UNIVERSITÄT BONN
Physikalisches Institut

Measurement of $e^-p \to e^-X$ differential cross sections at high Q^2 and of the structure function xF_3 with ZEUS at HERA

von

Alexander Kappes

The $16\,\mathrm{pb}^{-1}$ of e^-p data taken with the ZEUS detector during the 1998/99 running period of HERA have been used to measure neutral current single- and double-differential cross sections in deep inelastic e^-p scattering in the range $200\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $3.7 \cdot 10^{-3} < x < 1$ at a center-of-mass energy of 318 GeV. Comparing the cross sections obtained to theoretical predictions based on the Standard Model and CTEQ5D PDFs yields good agreement. Particularly, no excess of data events at high Q^2 is visible. A comparison of the measured e^-p cross sections to those from e^+p clearly reveals the influence of the Z-boson contribution for $Q^2 \gtrsim M_Z^2$. This dependence is exploited to determine the mass of the Z boson, M_Z, and by combining the e^-p and e^+p cross sections the parity violating structure function xF_3 is extracted for the first time in deep-inelastic ep scattering at ZEUS.

Post address:
Nussallee 12
53115 Bonn
Germany

BONN-IR-2001-16
Bonn University
December 2001
ISSN-0172-8733

UNIVERSITÄT BONN
Physikalisches Institut

Measurement of $e^- p \to e^- X$ differential cross sections at high Q^2 and of the structure function xF_3 with ZEUS at HERA

von

Alexander Kappes

aus

Frankfurt am Main

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen.

Angenommen am: 20. Dezember 2001

1. Referent: Prof. Dr. E. Hilger

2. Referent: Prof. Dr. N. Wermes

Contents

Chapter 1

Introduction

Looking up into the sky in a clear and moon-less night opens a glimpse into the far depth as well as into the distant past of our universe. Stars, barely visible to the naked eye, sent their light into space millions of years ago. Mankind always wanted to learn more about the seemingly endless darkness and the few bright sparkling spots in it and therefore people started to build optical telescopes, the first one in 1608. Over the centuries the art of telescope building was more and more refined and also extended to wavelengths outside the visible spectrum. Today, it is even possible to bring telescopes into space, where no atmosphere is disturbing the sight, or to combine several large telescopes on earth and by exploiting interference reach unsurpassed resolution. With modern instruments like the Hubble space telescope it is possible to observe galaxies and quasars over 10 billion light-years away, which are almost at the edge of our universe as we understand it today, and hence disclose its origin. The cosmic-microwave background tells us of a long-gone time when atomic nuclei and electrons combined to atoms, just 300 000 years after the Big Bang[1]. But cosmologists want to go even further back into the past. And indeed, in todays high-energy physics experiments at accelerators here on earth it is possible to follow our universe back to a time when it was unimaginably hot and dense and consisted only of quarks, gluons and leptons, $1 \cdot 10^{-12}$ seconds after the Big Bang.

The question of the origin of our universe always has been one of the great mysteries of mankind and a stimulus for technological progress. Another big question has been about the fundamental building blocks and forces that form our world and last but not least our very selves. During recent decades it has become clear that these two questions are highly intertwined and particle physics combines them in a unique way.

Particle physics deals with the fundamental building blocks of matter and the forces between them. However, the perception of fundamentality changed drastically over the years. At the end of the 19th century the notion of the structure of matter was still rather incomplete. Though chemists had measured the weights of the known elements and Mendeleyev had produced his famous periodic table, the concept of atoms was not established yet. But then in a rather short

[1] Though today widely accepted as the best theory of the outset of our universe, the Big Bang model is not the only one and its implications like inflation are subject of intense discussions among cosmologists.

period of 80 years the complex standard model (SM) of particle physics as we know it today was developed with its leptons, quarks, gluons and the various forces that act between them.

In the history of particle physics, scattering of (so far) point-like particles like electrons off nucleons has always played an important role in our growing understanding of the structure of the nucleon. In contrast to reactions between two composite objects like in proton-proton scattering, where the complicated convolution of the two structures has to be taken into account, lepton-nucleon scattering gives rise to easier interpretations. In 1956 McAllister and Hofstadter at Stanford were able to measure the charge radius r of the proton by shooting electrons with an energy of 188 MeV onto a hydrogen target. This elastic ep scattering yielded $r = (0.74 \pm 0.24)$ fm. Only 10 years later, in the late 1960s, the two miles long linear accelerator at SLAC was able to produce electrons to energies of 18 GeV and opened a completely new energy domain for ep-scattering experiments. Inelastic scattering of these then high-energy electrons off protons lead to the experimental discovery of quarks and cleared the way to the invention of QCD in the 1970s.

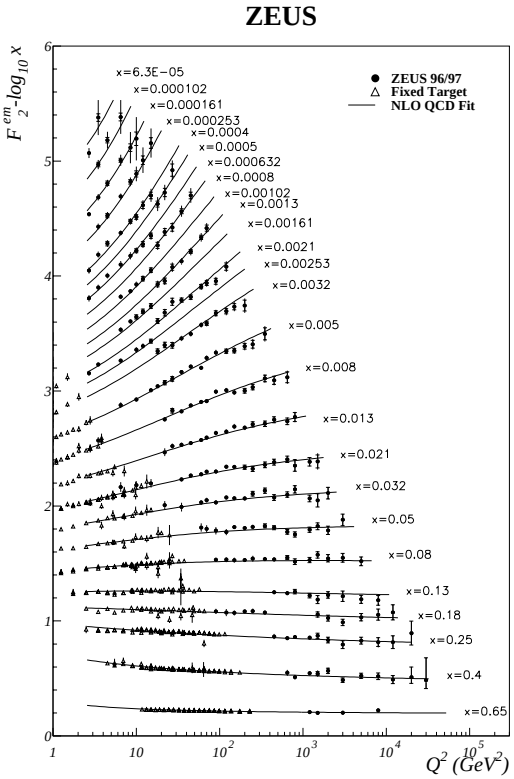


Figure 1.1: *The structure function F_2 as a function of Q^2 in bins of x.*

Until the early 1990s, electrons were always scattered from fixed targets which limited the center-of-mass energy to a few 10 GeV at most. Then, in the 1980s HERA was planned and built which, for the first time, accelerated both electrons/positrons and protons to energies of 27.5 GeV and 820 GeV, respectively, yielding in their collision a center-of-mass energy of 300 GeV, 20 times higher than ever before[2]. One of the major tasks for the collider experiments at HERA has been and will be the detailed investigation of the proton structure. This is primarily accomplished by the measurement of differential cross sections and the extraction of structure functions from the data. Already in the early HERA running, where positrons were collided with protons, the two experiments H1 and ZEUS observed a rather unexpected strong rise in the proton structure function F_2 towards low x that could not be accommodated by most models existing at that time [1]. With improved statistics the measurement of the structure function F_2 has reached a very high precision in a wide range of the kinematic plane of Q^2 and x, where Q^2 is the negative square of the four-momentum transfer between electron and proton and x is the proton momentum-fraction of the struck quark. The high precision of the measurements can be seen in Fig. 1.1, collecting all the ZEUS data obtained from positron-proton scattering [2–7].

[2] Since 1998, HERA accelerates protons to 920 GeV, yielding an ep center-of-mass energy of 318 GeV.

Then, in 1996 an observation by both experiments H1 and ZEUS in parallel caused particular excitement. In the high Q^2-region above $20\,000\,\mathrm{GeV}^2$ more events were found than expected (Fig. 1.2) [8]. The question was: Is this a sign for new physics beyond the SM, e.g. leptoquarks or SUSY, or a detector effect or just a fluctuation? The following years of data taking saw a decrease of the excess as no further events in the highest Q^2 region were found. However, it was not possible to conclusively answer this question. Hence, when HERA switched from positron- to electron-running in 1998, there was great eagerness to find out whether a similar effect could be observed in e^-p scattering, the more so as for $Q^2 \gtrsim 10\,000\,\mathrm{GeV}^2$ the e^-p cross section is considerably higher than that of e^+p.

