated in bins of detector level η^{lab}. For each bin a straight line is fit which is then used to correct all jets that are reconstructed in that pseudorapidity bin.

Chapter 6

Event Selection

6.1 Electron Identification

The primary signature that distinguishes NC DIS from other processes is the scattered electron in the calorimeter. To find it, the Sinistra electron finder (see section 5.3) is run over all energy deposits in the calorimeter. If a candidate for an electron has been found (with a probability of at least 0.9) the event is accepted and the energy deposits belonging to that electron are removed from the further analysis.

To ensure a high purity and efficiency of the electron finder a high electron energy ($E_e > 10\,\mathrm{GeV}$) of the electron is necessary. However a direct cut on the electron energy is not necessary as the cut $y_e < 0.6$ which is introduced in section 6.2 cuts these events as well.

To stay away from the beam pipe region, where a bad reconstruction of the electron may occur, a box cut on the electron position in the RCAL is applied:

$$|x_e| > 13\,\mathrm{cm} \quad \text{or} \quad |y_e| > 7\,\mathrm{cm} \tag{6.1}$$

If the electron is too close to the hadronic system there can be problems in separating the energy deposited by the electron from other energy deposits. To reduce this effect an upper limit is imposed on the amount of energy around the electron (within a radius of 1 in η-ϕ space) that is not associated with the electron:

$$E_{had,cone}/E_e < 0.1 \tag{6.2}$$

If the electron is scattered at low angles with high energy it loses hardly any $E-p_z$. As there is a finite resolution of the calorimeter we may end up measuring a higher $E-p_z$ then kinematically allowed which distorts reconstruction of the Breit frame. To avoid that these events are cut out:

$$(E-p_z)_{ele} < 54\,\mathrm{GeV} \tag{6.3}$$

6.2 Kinematic Selection

All the cuts described in this section describe the kinematic range analyzed. This means that they can be directly reconstructed from the four momentum of

the virtual photon, which in turn means they can be directly localized in the x-Q^2 plane. For the hadron level selection they are all applied to the corresponding hadron level variables (see section 6.6). For reconstructing these quantities there are different methods applied, as described in section 5.4.

The events are restricted to be in the kinematic range of

$$10\,\mathrm{GeV}^2 < Q^2 < 5000\,\mathrm{GeV}^2 \tag{6.4}$$

The lower end of the kinematic region was set to allow for a reasonably good reconstruction of the event (and agreement with the LO MC shower as shown in section 7.4) and the upper cut was chosen to exclude low statistics regions.

Good reconstruction of jets requires a reasonably good position resolution in the calorimeter. In the Breit frame the calorimeter geometry is different from event to event. The boost sends the proton remnant down the positive z direction and makes the current jet[1] down the negative z-direction. If there are only a few cells between the remnant and the current jet these are blown up to fill half the phase space, giving a bad granularity [4]. Therefore a cut on the angle of the current jet is introduced:

$$\cos\gamma_{had} < 0.7 \tag{6.5}$$

To remove fake electron candidates (mostly highly energetic pions in the FCAL decaying into two photons) and improve the electron finding inefficiency for low energy electrons, an upper cut on the Bjorken y reconstructed with the electron method is done:

$$y_e < 0.6 \tag{6.6}$$

To ensure a sufficient energy in the hadronic system which is necessary to use the double angle method, a lower cut on the Bjorken y reconstructed with the Jaquet-Blondel method is done (y_{jb} is reconstructed from the hadronic variables and scales directly with the hadronic $E - p_z$):

$$y_{jb} > 0.04 \tag{6.7}$$

6.3 Cuts on Global Quantities

Four momentum is assumed to be conserved in all processes. The only energy and momentum that is supposed to escape undetected is associated with the proton remnant going down the beam pipe. It carries energy and longitudinal momentum, but no $E - p_z$. Therefore $E - p_z$ should be conserved ($= 2 \cdot 27.52\,\mathrm{GeV}$), permitting the requirement:

$$40\,\mathrm{GeV} < E - p_z < 60\,\mathrm{GeV} \tag{6.8}$$

The lower cut mainly removes proton beam gas interactions, whereas the upper cut removes cosmic events superimposed on normal ep events.

Transverse momentum conservation requires no total transverse momentum in the final state. Due to the finite resolution of the calorimeter one will always

[1] An imaginative jet going the direction a quasi-free elastically scattered quark would go in a one jet event. For the events studied there are typically no particles going in that direction at all.

measure some transverse momentum. The resolution of the calorimeter scales approximately with the square root of the transverse energy. Taking the ratio of transverse momentum (summed as vector) and square root of transverse energy (summed as scalar) gives the deviation from 0 with respect to the resolution. Introducing an upper cut mainly reduces events with one cosmic particle going through the detector simultaneous with the *ep* interaction:

$$p_T/\sqrt{E_T} < 3\sqrt{\text{GeV}} \tag{6.9}$$

The coordinate system is set up so that the nominal interaction point is at its origin. The vertex (point of interaction) should be somewhere nearby. If the interaction happened upstream or downstream of the nominal interaction point it is more likely to originate from an interaction of the beam with residual gas in the beam pipe. Therefore the reconstructed z-position is restricted:

$$-50\,\text{cm} < z_{vertex} < 50\,\text{cm} \tag{6.10}$$

At HERA we do not only measure normal DIS events, but also diffractive events. These events are not simulated by the Monte Carlo or the NLO and are therefore cut out. One characteristic of these events is that there is no color connection and therefore no particles between the proton remnant and hadronic systems. To cut these events out we require a minimum pseudorapidity (in the lab frame) of the most forward energy deposition:

$$\eta_{max} > 2.5 \tag{6.11}$$

6.4 Jet Selection

The jet finding is done on groups of energy deposits in the calorimeter called ZUFOs (see section 5.5) which are made massless by scaling the energy to match the momentum. The ZUFOs are first boosted to the Breit frame (see sections 1.3.2 and 5.4.4) and then subjected to the jet finder. The jet finder uses the longitudinally invariant K_T algorithm (see section 1.3.3).

To make sure the jets are fully contained in the calorimeter they are boosted back to the lab frame to check their impact position. A cut on the pseudorapidity is applied:

$$-1 < \eta_{lab}^{jet} < 2.5 \tag{6.12}$$

As the hadronic jets have to go through a sizeable amount of matter before hitting the calorimeter they lose a fraction of their energy. This is corrected for using the simulation. To avoid insignificant energy deposits being blown up to high energies a minimum uncorrected transverse jet energy is required:

$$E_T^{jet,uncor} > 3\,\text{GeV} \tag{6.13}$$

As the goal is to measure jets coming from partons associated with the hard part of the scatter a minimum transverse energy is required.[2] This also has experimental advantages as high energy jets are easier to measure than low energy jets. Therefore the following cut was chosen:

$$E_T^{jet} > 5\,\text{GeV} \tag{6.14}$$

[2]Soft partonic radiation tends to happen at low transverse energies. Therefore a high transverse energy indicates a hard process.

Only the events where at least two (three) jets satisfy above conditions are kept for the dijet (trijet) sample. Another cut on the invariant mass of the dijet (trijet) system is introduced to ensure stable behavior of the NLO calculation (see section 3.6):

$$m_{jets} > 25\,\text{GeV} \tag{6.15}$$

6.5 Trigger Selection

At ZEUS we use a three level trigger system to preselect events and reduce the rate. This means that we have to deal with a number of problems:

1. The trigger does not only cut background, but also physics events. This makes certain regions of phase space inaccessible and gives other regions an acceptance that is difficult to understand. To avoid that one focuses on a few trigger slots and a region of phase space that gives these slots a good acceptance.

2. There are simulated events that survive all cuts, but are not selected by the trigger. For the data these are missing, but for the simulation they are still in the sample. By selecting exactly the same trigger bits for data and simulation these can be removed from the Monte Carlo.

3. Some slots are prescaled, i.e. they don't take every event they fire on, but only every n-th event (n is called the prescale). The effective luminosity for prescaled slots is different, which is best dealt with by giving different weights for these events. If a large fraction of events comes from prescaled slots this means that the statistical error becomes larger since it corresponds to the effective luminosity.

This analysis uses a combination of three different TLT slots: The photoproduction dijet trigger (HPP14), the medium Q^2 DIS trigger (DIS03) and the fully inclusive, but prescaled DIS trigger (DIS01). Each of these triggers has its own configuration of FLT and SLT slots for which it works reliably. When using one of these slots it is necessary that at least one of the corresponding FLT and SLT slots took the events. Therefore the slot description of the FLT and SLT slots includes which TLT slot relies on them.

6.5.1 FLT selection

For this analysis several FLT slots are used. They can be vetoed by signals from the C5, Veto Wall, SRTD or CTD. If there is no veto at least one of them has to fire for the event to be taken:

- The slot FLT40 takes the event if the electromagnetic energy in the calorimeter exceeds 15 GeV.

- The slot FLT41 takes the event if the total transverse energy in the calorimeter exceeds 21 GeV.

- The slot FLT42 is only used for HPP14. It takes the event if there is a good track and one of the following is true:

  - the total energy in the calorimeter is greater than 15 GeV
  - the total electromagnetic energy in the calorimeter is greater than 10 GeV
  - the total electromagnetic energy in the BCAL is greater than 3.4 GeV
  - the total electromagnetic energy in RCAL is greater than 2.0 GeV.

- The slot FLT43 takes the event if the total transverse energy in the calorimeter is greater than 11.5 GeV and a good track is found.

- The slot FLT44 is not used by HPP14. It takes the event if the electromagnetic energy in RCAL (BCAL) is greater than 3.4 GeV (4.8 GeV).

- The slot FLT46 is not used by HPP14. It takes the event if there is an isolated electromagnetic deposit in the RCAL. In addition it requires one of the following:

  - the total electromagnetic energy in the RCAL is greater than 2 GeV.
  - there is SRTD data and a good track.
  - there is a good track and a total energy greater than 18 GeV.

6.5.2 SLT selection

For the second level trigger there are two slots used depending on whether a DIS or HPP slot is used for the TLT:

- The slot DIS06 takes the event if $E - p_z + 2 \cdot E_\gamma^{(lumi)} > 29\,\text{GeV}$ and one of the following:

  - electromagnetic energy in RCAL greater than 2.5 GeV
  - electromagnetic energy in BCAL greater than 2.5 GeV
  - electromagnetic energy in FCAL greater than 10 GeV
  - hadronic energy in FCAL greater than 10 GeV

- The slot HPP01 takes the event if all of the following is true:

  - the vertex is reconstructed with $\|z\| < 60\,\text{cm}$ or no vertex is reconstructed.
  - there is at least one vertex track
  - $E - p_z > 8\,\text{GeV}$
  - the transverse energy (except for the innermost FCAL ring) is greater than 8 GeV.
  - $E - p_z > 8\,\text{GeV}$ or $p_z/E > 0.95$

6.5.3 TLT selection

For this analysis three TLT slots are used and the event is taken if one of the slots takes the event. To improve the selection quality for each slot (except the inclusive DIS01) the cut done online in the TLT ir repeated with the offline quantities at a higher cut value (thereby overriding the online cut value):

- The slot HPP14 is a dijet PHP trigger. It requires at least two jets found in the lab frame with an E_T of at least 2 or 2.5 GeV (depending on the trigger configuration: at the beginning of the 98–00 it was set to 2 GeV and later on it was raised to reduce the rate). For this analysis and this slot it is required that the two jets with the highest transverse energy each have an uncorrected transverse energy in the lab of at least 5 GeV.

- The slot DIS03 is the inclusive DIS trigger for high Q^2. It requires an electron found with at least 4 GeV energy outside of a circle with radius 35 cm centered around the beam pipe. For this analysis and this slot it is required that the electron is outside a circle of 36 cm on the face of RCAL.

- The slot DIS01 is the inclusive DIS trigger. It requires an electron found with at least 4 GeV of energy outside of a region of $24 \times 12\,\mathrm{cm}^2$ centered around the beam pipe. Depending on the trigger configuration it has a prescale (see section 6.5) of 1, 10 or 100 resulting in an integrated luminosity of $11.37\,\mathrm{pb}^{-1}$ (This number has been calculated by summing up the ratios of integrated luminosity and prescale for the individual runs). In this analysis it is is used as a fall-back trigger for events that were not selected by the other two TLT slots, i.e. if an event is taken by either HPP14 or DIS03 it is filled with a weight of $1/82.2\,\mathrm{pb}^{-1}$ (corresponding to the full luminosity) and if it is not taken by those slots but taken by DIS01 it is filled with a weight of $1/11.37\,\mathrm{pb}^{-1}$ (corresponding to the prescaled luminosity).

6.6 Hadron Level Selection

Up to now the cuts have been described as they are implemented at detector level. Many of these cuts are rather specific, dealing with one minor effect and are better dealt with as an inefficiency of the detector. For the actual definition of the hadron level cross section a smaller subset of these cuts should be chosen with each cut having a clear physical meaning. This section contains a complete listing of all the cuts applied at hadron level.

The cuts in the kinematic plane are the most general, so they can all be extended to the hadron level:[3]

$$10\,\mathrm{GeV}^2 < Q^2 < 5000\,\mathrm{GeV}^2 \tag{6.16}$$

$$0.04 < y < 0.6 \tag{6.17}$$

$$\cos\gamma_{had} < 0.7 \tag{6.18}$$

[3]These variables are directly calculated from the four momentum of the hadron level photon assuming nominal beam energies ignoring that for ISR events the actual beam energy is reduced by the energy of the emitted photon. However this effect is not present after the QED corrections introduced in section 3.8.

The cut on the pseudorapidity of the jets should also be applied at the hadron level. It cuts out a large amount of jets which would have to be estimated using theoretical predictions. The jets going into the far forward direction are also theoretically not well understood, making this extrapolation even more suspicious:

$$-1 < \eta_{lab}^{jet} < 2.5 \tag{6.19}$$

The requirement of a minimum transverse energy is as much a theoretical as an experimental requirement. For the experiment a high transverse energy is required to reduce the influence of noise. For the theory it is required to separate hard from soft processes:

$$E_T^{jet} > 5\,\mathrm{GeV} \tag{6.20}$$

The invariant mass cut is perscribed by theory (see section 3.6) and therefore has to be applied at the hadron level:

$$m_{jets} > 25\,\mathrm{GeV} \tag{6.21}$$

The definition of the hadron level does not only consist of these cuts, but also on the definition of what a hadron is and which processes are actually considered. This will be described in sections 2.6 and 3.8.

Chapter 7

Analysis Method and Data Correction

7.1 Cross Section Definition

When evaluating scattering experiments the outcome is always just the probability of a certain result to happen. In classical predictions (fully deterministic: each hit position corresponds to a specific result) this probability directly scales with the area that can be hit to cause a certain result. Therefore all reaction probabilities are defined as cross sections (the area that can be hit for the reaction to happen) although in modern theories the outcome doesn't depend on the impact position.

When conducting an experiment the number of observed events is (within statistics) directly proportional to the cross section. The proportionality constant is the integrated luminosity $\mathcal{L}$ (in the classical interpretation the density of scattering targets). Although the luminosity could, in principle, be calculated from the beam geometry, it is more precise to measure a well understood cross section and then calculate the luminosity from that.

From the above follows the general definition of a cross section:

$$\sigma = \frac{N_{event}^{data}}{\mathcal{L}} \tag{7.1}$$

When conducting an experiment there are some obstacles to using eqn. 7.1 directly. Neither efficiency e (percentage of generated events detected) nor purity p (percentage of correctly detected events) is at 100%. To correct for inefficiencies and impurities the formula is modified (using a correction factor c):

$$\sigma = c \cdot \frac{N_{event}^{data}}{\mathcal{L}} = \frac{p}{e} \cdot \frac{N_{event}^{data}}{\mathcal{L}} \tag{7.2}$$

To determine the purities and efficiencies a Monte Carlo model is used. The model gives access to the true (hadron level) and reconstructed (detector level) values. By counting how many events are accepted at the detector level (det), at the hadron level (had) and at both levels ($det \oplus had$) the purities and efficiencies

can be calculated:

$$p = \frac{det \oplus had}{det} \tag{7.3}$$

$$e = \frac{det \oplus had}{had} \tag{7.4}$$

As explained in section 3.8 the cross section has to be corrected for QED effects yielding the final formula:

$$\sigma = c_{QED} \cdot c \cdot \frac{N_{event}^{data}}{\mathcal{L}} = c_{QED} \cdot \frac{had}{det} \cdot \frac{N_{event}^{data}}{\mathcal{L}} \tag{7.5}$$

7.2 Calculation of the Statistical Error

This section only describes the calculation of statistical errors. For the calculation of the systematic errors see section 8.1. For a simple counting experiment that results in n events seen, the statistical error is $\Delta n = \sqrt{n}$ (Poisson statistic). This is no longer true if the events have different weights. (This can happen by prescaling (sec. 4.9 and 6.5), prefiltering (sec. 7.3), reweighting 7.3.1, ...) For a sample where the event i has the weight w_i the formulas are

$$n = \sum w_i \qquad \Delta n = \sqrt{\sum w_i^2} \tag{7.6}$$

which is equivalent to Poisson statistic for $w_i = 1$.

Often several values n_i with errors Δn_i have to be combined into one value $\hat{n}$ using a function f:

$$\hat{n} = f(n_1, n_2, n_3, \ldots) \tag{7.7}$$

If the quantities n_i are uncorrelated the error can be calculated using error propagation:

$$\Delta\hat{n} = \sqrt{(\Delta n_1)^2 \cdot (\frac{\partial f}{\partial x_1})^2 + (\Delta n_2)^2 \cdot (\frac{\partial f}{\partial x_2})^2 + (\Delta n_3)^2 \cdot (\frac{\partial f}{\partial x_3})^2 + \cdots} \tag{7.8}$$

For correlated quantities this can not be done directly. One possible way to calculate the error in this case is to convert the correlated n_i into uncorrelated m_i and then apply the error propagation to them. The typical example of that is the correction factor c introduced in section 7.1:

$$c = \frac{had}{det} = \frac{det \oplus had + \overline{det} \oplus had}{det \oplus had + det \oplus \overline{had}} \tag{7.9}$$

Correlated quantities appear in this analysis due to two reasons: analysis of events at detector and at hadron level and analysis as events as part of the dijet and the trijet sample. The proper error propagation is done by a program that employs the algorithm outlined in the last two paragraphs without manual intervention.

7.3 Monte Carlo Samples Used

Two Monte Carlo simulations are used for correcting the data to the hadron level: ARIADNE and LEPTO. For both data has been generated using two different prefilters:[1] $Q^2_{obs} > 3\,\mathrm{GeV}^2$ and $Q^2_{obs} > 10\,\mathrm{GeV}^2$. For LEPTO (ARIADNE) these samples correspond to $29.7\,\mathrm{pb}^{-1}/93.2\,\mathrm{pb}^{-1}$ ($12.1\,\mathrm{pb}^{-1}/182.6\,\mathrm{pb}^{-1}$). For each Monte Carlo these samples are combined according to luminosity. Combining according to luminosity means that for each event the weight is chosen to be the inverse luminosity of the sample it could belong to, i.e. a LEPTO event with $Q^2_{obs} = 5\,\mathrm{GeV}^2$ would be assigned the weight $1/29.7\,\mathrm{pb}^{-1}$ (it can only come from the low Q^2_{obs} sample), whereas an event with $Q^2_{obs} = 15\,\mathrm{GeV}^2$ would be assigned the weight $1/122.9\,\mathrm{pb}^{-1}$ (it can come from either the low or the high Q^2_{obs} sample).

7.3.1 Reweighting Monte Carlo Events

The cross section calculation presented above is based on the assumption that the simulation is a perfect description of the processes happening inside the experiment. To validate that a couple of comparisons between data and simulation are made at the detector level. Any discrepancy can have two causes:

1. a bad simulation of the detector. This is best solved by correcting the simulation to be a closer match to the data.

2. a bad simulation of the underlying processes. This could be solved by improving the physical model used. In practice it is more convenient to change the weight for the events in the Monte Carlo sample (reweight the Monte Carlo).

In this analysis reweighing is used to improve the Q^2 dependence of the simulation. To the ratio of data to Monte Carlo as a function of $\log Q$ a line is fit (fig. 7.1). For each Monte Carlo event this fit result is then used to look up a reweighting factor that corresponds to the true Q^2 of that event. The weight of the event is then divided by that factor. The fitted lines are:

$$f = 1.305 - 0.202 \cdot \log_{10}(Q^2/\mathrm{GeV}^2) \qquad \text{(LEPTO)} \qquad (7.10)$$

$$f = 0.365 + 0.478 \cdot \log_{10}(Q^2/\mathrm{GeV}^2) \qquad \text{(ARIADNE)} \qquad (7.11)$$

7.4 Detector Level Comparison

As described in section 7.1 the Monte Carlo will be used to correct the measured rate from the detector to the hadron level. This will only work if the simulation describes the real processes correctly. Any deviation will introduce a systematic shift of the cross section. One way to check how good a job the Monte Carlo does, is to compare detector level distributions for both data and simulation (control plots).

[1] These prefilters are based on Q^2_{obs} which is calculated from the difference of the four momenta of incoming and outgoing lepton. For events without initial or final state radiation it is the same as Q^2.

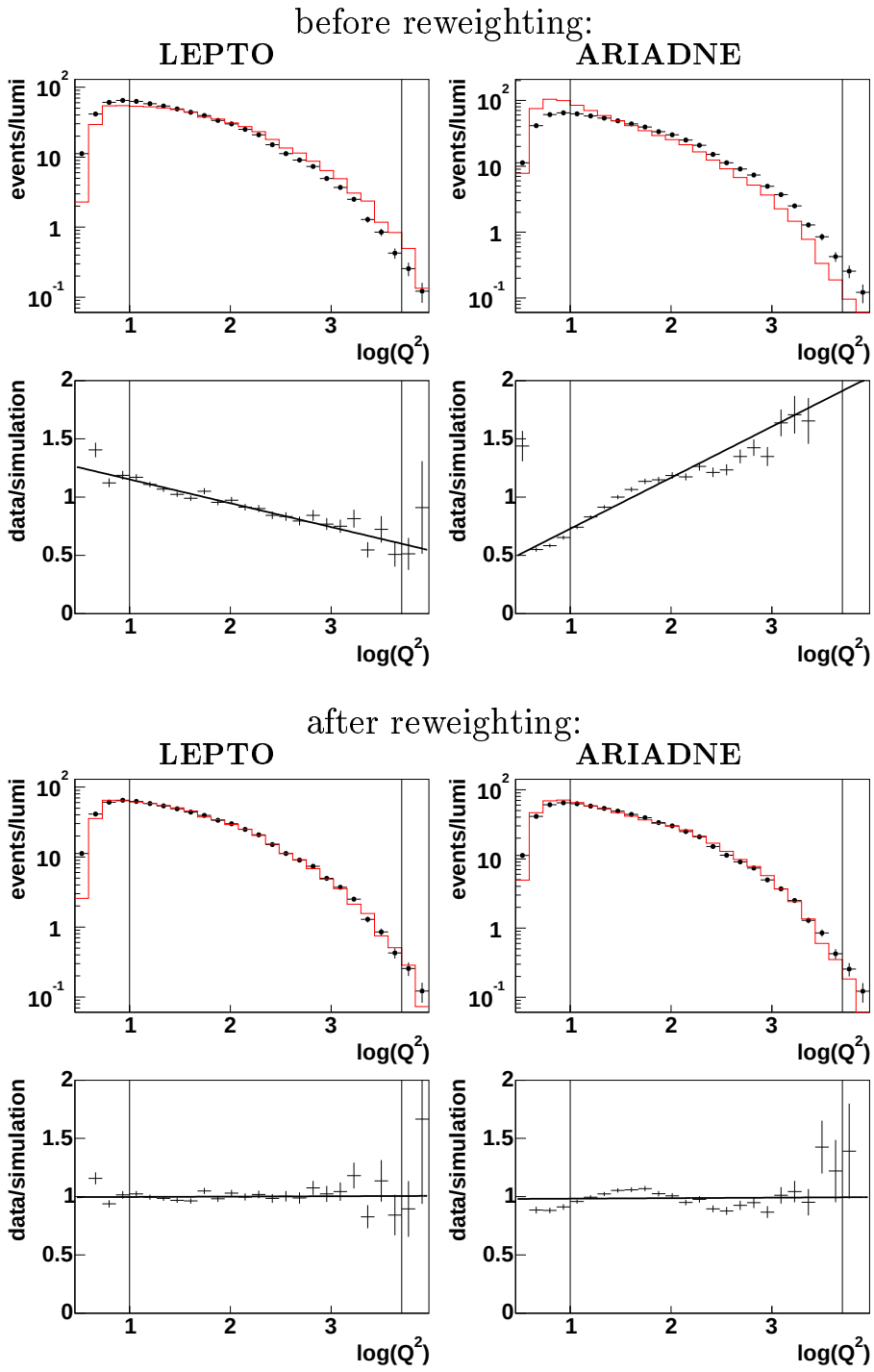

Figure 7.1: *Detector level comparisons for Q^2 without (top) and with reweighting (bottom) for both LEPTO (left) and ARIADNE (right). Shown are both absolute predictions and the ratio data to Monte Carlo.*

For this the Monte Carlo distributions are area normalized[2] to the data. As deviations are specifically interesting at the cut boundaries, the distributions are shown beyond the cut boundaries (with the cut boundaries indicated as a line). The control plots are shown in appendix A. It can be seen that the agreement for the basic kinematic variables are reasonable for both Monte Carlo models. However ARIADNE gives a much worse description of the E_T and m_{jet} distribution than LEPTO does. Therefore LEPTO is used as the default Monte Carlo for unfolding and ARIADNE is only used as a systematic check of the model dependencies.

7.5 Resolutions

The measurement of a variable x results in a measured value x^{det} (detector level) that is different from the true value x^{had} (hadron level) which the variable has in the underlying process. The spread of x^{det} around x^{had} is called the resolution and is one of the basic parameters that determines the quality of the measurement:

- When binning a cross section in x, the assignment of events to bins can be different at hadron and detector level (bin migration). This creates impurities and inefficiences in the measurements. To keep the purities and efficiencies reasonably high one normally makes the bins at least three to four times as wide as the resolution.
- When applying a selection cut on x^{det} this selection will be very pure and efficient for events with x^{had} values far away from the cut. Events with x^{had} values close around the cut will have x^{det} values that are randomly above or below the cut causing the events to be randomly accepted and rejected reducing efficiency and purity. To judge what is far and what is close the resolution is the scale.

To measure the resolution of a variable x one calculates the distribution of the deviation of measured from true values and then fits a Gaussian to that distribution. The width of that Gaussian is then taken as the resolution. Both the absolute $(x^{det} - x^{had})$ and the relative $((x^{det} - x^{had})/x^{had})$ resolution can be measured. Which of the two is better depends on whether the resolution is expected to scale with x or not.

The resolution is a function of the event sample chosen to calculate it. The event sample itself can be chosen rather arbitrarily within certain constraints. The only major constraint on the event sample is that cuts on variables (strongly) correlated with x should only be done on either the hadron or the detector level (otherwise a serious bias is introduced). A minor constraint is that some detector level cuts might be required to make sure that x is defined and well reconstructed at the detector level. Additionally it can prove useful to restrict x to a small range to get the resolution in that range.

For the resolution plots shown here (fig. 7.2), all detector level cuts (see chapter 6) were applied with a few exceptions. If only a certain range in x is interesting that cut is done at the hadron level and all cuts on variables correlated with x are also done at the hadron instead of the detector level

[2] multiplied with a factor so that it has the same integral

to avoid cutting on correlated variables at both hadron and detector level.[3] For kinematic variables all kinematic cuts (see section 6.2) were applied at the hadron instead of the detector level. For jet variables all jet cuts including the invariant mass cut (see section 6.4) were applied at the hadron instead of the detector level. For most variables the relative resolution is shown, except for η^{jet} where the relative resolution isn't applicable (η can be 0 or negative making a division impossible). For the jet variables E_T and η a matching between detector level and hadron level jets is done. This matching requires the jets to be closer than 0.75 in η-ϕ-space and the hadron (detector) level jets to be separated by 1.5 in η-ϕ-space from all other detector (hadron) level jets. After the matching all hadron level jets are removed that don't pass all jet cuts ($E_T > 5\,\text{GeV}$ and $-1 < \eta < 2.5$). For the invariant jet mass it is required that at least two (three) jets are found at detector level that fulfill all jet cuts.

7.6 Acceptance Correction Factors, Purities and Efficiencies

For each hadron level cross section to be calculated a set of acceptance correction factors (called c in section 7.1) has to be calculated. To understand them better it is also useful to look at the purities and efficiencies from which they are derived. This can be seen in fig. 7.3 and 7.4. It can be seen that the efficiencies and purities are generally lower for the trijet than for the dijet sample. This is due to the additional migrations introduced when selecting the third jet.

For low E_T (esp. for dijets) of the leading jet (and with reduced effect for the next to leading jet) the efficiency and purity drops while the correction factors increase. This is most likely due to the jets being forced into extreme configurations to pass the m_{jets} cut. For the most forward η bin the efficiency drops, which is most likely due to this bin lying outside the acceptance of the CTD. For very forward and backward jets the efficiency and purity drops which is also likely due to the jet configuration for these bins.

[3] An alternative approach would be to apply all cuts just at the detector level to avoid removing cuts at the detector level that are needed to clean up the sample and improve the resolution. This has also been tried and yields slightly smaller resolutions. However these results could also be influenced by the shape of the Monte Carlo distribution, which can have an effect similar to a cut at the hadron level.

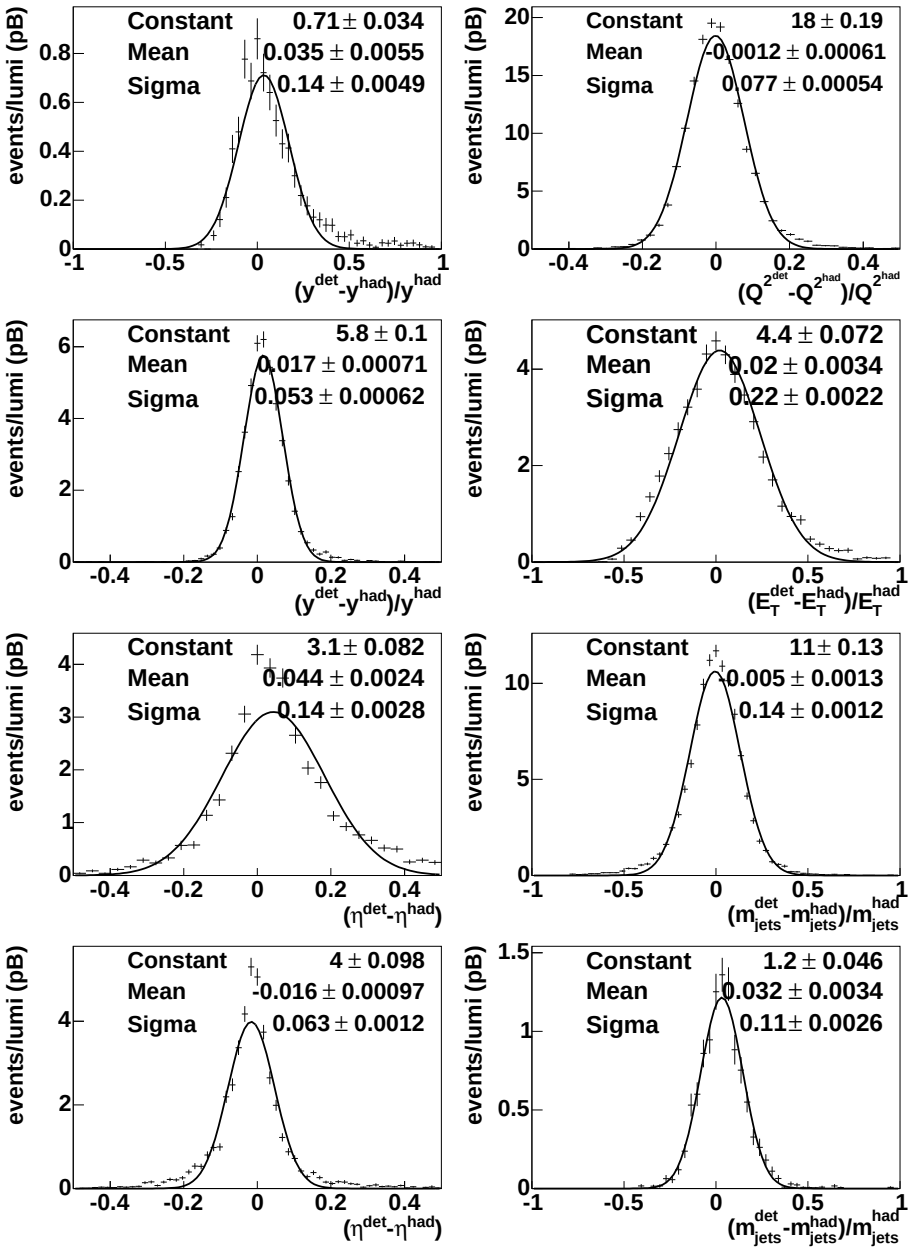

Figure 7.2: *Resolutions for different variables. The resolutions are comparable for dijet and trijet events (except for invariant mass), so only the dijet resolutions are shown. The resolutions are calculated in the regions near the hadron level cut boundaries (left to right, top to bottom):* $y < 0.10$, $Q^2 < 40\ GeV^2$, $y > 0.5$, $E_T < 7\ GeV$, $\eta > 2.3$, $m_{dijet} < 30\ GeV$, $\eta < 0.5$, $m_{trijet} < 30\ GeV$

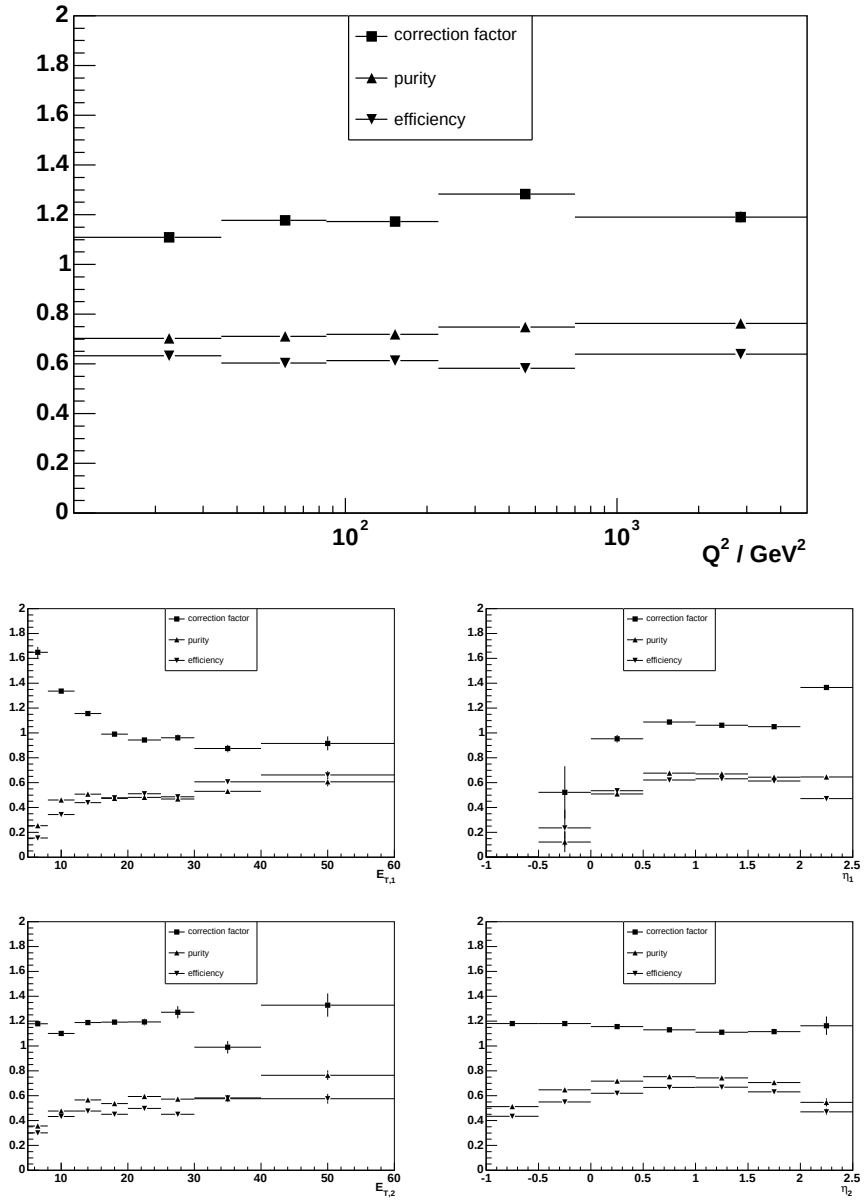

Figure 7.3: *Correction factors for dijet variables. Shown are the correction factors, purities and efficiencies in the same bins as the cross sections are calculated.*

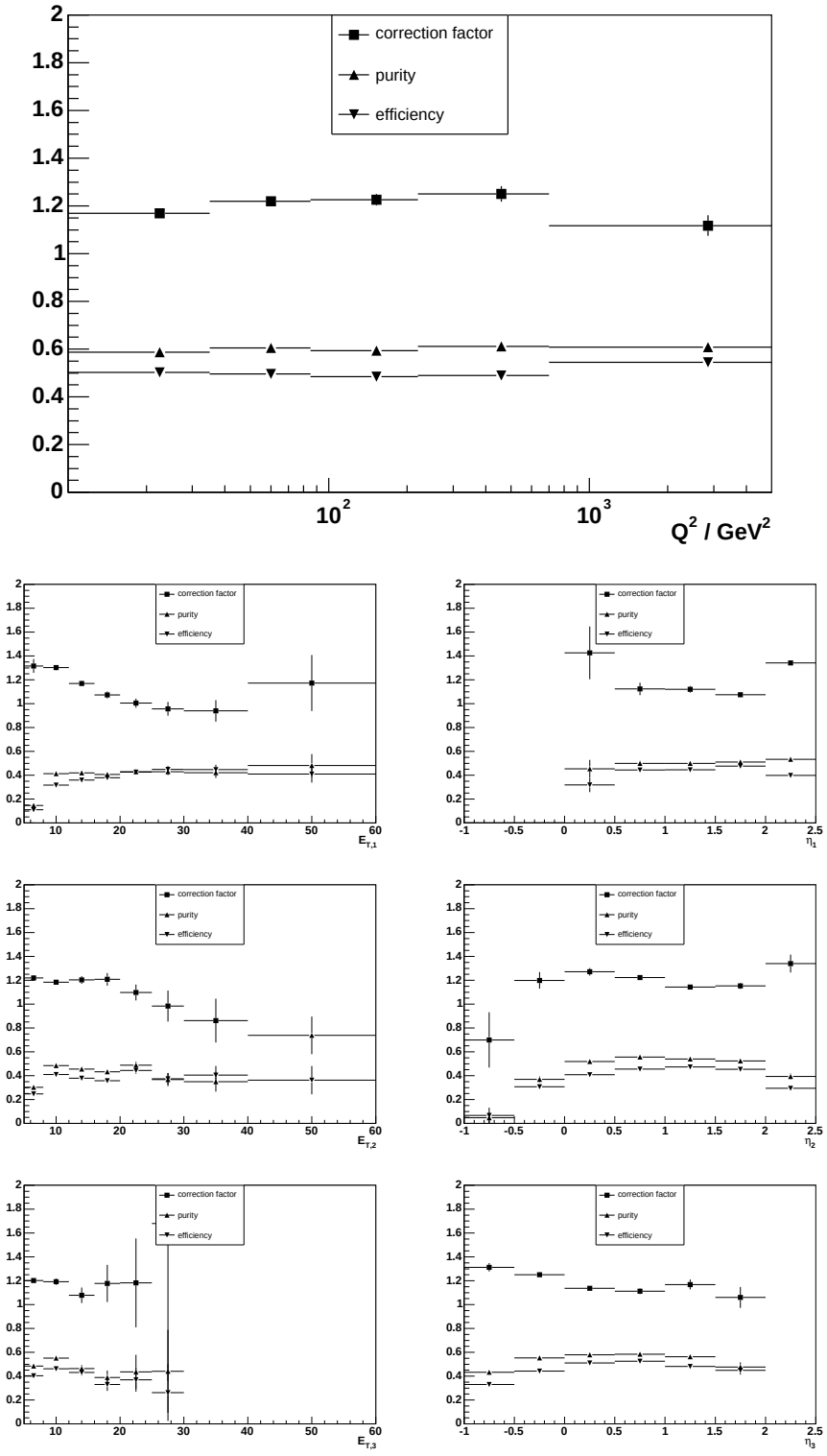

Figure 7.4: *Correction factors for trijet variables (as fig. 7.3).*

Chapter 8

Results

8.1 Systematic Uncertainties

Systematic uncertainties is the general term for all uncertainties in the measurement that come from assumptions made that were not completely true. Once such an assumption has been determined, a test can be devised for estimating the error possibly introduced. The analysis is repeated with slight modifications that correspond to (reasonable) variations of the tested assumption. The difference in result is then taken to be the systematic error. All systematic errors are then added in quadrature to give the total systematic error.[1]

As a Monte Carlo simulation is used for acceptance corrections a mismeasurement of a quantity will not directly lead to a systematic error, but instead systematic errors are introduced if the simulation fails to describe this (mis)measurement or fails to describe the underlying process.

Most systematic errors come from applying cuts on distribution that are not fully reproduced by the Monte Carlo. At each cut there is a number of events migrating over the cut boundary. If the Monte Carlo describes these migrations perfectly there is no systematic error. However for many variables this description is imperfect, introducing a systematic error. This is treated by moving the cut at the detector level up and down (while keeping it the same at the hadron level). Normally the cut is moved up and down by one times the resolution in that quantity. That has been done for the following cuts:

- $38 \mp 1\,\mathrm{GeV} < E - p_z < 60 \pm 1.5\,\mathrm{GeV}$
- $(E - p_z)_e < 54 \pm 1\,\mathrm{GeV}$
- $\eta > -1 \pm 0.1$
- $\eta < 2.5 \pm 0.1$
- $\cos\gamma_h < 0.7 \pm 0.2$
- $y_{jb} > 0.04 \pm 0.0056$

[1] The term total systematic error might be confusing, as this error only contains the systematic errors the author checked, but not necessarily all the systematic errors having a significant effect.

- $y_{el} < 0.6 \pm 0.03$
- $-50 \pm 5\,\text{cm} < z_{vertex} < 50 \pm 5\,\text{cm}$
- $E_T > 5 \pm 1.1\,\text{GeV} \quad m_{jets} > 25 \pm 4\,\text{GeV}$

There are also a number of systematic effects that are of a more general nature and need special tests:

- The uncertainty due to the luminosity measurement is 2.25%. This error comes from the way the luminosity is measured and was not checked in this analysis. This error is not included in the cross sections.
- To estimate the bias from the underlying physics model, the analysis has been repeated once using Ariadne in place of Lepto for acceptance corrections.
- To estimate the bias introduced by reweighing the Monte Carlo to match the data, the analysis has been repeated without reweighting the Monte Carlo.
- To estimate the uncertainties coming from using the double angle method, the analysis has once been repeated using only the electron method. A corresponding test of using only the double angle method to estimate the uncertainty from using the electron method has not been done, as the double angle method is known to behave very poorly at low Q^2.
- The absolute scale of the energy measurement is only known within 1–3%. To estimate that error the jet energy in the data (but not the Monte Carlo) is varied up and down by that uncertainty.

The effects of each systematic on the measured cross section can be seen in appendix B. Depending on the cross section studied, the biggest systematic errors are due to:

- the energy scale uncertainty
- the cut on the E_T and invariant mass
- the model dependency
- the forward η cut

8.2 Theoretical Uncertainties

Not only the data but also the theoretical prediction has some uncertainty associated with it. These can be divided into two groups:

1. Uncertainties due to parameters fit to the data. This are the uncertainties coming from the value of α_s and from the PDF used. These uncertainties can be directly evaluated by varying the parameters by the uncertainty of the fit. The evaluation of this uncertainty is still ongoing and will be available in [3].

2. Uncertainties due to the way the calculations are carried out. This is the renormalization and factorization scale uncertainty. For these uncertainties only a rough estimate can be given.

The main theoretical uncertainty in this analysis is the renormalization scale uncertainty. If all orders in α_s were included in the calculation the cross section prediction should not depend on the renormalization scale. Therefore a dependence on μ_r can be taken as an indication of the size of the error coming from missing higher order effects. It is convention to estimate this error by varying the renormalization and the factorization scale at the same time up or down by a factor of 2.

8.3 Presentation of the Cross Section

The cross sections given in this chapter are given as differential cross sections. For that the cross sections have been divided up into bins, the cross sections are normalized to bin width and plotted at the bin center. *No* bin-center corrections have been applied to avoid biases in the cross section ratio. This procedure is consistent as it has been done the same way for data and theory predictions, but it can give a wrong impression of the shape of the differential cross section. The bin borders are given in table 8.1.

E_T	(GeV)	5	8	12	16	20	25	30	40	60
η		-1	-0.5	0	0.5	1	1.5	2	2.5	
Q^2	(GeV^2)	10	35	85	220	700	5000			

Table 8.1: *The boundaries for the cross section bins for the differential variables.*

The data are plotted as points with inner and outer error bars. The inner error bars indicate the statistical error (including the statistical error from the Monte Carlo) and the outer error bars the combined statistical and systematic errors (except for the calorimeter energy scale). The error coming from the calorimeter energy scale is shown in form of a yellow band connecting all data points (as it is correlated over all bins and therefore effects more the normalization and less the shape of the cross section). For presentation reasons some cross sections are scaled by a power of 10. This is indicated next to the cross section.

On the plots for the absolute cross section the theory prediction is given as a single line. For all cross sections also the ratio of data to theory is shown, allowing for easy comparison. On these plots the renormalization scale uncertainty is shown in form of a hatched area.

8.4 Measurement of the Cross Section in E_T

The jet finding is done in terms of the variables E_T, η and ϕ, making it sensible to look at the cross section as a function of the transverse energy E_T of the jets. The transverse energy is preferred over the energy as it is insensitive to a longitudinal boost. These longitudinal boosts occur due to the struck quark carrying different momentum fractions ξ from event to event (note: ξ is

different from the Bjorken scaling variable x, so these boosts also exist in the Breit Frame).

As mentioned in section 1.3.2 one of the advantages of using the Breit frame is that a pure one jet event would end up with no transverse energy. Therefore any presence of transverse energy is a sign of higher order effects. As soft partonic radiation occurs at low transverse energies a high transverse energy also indicates that a jet is originating from a hard process.

When studying multijet events this interpretation is a little problematic: If one of the jets emits soft partonic radiation it will have low transverse momentum with respect to the jet, but comparatively high transverse energy with respect to the z-axis. This is however not as bad as it sounds as:

1. Soft radiation tends to be collinear with the jet axis. Only sufficiently hard radiation will be separated enough to form its own jet.

2. The transverse energy is limited by the jet energy which in turn is limited by the hardness of the process. Thereby the hardness of the jet still has an influence on the transverse energy.

The measured cross section has been compared to the NLO predictions for both dijets (Fig. 8.2) and trijets (Fig. 8.3) with the jets ordered in E_T and only the two (three) jets with the highest E_T being considered. In general the NLO describes the data within errors. At low E_T of the first jet, (esp. for dijet events) there is some excess of the data over the Monte Carlo. This might be due to the fact, that the jets have to be in an extreme configuration to generate a high invariant mass at such a low E_T for all jets. It can also be doubted that the theoretical uncertainty at that point is as small as indicated by the size of the renormalisation scale uncertainty.

8.5 Measurement of the Cross Section in η

The pseudorapidity η is another fundamental jet variable. Although η is not insensitive to longitudinal boosts, differences in η are invariant under longitudinal boosts. *eta* corresponds to the polar angle and can therefore be associated with a position in the detector. A jet that is more backwards (small η) is more likely to come from the hard interaction, whereas a forward jet (big η) can also be influenced by the interplay of struck parton and remnant. Studying the jet production as a function of η is a test if the processes in jet production are well described.

The measured cross section has been compared to the NLO predictions for both dijets (Fig. 8.4) and trijets (Fig. 8.5) with the jets ordered in η and only the two (three) jets with the highest E_T being considered. This is done to avoid ambiguities in dijet events were the jets are balanced in E_T. In general the NLO describes the data within errors. The scale uncertainty for the leading jet is likely to underestimate the theoretical uncertainty. This decrease is due to different shapes of the NLO predictions for different scales. The point where it gets minimal is the crossing point for the NLO predictions.

8.6 Measurement of the Cross Section in Q^2

The general DIS cross section is often expressed as a function of the Lorentz invariant variables x and Q^2 (see section 1.2.2). For multijet events x loses its direct meaning as the momentum fraction carried by the struck parton. Q^2 denotes the momentum transfer (squared) from the electron to the parton. It is a measure of the hardness of the event and governs the dynamics of the outgoing partons.

The measured cross section has been compared to the NLO predictions for both dijets and trijets (Fig. 8.6). In general the NLO describes the data well within errors showing that the NLO is capable of describing the data over a wide range of scales.

8.7 Measurement of the Cross Section Ratio

Due to the good agreement between the measured and predicted cross section as a function of Q^2 a more stringent test has been devised. For this the ratio of the cross sections for trijets to dijets has been calculated. In the cross section ratio both experimental and theoretical errors cancel out resulting in a more precise cross section.[2]

The measured cross section has been compared to the NLO predictions for both dijets and trijets (Fig. 8.1 a). Over the complete Q^2 range the NLO gives a good description of the data. Also the shape is reasonably well described.

8.8 Measurement of $\alpha_s(M_Z)$

The trijet to dijet cross section ratio is an observable that is expected to scale with α_s. As the NLO gives a good description of the measurement this observable has been used to extract a value of $\alpha_s(M_Z)$ for each bin (fig. 8.1 b). From the values for the individual bins an average $\alpha_s(M_Z)$ for the complete measurement has been calculated.

To extract the $\alpha_s(M_Z)$ value for one Q^2 bin the theoretical predictions by NLOJET have been calculated for five different values of $\alpha_s(M_Z)$ (using the CTEQ4 PDFs that have been extracted using these five different values). The theoretical prediction is then available as function of $\alpha_s(M_Z)$ (see fig. 8.7) and parameterized by fitting a line to the points. To make use of the more recent CTEQ6 PDF this line is then scaled up or down to match the prediction using CTEQ6. This line is then used to map the measured cross section to a measured value of $\alpha_s(M_Z)$ for that bin in Q^2.