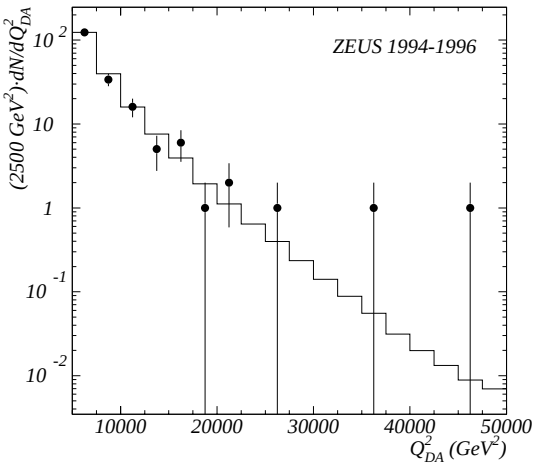


Figure 1.2: *Number of NC DIS events in a bin as a function of Q^2 for the 1994–96 e^+p ZEUS data.*

But with the availability of both e^-p and e^+p data another interesting and important measurement can be performed. The difference in the cross section of e^-p and e^+p scattering is caused by the parity-violating structure function xF_3. In order to measure this structure function for the first time in deep-inelastic ep scattering, high statistics samples in the high-Q^2 regime of both charge signs are mandatory.

This thesis deals with the data from electron-proton scattering collected by the ZEUS detector in 1998/99 at a center-of-mass energy of 318 GeV. The aim of the analysis is the measurement of single- and double-differential cross sections, the extraction of the structure function xF_3 as well as a determination of the mass of the Z boson.

The outline of this thesis is as follows: In the **second chapter**, the theoretical basics of particle physics relevant for this analysis are discussed. At first a brief introduction to the SM of particle physics and to the structure of the proton within this framework is given. Then, deep-inelastic scattering (DIS) is treated together with the notion of structure functions. The chapter ends with a discussion of diffractive reactions and various background processes relevant for this analysis.

The **third chapter** presents the experimental setup used. First, the HERA accelerator and the ZEUS detector are described, whereby detector components important for this analysis are discussed in separate sections. Next, concept and setup of the ZEUS trigger system are explained. The chapter ends with a section on event generators and detector simulation.

The subject of the **fourth chapter** is the event reconstruction. After the introduction of some pre-corrections, applied to energy deposits in the calorimeter, the calorimeter alignment (details in Appendix A) and the event-vertex reconstruction are discussed. The following section describes in detail the reconstruction of the scattered electron, containing identification, position reconstruction and energy corrections. Then, the reconstruction of the hadronic final state is

discussed, where different methods are introduced and compared. The chapter winds up with a comparison of the performance of different reconstruction methods of the kinematic variables.

The **fifth chapter** at first introduces the data and Monte Carlo (MC) samples used for this analysis. Next, the on- and offline event selections are discussed and detailed information on the offline cuts is given. Finally, data and MC are compared in various distributions in order to verify that the data is well described by the MC. Emerging deviations are discussed and where available explanations and corrections are presented.

The **sixth chapter** starts with a general discussion on the incorporation of systematic effects. Then the systematic effects relevant for this analysis are discussed.

The **seventh chapter** finally focuses on the results of the analysis. The chapter is divided into sections on the measurements of single- and double-differential cross sections. Each section contains, in addition to the results, a detailed description of binning, unfolding, efficiencies, purities and systematic checks. The e^-p cross sections obtained are then used in combination with the corresponding e^+p cross sections from [9, 10] to measure the Z mass and to extract the parity violating structure function xF_3.

A summary of this thesis is found in the **eighth chapter**.

Appendix A contains a detailed description of the calorimeter alignment performed in the course of the thesis work and **Appendix B** shows examples of NC DIS events. **Appendix C** displays the individual contributions of the systematic checks to the various cross sections, and **Appendix D** collects tables of the measured cross sections and of the extracted structure function xF_3.

Preliminary results of the analysis presented in this thesis have been published in contributed papers to the 1999 EPS conference in Tampere [11] and the 2000 ICHEP conference in Osaka [12]. The results were also presented by me at the 2000 DIS conference in Liverpool and have appeared in the corresponding proceedings [13]. A publication of final results by the ZEUS collaboration is planned for the near future.

Before starting on the analysis of the 1998/99 e^-p data I intensely worked half a year on the test setup for the readout chips and silicon strip detectors of the ZEUS micro vertex-detector (MVD) that was recently installed in the ZEUS detector. Together with Margherita Milite, another PhD student, I was responsible for the setup of the readout system and was able to perform first measurements that largely verified the expected performances of the components. During my stay at DESY I was for one year responsible for the quality monitoring of the recorded data from the detector. This task comprised also the coordination of the monitoring activities and the communication of upcoming problems to the respective experts within the collaboration. As a member of the ZEUS collaboration I participated in data-taking shifts. Throughout my PhD period I have been responsibly administrating the ZEUS-Bonn computer cluster. Today, the cluster, for which I have been decisively involved in the setup of the Linux server-client architecture, consists of 16 Linux-PC clients with a dedicated server and 7 Digital Alpha-stations.

Chapter 2

Theory

The objective of this chapter is the introduction to the theoretical basics of deep-inelastic scattering (DIS), where whenever necessary links to more detailed literature are given. The first part comprises a short introduction to the Standard Model (SM) and the electroweak and strong interactions contained therein. In the second part, DIS is treated in detail and the concept of structure functions is explained. The end of this chapter forms a section about background processes relevant for this analysis. Throughout this thesis $\hbar = c = 1$ is used.

2.1 The Standard Model (SM)

The SM contains our current best knowledge about the building blocks of matter and the mediating forces between them. Table 2.1 shows a summary of the particles and forces contained within the SM. The SM is based on group theory and the idea of symmetry that are closely related to each other. According to Noether's theorem, each continuous symmetry of nature corresponds to a conservation law and vice versa, e.g. invariance of physics laws under translation in time and space corresponds to conservation of energy and momentum, respectively. For example, looking at a wavefunction Ψ under a global phase transformation $\Psi(x) \rightarrow e^{i\alpha}\Psi(x)$ (α being a real constant) reveals that the Lagrangian

$$\mathcal{L} = i\overline{\Psi}\gamma_\mu\partial^\mu\Psi - m\overline{\Psi}\Psi \ , \tag{2.1}$$

leading to the Dirac equation, is invariant under such a transformation. The transformations $e^{i\alpha}$ form a group called $U(1)$, one says that $\mathcal{L}$ is invariant under $U(1)$ transformations. Applying Noether's theorem yields a conserved current that can be identified with the electromagnetic charge current, i.e. charge is conserved because of $U(1)$ phase invariance.

Looking at symmetries this way yielded deep theoretical insight into the structure of particle interactions and forms the basis for the theory of electromagnetic, weak and strong interactions. However, up to now the most prominent force in everyday life, the gravitation, refuses to be integrated into this scheme, though there has been progress throughout recent years. In group theory the SM can be represented by

$$SU_C(3) \times SU_L(2) \times U_Y(1) \ , \tag{2.2}$$

where the meaning of the different terms will be explained in the following sections.

	Leptons		Quarks	
el. charge [e]	-1	0	$^2/_3$	$-^1/_3$
1st fam.	electron (e)	elec. neutrino (ν_e)	up (u)	down (d)
2nd fam.	muon (μ)	muon neutrino (ν_μ)	charm (c)	strange (s)
3rd fam.	tau (τ)	tau neutrino (ν_τ)	top (t)	bottom (b)

	Forces				Higgs
	electromagnetic	weak		strong	
bosons	photon (γ)	$W^\pm$	Z	gluons (g)	H
mass of boson [GeV]	0	82	91	0	$\gtrsim 114$

Table 2.1: *Particles and forces contained in the Standard Model. Additionally to what is listed here, each lepton and quark has its corresponding anti-particle.*

2.1.1 Quantum electrodynamics (QED)

Quantum electrodynamics (QED) describes the interaction between electrically charged particles via the exchange of photons. The symmetry group of QED is the $U(1)$ group. It is invariant under local gauge transformations, implying that the Lagrangian, containing all information about the interaction, is invariant under a local gauge transformation $\Psi(x) \rightarrow e^{i\alpha(x)}\Psi(x)$ (a generalization of the $e^{i\alpha}$ phase in the previous section) [14], where $\Psi(x)$ is for example the wavefunction of an electron. The requirement of local gauge invariance introduces a massless vector field that can be identified as the photon. As the members of the $U(1)$ group, the phase transformations $e^{i\alpha(x)}$, commute with each other, the group is an Abelian group and hence the gauge bosons, i.e. the photons, do not interact with each other.