The error weighted mean over all bins is taken as the extracted value of $\alpha_s(M_Z)$. To estimate the systematic and theoretical uncertainties this process has been repeated for all systematic and theoretical deviations of the process (fig. 8.8). For the calculation of the final error the renormalization scale uncertainty has been symmetrized. This yields the following value:

$$\alpha_s = 0.1203 \pm 0.0010(stat)^{+0.0029}_{-0.0050}(sys) \pm 0.0047(theo) \qquad (8.1)$$

This value is compatible with other measurements and the world average (fig. 8.9).

[2] For the calculation of systematic and theoretical uncertainties of the ratio, the systematic deviations of the ratio have been used, instead of propagating the errors from the cross section.

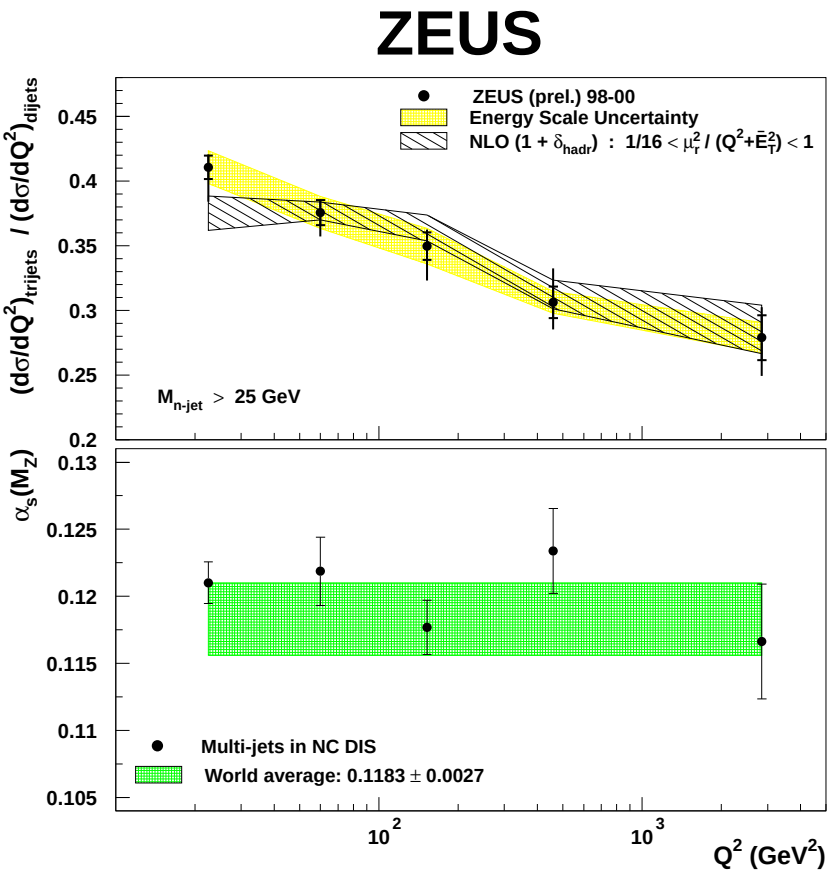

Figure 8.1: *a) The cross section ratio of trijets to dijets as a function of Q^2. It is compared to the theory prediction with the band indicating the renormalisation scale uncertainty. b) For each bin a value of $\alpha_s(M_Z)$ has been extracted and is compared to the world average.*

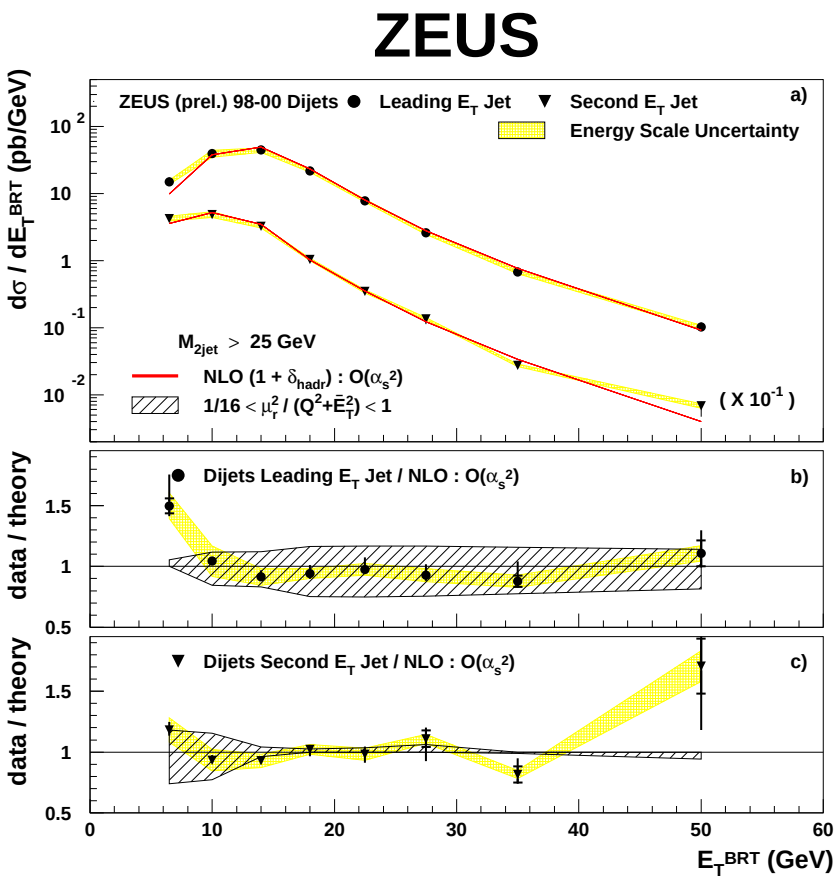

Figure 8.2: *The measured dijet cross section as a function of E_T of the first/second jet (ordered in E_T). Subplot a) shows the absolute cross section. The decrease in cross section at low E_T is a result of the invariant mass cut that has been applied to depopulate that region. Subplot b) and c) show the relative difference of measured and predicted cross section.*

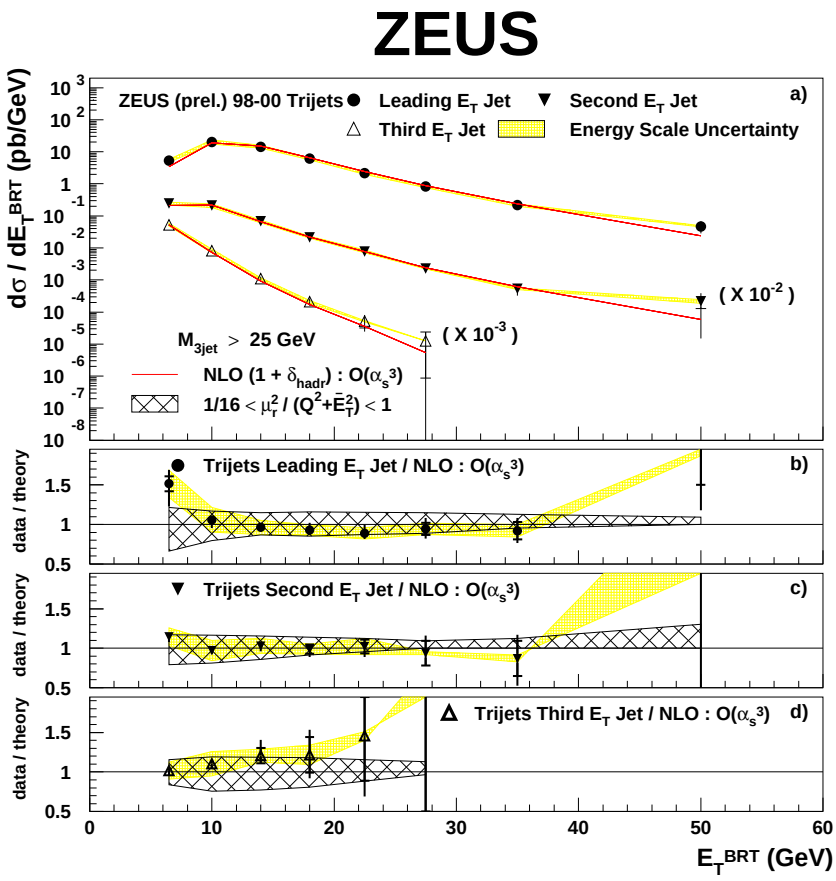

Figure 8.3: *The measured trijet cross section as a function of E_T of the first/second/third jet (ordered in E_T). Subplot a) shows the absolute cross section. Subplot b), c) and d) show the relative difference of measured and predicted cross section.*

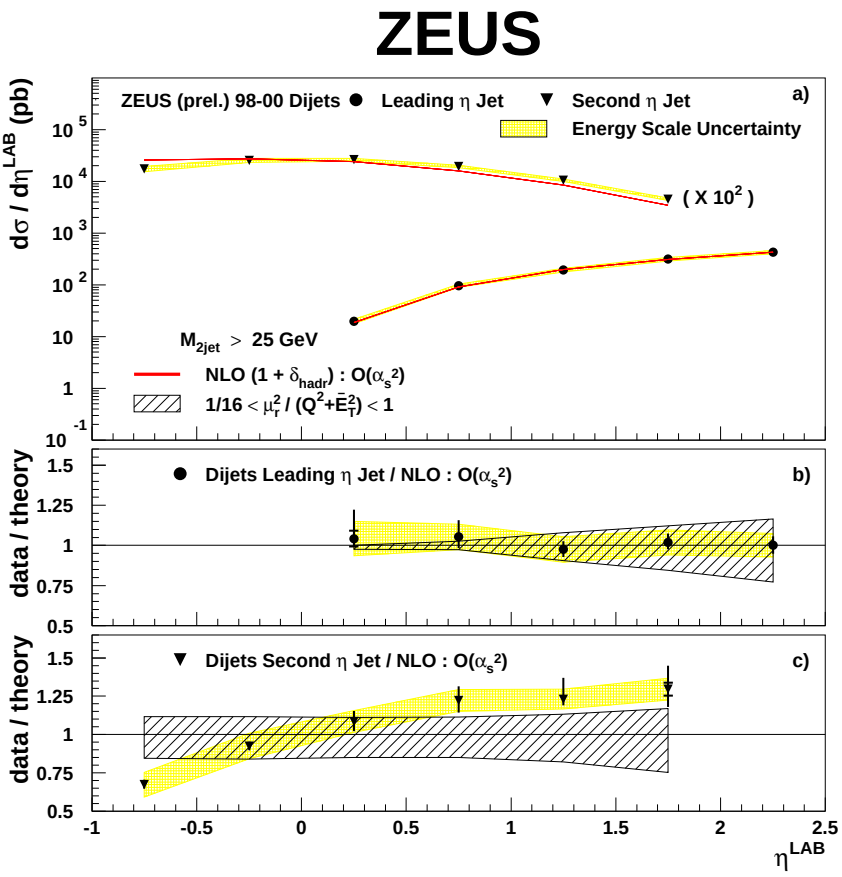

Figure 8.4: *The measured dijet cross section as a function of η of the first/second jet (ordered in η). Subplot a) shows the absolute cross section. Subplot b) and c) show the relative difference of measured and predicted cross section.*

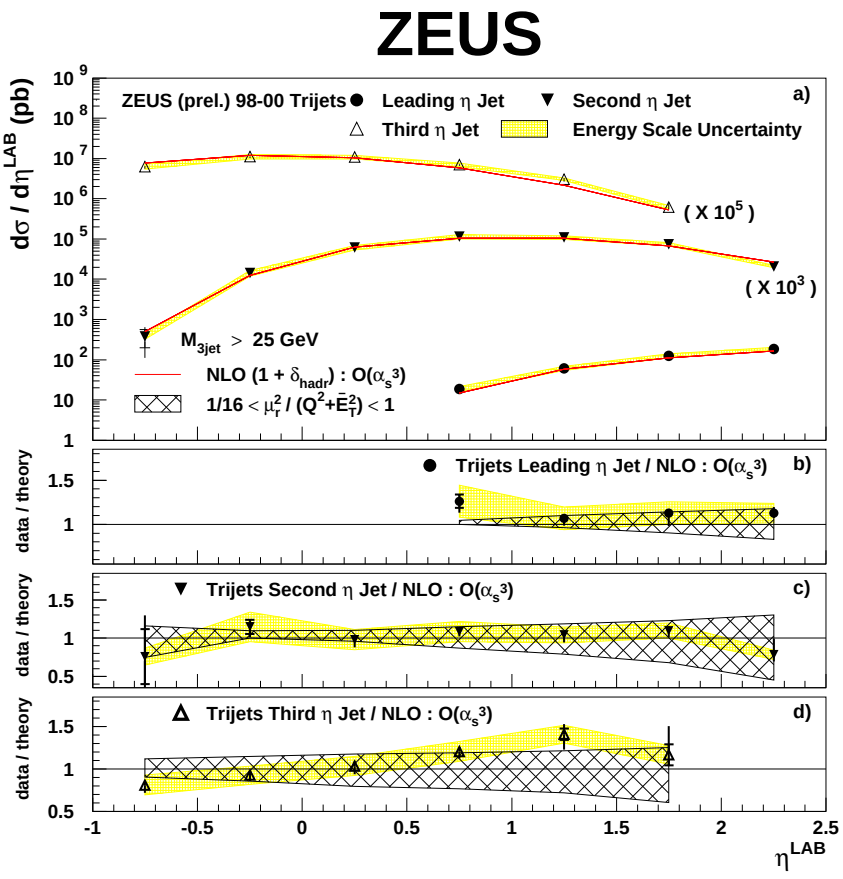

Figure 8.5: *The measured trijet cross section as a function of η of the first/second/third jet (ordered in η). Subplot a) shows the absolute cross section. Subplot b), c) and d) show the relative difference of measured and predicted cross section.*

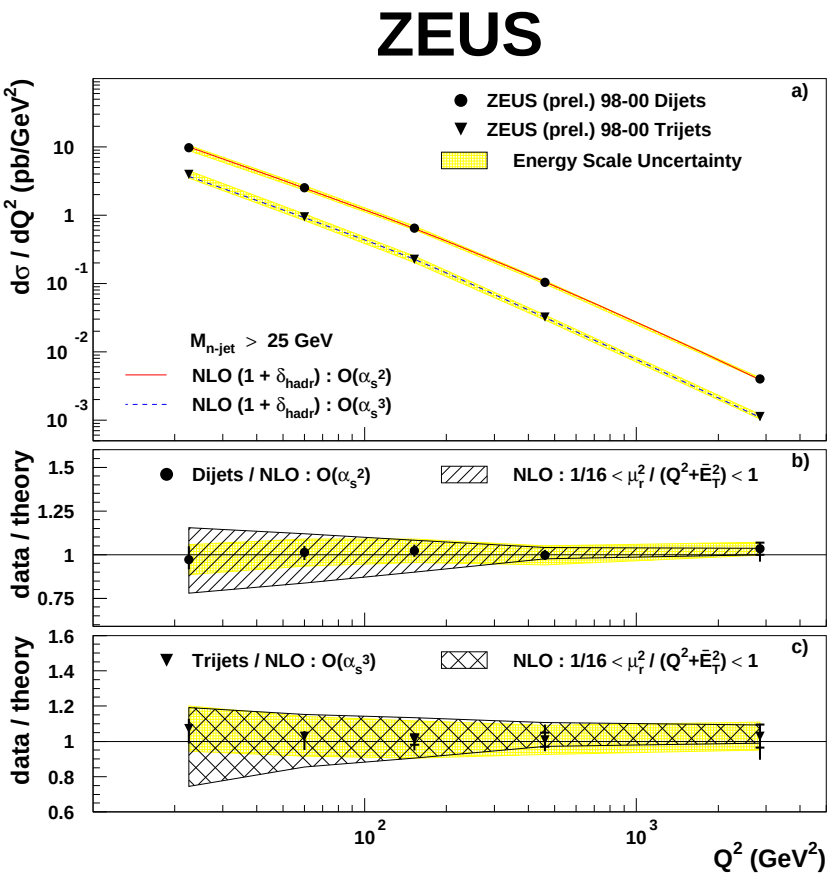

Figure 8.6: *The measured di- and trijet cross section as a function of Q^2. Subplot a) shows the absolute cross section. Subplot b) and c) show the relative difference of measured and predicted cross section.*

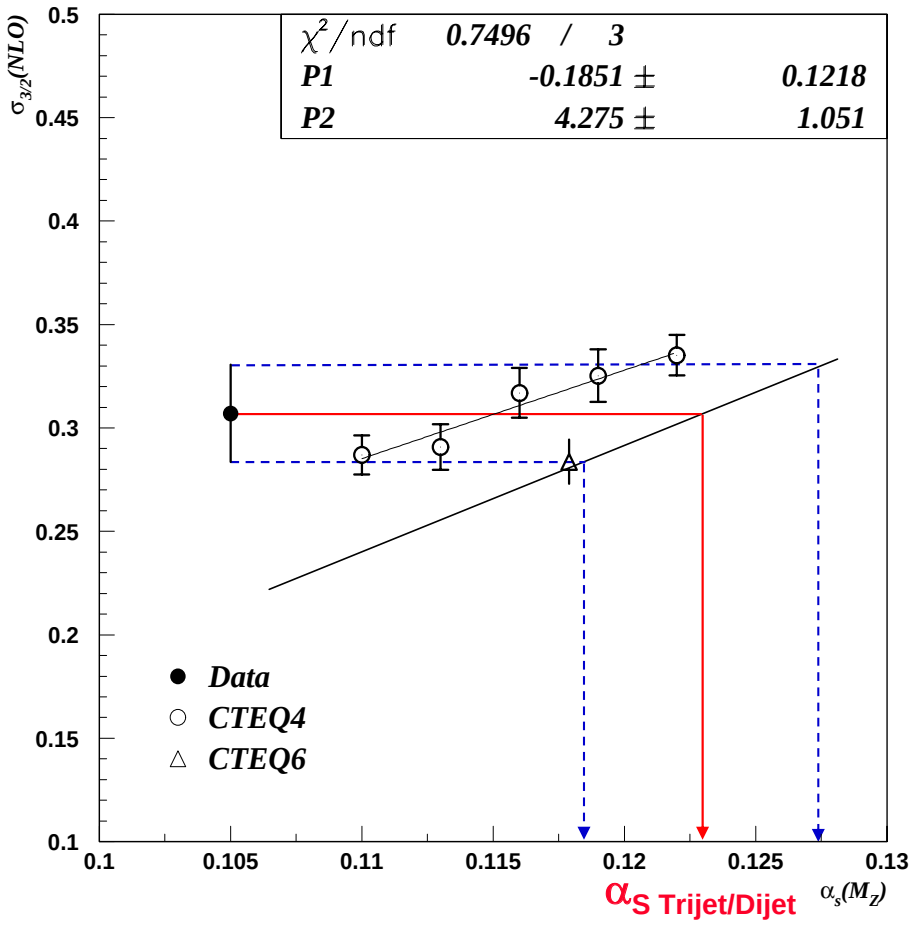

Figure 8.7: *The α_s dependence of the cross section ratio for an individual bin. The points are the CTEQ4 prediction. The dependence is parameterized by fitting a line to it. The line is then scaled up or down to match the prediction using CTEQ6. This line is then used to map the cross section measurement in that bin to an α_s measurement in that bin.*

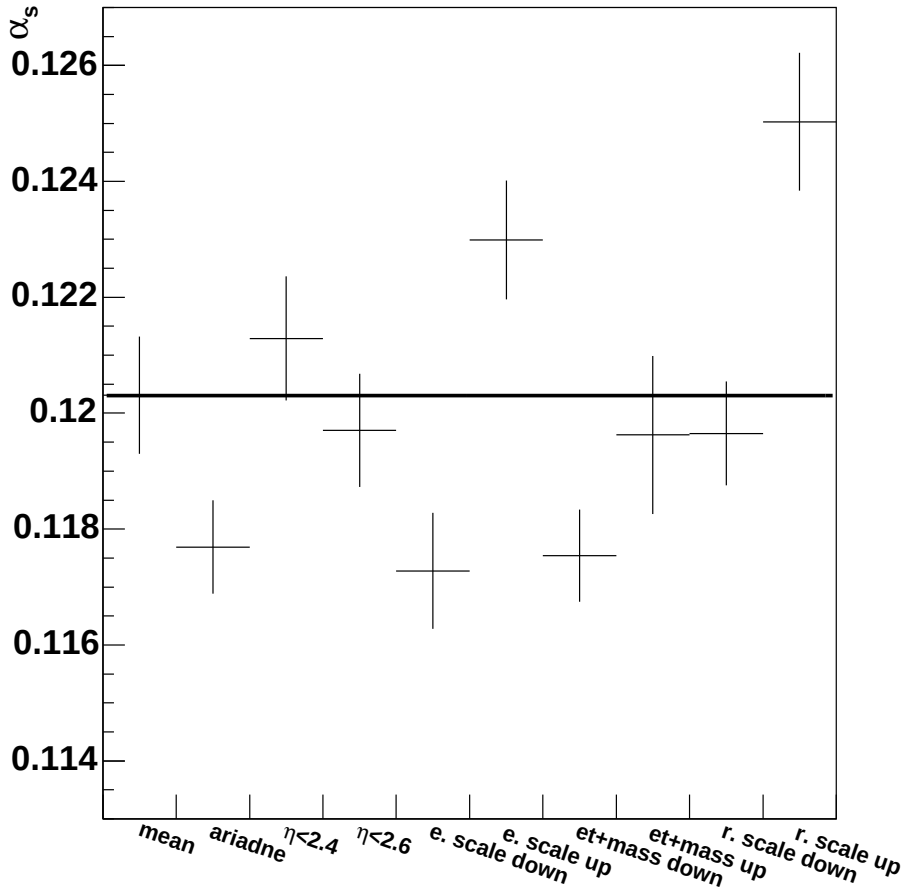

Figure 8.8: *The systematic and theoretical uncertainties for the extracted value of α_s. The horizontal line indicates the exacted value of α_s. For the calculation of the final error the renormalization scale uncertainty has been symmetrized.*

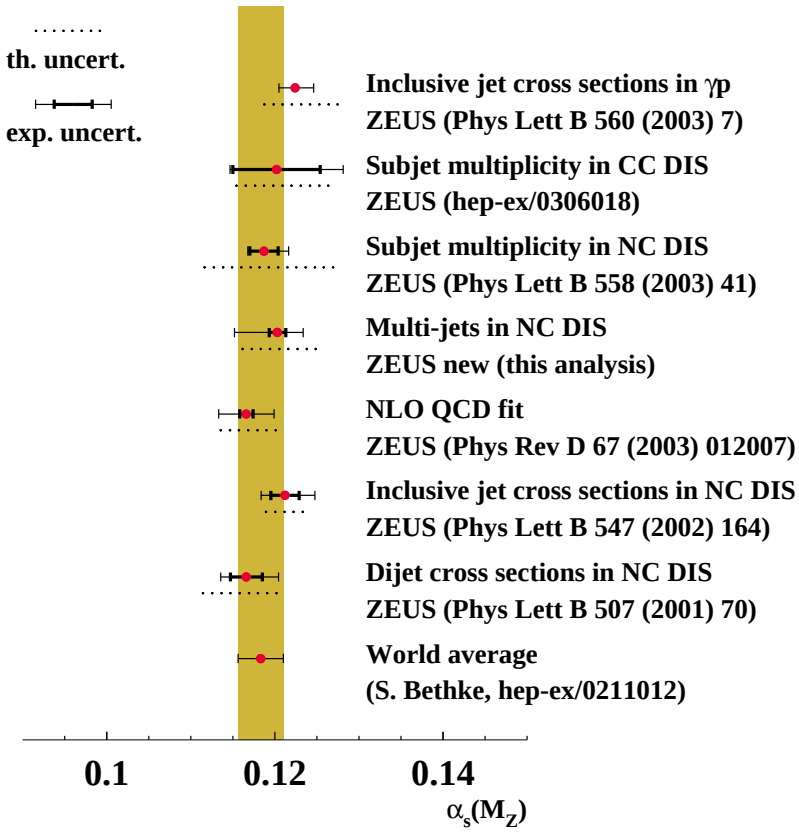

Figure 8.9: *The measured α_s value from this analysis compared to other measurements and the world average.*

Chapter 9

Conclusions

Measurements of differential cross sections in the variables E_T^{jet}, η^{jet} and Q^2 have been presented for a neutral current, multijet DIS event sample of $82.2\,\mathrm{pb}^{-1}$ of ZEUS data taken in the 98–00 running. These have been compared to NLO predictions generated using the NLOJET program. The NLO has been found to describe the data within error. As a more precise test the cross section ratio of trijet to dijet events has been calculated in Q^2. Again an agreement between data and theory prediction was shown. From this cross section ratio a value of α_s was extracted using the α_s dependence as predicted by the CTEQ4 PDF. The value measured is in agreement with the world average:

$$\alpha_s = 0.1203 \pm 0.0010(stat)^{+0.0029}_{-0.0050}(sys) \pm 0.0047(theo) \tag{9.1}$$

The biggest error for this measurement comes from the renormalization scale uncertainty, which can probably be reduced, when NNLO calculations are available. On the experimental side the biggest error is the energy scale uncertainty, which could be reduced by understanding the energy scale of the calorimeter better. Another big error source is the incorrect description of the E_T (and invariant mass) spectrum of the jets, an improved (or better tuned) Monte Carlo model could reduce this effect. An alternative solution to these two problems could be presented by the HERA II program which aims at collecting a high luminosity sample of DIS data. This sample would allow to measure at higher E_T (and Q^2), which would lead to a lower scale uncertainty in both theory and experiment.