Cross sections of interactions can be calculated via Feynman diagrams [15, 16], which describe their amplitudes. These diagrams can be classified by their order, i.e. by the number of vertices they contain. Unfortunately, there exists an infinite number of such diagrams, as the number of vertices is not limited, so that the full higher-order corrections to a cross section can never be calculated[1]. Fortunately, the coupling strength at a QED vertex is proportional to $\sqrt{\alpha(\mu)}$ (μ is the renormalization scale of the process and will be discussed later in this paragraph), with $\sqrt{\alpha(0)} \approx \sqrt{1/137}$, and is much smaller than 1 for all experimental accessible μ. Hence, perturbation theory and with it the Feynman calculus can be applied. However, there is another problem in the form of singularities, originating from logarithmically diverging integrals over particle momenta circulating in loop diagrams like those depicted in Fig. 2.1. This problem can be overcome by introducing a cut-off scale μ. In the so-called renormalization, all interactions that happen on a time scale much smaller than $1/\mu$, i.e. in particular the loop divergencies, are

[1] The order to which the cross section is calculable depends strongly on the process. The highest-order calculation done so far was performed for the anomalous magnetic-moment of the muon and contains all Feynman diagrams up to the eighth order in electromagnetic and weak couplings. This calculation was necessary to comply with the precision of the recent $g-2$ experiment [17]. For most other processes, however, the highest order calculated is much lower.

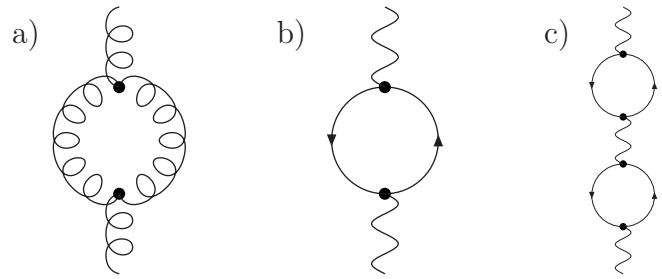


Figure 2.1: *Three examples for loop diagrams.*

absorbed into the coupling $\alpha(\mu)$ [18], which is therefore called a running coupling strength. The scale μ is called renormalization scale. In QED, the dependence of α on the renormalization scale is small.

Obviously, the gauge-theory approach of QED leads to problems when transferred to the weak interaction, as here the $W^{\pm}$ and Z gauge bosons with masses of 82 GeV and 91 GeV cannot be described by a massless vector-field like the photon in QED. On the other hand, an ad hoc introduction of mass terms for the gauge bosons would first spoil gauge invariance and, far worse, would lead to unrenormalizable divergences rendering such a theory rather useless [14]. The solution to this problem is called "spontaneous symmetry-breaking" and will be explained in the next section.

2.1.2 Spontaneous symmetry breaking and the weak interaction

In 1961 Glashow proposed to consider the electromagnetic and weak force not as two separate forces but rather as the different manifestations of a more fundamental combined electroweak force [19]. The obvious difference in strength between them was explained as being due to the large masses of the weak gauge bosons. However, the origin of these large masses was unclear until 1967, when Weinberg and Salam implemented the so-called Higgs mechanism into the electroweak theory [20, 21].

The Glashow-Weinberg-Salam (GWS) theory of the weak force assumes that the coupling of quarks and gluons to the weak bosons $W^{\pm}$ and Z is a mixture between a vector and an axial-vector coupling, where the axial-vector part violates parity [18]. For $W^{\pm}$ the coupling is $V - A$ and hence the vector and axial-vector parts are of equal strength, whereas for Z the relative and absolute strength depends on the type of coupling fermions. This special structure of the coupling has the consequence that the charged weak current (W bosons) couples only to left-handed particles (or right-handed anti-particles)[2], whereas the neutral weak current (Z boson) also involves right-handed particles. The underlying symmetry group is called $SU_L(2) \times U(1)$, where the subscript L refers to the fact that only left-handed particles are involved. However, this symmetry is spontaneously broken which finally leads to two massive charged bosons ($W^{\pm}$),

[2] Strictly speaking, this statement is only true for massless particles, as only under this condition the projection operator $1-\gamma^5$ that is applied to the fermion states picks out a helicity eigenstate. However, for energies $\gg M$, the mass M of the particle can be neglected and $1-\gamma^5$ is equal to a helicity operator.

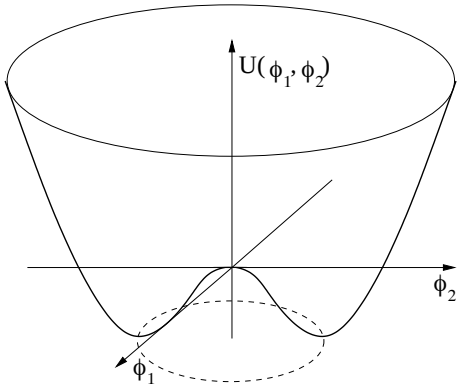


Figure 2.2: *Potential function of the Higgs field* $\phi = \phi_1 + i\phi_2$.

a massive neutral boson (Z) and a massless neutral boson (photon). This will be discussed in more detail in the next paragraph.

The concept of spontaneous symmetry-breaking was developed in the middle of the 1960s [22–24] and it was applied to the electroweak interaction by Weinberg and Salam in 1967. In a simpler form[3], this concept involves a complex scalar field $\phi = \phi_1 + i\phi_2$, which is also called the Higgs field. The potential energy $U(\phi_1, \phi_2)$ of this field is displayed in Fig. 2.2. The function is perfectly symmetric under rotation about the U axis if the Higgs field is located at the unstable equilibrium at $\phi_1 = \phi_2 = 0$. However, this symmetry gets spontaneously broken when the field falls into one of its ground or vacuum states, depicted by the dashed circle. By applying local gauge invariance to ϕ, again a gauge field A_μ is introduced, though this time together with an unwanted massless scalar boson, also called Goldstone boson [18]. However, this complication can be overcome by exploiting local gauge invariance and chosing an appropriate gauge for ϕ, leading to the disappearance of the Goldstone boson and the acquirement of a mass by A_μ.

If applied to $SU(2) \times U(1)$ with four fields A^i_μ the Higgs-mechanism yields three Goldstone bosons [14] that are absorbed by three of the fields. Hence, through spontaneous symmetry breaking it is possible to obtain masses (without spoiling gauge invariance) for the originally massless fields, which in turn form the heavy weak-bosons $W^\pm$ and Z.

2.1.3 Quantum chromodynamics (QCD)

Quantum chromodynamics (QCD) describes the interaction between quarks. In contrast to QED, QCD has $SU(3)$ as the gauge group where the quantum number charge is replaced by 3 color charges (green, blue and red). The strong color force is mediated by 8 gluons which are, in contrast to the photon in QED, themselves (color) charged and can therefore not only interact with the quarks but also with each other. This follows directly from the non-Abelian character of the $SU(3)$ gauge group. In QCD, the non-Abelian character also leads to the phenomenon

[3] The following description is given for a $U(1)$ rather than a $SU(2) \times U(1)$ gauge field, which would lead to a ϕ doublet instead of a singlet. As the line of argument stays the same, the simpler approach for a $U(1)$ gauge field is chosen.

that quarks can never be observed as free particles (confinement)[4] as the energy between two bound quarks, stored in the gluon flux, rises with the distance between them. If the energy in the gluon band gets large enough the gluon band rips and a quark anti-quark pair is produced, thus leading to two bound $q\bar{q}$ systems. On the other hand, the strong coupling becomes small asymptotically for large energy scales. This phenomenon is referred to as "asymptotic freedom".

Free observable particles have to be colorless, to be more precise they have to form a color singlet, and they occur in two configurations: quark triplets, which are called baryons and quark anti-quark doublets which are called mesons. It can be shown that these two quark combinations (qqq and $q\bar{q}$) have relatively large negative binding energy, whereas the combination of e.g. two quarks (qq) yields only weakly binding or even repelling forces [15].

2.1.3.1 Perturbative QCD (pQCD)

In contrast to QED, whose coupling strength α is small and (within the range accessible to experiments) only weakly depending on the renormalization scale, in QCD the strong coupling strength $\alpha_s(\mu)$ gets large for small scales and is of the order of 1 for $\mu \approx 1\,\mathrm{GeV}$, i.e. in regions of the ep kinematic plane accessible with HERA. The scale dependence of $\alpha_s(\mu)$ is determined by:

$$\alpha_s\left(\mu\right) = \frac{12\pi}{\left(11n - 2n_f\right)\ln\left(\frac{\mu^2}{\Lambda^2}\right)} \ , \tag{2.3}$$

where n is the number of colors, n_f is the number of "active" flavors (at HERA 3–5, depending on the scale μ) and $\Lambda \approx 200\,\mathrm{MeV}$. The latter denotes the energy at which the coupling gets too strong and perturbation theory breaks down.

In ep reactions the scale is defined by the momentum transfer between electron and proton at the proton vertex. The range of this analysis is bounded below by $185\,\mathrm{GeV}^2$ and hence pQCD is applicable here.

2.1.3.2 Factorization theorem of QCD

Another difficulty in the calculation of QCD processes arises from diverging, collinear gluon radiation. These "soft", long-range parts of the strong interaction are not calculable in QCD. This problem can be solved by factorizing out and absorbing the diverging parts of the soft interactions into the description of the incoming hadrons, where the latter then has to be determined from experiment. The cut-off parameter is called the factorization scale μ_f. With this method, the cross section $\sigma(P_1, P_2)$ of two colliding hadrons P_1 and P_2, depicted in Fig. 2.3, can be written as

$$\sigma(P_1, P_2) = \sum_{i,j} \iint dx_1 dx_2 \ f_{i/1}(x_1, \mu^2, \mu_f) \cdot \hat{\sigma}_{ij}\left(x_1, x_2, \alpha_s(\mu), \frac{\mu}{Q}, \frac{\mu_f}{\mu}\right) \cdot f_{j/2}(x_2, \mu^2, \mu_f) \ , \tag{2.4}$$

[4] In general, a non-Abelian gauge group does not necessarily lead to confinement, e.g. spontaneously broken $SU(2)_L$. Even in QCD there would be no confinement if more than 16 flavors existed (see (2.3)).

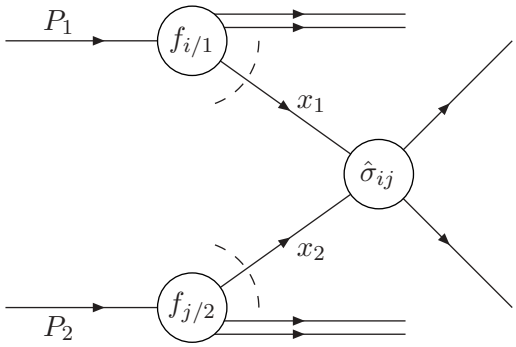


Figure 2.3: *Interaction of two hadrons P_1 and P_2. All soft interactions occurring left of the dashed arcs are absorbed into the probability functions $f_{i/1}$ and $f_{j/2}$, which describe the probability to find a parton i, j with momenta $x_{1,2}$ in hadrons $P_{1,2}$. $\hat{\sigma}_{ij}$ is the cross section of the hard interaction of the two partons with proton-momentum fractions x_1 and x_2.*

where x_1 and x_2 are the momentum fractions of the partons involved in the hard interaction, $f_{i/1,2}$ are the probability functions of the hadrons, describing the probability to find a parton i with momentum $x_{1,2}$ in hadron $P_{1,2}$, and μ is the renormalization scale. A priori, μ and μ_f do not have to be equal, but setting $\mu_f = \mu$ is a common choice. It can be proven, that in perturbation theory, factorization holds to all orders [25].

Apart from the diverging parts of the long-range interactions, also their finite parts can be (partially) absorbed into the parton distributions, leaving it up to the user which of these so-called renormalization schemes he chooses. Common schemes are the $\overline{\text{MS}}$ (modified minimal subtraction) [26] and the DIS [27] scheme, where the former is used more frequently. For ZEUS analyses, the DIS scheme is preferred, in which all finite contributions of order α_s or higher are absorbed into the description of the hadron (proton). This has the appealing consequence that the structure function F_2, to be discussed in Chap. 2.3.1, keeps its simple form in next-to-leading order (NLO). However, this does not hold for xF_3 which in NLO contains an integral over x in addition to the leading order (LO) term.

2.1.3.3 DGLAP splitting functions

In QCD a quark can radiate gluons and a gluon can split into a quark anti-quark pair. Due to the non-Abelian character of QCD a gluon can also split into two other gluons. The various splitting graphs of quarks and gluons are depicted in Fig. 2.4. As already pointed out in the previous section, these soft or long-distance interactions are absorbed into the structure functions. Hence, these structure functions cannot be calculated from first principles as they now contain non-perturbative contributions. However, the dynamic process of quark and gluon evolution can be described by the so-called DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equation, developed independently by several groups in the 1970s of the last century [28]:

$$t\frac{\partial}{\partial t}\begin{pmatrix} q_i(x,t) \\ g(x,t) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi} \sum_{q_j,\overline{q}_j} \int_x^1 \frac{d\xi}{\xi} \times \begin{pmatrix} P_{q_iq_j}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{q_ig}\left(\frac{x}{\xi},\alpha_s(t)\right) \\ P_{gq_j}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{gg}\left(\frac{x}{\xi},\alpha_s(t)\right) \end{pmatrix} \begin{pmatrix} q_j(\xi,t) \\ g(\xi,t) \end{pmatrix}, \qquad (2.5)$$

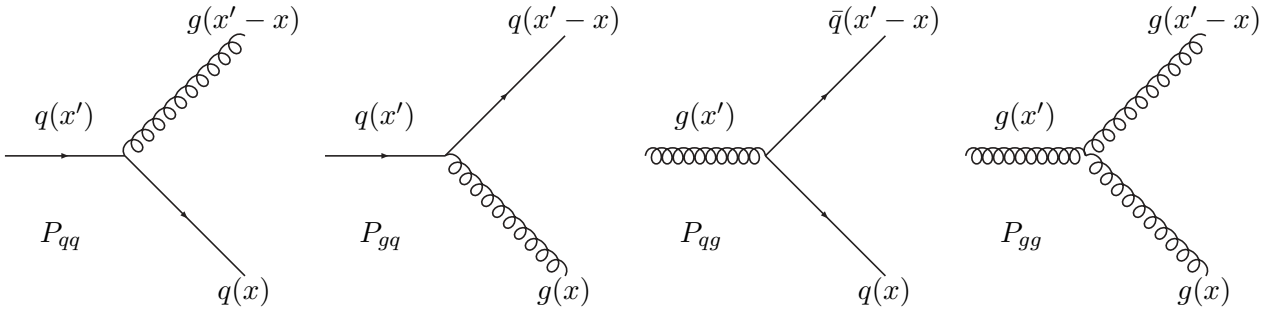


Figure 2.4: *Splitting graphs in lowest order of α_s.*

where t replaces the scale μ^2 at which the process is calculated and q_i and g are the probability or parton-density functions (PDFs) of quarks and gluons, respectively. Each splitting function P_{ab}[5] is calculable as a power series of α_s [25], where in LO P_{ab} is independent of α_s.

The splitting functions have been calculated to NLO in the $\overline{\text{MS}}$ scheme. However, only the LO is readily interpretable. To this order, the splitting function $P_{ab}(z)$ denotes the probability of finding a parton of type a in a parton of type b with a fraction z of the longitudinal momentum of the parent parton and a transverse momentum squared much smaller than μ.

Equation (2.5) is a powerful tool as it enables us to calculate the x dependence of a structure function at any desired Q^2 value, provided that one knows the x distribution at some value Q_0^2. The x distribution at Q_0^2 cannot be derived from first principles and has to be obtained from measurements.

2.1.3.4 The hadronic final state

After the hard scattering process the developing shower is perturbatively calculated down to an energy (for single particles) of the order of GeV. The following fragmentation of the generated quarks and gluons into hadrons occurs typically at a scale $\mu \approx 1/\Lambda$, i.e. with $\alpha_s \approx 1$. Therefore, this process is not calculable in perturbation theory. Fortunately, this hadronization process occurs at a much later time scale than the hard scattering process and therefore cannot alter its probability. Hence, the hard scattering process remains calculable in perturbation theory. For the hadronization process phenomenological models have to be applied.

2.2 The proton

The proton was observed for the first time by Wilhelm Wien in 1898 while performing experiments with ionized streams of gas atoms and molecules. However, he only realized that he observed a positive particle with the mass of the hydrogen atom, it was not clear to him that he observed its nucleus. Only 22 years later in 1920, Ernest Rutherford, who had experimented with alpha particles, came to the conclusion that what appeared to be hydrogen nuclei emerging

[5] Note that in LO α_s $P_{q_i q_j}$ is 0 for all $q_i \neq q_j$.