Chapter 10

Acknowledments

I would like to thank my advisor Klaus Wick for giving me the opportunity to work on this thesis. Thanks goes also to Wesley Smith for giving me additional support and to Juan Terron for having the answers when I didn't have them.

Biggest thanks go to Liang Li for doing the second analysis and debugging the analysis during countless nights. Special thanks go to Klaus Wick, Wesley Smith and Sabine Lammers for reading the complete thesis draft and submitting their comments in time. I also want to thank Sabine for many interesting and usefull discussions. In general the ZEUS Hamburg I group, the ZEUS Wisconsin group and all my other friends have made my past six years in Hamburg a very pleasant experience.

Appendix A

Control Plots

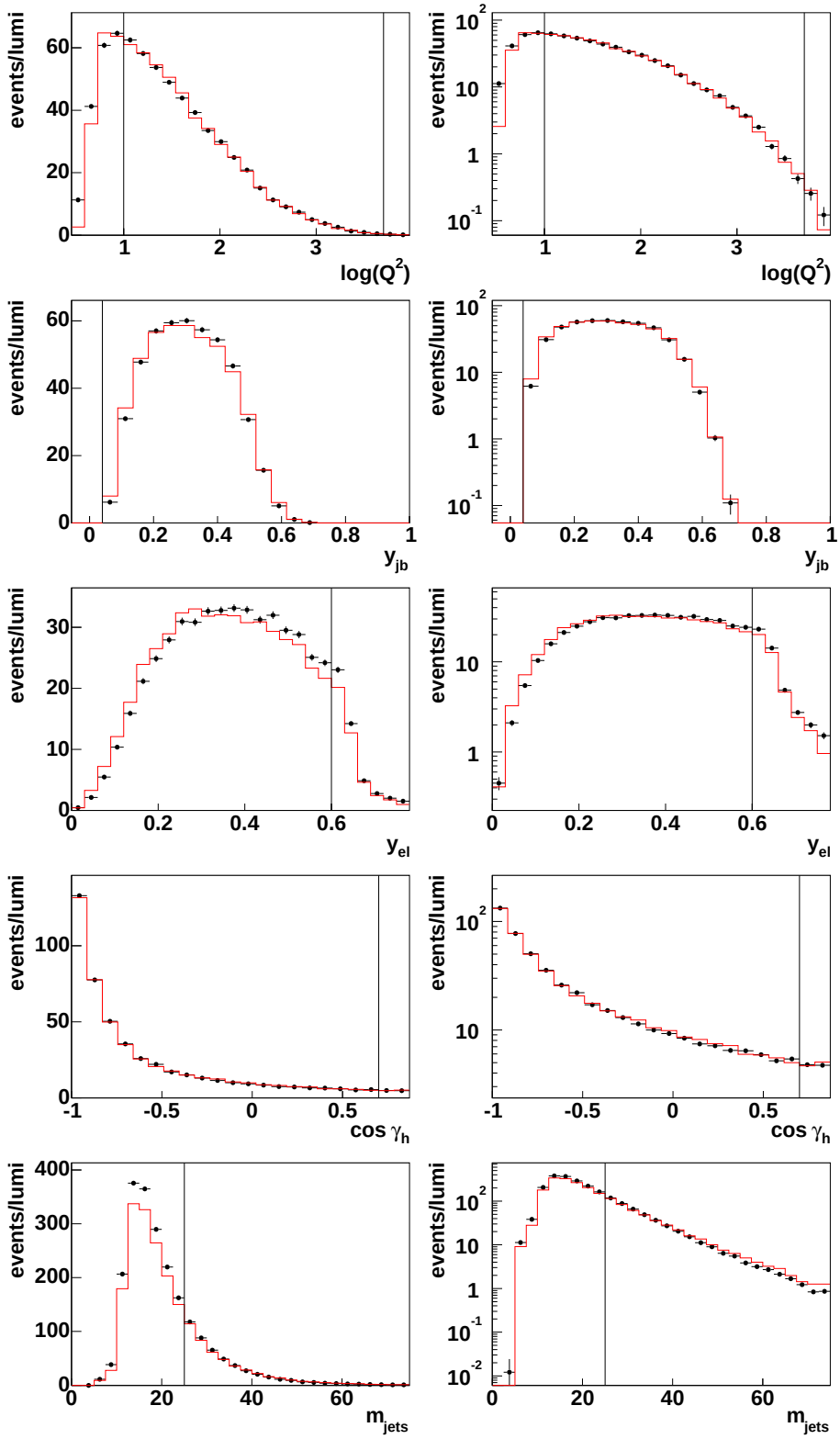

Figure A.1: *Detector level comparisons for the kinematic variables. The data are compared to LEPTO predictions. The Monte Carlo predictions are area normalized to the data wihtin the cut boundaries. Vertical lines indicate cuts applied to the sample. The distributions on the left are on linear scale and on the right on logarithmic scale.*