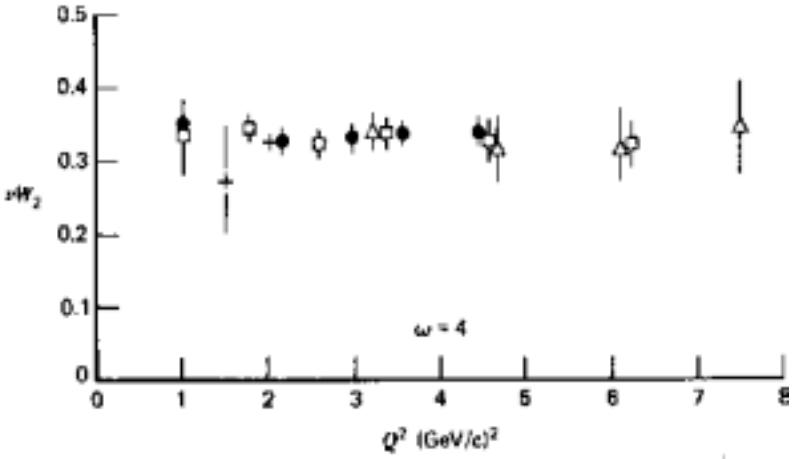


Figure 2.5. *Scaling behavior as seen by SLAC in 1968. The data were obtained by electron-proton scattering. The plot shows νW_2 as a function of Q^2 for $\omega = 4$, where ω is the inverse of the momentum fraction carried by the struck parton and νW_2 is the structure function. (taken from [14])*

from the collision of alpha particles with nitrogen atoms were indeed particles in their own right. He named them protons.

In the following decades the physicists attempted to resolve the structure of the proton. In order to calculate the amplitude of the scattering process of pointlike particles off the proton, the proton was treated as a pointlike Dirac particle, modified by so-called form factors that account for the spatial extension of the charge and the anomalous magnetic moment of the proton. Building on this theoretical background, McAllister and Hofstadter were able to determine the root-mean-square charge-radius $\sqrt{< r^2 >}$ of the proton in 1956 at the linear accelerator at Stanford university, California. In their experiment they collided electrons of $E = 188\,\mathrm{MeV}$ with protons in a hydrogen target and obtained $\sqrt{< r^2 >} = (0.74 \pm 0.24)\,\mathrm{fm}$.

In the 1960s, the new 2-mile linear accelerator at SLAC was able to accelerate electrons to energies between 7–17 GeV that were then directed on a fixed target. In 1969 an MIT-SLAC group investigated inelastic electron-proton scattering off a hydrogen target at then high four-momentum transfers [29, 30]. In contrast to expectation, the data (Fig. 2.5) showed only a weak Q^2-dependence.

Bjorken had anticipated this result based on a complex study, taking the proton to consist of pointlike particles. Independent of Bjorken, Feynman developed a proton model which is also based on pointlike particles or partons as he called them.

2.2.1 Static parton model of the proton

In the static or simple parton model, suggested by Richard Feynman in 1969, the proton consists of partons, where each parton carries a momentum fraction x ($x = \frac{1}{\omega}$) of the total proton momentum. According to this model, deep inelastic *ep*-scattering is the incoherent sum of the elastic scattering of electrons off quasi-free, static partons. Here, incoherent means that the electron scatters off a single parton and that the other partons within the proton, also called spectators, are not affected by this process. Quasi-free refers to the circumstance that the partons do not interact with each other during the scattering process. This is guaranteed by the relativistic speed of the proton with respect to the electron, which leads to a time dilation inside the proton in the reference system of the electron. Hence, the interaction speed of the partons

with each other is greatly reduced and the electron can scatter from a single parton without interference from the others. Finally, static means that the partons cannot be generated or destroyed.

This model was able to explain the SLAC data, however it failed to account for the scaling violation observed by experiments at CERN and Fermilab [31] in the meantime. Only with the theory of QCD, developed in the 1970s, could this scaling violation be explained. This led to the dynamic parton model.

2.2.2 Dynamic parton model of the proton

The dynamic parton model is based on QCD, hence, the quarks inside the proton can interact with each other via the exchange of gluons. Therefore, in contrast to the static parton model, the momentum of a quark can be altered via the radiation of a gluon or a quark-antiquark pair can be generated by a gluon (Fig. 2.4), leading to a much more complex picture of the proton. Apart from the three so-called valence quarks (two u and one d quark) that are mainly probed at low momentum transfer and high x ($x \gtrsim 0.1$), with increasing momentum transfer one finds a much larger number of sea quarks that are dynamically generated via the splitting processes described above. Consequently, the sea quarks dominate at low x values, producing the strong rise in the PDFs (Fig. 2.7) towards lower x. However, the sea-quark PDF (xS in Fig. 2.7) is not expected to rise indefinitely, as at some point recombination of quarks and antiquarks has to set in, forcing the PDFs to saturate.

2.3 Deep inelastic scattering (DIS)

DIS has always played an important role in the development and establishment of the SM and our understanding of the structure of nucleons. Scattering pointlike particles (e.g. electrons or neutrinos) from a complex object like the proton, has the advantage that the leptonic side of the reaction is well known and calculable within electroweak theory and that the lepton itself has no substructure that could obscure that of the complex object.

Electrons can interact with the quarks inside the proton in two ways. In the first case, a photon γ or a Z boson is exchanged (Fig. 2.6 a). Here, k and k' are the four-momenta of the incoming and outgoing electron, p is the four-momentum of the incoming proton, x is the momentum fraction of the quark struck by the electron and q is the four-momentum transfer between electron and quark. As the gauge bosons involved in this reaction carry no charge, this class of reactions is called neutral current (NC). In the second case (Fig. 2.6 b), a $W^{\pm}$ is exchanged where this class is referred to as charged current (CC) for obvious reasons. Here, the outgoing lepton is an electron neutrino.

As in both NC and CC reactions the struck quark and the proton remnant carry color, a color flux is observed between the jet and the proton remnant leading to the production of particles in the corresponding region of the hadronic final state.

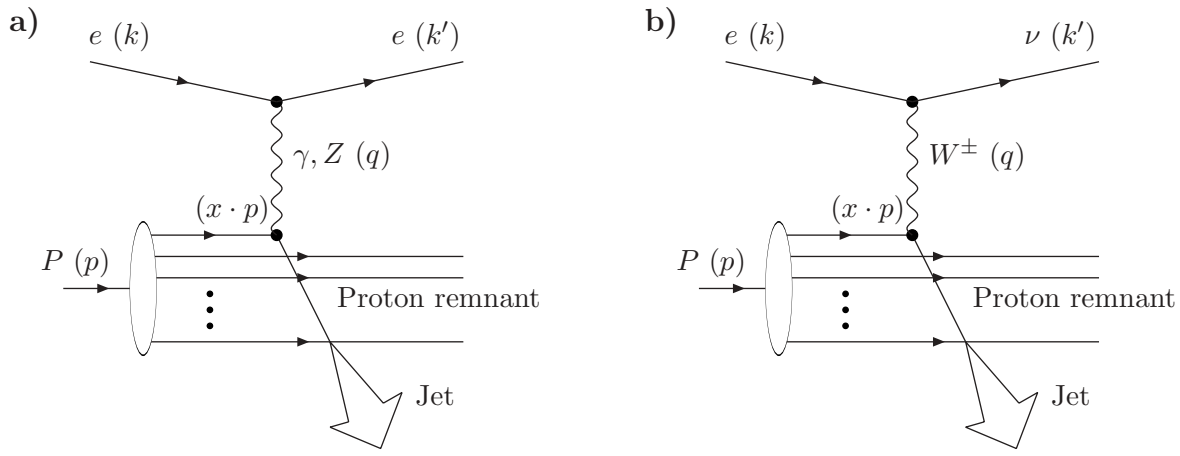


Figure 2.6: *Lowest order Feynman diagrams for electron-proton scattering for (a) neutral current and (b) charged current.*

For NC- and CC-reactions the following variables can be defined:

$$Q^2 = -q^2 = -\left(k - k'\right)^2 \tag{2.6}$$

$$x = \frac{Q^2}{2\,p \cdot q} \tag{2.7}$$

$$y = \frac{p \cdot q}{p \cdot k} \tag{2.8}$$

$$s = (k + p)^2 \tag{2.9}$$

$$Q^2 = s \cdot x \cdot y \ . \tag{2.10}$$

Q^2 is the virtuality of the exchanged boson and y can be interpreted (in this simple picture) as the energy fraction that is transfered from the electron to the quark in the proton rest-frame. $\sqrt{s}$ is the center-of-mass energy. For fixed $\sqrt{s}$ and for an inclusive measurement of the hadronic final-state, two of the four variables suffice to describe the reaction, with x and Q^2 being a common choice.