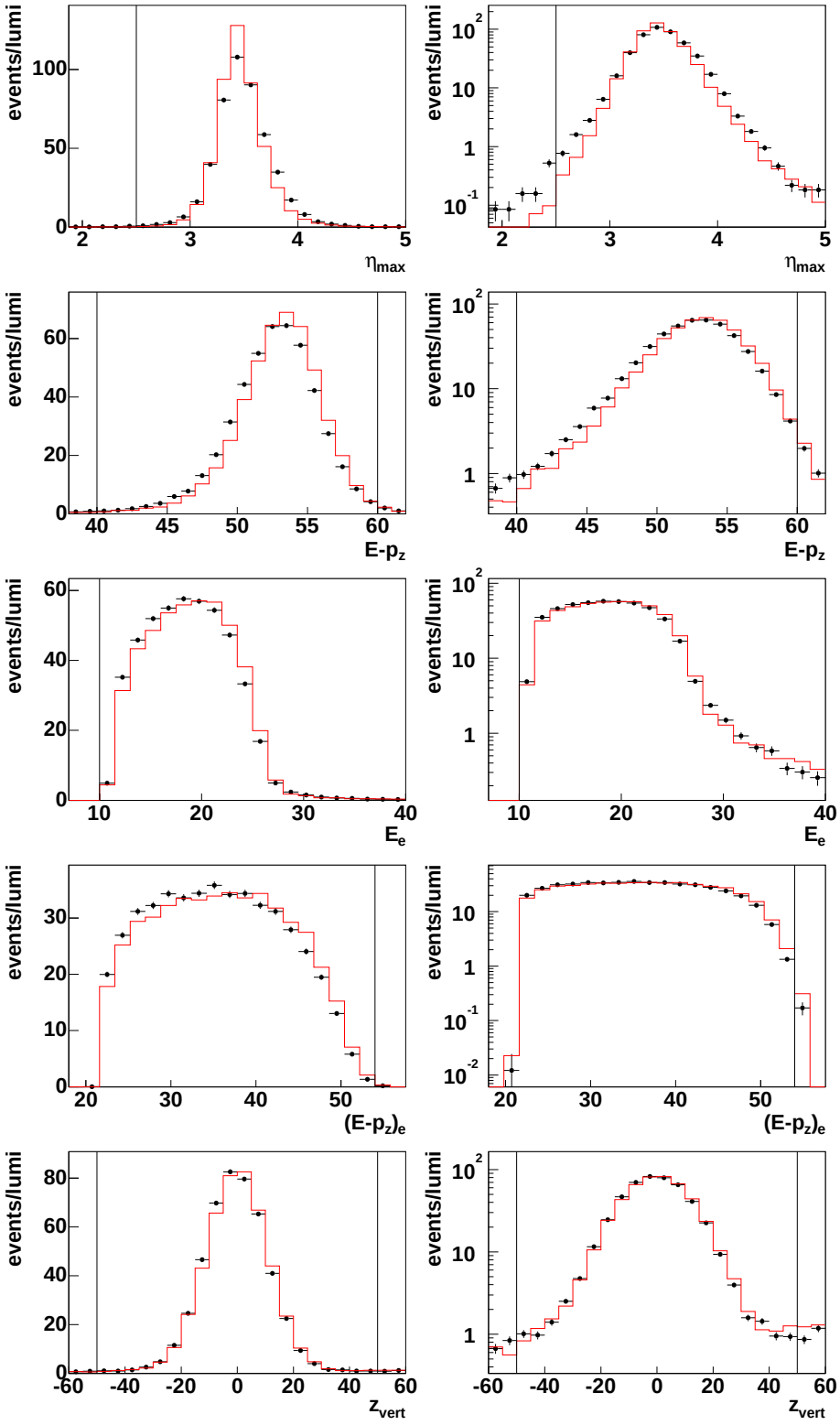

Figure A.2: *As fig. A.1 but for other event variables.*

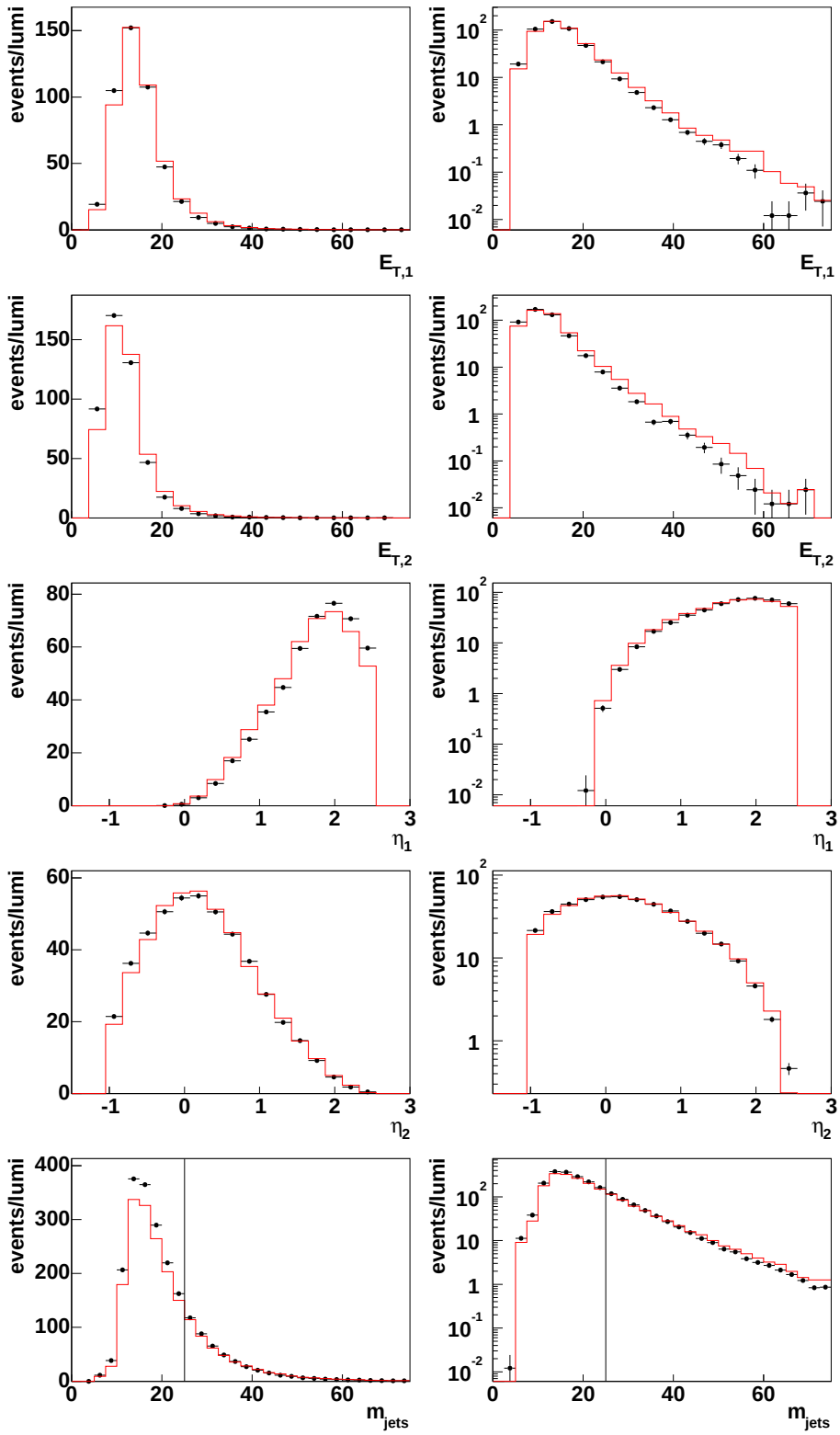

Figure A.3: *As fig. A.1 but for the jet variables.*

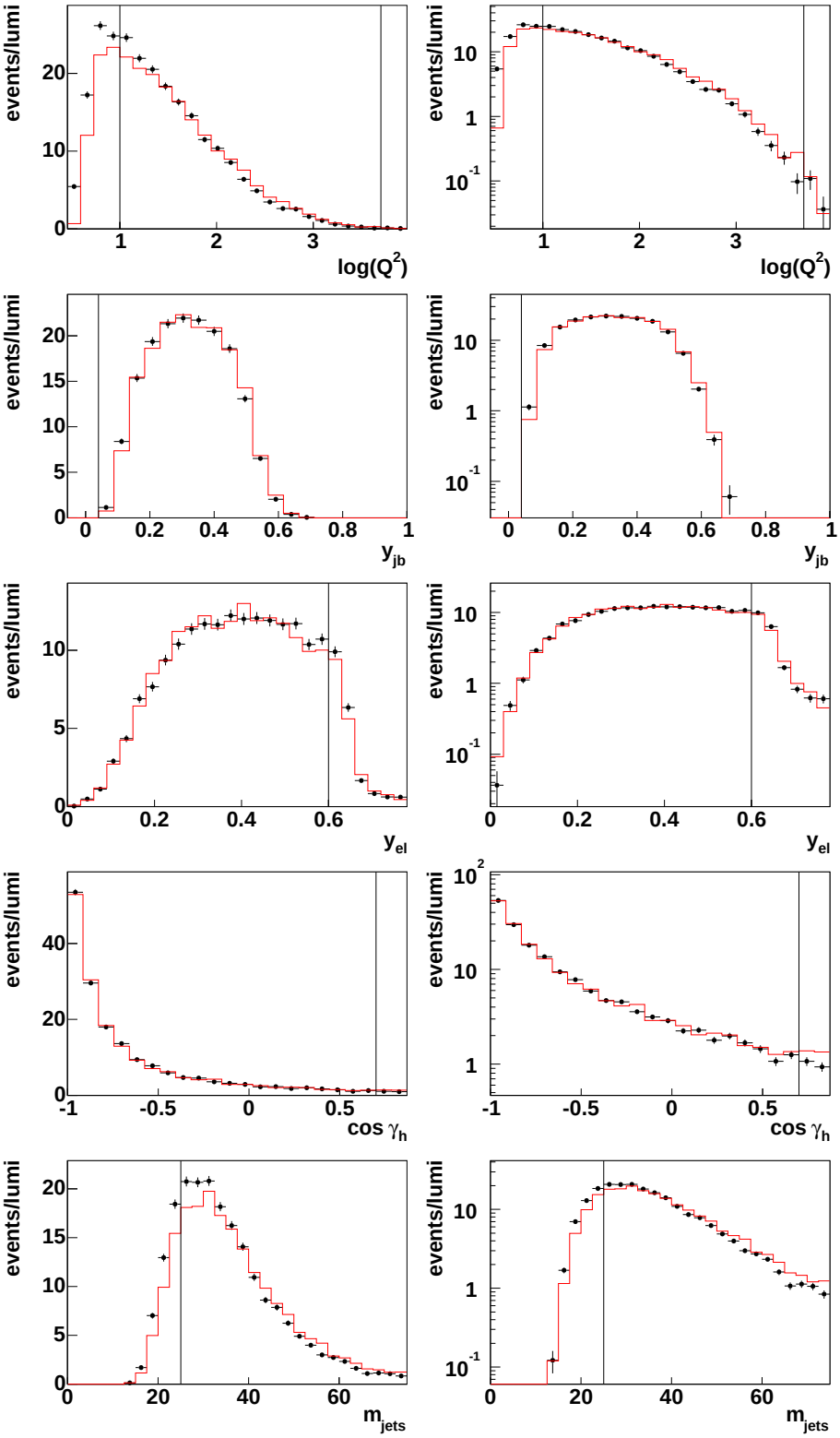

Figure A.4: *As fig. A.1 but for trijet events.*

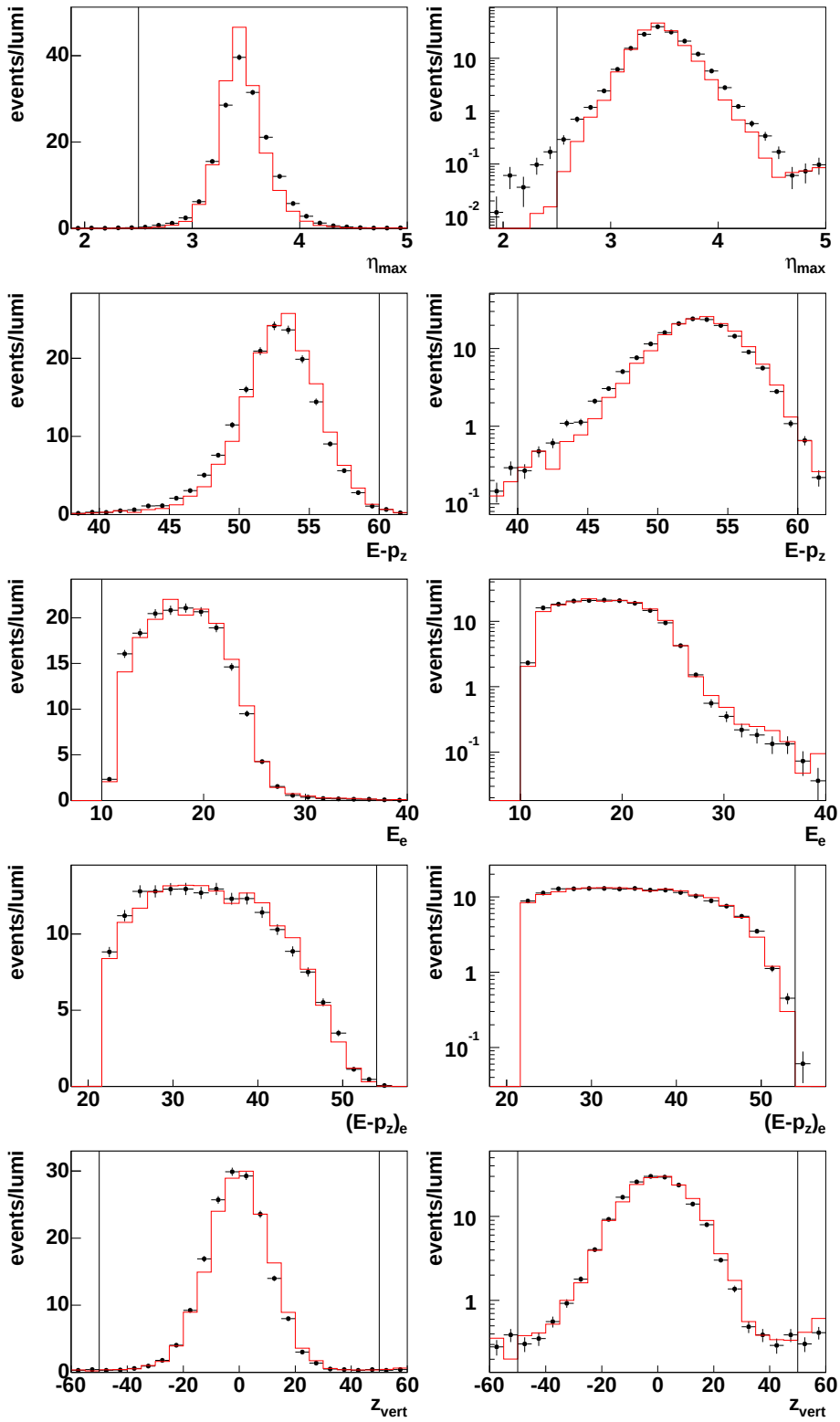

Figure A.5: *As fig. A.2 but for trijet events.*

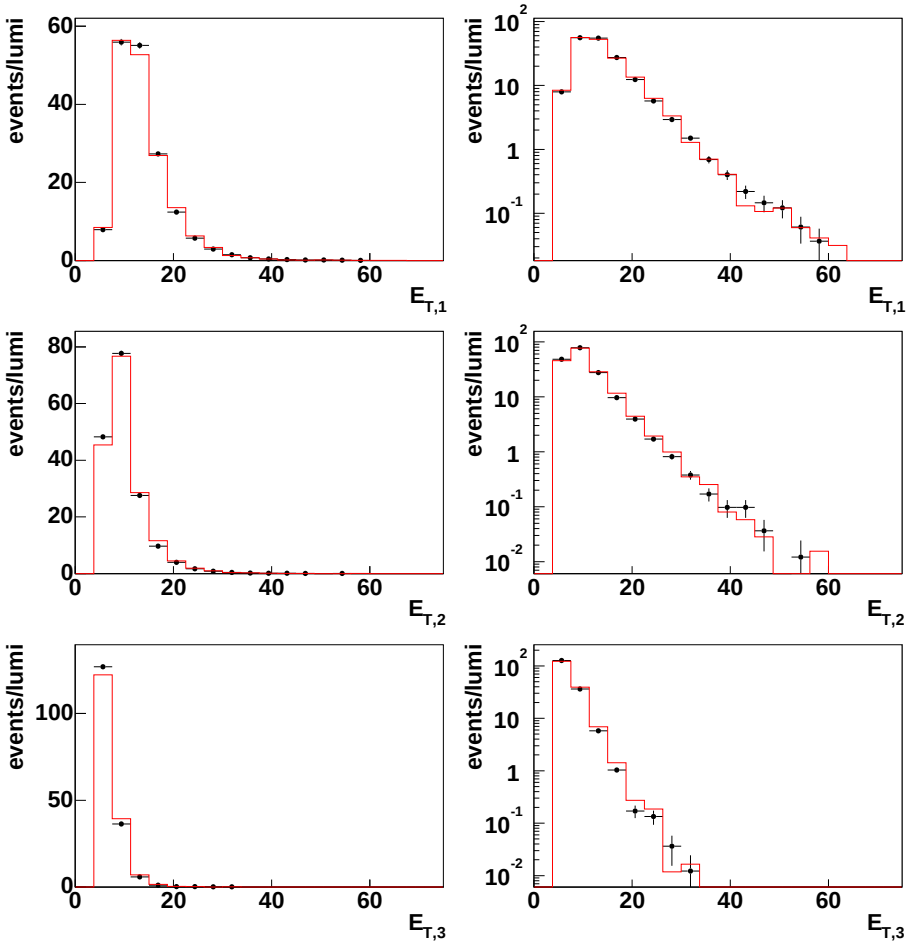

Figure A.6: *As fig. A.3 but for trijet events. Continued in fig. A.7*

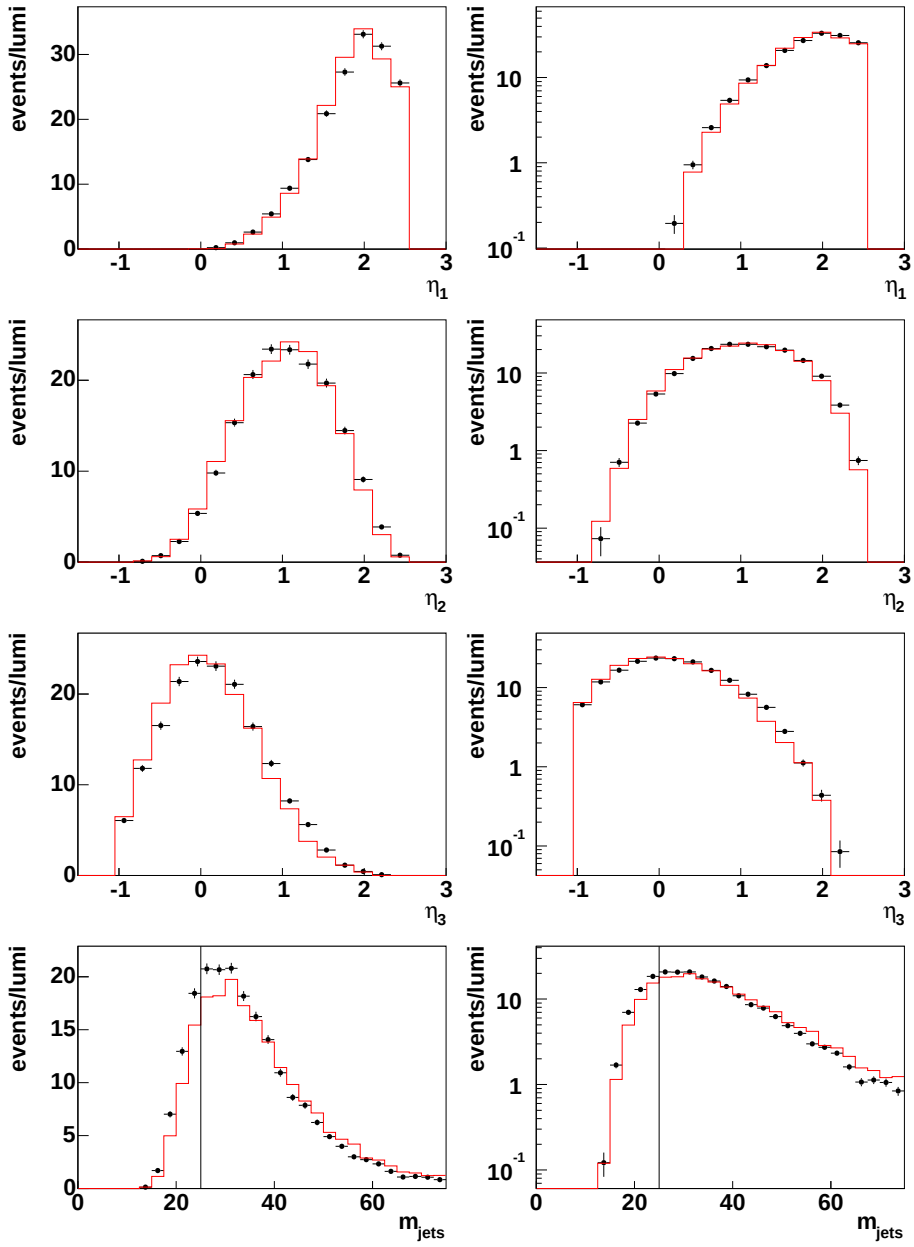

Figure A.7: *Continued from fig. A.6.*

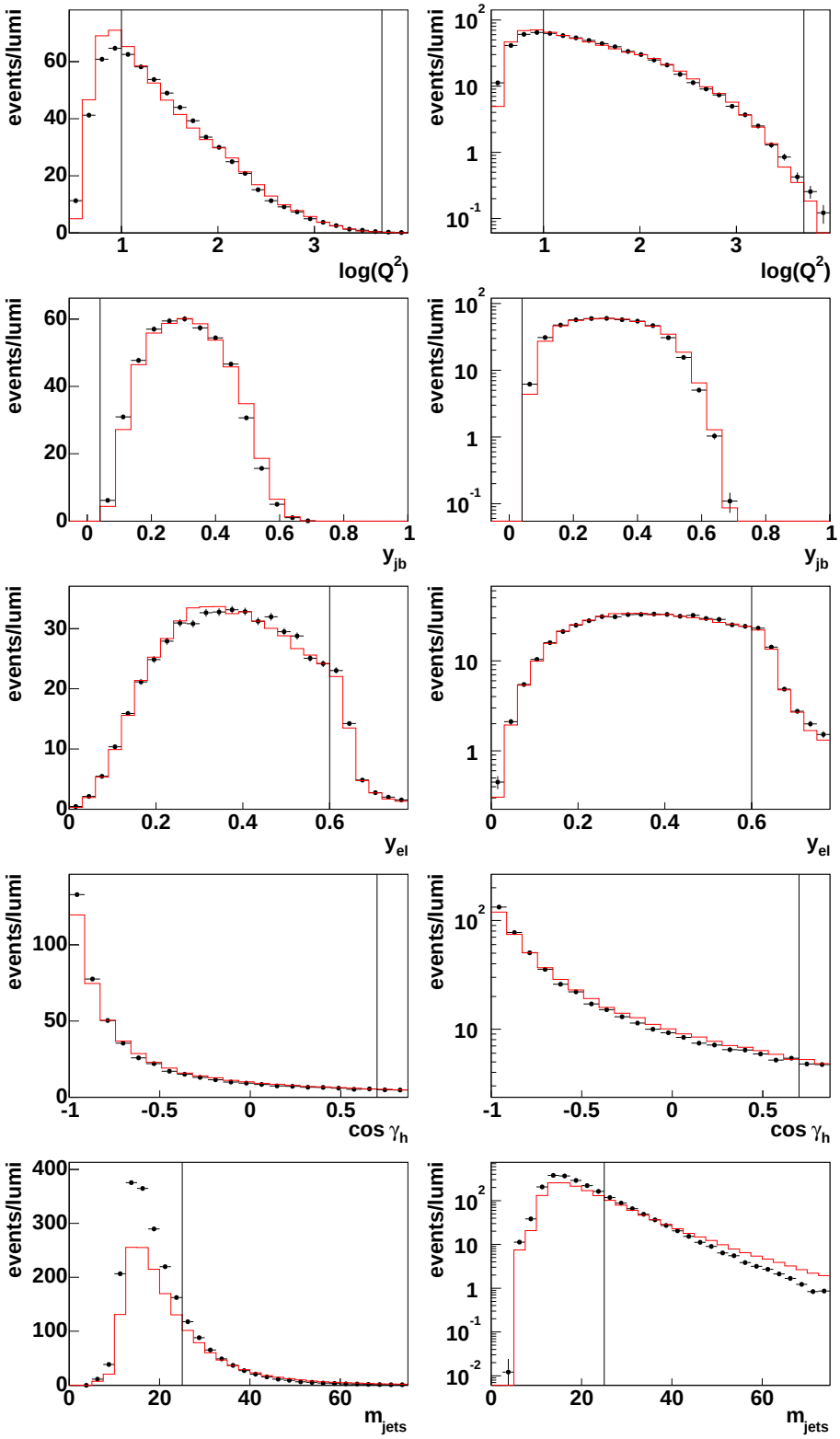

Figure A.8: *as fig. A.1, but for ARIADNE.*

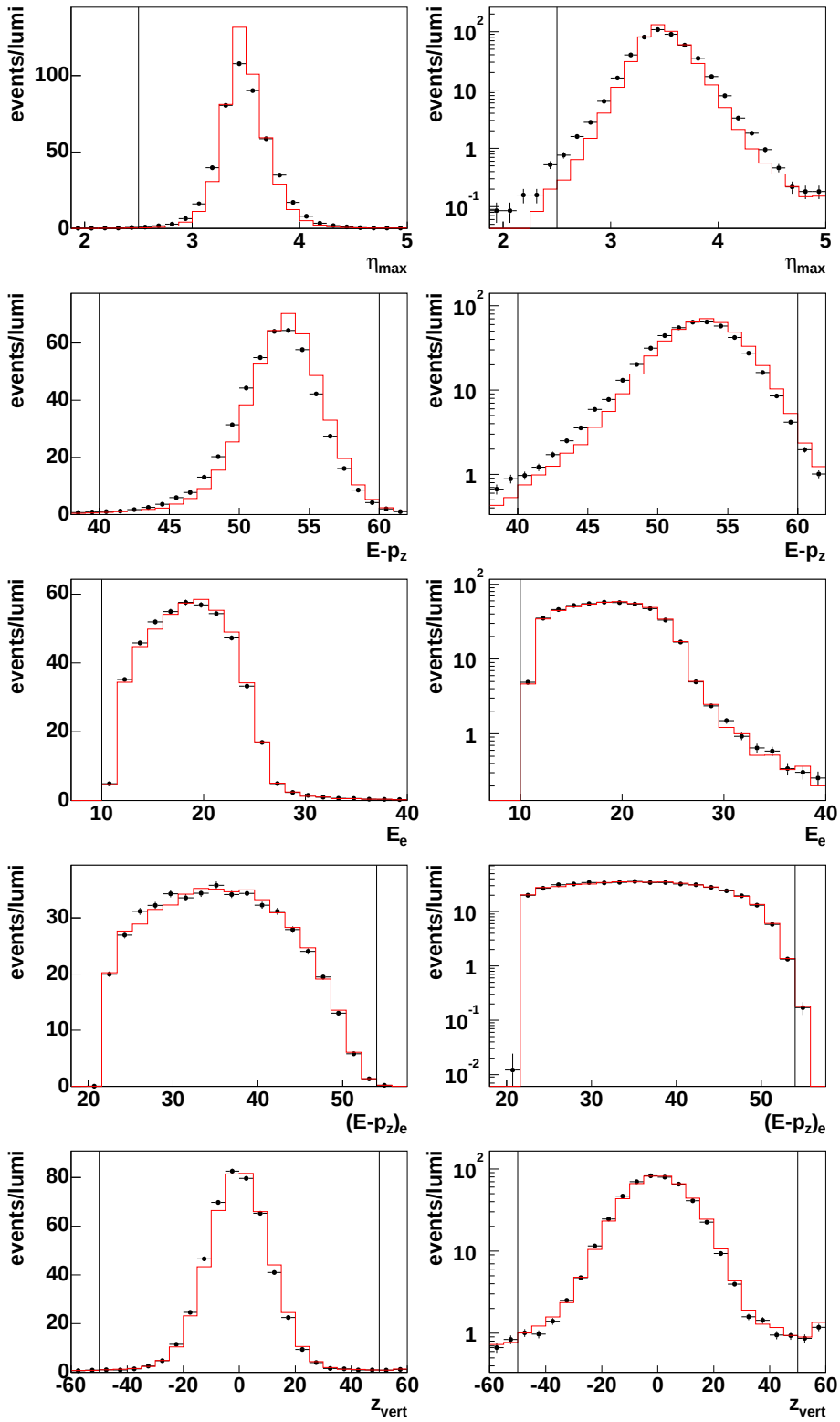

Figure A.9: *as fig. A.2, but for ARIADNE.*

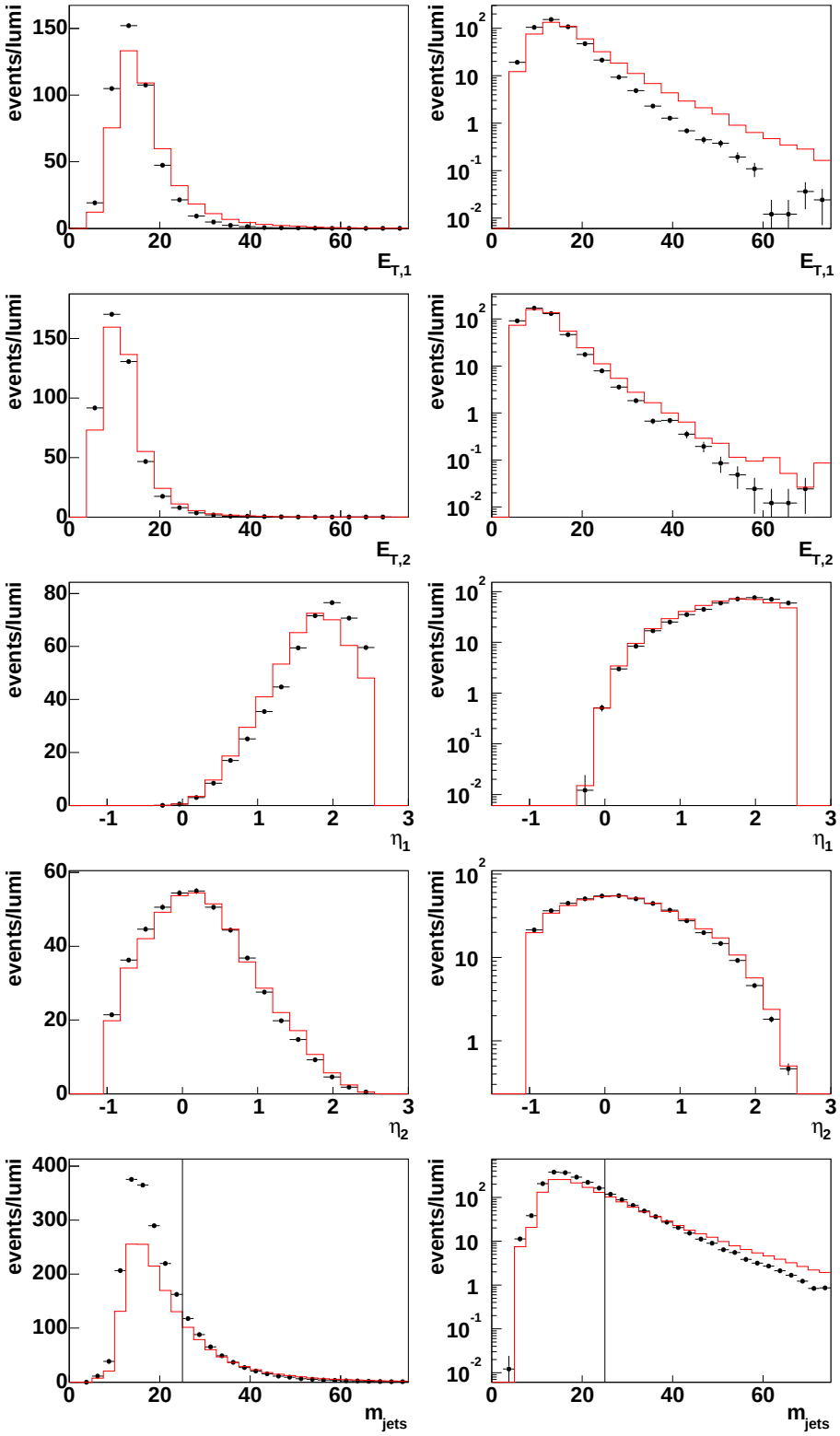

Figure A.10: *as fig. A.3, but for ARIADNE.*

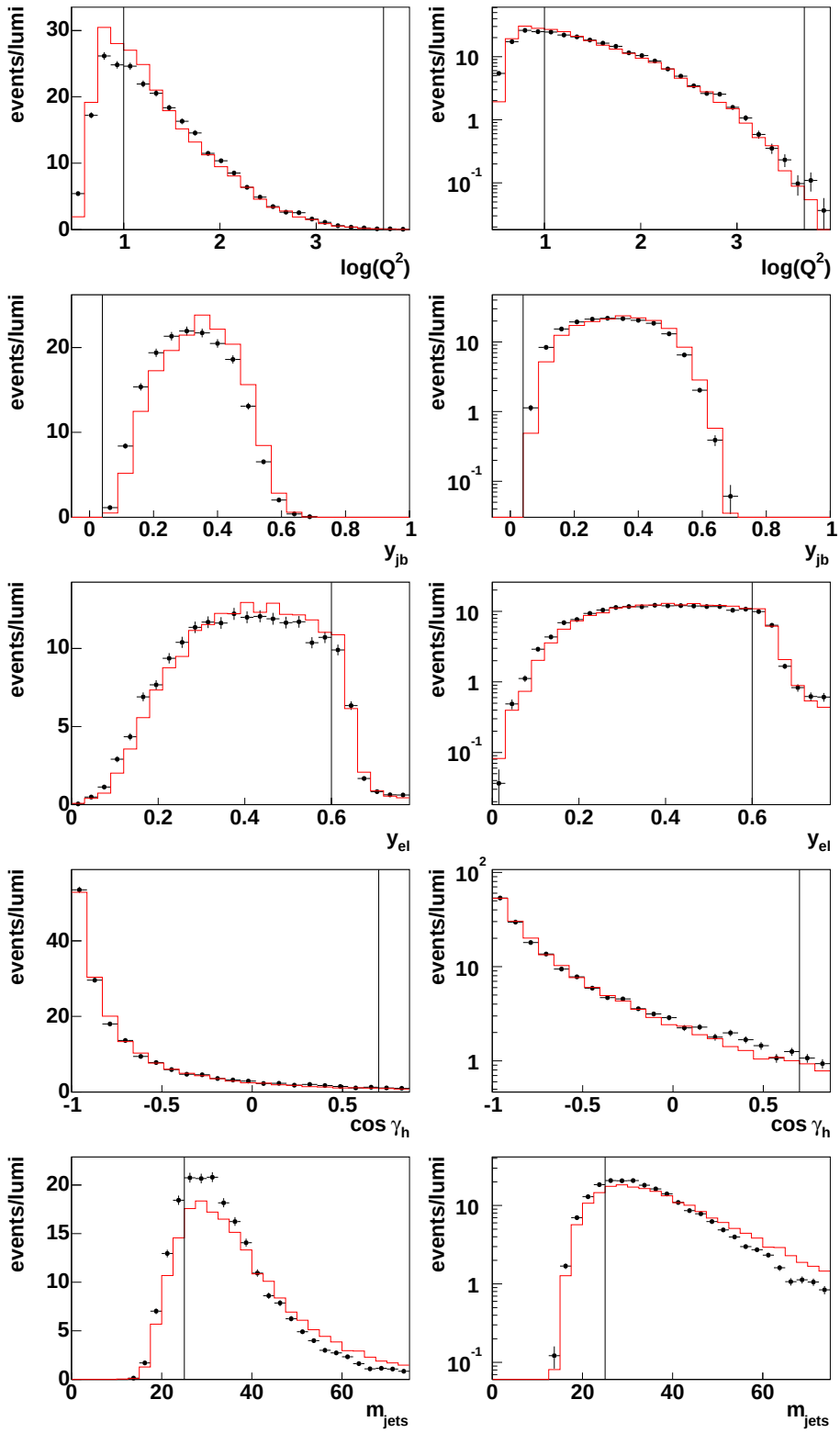

Figure A.11: *as fig. A.4, but for ARIADNE.*

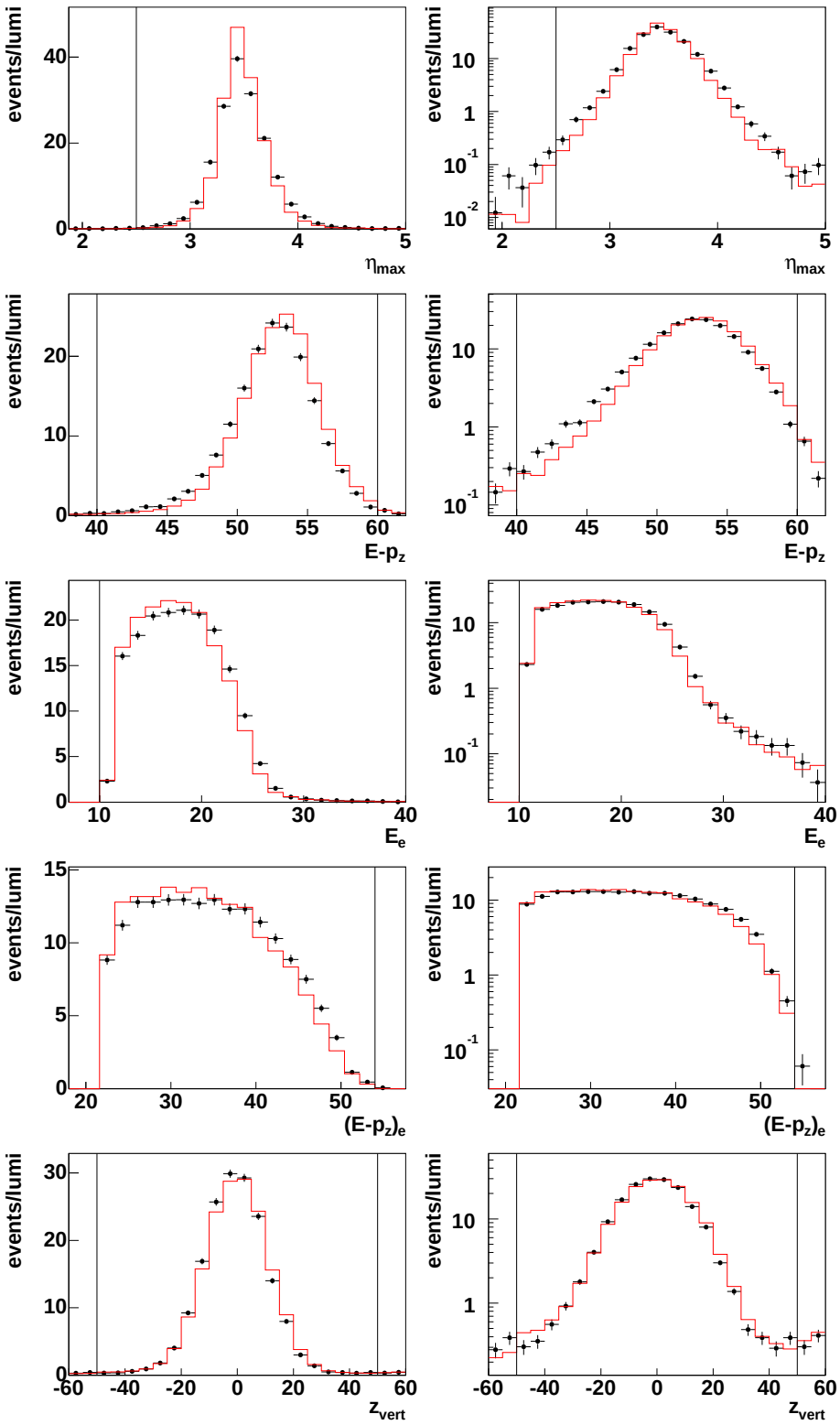

Figure A.12: *as fig. A.5, but for ARIADNE.*

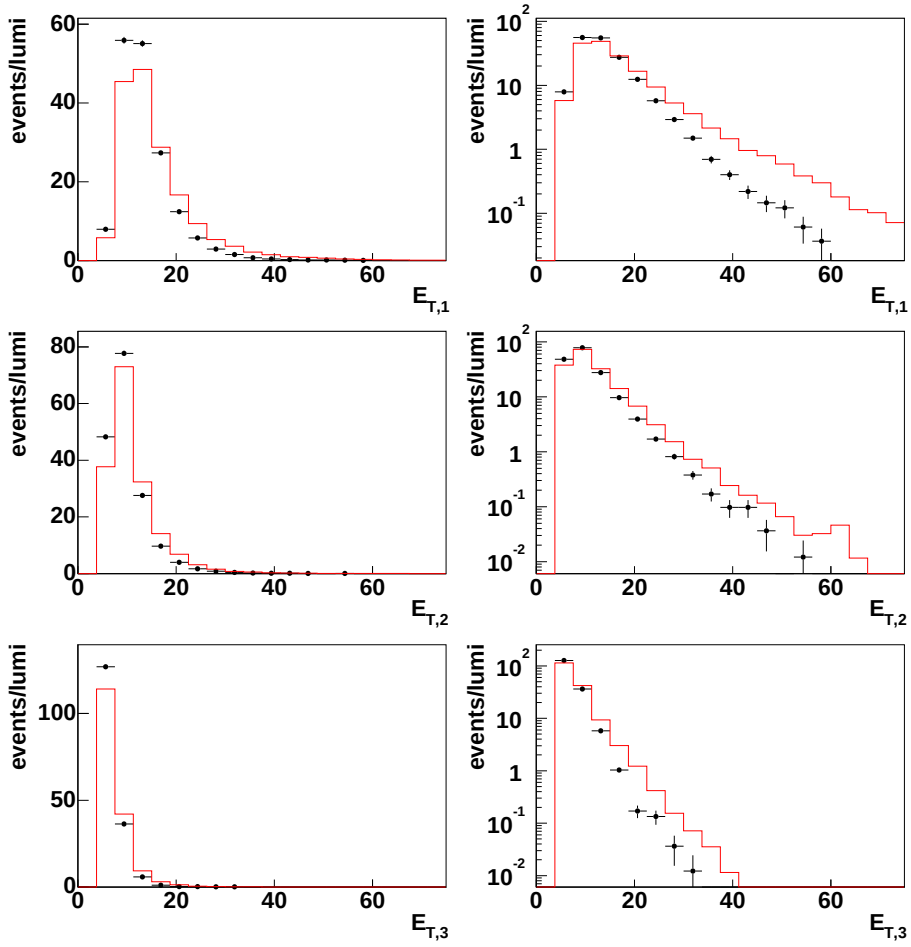

Figure A.13: *as fig. A.6, but for ARIADNE. Continued in fig. A.14.*

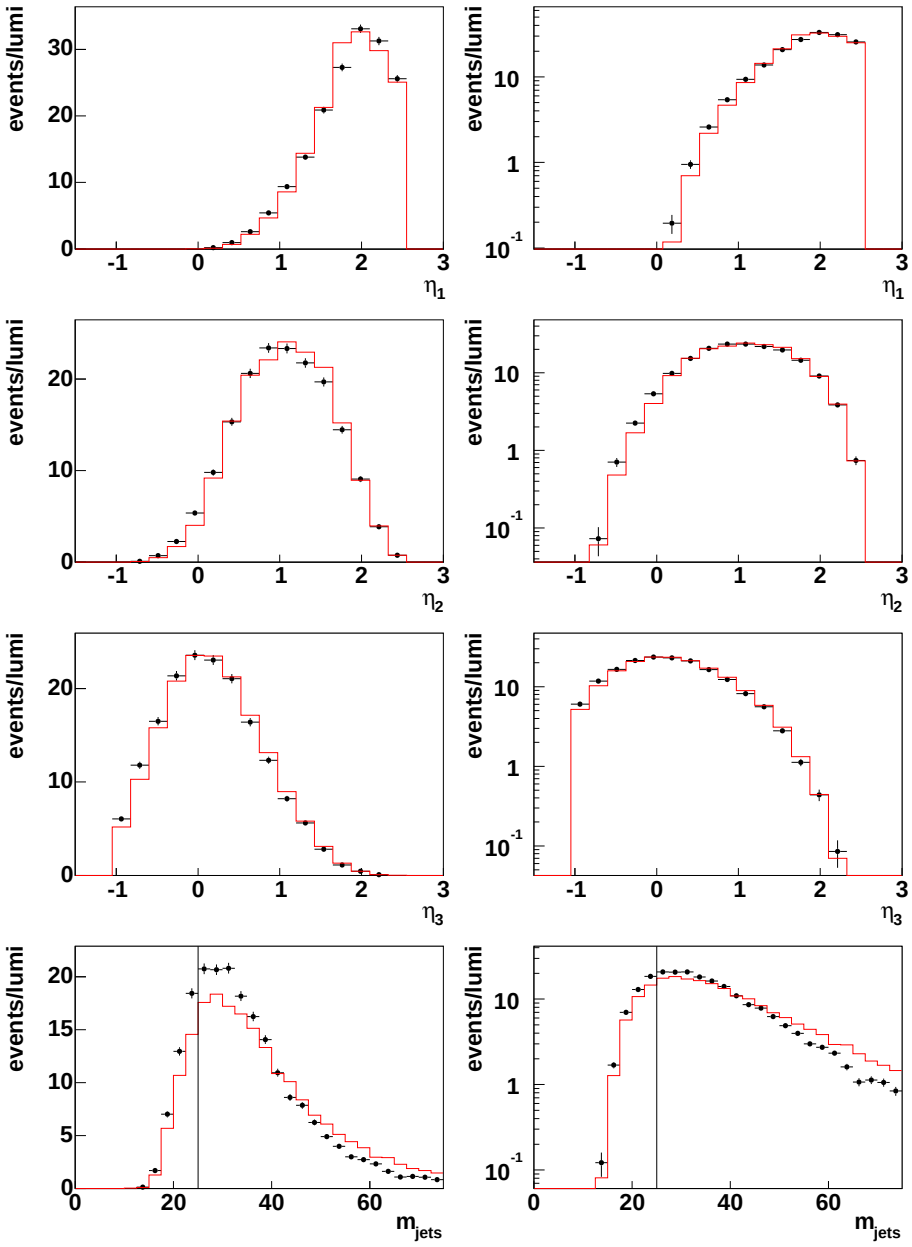

Figure A.14: *as fig. A.7, but for ARIADNE. Continued in fig. A.13.*

Appendix B

Systematic Deviations

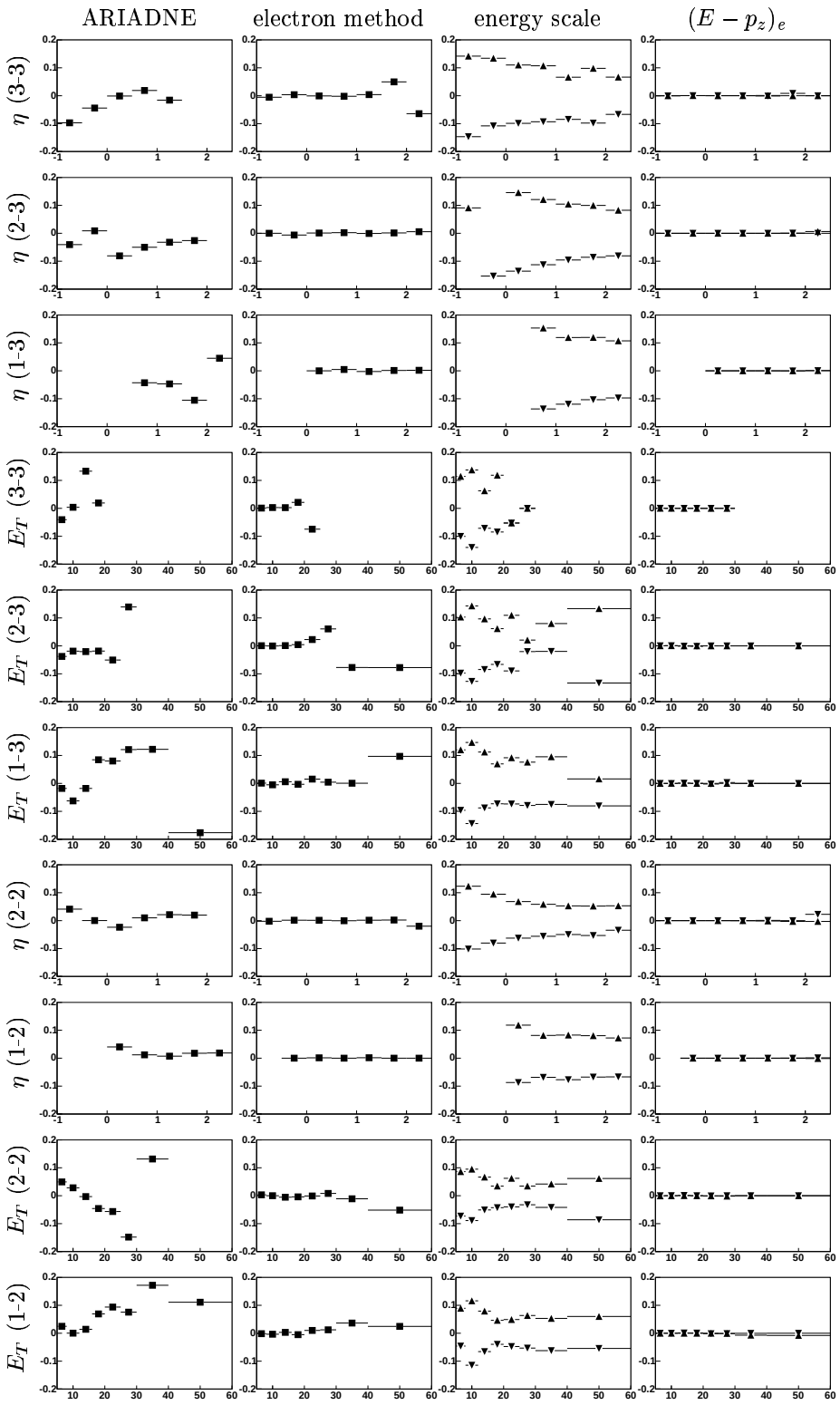

Figure B.1: *Systematic uncertainties of the jet variables for specific systematic tests.*

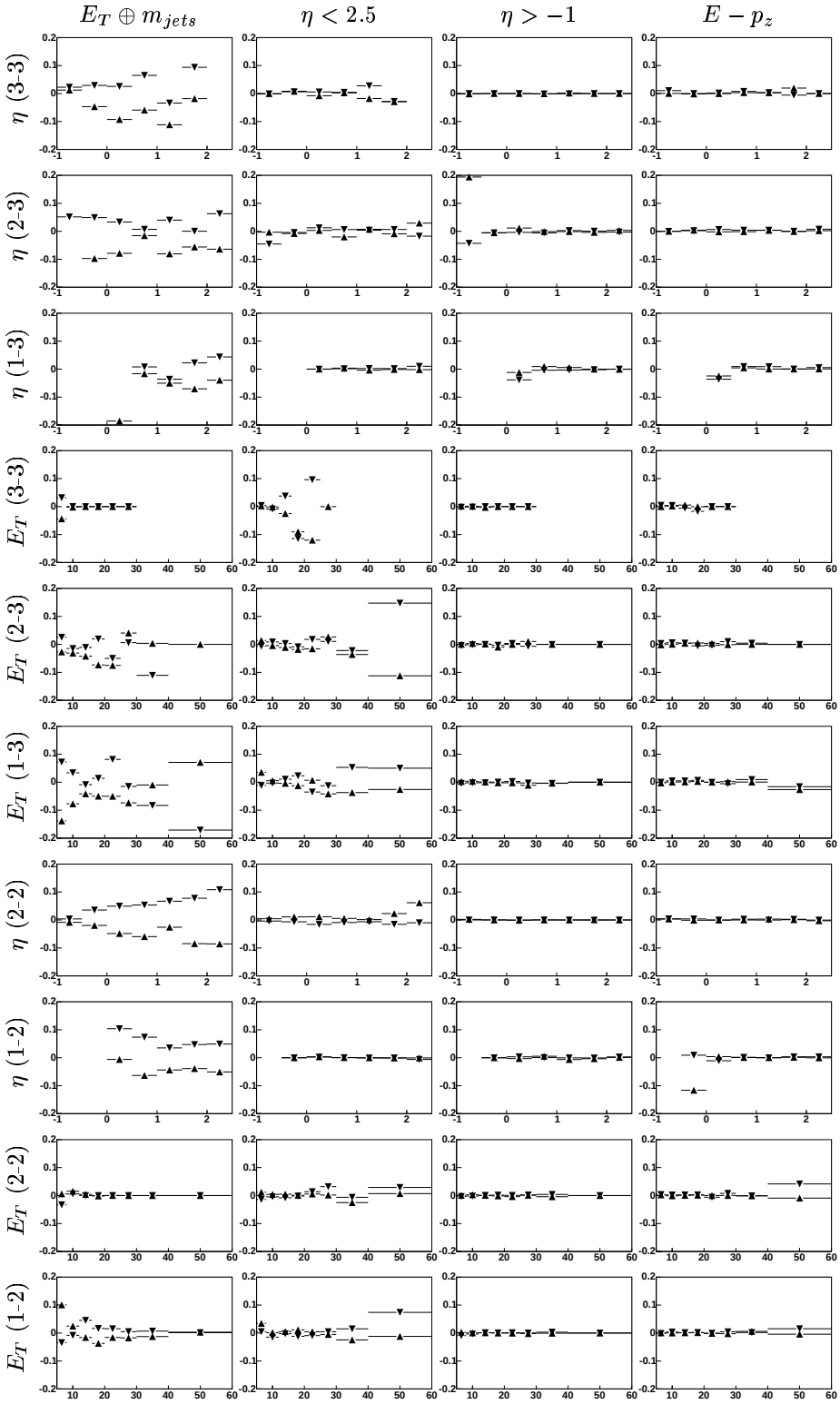

Figure B.2: *Systematic uncertainties of the jet variables for specific systematic tests.*

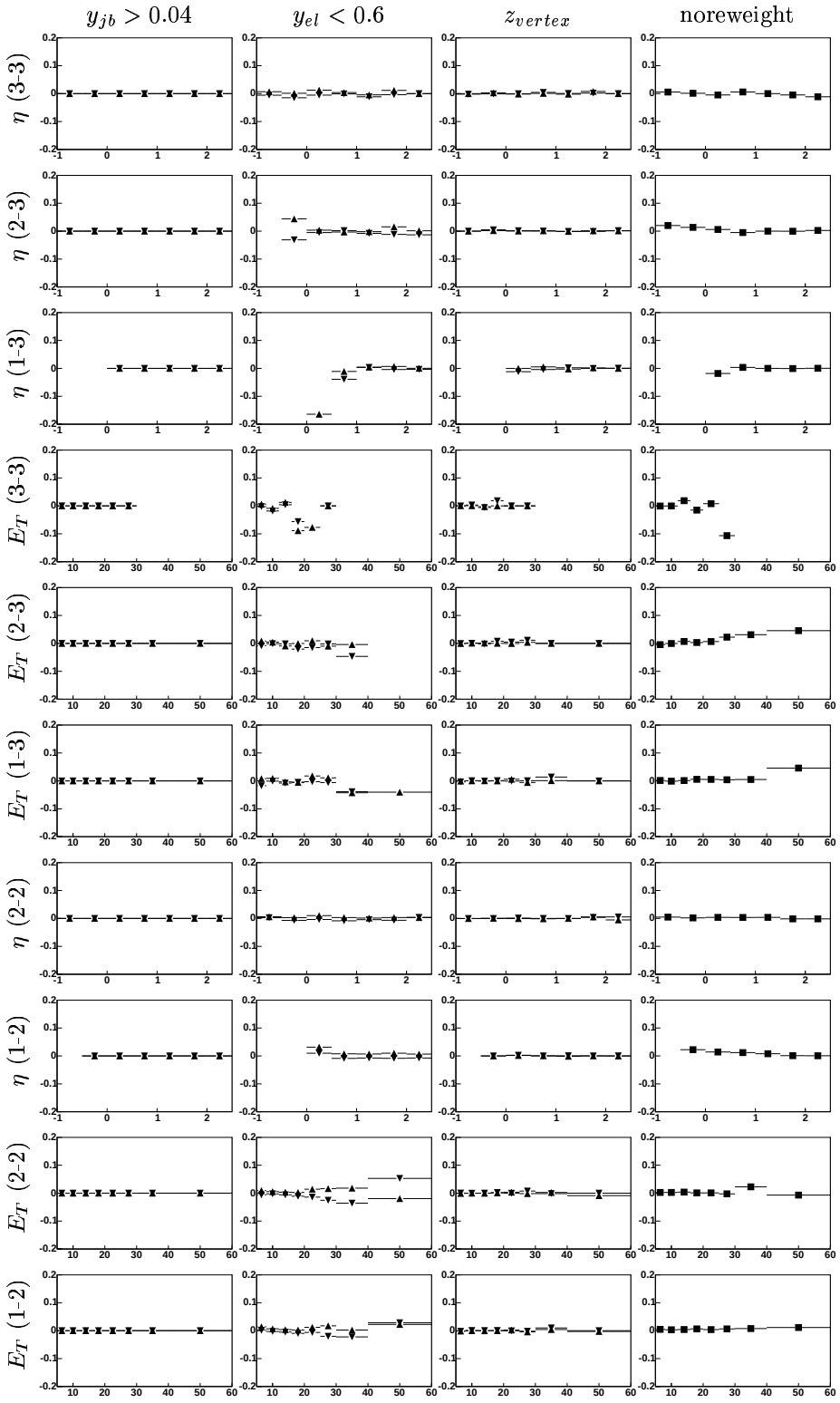

Figure B.3: *Systematic uncertainties of the jet variables for specific systematic tests.*

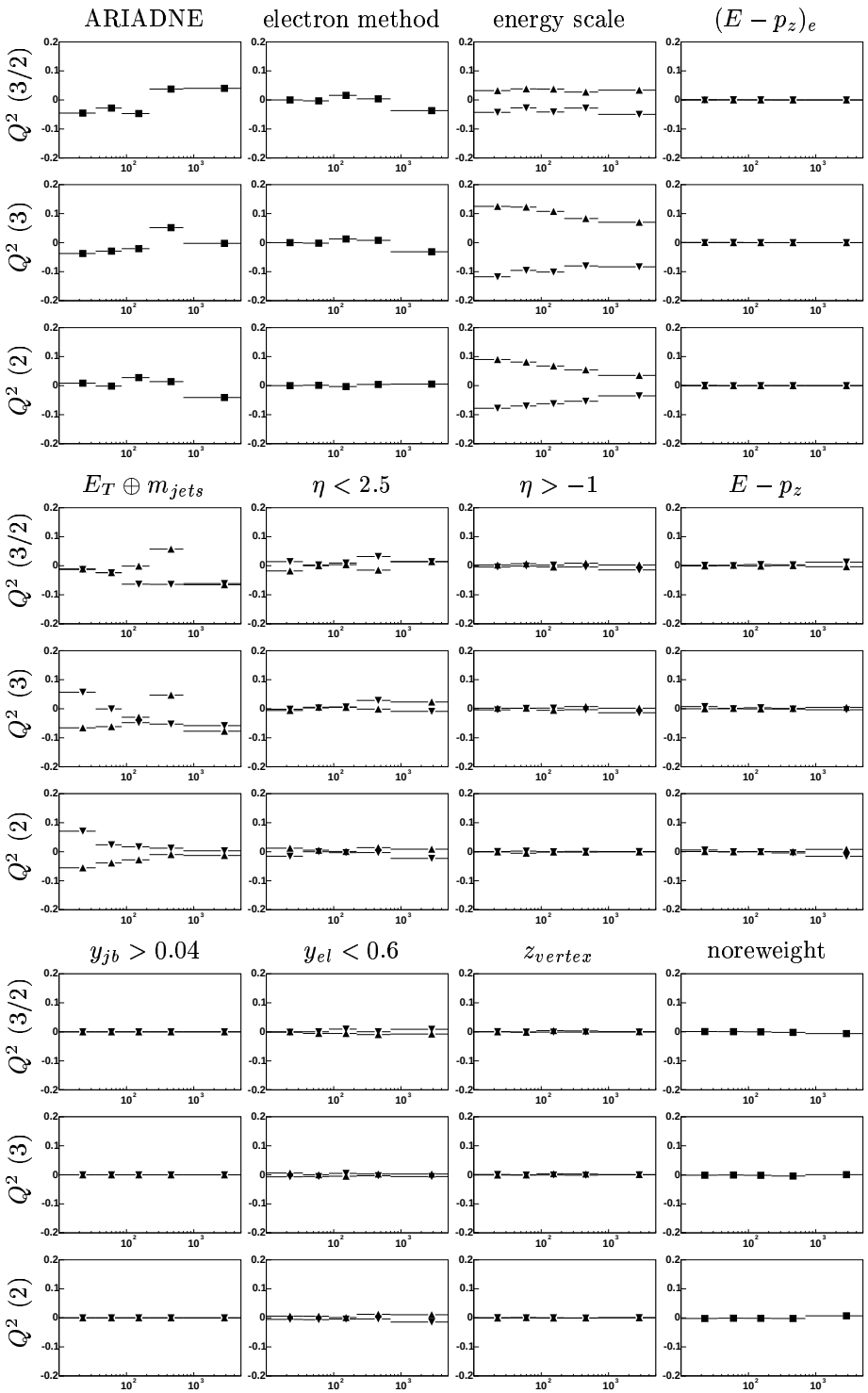

Figure B.4: *Systematic uncertainties of the Q^2 cross sections for specific systematic tests.*

Appendix C

The Rebuilt 6m Tagger

C.1 The New Lumi System

After the 98–00 running period a major upgrade of the HERA machine was done. The purpose of the upgrade was to get a higher instantaneous luminosity and polarized beams. Due to the higher luminosity there was an increased level of radiation hitting the photon calorimeter, making a shielding of the photon calorimeter necessary. Due to that shielding the efficiency of the calorimeter dropped and is not as well determined. To offset that problem the photon calorimeter is augmented by the pair spectrometer that measures the energy spectrum of photons that do pair production in the beam exit window. Although this gives an independent measurement of the luminosity, it doesn't help in determining the efficiency. To solve that the 6m tagger was built. By measuring the electrons and the correlation with the photons one can measure the efficiency of the spectrometer and calorimeter and also cross-calibrate them.

The 6m tagger is a very compact spaghetti calorimeter ($85 \times 23 \times 100\,\mathrm{mm}^3$). It has been built and tested in 2000/2001 before installation at ZEUS. The tests showed a good agreement between data and simulation and an overall satisfactory performance. A more detailed description of these tests and of the detector can be found in [2]. The original design included a water cooling as the GIR5 magnet in which it was installed was expected to reach temperatures of up to 80°C. A leak in this cooling was found in summer 2002 and the cooling was turned off. The tagger showed a continuous decrease in light yield over the rest of the year. When it was removed in spring 2003 it was found to have completely rusted away. This made it necessary to rebuild it.

C.2 Rebuilding Process

The first step of rebuilding was to take the old detector completely apart and clean all the tungsten plates it has been constructed from. For cleaning the plates were first washed in citric acid and afterwards cleaned with a wire brush. During this process several plates broke and had to be replaced with plates from the manufacturing sample. These plates were designed for different spacing of the fibers, so that it is not possible to get any fibers between dummy and normal plates. This results in two layers of dead fibers, one at each end of the detector.

As during the original building process the plates were then stacked and the fibers moved in from the back. As the plate surfaces became more rough (due to the cleaning process) a considerable number of fibers were damaged when inserted or could not be inserted at all. Finally the front of the detector was painted white to increase reflections and the detector was wrapped in black foil to shield it from outside light.

C.3 Testbeam Setup

The tagger had to undergo a testing program at the DESY testbeam. This was in order to have a valid calibration for the later installation at ZEUS and to prove that the tagger will work although the tungsten plates are in a bad condition.

The tagger was installed at a 5° angle towards the beam. To trigger the electrons two crossed 2 cm trigger counters were used. For the fine scan a scintillator plate with a 3 mm hole was used as veto. The gate length was set to approximately 68 ns, which is the gate length used by the ZEUS lumi readout.[1]

C.4 Monte Carlo Tuning

The original Monte Carlo simulation [2] had a number of shortcomings which had to be overcome to get a reasonable agreement between data and simulation for this testbeam:

- The original simulation didn't contain any broken fibers. To overcome this an approximate map of broken fibers has been generated and the signal response of these fibers has been set to 0. For generating this map, the fibers have been lit from the back and photographed from the front. Each fibers that was dark was counted as bad.
- The original simulation didn't contain an ADC digitization. For this the signal response has been scaled to the correct range and then rounded to the nearest quarter of a pC (4 ADC counts are 1 pC).
- The original simulation didn't contain random noise. For this the signal distribution in cells far away from the beam was averaged and then used as a probability distribution for a factor being added to each individual cell.
- The original simulation was not subjected to the calibration. This time it had to undergo the same calibration procedure as the detector (see section C.5).
- The new data gives a wider energy distribution than the simulation. This is probably due to the mass loss of the plates. This has not been compensated, but the calibration has been altered to treat data and simulation differently.

[1] The gate length is the integration time of the ADC. The Lumi readout measures an event every 96 ns. The readout integrates the current over time to measure the total charge coming from the photomultiplier. To reduce the influence of noise this integration is restricted to the time interval with a relevant signal.

C.5 Calibration

In response to the different shower behavior of original and new detector the original calibration procedure [2] had to be altered. It still uses only 5 GeV events and only uses the events that have their maximum energy in one particular cell to calibrate that cell, but the selection cuts are different. In addition some of the variables have been corrected for the simulation to match the data (Fig. C.1).

The cuts are:

- the overall size (RMS of energy deposits in tagger) to be between 5.5 mm and 18 mm.
- the overall energy (sum of energy deposits in tagger) to be between 44 pC and 59 pC (8 pC equal approximately 1 GeV).
- the ratio of the energy in the highest energy cell to the energy in 3×3 cells (centered around the highest energy cell) to be between 0.50 and 0.68. For the simulation this has been scaled down by a factor of 0.850 and then shifted up by 0.020 to match the data.
- the energy difference between 5×5 and 3×3 cells (centered around the highest energy cell) to be between -1 pC and 12 pC. For the simulation this has been scaled up by a factor of 1.257 and then shifted up by 0.228 pC to match the data.
- the energy difference between all and 5×5 cells (centered around the highest energy cell) to be between -5 pC and 8 pC. For the simulation this has been scaled up by a factor of 1.497 and then shifted down by 0.978 pC to match the data.
- the highest absolute value of charge collected in a single cell at least seven cells away from the highest energy cell to be less than 2.125 pC (corresponds to 8–9 ADC counts).

The calibration spectrum of a cell is the energy spectrum in that cell for the events that have that cell as their highest energy cell. To the calibration spectrum a Gaussian is fit (Fig. C.2–C.5). The ratio of 20 pC (the desired signal) to the fit mean value is taken as an additional calibration factor for that channel. For the simulation this is multiplied by an additional factor of 1.13956 to make up for different shower behavior.

To make up for the non-uniformities from the broken fibers the mean energy is determined as a function of the corrected reconstructed position[2] (Fig. C.6). The ratio of the beam energy to this energy is used as a non-uniformity correction factor.

C.6 Energy Scan

To test the resolution and linearity of the energy reconstruction the beam energy has been altered from 1 to 6 GeV in 1 GeV steps. For each energy a Gaussian was fit to the energy spectrum after some cuts (Fig. C.7 and Fig. C.8).

[2] From now on I will always speak of position reconstruction when I mean the corrected position reconstruction.

For determining the linearity a straight line is fit to the measured mean as a function of the beam energy. The relative difference between the data points and the fit is the nonlinearity of the detector (Fig. C.9 and Fig. C.10). It can be seen that for data the linearity is at a 1% level over the whole range and at a 0.2% level above 3 GeV. For the simulation it is compatible with no nonlinearity.

For determining the resolution the ratio of fit sigma to fit mean was calculated as a function of the beam energy. To this the function

$$\Delta E_{da} = E \cdot (16.5\%/\sqrt{E} \oplus 5.0\%/E)$$

was fit to the testbeam data and the function

$$\Delta E_{mc} = E \cdot (17.1\%/\sqrt{E} \oplus 1.7\%/E)$$

was fit to the simulated data.

C.7 Fine Scan

To test the position reconstruction and the position dependence of the energy reconstruction three fine scans were done. For this fine scan the veto trigger with the 3 mm hole was used and moved horizontally over the detector. The scans were done in the middle, at the upper and the lower edge of the central row of cells.

To test the uniformity of the energy reconstruction the mean energy has been calculated as a function of the trigger position (Fig C.13). The spectrum of these mean energies gives a good estimate of the actual non-uniformity of the detector (Fig. C.14). Considering the size of the uncertainties in the mean energy, one can say that the non-uniformities are within 1%.

The reconstructed x-position was determined as a function of the trigger position. To this a straight line is fit and the absolute deviation from the fit is determined (Fig. C.15 to Fig. C.17). This leads to a nonlinearity of the position reconstruction of less than 0.2 mm. From the RMS at one position one gets an estimate of the position resolution at that position. The spectrum of all these resolutions (Fig. C.18) gives the overall resolution of 1.6 mm. These tests are only rough estimates of the performance of the position reconstruction, as they are influenced by nonlinearities of the position reconstruction. The simulation yields a much better position resolution of 0.6 mm which is not understood up to now. This could come from the more collimated showers in the simulation.

C.8 Other studies

At ZEUS the electrons were diverted onto the tagger by a dipole magnet which leads to a direct correlation of energy, angle and position. To study this energy scans in the center of the detector have been done for angles from 2° to 8° in 1° steps (at ZEUS we have the angles 3–7°). For each of them the same evaluation as for the energy scan has been done. It shows a dependence of the gain (Fig. C.19) and of the resolution (Fig. C.20) on the angle, but are not as strong as for the original detector [2].

To test if a lead shielding would affect the resolution of the tagger three runs at 5 GeV have been taken and a Gaussian has been fit to each energy spectrum, showing no different in the width.

C.9 Summary of Testbeam Measurements

After water damage the 6m tagger has been taken apart, the tungsten plates have been cleaned and a new tagger has been built using the old plates and a set of new fibers. This tagger has been calibrated and tested at the DESY testbeam.