2.3.1 Structure functions

In general, the cross section of lepton-proton scattering can be factorized by

$$d\sigma \propto L^{\mu\nu} W_{\mu\nu} \ , \tag{2.11}$$

where $L^{\mu\nu}$ represents the leptonic tensor, describing the upper, well known leptonic vertex of Figs. 2.6 a,b, and $W_{\mu\nu}$ is the hadronic tensor, parameterizing our (partial) ignorance of the lower part of the diagrams[6]. The most general form of $W_{\mu\nu}$ for unpolarized protons, generated from

[6] Note that the picture of the interaction between lepton and proton given by (2.11) is based on the intention that one wants to probe the structure of the real proton with a virtual probe, i.e. the photon. This could also be turned around and the proton could be viewed as a source of virtual particles (quarks and gluons) that are used to probe the structure of the photon. However, this point of view is less well investigated up to now.

the only two independent four-momenta at the hadronic vertex, p and q, can be written as [32]

$$W_{\mu\nu} = -g_{\mu\nu}W_1 + \frac{p_\mu p_\nu}{M^2}W_2 - i\frac{\epsilon_{\mu\nu\alpha\beta}p^\alpha q^\beta}{2M^2}W_3 + \frac{q_\mu q_\nu}{M^2}W_4 + \frac{(p_\mu q_\nu + q_\mu p_\nu)}{M^2}W_5 + i\frac{(p_\mu q_\nu - p_\nu q_\mu)}{2M^2}W_6 \ , \quad (2.12)$$

where p is the four-momentum of the incoming proton, q is the four-momentum transfer between lepton and proton and M is the proton mass. The imaginary unit i in the W_3 and W_6 terms follows from the hermitian nature $W_{\mu\nu} = W^*_{\nu\mu}$ of the electromagnetic and weak currents. The W_i's are Lorentz scalar functions of q^2 and $q \cdot p$ and are called structure functions, as they tell about the structure of the proton.

If only photons are exchanged, the leptonic tensor $L^{\mu\nu}$ is symmetric and hence, the antisymmetric term W_6 in (2.12) drops out if (2.12) is inserted into (2.11). Also the W_3 term is absent as it violates parity (see (2.17)). Additionally, the conservation of the electromagnetic current requires $q^\mu W^{\text{em}}_{\mu\nu} = q^\nu W^{\text{em}}_{\mu\nu} = 0$ [14]. As a result it follows that

$$W^{\text{em}}_4 = \frac{p \cdot q}{q^2}W^{\text{em}}_2 + \frac{M^2}{q^2}W^{\text{em}}_1 \quad \text{and} \quad (2.13)$$

$$W^{\text{em}}_5 = -\frac{p \cdot q}{q^2}W^{\text{em}}_2 \ . \quad (2.14)$$

Thus, we are left with only two independent structure functions W_1 and W_2, and therefore (2.12) reduces to

$$W^{\text{em}}_{\mu\nu} = -W^{\text{em}}_1\left[g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}\right] + \frac{W^{\text{em}}_2}{M^2}\left[\left(p_\mu - \frac{q_\mu q \cdot p}{q^2}\right)\left(p_\nu - \frac{q_\nu q \cdot p}{q^2}\right)\right] \ . \quad (2.15)$$

Introducing the photoabsorption cross section with its two independent cross sections σ_S and σ_T for photons with helicity 0 (longitudinal, scalar) and ± 1 (transverse), respectively, it can be shown that [33]

$$W^{\text{em}}_1 \propto \sigma_T \qquad \text{and} \qquad W^{\text{em}}_2 \propto (\sigma_T + \sigma_S) \ . \quad (2.16)$$

Hence, W_1 depends only on the cross section for transversally polarized photons, whereas W_2 has contributions from both.

For the exchange of Z or $W^\pm$ bosons, the leptonic tensor $L^{\mu\nu}$ is not symmetric anymore and hence, W_6 does not vanish in (2.11). Additionally, W_3 is present as the weak force violates parity and W_4 and W_5 are not constrained, since the weak current is not conserved. But, W_4, W_5 and W_6 yield contributions to the cross section of the order of the lepton mass only [33] and can therefore be neglected for the energy range investigated in this analysis. In analogy to photon exchange it can be shown that for W/Z exchange

$$W^{\text{weak}}_1 \propto (\sigma_R + \sigma_L) \ , \quad W^{\text{weak}}_2 \propto (\sigma_R + \sigma_L + 2\sigma_S) \quad \text{and} \quad W^{\text{weak}}_3 \propto (\sigma_R - \sigma_L) \ , \quad (2.17)$$

where σ_R, σ_L and σ_S are the W/Z-absorption cross sections for right-, left-handed and scalar W/Z. Note the presence of the parity-violating W^{weak}_3 term. For the electromagnetic interaction $\sigma_R = \sigma_L$ due to parity invariance and thus $\sigma_T = \frac{1}{2}(\sigma_R + \sigma_L)$ and $W^{\text{em}}_3 = 0$.

For deep inelastic neutral-current electron-proton scattering the double-differential cross section can be written as

$$\frac{d^2\sigma^{\mathrm{NC}}(e^{\pm}p)}{dx\,dQ^2} = \frac{4\pi\alpha^2}{x\,Q^4}\left[y^2 x F_1^{\mathrm{NC}}(x,Q^2) + (1-y)\,F_2^{\mathrm{NC}}(x,Q^2) \mp \left(2y - y^2\right) x F_3^{\mathrm{NC}}(x,Q^2)\right] \ , \quad (2.18)$$

where $F_i \propto W_i$. Note that the sign of xF_3^{NC}, in contrast to F_2^{NC} and F_L^{NC}, depends on the charge of the incoming lepton. With the introduction of a new structure function $F_L^{\mathrm{NC}} = F_2^{\mathrm{NC}} - 2xF_1^{\mathrm{NC}}$, (2.18) can be written as

$$\frac{d^2\sigma^{\mathrm{NC}}(e^{\pm}p)}{dx\,dQ^2} = \frac{2\pi\alpha^2}{x\,Q^4}\left[Y_+ F_2^{\mathrm{NC}}(x,Q^2) \mp Y_- x F_3^{\mathrm{NC}}(x,Q^2) - y^2 F_L^{\mathrm{NC}}(x,Q^2)\right] \ , \quad (2.19)$$

where $Y_\pm = 1 \pm (1-y)^2$. $F_L \propto \sigma_S$ is called the longitudinal structure function. In LO of α_s the quarks inside the proton do not radiate gluons and have therefore no transverse momentum in the infinite momentum frame. As a consequence, they cannot absorb longitudinally polarized photons and hence $F_L = 0$ and $F_2 = 2xF_1$, with the latter also known as the Callan-Gross-relation [14].

Structure functions parameterize our (partial) ignorance of the proton structure. In LO α_s, F_2 and xF_3 lend themselves to an easy interpretation:

$$F_2(x,Q^2) = x \sum_{\text{f=flavors}} A_f(Q^2)\left[q_f(x,Q^2) + \overline{q}_f(x,Q^2)\right] \ , \quad (2.20)$$

$$xF_3(x,Q^2) = x \sum_{\text{f=flavors}} B_f(Q^2)\left[q_f(x,Q^2) - \overline{q}_f(x,Q^2)\right] \ , \quad (2.21)$$

i.e. they are linear combinations of the PDFs $q_f(x,Q^2)$ with prefactors $A_f(Q^2)$ and $B_f(Q^2)$, that are independent of x. Note that F_2 is the sum of quarks and antiquarks, whereas xF_3 is the difference between quarks and antiquarks. Hence, measuring xF_3 probes the valence quarks inside the proton. The prefactors $A_f(Q^2)$ and $B_f(Q^2)$ are defined as

$$A_f(Q^2) = q_f^2 - 2v_e v_f q_f P_z + (v_e^2 + a_e^2)(v_f^2 + a_f^2)P_z^2 \quad (2.22)$$

$$B_f(Q^2) = -2a_e a_f q_f P_z + 4v_e a_e v_f a_f P_z^2 \quad (2.23)$$

with

$$P_z = \frac{1}{4\sin^2\theta_W \cos^2\theta_W} \cdot \frac{Q^2}{Q^2 + M_Z^2} \ . \quad (2.24)$$

Here, q_f is the electric charge of the struck quark, $v_{e/f}$ and $a_{e/f}$ are the vector and axial-vector coupling-constants for the lepton e and the quark f, respectively, θ_W is the weak mixing angle and M_Z is the Z-boson mass.

As already mentioned in Chap. 2.1.3.3, the PDFs $q(x,Q^2)$ cannot be derived from first principles and hence, have to be determined from experiments. The parameterization of these PDFs is discussed in the next chapter.

2.3.2 Parameterization of structure functions

The parameterization of structure functions is a difficult task, as measurements from lots of experiments at different energy scales have to be combined into a fit. This fit is based on the DGLAP evolution equations, discussed in Chap. 2.1.3.3. The parton density functions, including the gluon, are parameterized as a function of x at a fixed Q_0^2 value of a few GeV (sometimes even below 1 GeV). Then, the evolution equations are applied to evolve the parton densities to higher Q^2 values. Note that the results of the fits depend on the renormalization scheme used. The most common schemes are the $\overline{\text{MS}}$ and the DIS scheme, with the latter exclusively used in this analysis. Over the years, several groups specialized in the extraction of parton densities.

2.3.2.1 CTEQ

The CTEQ group[7] (Coordinated Theoretical-Experimental Project on QCD) consists of roughly 20 members, both theorists and experimentalists. Their version of PDFs used for this analysis is CTEQ5D [34][8], where the 'D' refers to the DIS scheme. The fit includes DIS data-sets from BCDMS, NMC, H1, ZEUS, CCFR and E665, Drell-Yan data from E605 and E866, W-lepton-asymmetry data from CDF and inclusive jet data from DØ and CDF, where the ZEUS data comprises only the early results obtained until 1995. The different parton distributions are parameterized as follows [36]:

$$d_v(x,Q_0) = A_0 x^{A_1} \ (1-x)^{A_2} \ \left(1 + A_3 x^{A_4}\right) \tag{2.25}$$

$$u_v(x,Q_0) = A_5 x^{A_6} \ (1-x)^{A_7} \ \left(1 + A_8 x^{A_9}\right) \tag{2.26}$$

$$\left(\overline{u} + \overline{d}\right)(x,Q_0) = A_{10} x^{A_{11}} (1-x)^{A_{12}} \left(1 + A_{13} x^{A_{14}}\right) \tag{2.27}$$

$$\left(\overline{d} - \overline{u}\right)(x,Q_0) = A_{15} x^{A_{16}} (1-x)^{A_{17}} \left(1 + A_{18}\sqrt{x} + A_{19} x\right) \tag{2.28}$$

$$S(x,Q_0) = A_{20} x^{A_{21}} (1-x)^{A_{22}} \left(1 + A_{23} x^{A_{24}}\right) \tag{2.29}$$

$$g(x,Q_0) = A_{25} x^{A_{26}} (1-x)^{A_{27}} \left(1 + A_{28} x^{A_{29}}\right) \ , \tag{2.30}$$

where A_i are the fit parameters, u_v and d_v denote the valence quarks, S the sea quarks without $\overline{u}$ and $\overline{d}$ and g the gluon distribution. For the fit only data points with $Q^2 \geq 4\,\text{GeV}^2$ are included. Q_0 is fixed to 1 GeV.

2.3.2.2 MRST

The MRST group[9] (Martin, Roberts, Stirling and Thorne) [37] uses a similar approach as the CTEQ group, but for their 1999 global analysis[10] they use a more complex gluon parameterization in order to obtain a better fit to the new HERA data from 1996/97 [39].

$$xg(x,Q_0) = A_g (1-x)^{\eta_g} (1 + \epsilon_g x^{0.5} + \gamma_g x) x^{\delta_g} - A_- (1-x)^{\eta_-} x^{-\delta_-} \ , \tag{2.31}$$

where A_g, η_g, ϵ_g, γ_g, δ_g, A_-, η_- and δ_- are free parameters to be determined from the fit.

[7] `http://www.phys.psu.edu/~cteq`
[8] Lately, CTEQ6 has been released which includes now also the 1996/97 data from H1 and ZEUS [35].
[9] `http://durpdg.dur.ac.uk/hepdata/pdf`
[10] In the meantime a new set of PDFs has been released which includes the 1996/97 H1 and ZEUS data [38].

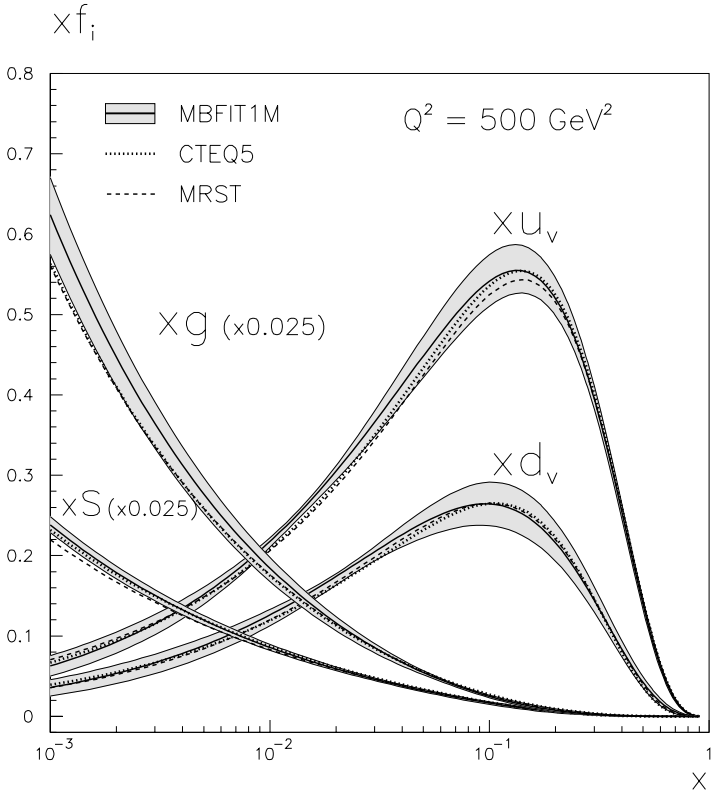


Figure 2.7. *Parton density functions obtained from a QCD fit to HERA and fixed target structure function data for $Q^2 = 500\,GeV^2$. xu_v and xd_v denote the valence quarks, xg the gluon and xS the sea quark and anti-quark distributions. The shaded band reflects the uncertainty on the PDFs and is obtained via the fit from the errors on the data points. The fit result is compared to CTEQ5D and MRST distributions. Note that the PDFs for sea quarks and gluons are both scaled down by a factor 40.*

2.3.2.3 ZEUS NLO QCD fit

Though several PDF parameterizations existed for quite some time, the uncertainty on these distributions has not been an issue until recently. With the availability of more and more precision data and in view of the upcoming Tevatron II and LHC era, the need for scrutiny of these errors becomes more and more obvious. The ZEUS NLO QCD fit [40, 41] has been one of the first QCD analyses that attempted to provide realistic errors on the parton distributions and to make them easily accessible for the user. The analysis uses the same kind of parameterization as the CTEQ and MRST groups and the input scale of the DGLAP evolutions (Chap. 2.1.3.3) is set to $Q_0^2 = 4(7)\,\mathrm{GeV}^2$ for [40] ([41]). Figure 2.7 shows the parton density functions obtained for $Q^2 = 500\,\mathrm{GeV}^2$ together with their uncertainty bands. Good agreement between all parameterizations is observed. Note that the sea-quark and gluon distributions are scaled down by a factor 40. For $x = 10^{-1}$ the ratio between the number of valence quarks $(u + d)$, sea quarks (xS) and gluons (g) is $0.74 : 0.8 : 1.2$ while for $x = 10^{-2}$ it is $0.65 : 3.2 : 6.8$. Thus, for $x \lesssim 0.1$ and high Q^2 the leptons are dominantly scattered from sea quarks rather than valence quarks, and by far the most frequent partons inside the proton are not quarks but gluons.