Due to broken fibers the new detector showed a non-uniformity that could be reproduced by the simulation. After correcting for these the tagger has been shown to have an energy resolution of

$$\Delta E_{da} = E \cdot (16.5\%/\sqrt{E} \oplus 5.0\%/E)$$

and a non-uniformity of 1%. This is comparable to the performance of the Monte Carlo and of the old detector. The new detector has wider showers than the simulation which is not understood, but might be due to a material loss of the tungsten from the cleaning.

The position resolution of the detector is less than 1.6 mm and the nonlinearity is less than 0.2 mm. The simulation (0.6 mm) underestimates the position resolution by far.

The dependence of the gain on the angle (and thereby the position) is not fully understood, as it is different from the original detector and the reasons are not known.

For cross calibration and acceptance tests of the other Lumi components these resolutions should be good enough, as they are still better than their energy resolutions ($< 20\,\%$ and $< 6\,\mathrm{mm}$). The tagger can also be used to determine the kinematics of photoproduction events at high W, but the exact requirements of these measurements are not known.

C.10 Acknowledgments for the Testbeam

I would like to thank Katarzyna Wichmann for taking over as coordinator for the time after I leave ZEUS and the 6m tagger group for helping me out from time to time when I needed a hand. Thanks goes also to Helmut Krause for repairing our video camera when it broke. The biggest thanks go to Rainer Peter Feller, Sabine Wedam Lammers and Liang Li for sitting in with me at the testbeam for many shifts. Special thanks goes to Sabine who's fascination and interest in this project reminded me how much fun these projects actually are.

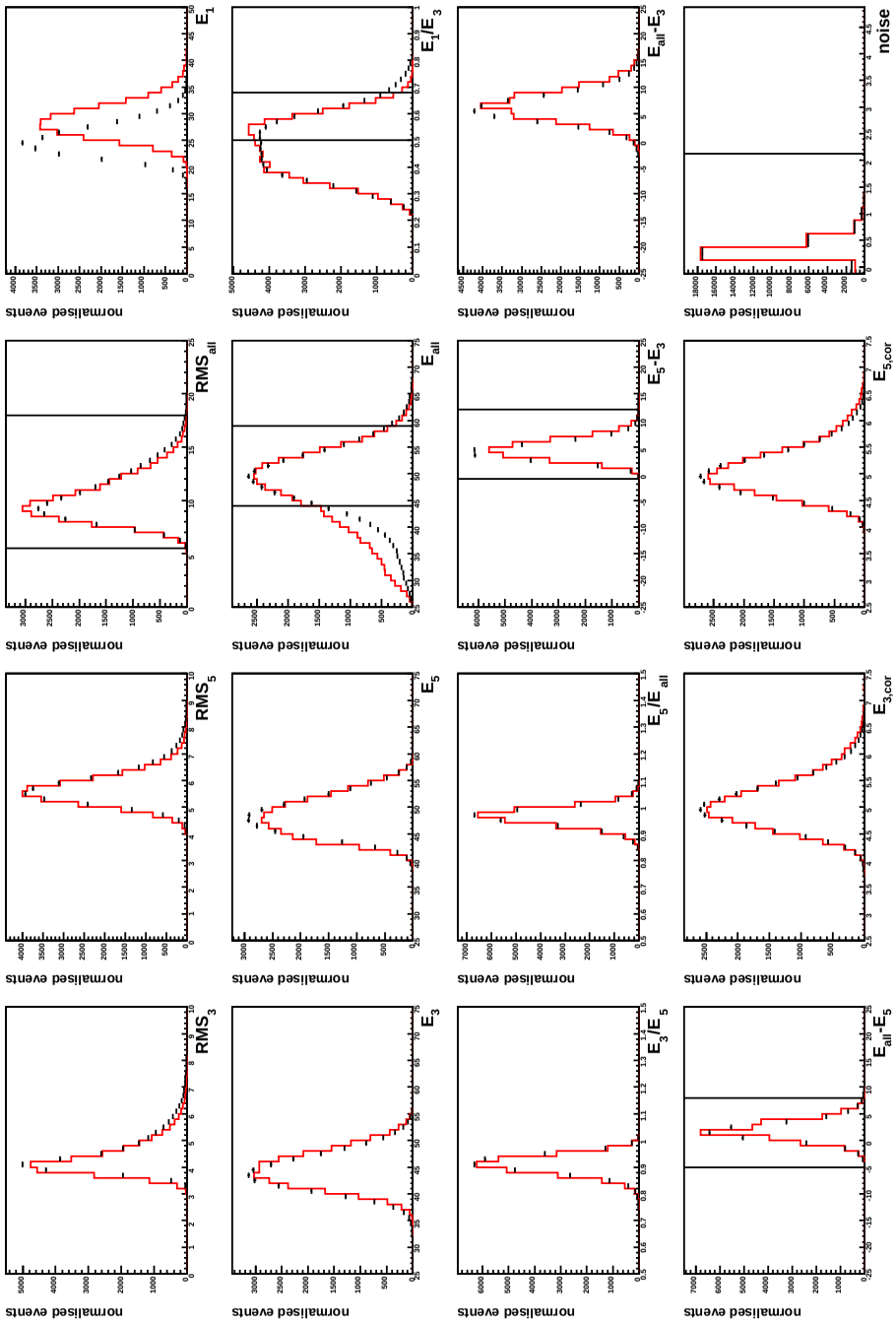

Figure C.1: *Control plots for the second calibration step. Data has been reweighted to match the position distribution of the Monte Carlo (which has a flat input distribution). The Monte Carlo has been scaled to match the event rate of the data. The vertical lines indicate the cut boundaries. For the energy ratios and differences the Monte Carlo variables have been scaled and shifted to match the data.*

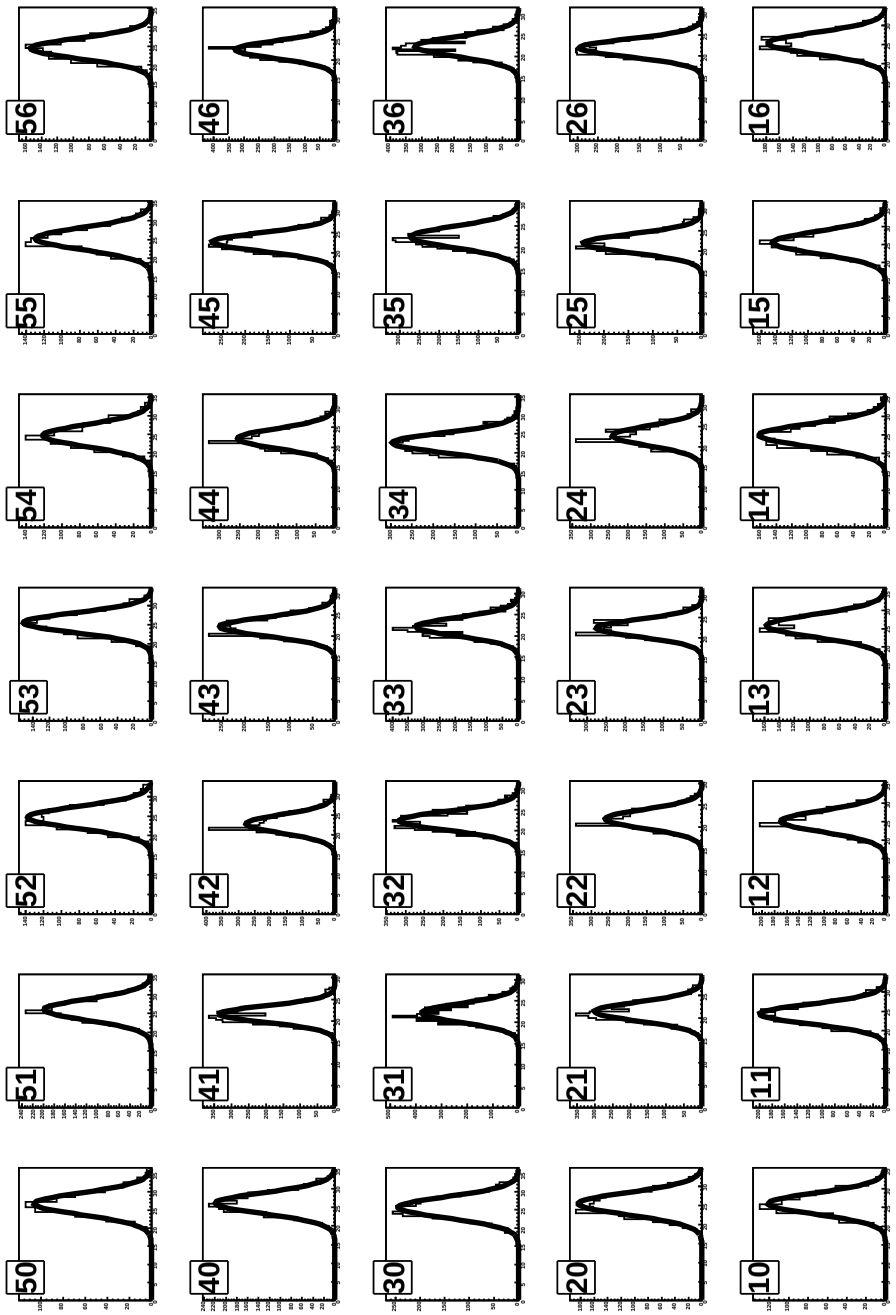

Figure C.2: *Calibration plots of the left side of the detector for data. Shown is the energy in the central cell for all events that center in that cell and survive the cuts.*

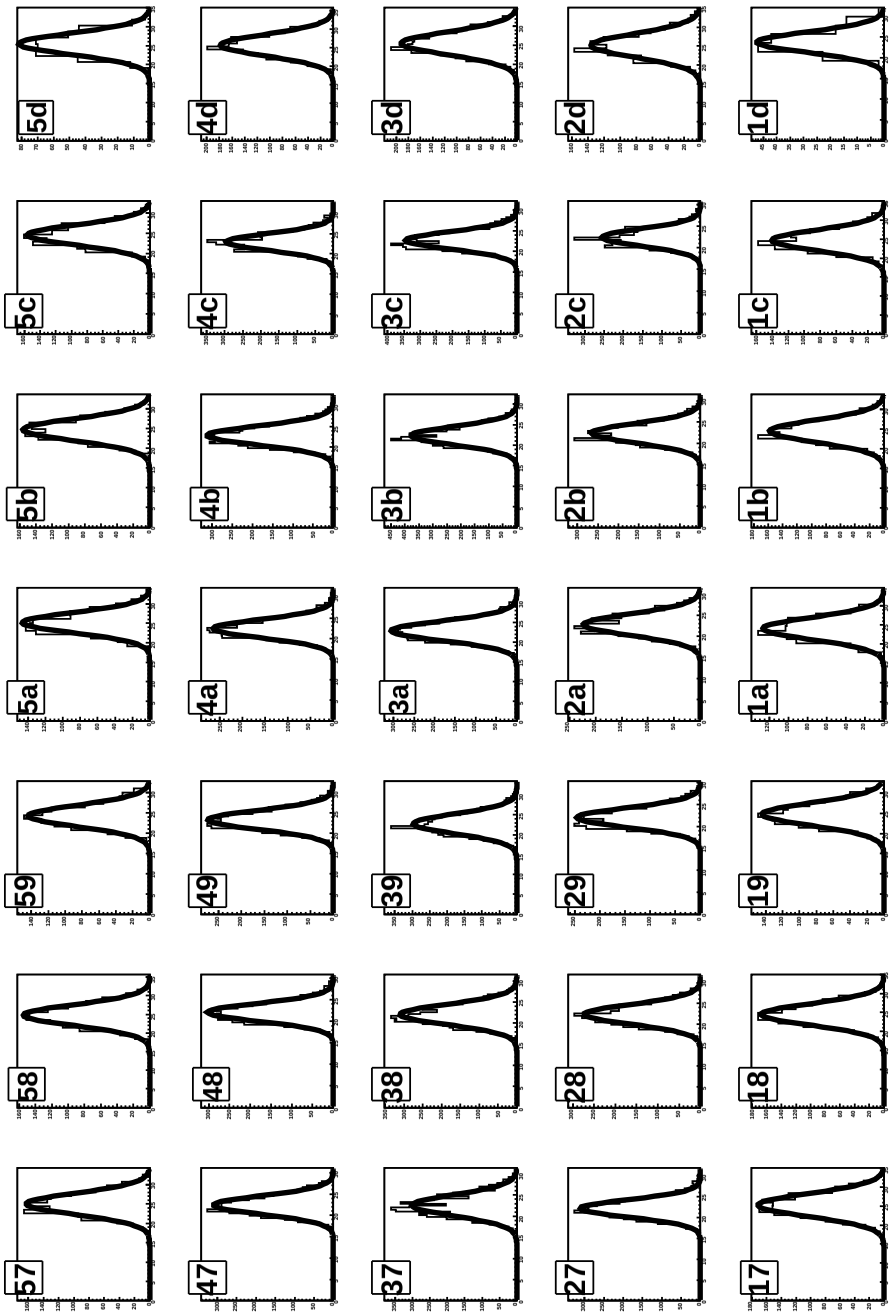

Figure C.3: *As figure C.2 but right half of detector.*

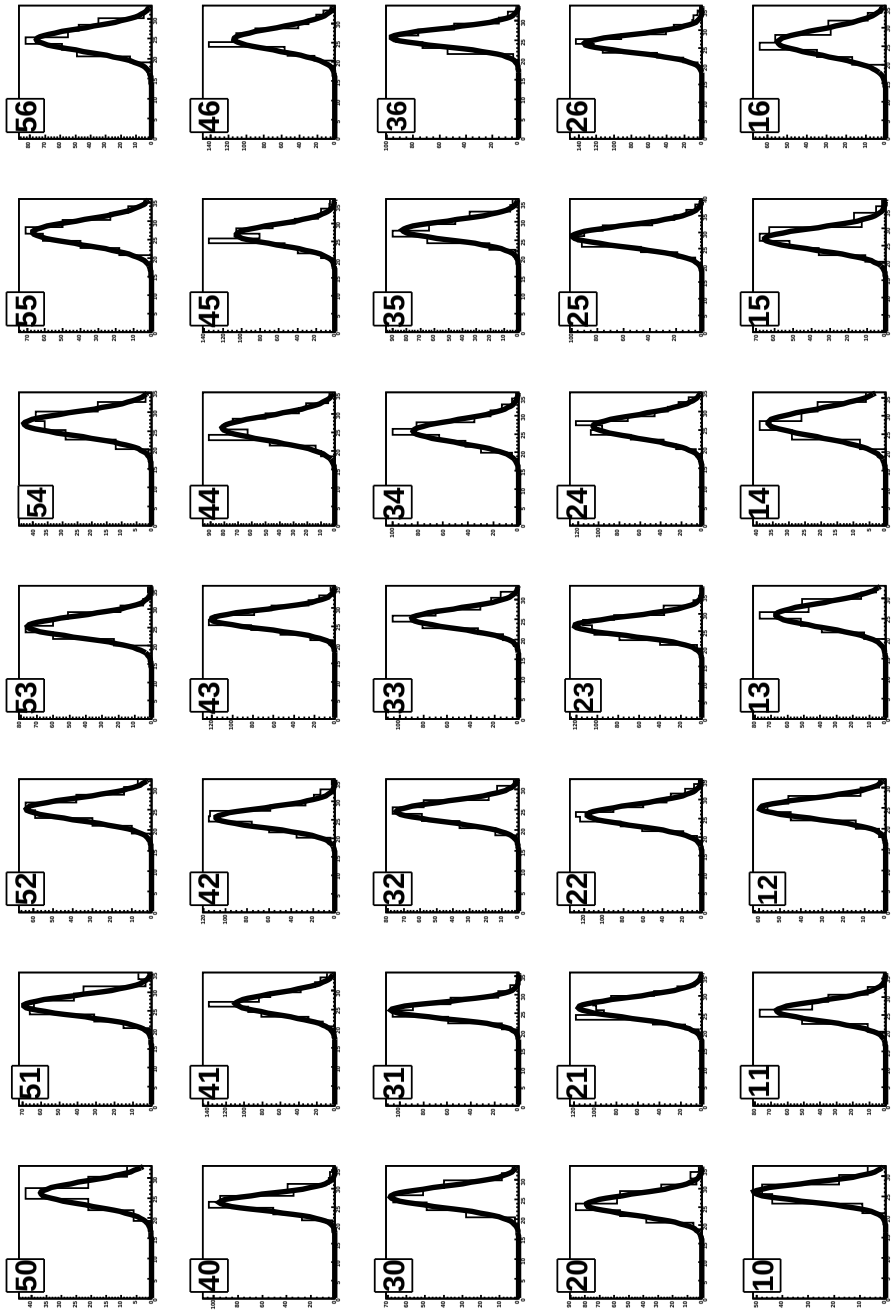

Figure C.4: *As figure C.2 but for Monte Carlo.*

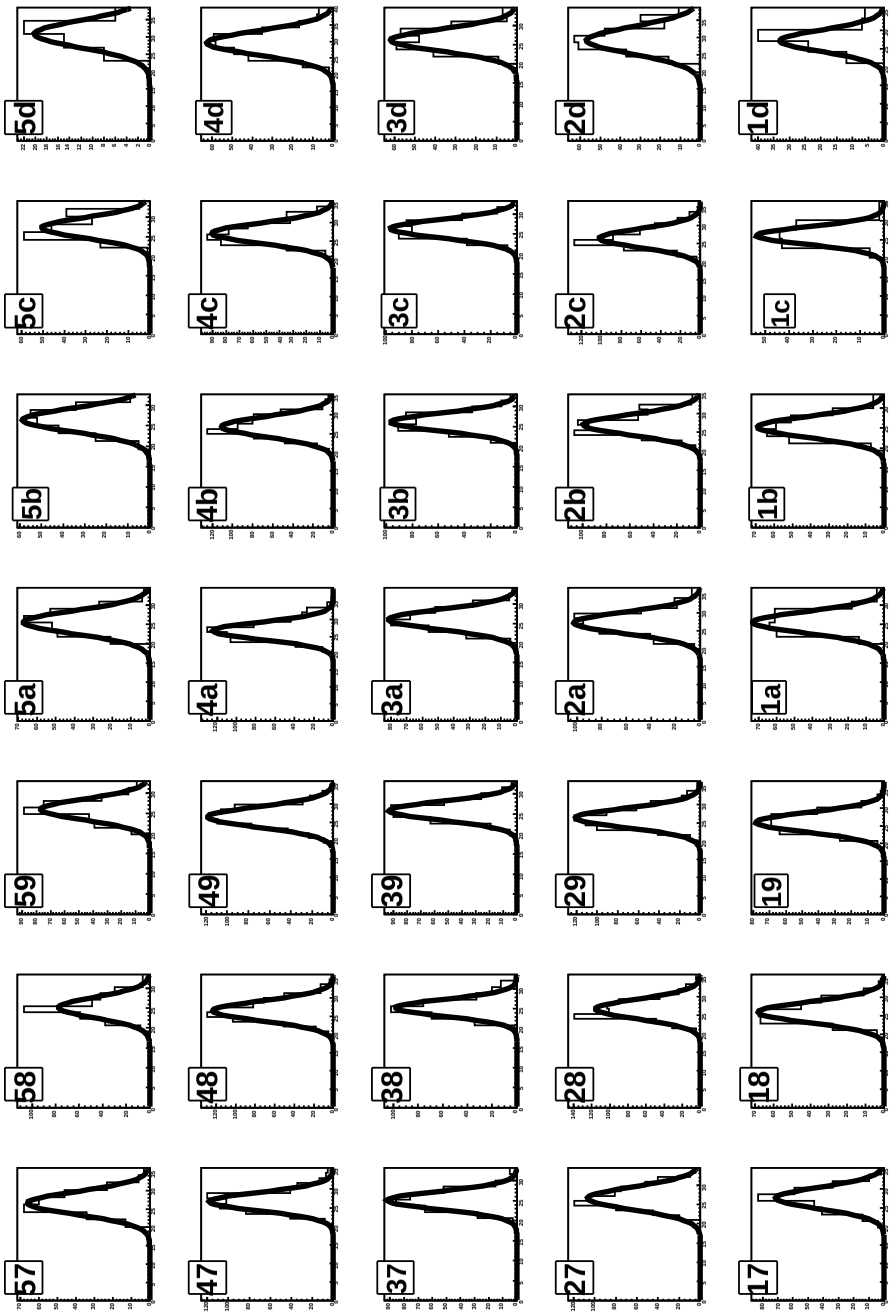

Figure C.5: *As figure C.4 but right half of detector.*

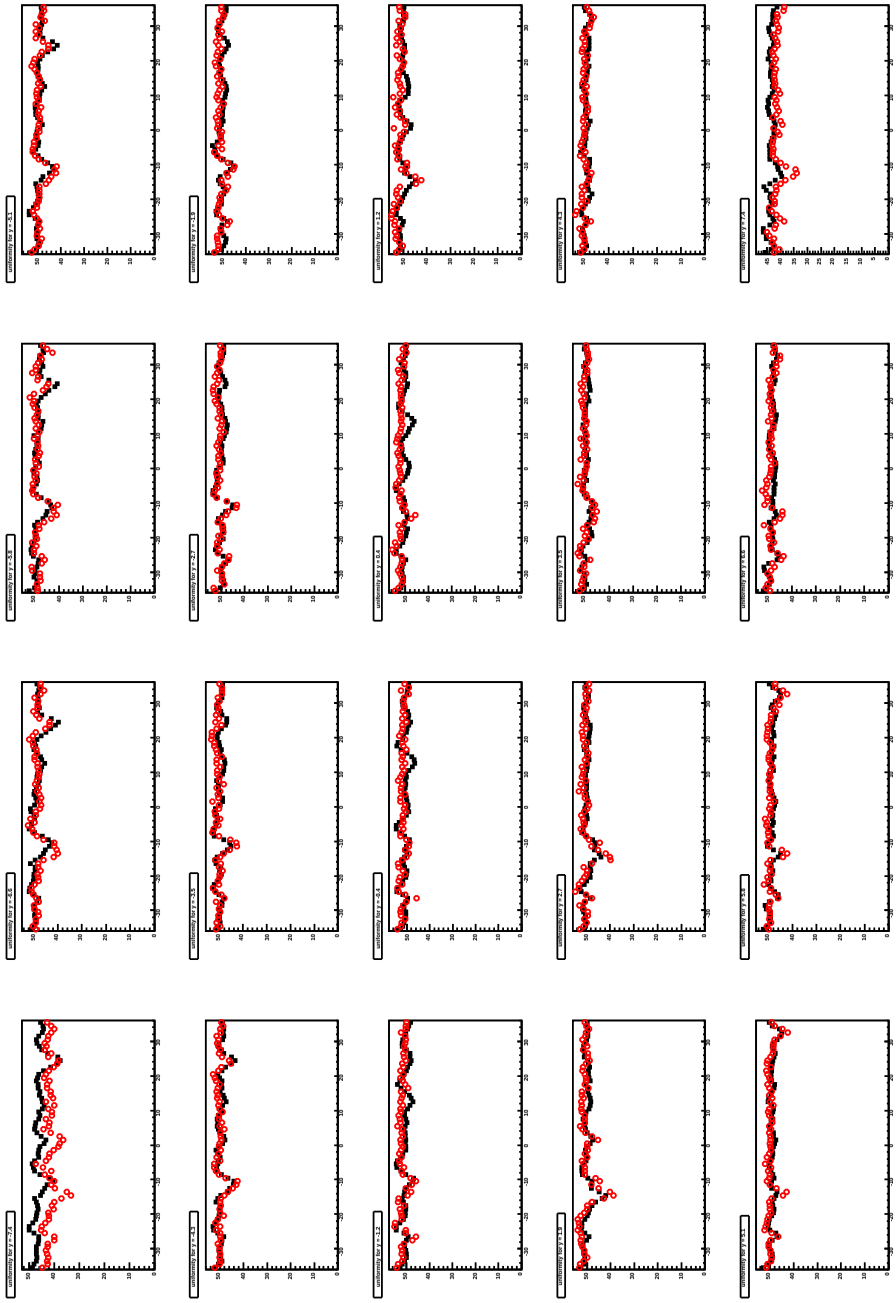

Figure C.6: *Non-uniformity for data (black solid boxes) and simulation (red open circles). Shown is the mean reconstructed signal in* 5×5 *cells as a function of the (reconstructed) position in x and y. These values have been used to correct the signal for non-uniformities.*

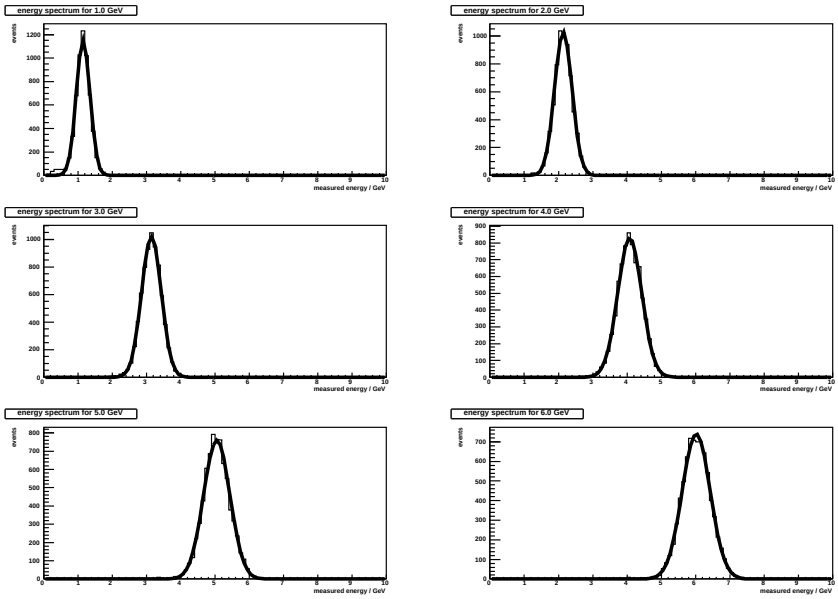

Figure C.7: *Energy spectrum of the events used for the energy scan for the six different beam energies. For each beam energy a cut on the overall shower size and the beam energy has been applied. In addition it was required that the highest energy cell is at least two cells away from the border.*

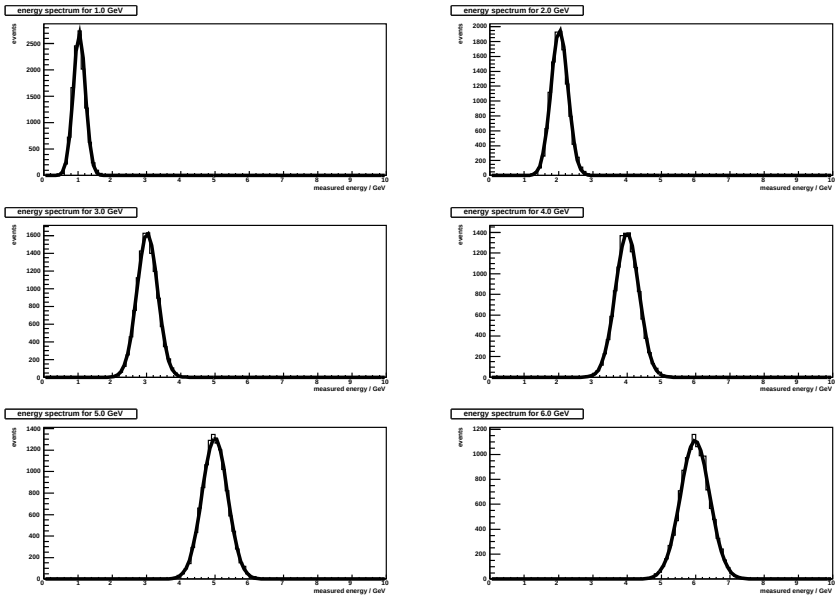

Figure C.8: *As figure C.7 but for Monte Carlo.*

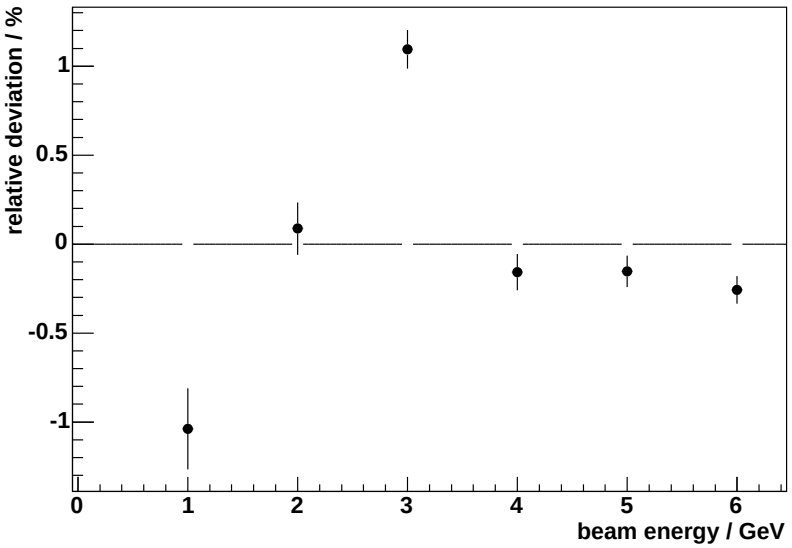

Figure C.9: *Deviation from linearity for the energy reconstruction in data. To the mean signal response, as a function of beam energy, a straight line has been fit and the relative difference is shown here.*

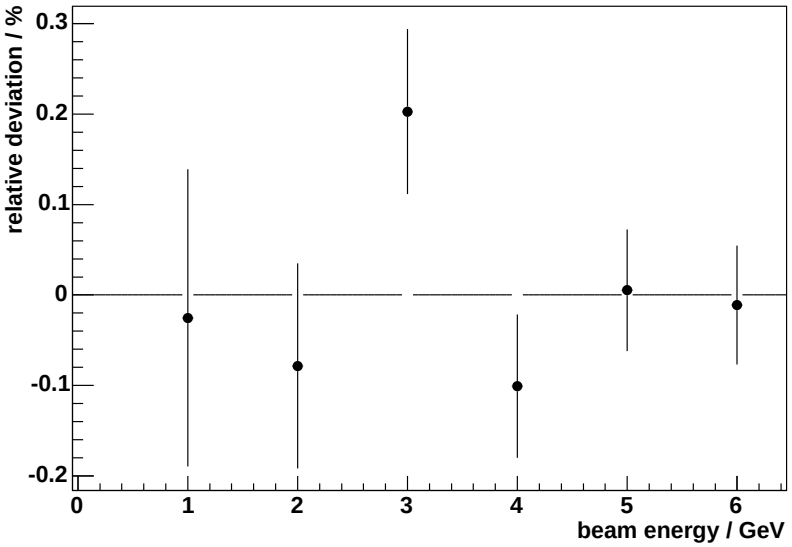

Figure C.10: *As figure C.9 but for Monte Carlo.*

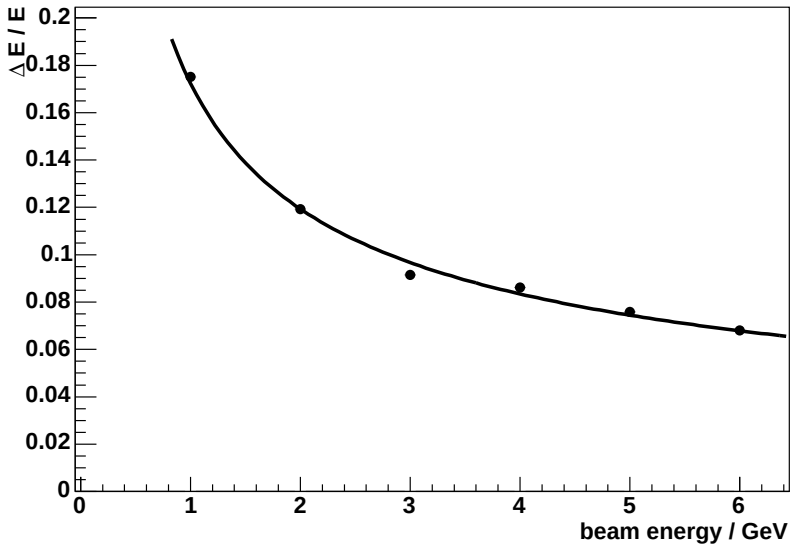

Figure C.11: *The energy resolution. From the fits to the energy spectra (Fig. C.7) the ratio of sigma to mean is calculated and shown as a function of beam energy. To this a statistical and a noise term (added in quadrature) are fit.*

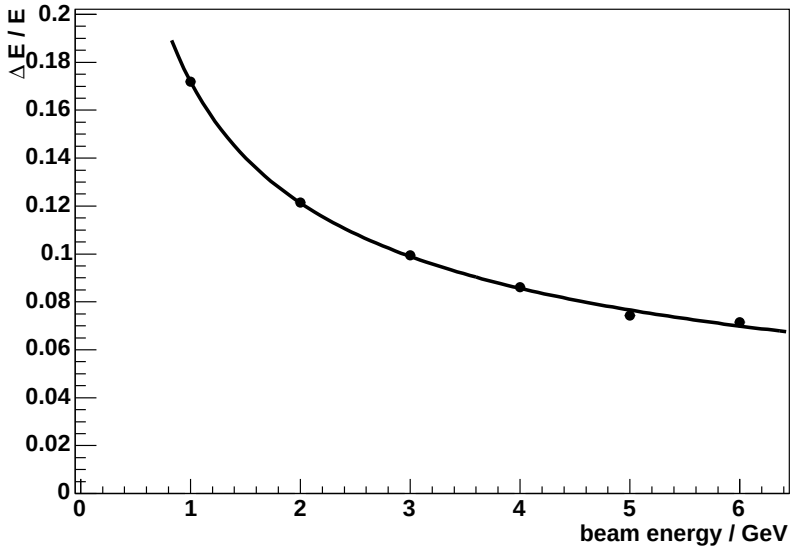

Figure C.12: *As figure C.11 but for Monte Carlo.*

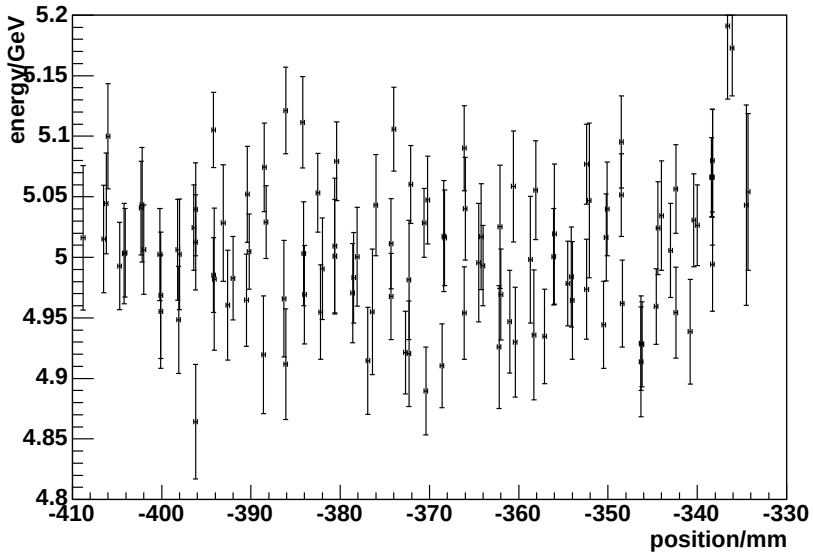

Figure C.13: *The energy uniformity from the three fine scans. The mean reconstructed energy as a function of the beam position is shown.*

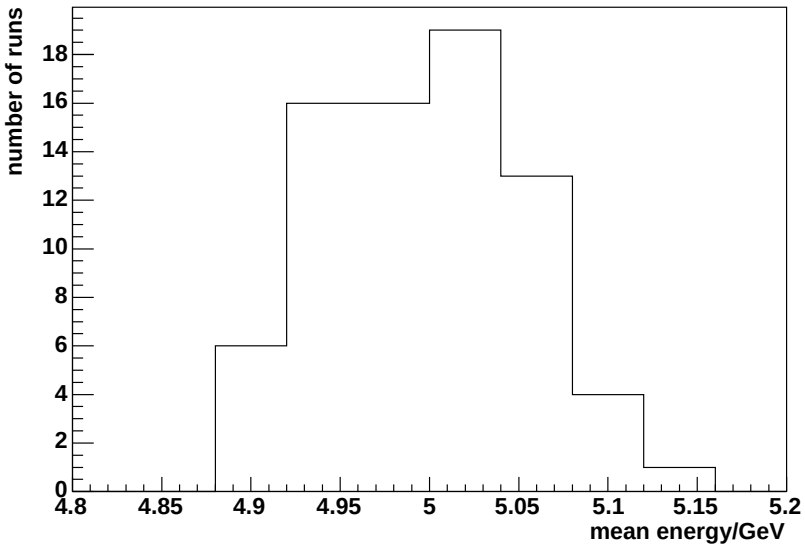

Figure C.14: *The energy uniformity from the three fine scans. The points from figure C.13 were filled into a histogram.*

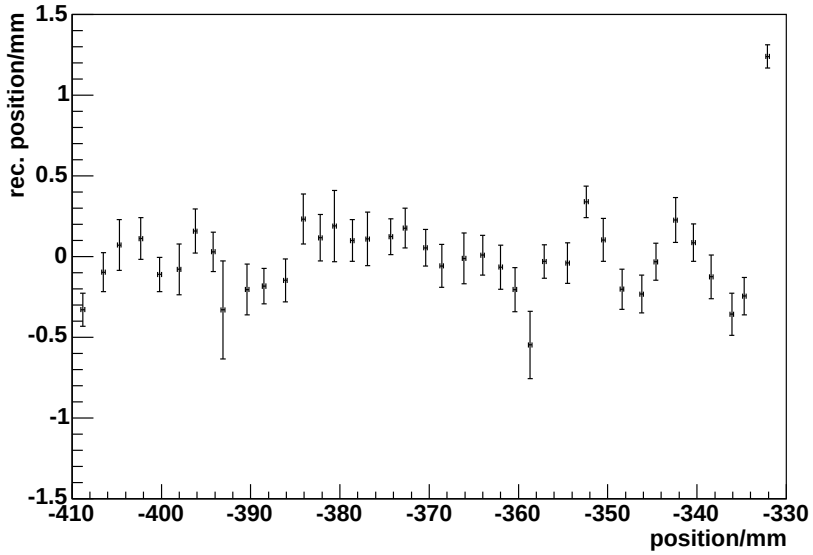

Figure C.15: *The deviations from linearity for the position reconstruction in x from the fine scan over the middle of the detector. The mean reconstructed position has been determined as a function of the beam position. To this a straight line has been fit and the absolute difference from that fit are shown.*

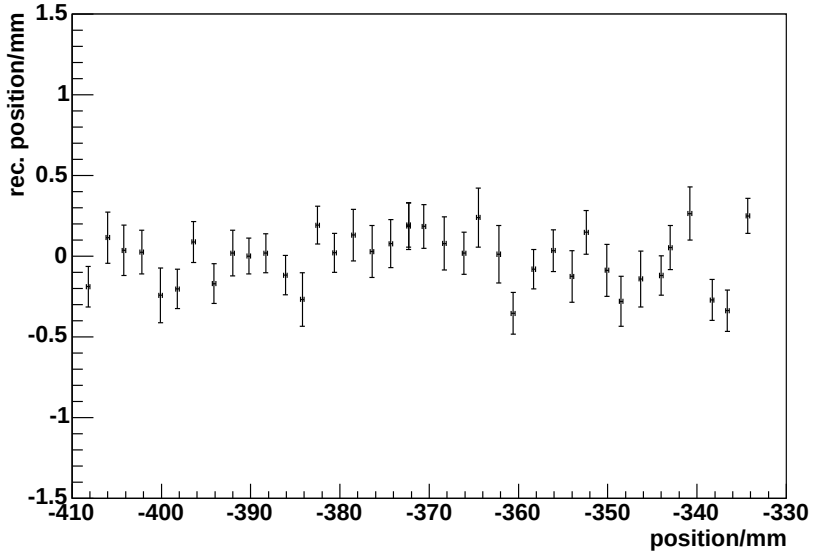

Figure C.16: *As figure C.15 but for the fine scan along the edge between middle cell row and the row above.*

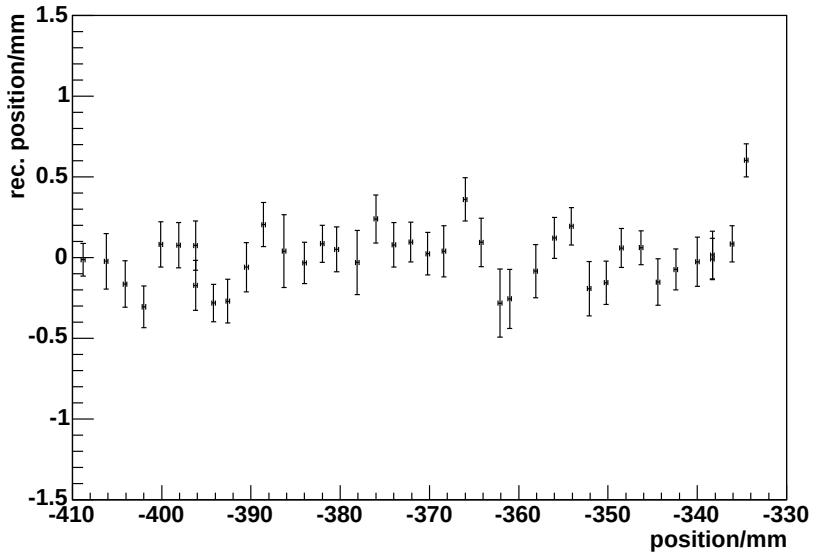

Figure C.17: *As figure C.15 but for the fine scan along the edge between middle cell row and the row below.*

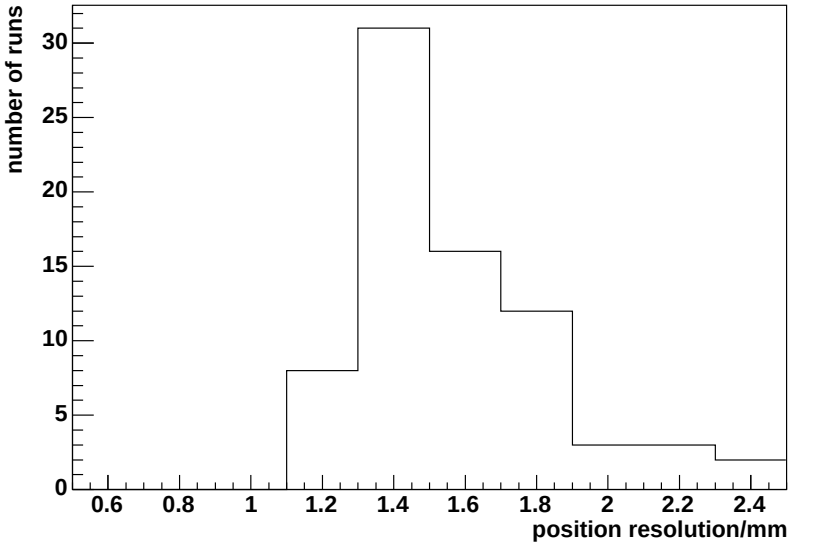

Figure C.18: *The position resolution in x from the three fine scans. For each beam position the RMS of the reconstructed x-position has been calculated and is taken as a resolution. The distribution of the resolutions is shown.*

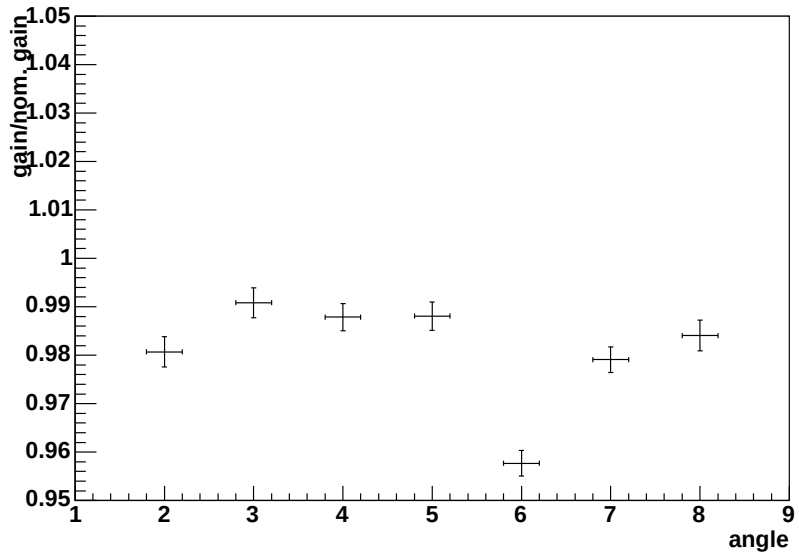

Figure C.19: *Gain as a function of the incident angle. For each angle an energy scan is done and the linear term of the linearity fit is shown.*

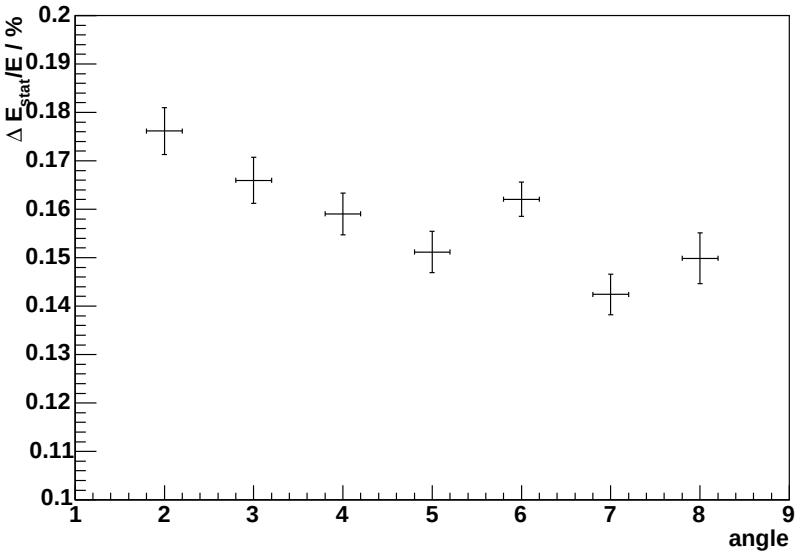

Figure C.20: *Statistical term of the resolution as a function of the incident angle. For each angle an energy scan is done and the statistical term of the resolution fit is shown.*

Bibliography

[1] ZEUS collab., U. Holm (ed), The ZEUS Detector Status Report (unpublished) available at `http://www-zeus.desy.de/bluebook/bluebook.html`

[2] N. Krumnack, diploma thesis, Testbeam Experiments and Monte Carlo Simulation of the ZEUS 6m Tagger, DESY thesis 01-033

[3] L. Li, Ph.D. thesis in preparation, Trijet in Neutral Current Deep Inelastic Scattering at HERA

[4] O. Lopez, Ph.D. thesis, Precise Determinations of the Strong Coupling Constant at HERA, DESY thesis 02-020

[5] N. Tuning, ZUFOs, ZEUS note 01-121

[6] D. Chapin, Ph.D. thesis, A Measurement of Dijet Production in Neutral Current Deep Inelastic Scattering with ZEUS at HERA, UMI-30-12414-mc, available at `http://amzeus.desy.de/ cflt/theses/Doug.ps.gz`

[7] G.F. Hartner, VCTRAK Briefing: Program and Math, ZEUS note 98-058

[8] M. Gell-Mann, Phys. Lett. **8** 214 (1964)
G. Zweig, CERN-8192/TH 401 (1964)
G. Zweig, CERN-8419/TH 402 (1964)

[9] M. Krawczyk et al., Proc. of the study for an *ep* facility for Europe, DESY79-48 391 (1979)

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

[11] J.D. Bjorken, Phys. Rev. **179** (1969) 1547
R.P. Feynman, Photon Hadron Interactions, W.A. Benjamin, New York (1972)

[12] L. Lönnblad, ARIADNE version 4, A Program for Simulation of QCD-Cascades Implementing the Colour Dipole Model, available at `http://www.thep.lu.se/ leif/ariadne/ariadne.pdf`

[13] G. Ingelman, A. Edin, J. Rathsman, LEPTO 6.5 A Monte Carlo Generator for Deep Inelastic Scattering available at `http://www3.tsl.uu.se/thep/MC/lepto/lepto-6.5.1.ps.gz`

[14] S. Catani, M.H. Seymour, A General Algorithm for Calculating Jet Cross Sections in NLO QCD, available at `http://hepwww.rl.ac.uk/theory/seymour/pubs/nlo.ps_gz`

[15] Z. Nagy, Z. Trócsányi, Multi-jet cross sections in deep inelastic scattering at next-to-leading order, hep-ph/0104315

[16] S. Eisenhardt, Ph.D. thesis, Measurement of the Strong Coupling Constant $\alpha(s)$ from Jet Rates Using the k_T Jet Algorithm at ZEUS, DESY thesis 98-038

[17] S. Catani et al., Longitudinally invariant k_T clustering algorithms for hadron-hadron collisions, Nucl. Phys., B406:187, 1993

[18] F. Halzen and A.D. Martin, Quarks and Leptons: An Introductory Course in Modern Particle Physics, John Wiley & Sons, Inc.

[19] R.P. Feynman, Photon Hadron Interactions, W.A. Benjamin, New York (1972)
K.H. Streng, T.F. Walsh and P.M. Zerwas, Z. Phys.C2 (1979) 237

[20] GEANT (CERN Long Writeup W5013; Application Software Group, Oct. 1994 Ed.)

[21] J. Pumplin et al., cteq6, JHEP, 0207:012, 2002

[22] S. Kananov, The noise study for the 1995 data in the ZEUS calorimeter, ZEUS Note 97-018
S.M. Wang, Study of Calorimeter Noise in the '95 Data, ZEUS Note 96-121
A. Savin, Study of Calorimeter Noise in the 1996 data, ZEUS Note 98-007
M. Martinez, Calorimeter Noisy Cell Study for 1998-2000 ZEUS Data, ZEUS Note 00-016

[23] R. Sinkus, T. Voss, DESY 96-264, Nucl. Inst. Meth **A 391**, 360 (1997)
A. López-Durán Viani, Ph.D. thesis, Measurement of the Proton Structure Function xF_3 in High Q^2 DIS Events Using ZEUS Data, DESY thesis 01-056

[24] H1 Coll., C. Adloff et al., Three-Jet Production in Deep-Inelastic Scattering at HERA, Phys. Lett. B515 (2001) 17-29 , 06/01