2.3.3 Radiative corrections to the cross section

Equation (2.19) only holds in lowest order QED. In order to fit the data, radiative corrections have to be taken into account originating either from real photons (Fig. 2.8 a – c) or from internal loops containing photons or other electromagnetically interacting particles (Fig. 2.8 d – f). The

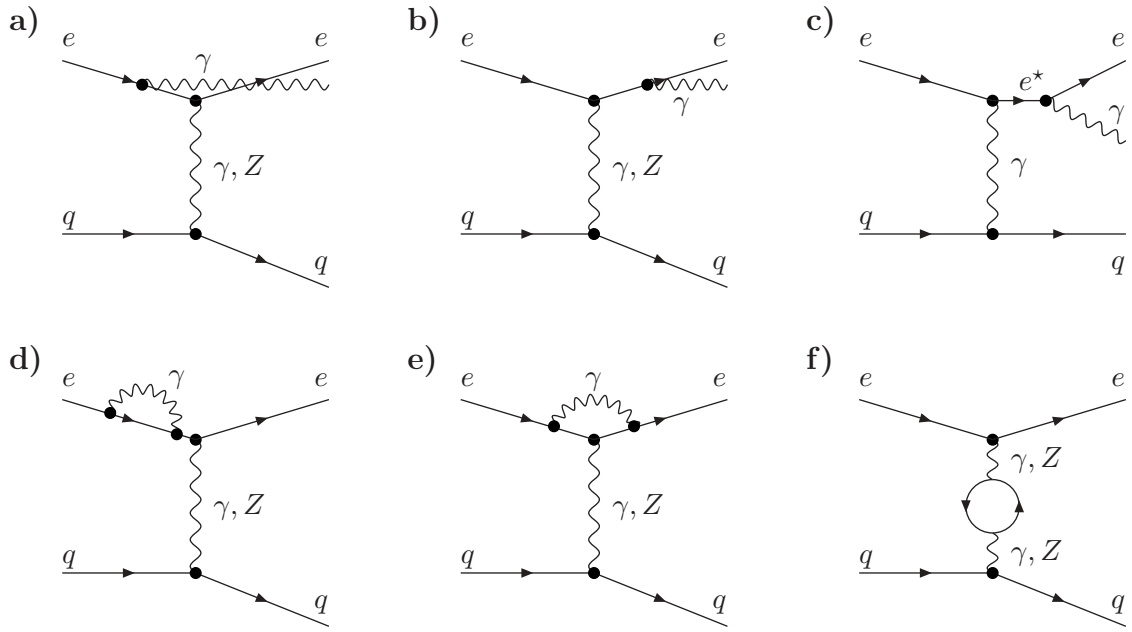


Figure 2.8: *Examples of radiative processes in DIS: (a) initial state radiation, (b) final state radiation, (c) QED-Compton scattering and (d)–(f) virtual loop diagrams.*

former are of the order $\mathcal{O}(\alpha^3)$ whereas the latter are of the order $\mathcal{O}(\alpha^4)$. Nevertheless, processes of $\mathcal{O}(\alpha^4)$ can contribute to $\mathcal{O}(\alpha^3)$ via interference with $\mathcal{O}(\alpha^2)$.

All photon radiations can occur from both the lepton and the quark line. However, the largest contributions come from real photons that are radiated from the lepton lines, i.e. Fig. 2.8 a – c [42]. In principal, the photon can be radiated in all directions with respect to the lepton line. But it turns out that in DIS the angular distribution has three poles, that are called initial-state radiation (ISR), final-state radiation (FSR) and QED-Compton scattering (QEDC) [43, 44]. In the first two cases a photon is radiated collinearly from the incoming or outgoing electron, respectively (Fig. 2.8 a and b). In the third case, the photon is radiated from the virtual electron which exchanges a quasi-real photon with the proton (Fig. 2.8 c).

All radiative corrections are either absorbed into the coupling (Fig. 2.8 e and f) and the mass terms (Fig. 2.8 d) or they are incorporated as a term δ_r into (2.19)

$$\frac{d^2\sigma^{\mathrm{NC}}}{dx\,dQ^2} = \left(\frac{d^2\sigma^{\mathrm{NC}}}{dx\,dQ^2}\right)^{\mathrm{Born}} \cdot (1+\delta_r) \ , \tag{2.32}$$

where the index Born denotes the cross section without radiative corrections.

2.4 Diffractive reactions

Diffractive events are characterized by the scattering of the exchanged photon off an object inside the proton with the quantum numbers of the vacuum, which is called pomeron. The pomeron has its origin in Regge theory [45], which is a phenomenological model that successfully describes

the hadronic cross sections in the low Q^2 region up to a few GeV^2 but fails at medium or high Q^2. Nevertheless, diffractive events with a clear event topology, explained in the next paragraph, occur up to $Q^2 \approx 1000\,\mathrm{GeV}^2$.

At first, diffractive events are part of the inclusive NC DIS cross section. However, they have a different event topology as there exists no color flux between the proton and the colorless pomeron. Hence, there is a region between the proton and the rest of the hadronic final state where no particles are produced. The percentage of diffractive events (with a clear event topology) in the total NC DIS cross section decreases with increasing Q^2. In the Q^2 region around $200\,\mathrm{GeV}^2$ the fraction amounts to roughly 10%. In this analysis the amount of diffractive events with a distinct event topology rather than the absolute differential cross section is relevant (Chap. 5.2.2).

2.5 Background processes

As the NC DIS event-selection chain relies heavily on the identification of the scattered electron, a background process to NC DIS has to generate an isolated electromagnetic cluster in the calorimeter with a transverse momentum of $p_t \gtrsim 5\,\mathrm{GeV}$, and if the cluster direction falls inside the acceptance of the CTD additionally a matching track. These strong requirements can only be met if the background process either has a cross section much larger than that of NC DIS or contains a (isolated) high-p_t electron. Studies [9] showed that only two processes give a significant contribution background to the NC DIS sample. They are discussed in the next two sections.

2.5.1 Photoproduction

In general, photoproduction (PHP) events feature a real photon as the exchange boson, i.e. $Q^2 = 0\,\mathrm{GeV}^2$. However, ZEUS assigns to PHP all those events where the scattered electron escapes the detector through the rear beam-pipe, which limits Q^2 to $\lesssim 3\,\mathrm{GeV}^2$. Hence, no scattered electron can hit the calorimeter. In order to be accepted as an NC DIS event the final state must fake an isolated high-p_t electron. As the exchanged four-momentum between electron and proton is very small, this happens only very sporadically, and only the huge cross section of PHP which is many orders of magnitude higher than the NC DIS cross section for $Q^2 > 100\,\mathrm{GeV}^2$ leaves a significant contribution after all selection cuts.

There exist two sorts of PHP events, direct (Fig. 2.9 a) and resolved/hadronic (Fig. 2.9 b). In direct PHP the photon interacts as a whole with the parton inside the proton, whereas in resolved PHP the photon fluctuates into a $q\bar{q}$ pair. If this fluctuation lives long enough it can form a bound $q\bar{q}$ state (vector meson) that interacts afterwards with the partons inside the proton (hadronic PHP). This interaction can be described with the VDM (Vector Dominance Model) [46].

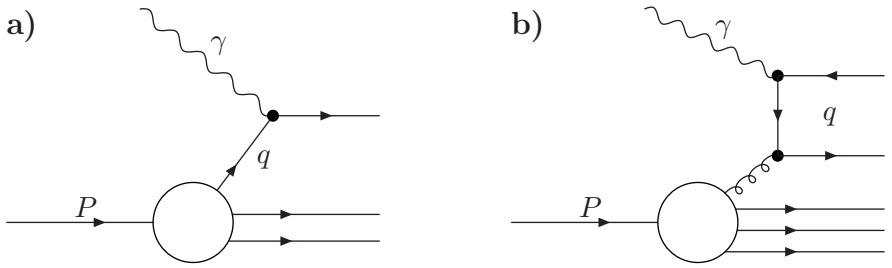


Figure 2.9: *Lowest order Feynman graphs for (a) direct and (b) resolved/hadronic PHP.*

2.5.2 QED-Compton scattering

Together with initial- and final-state bremsstrahlung, QED-Compton events (Fig. 2.8 c) [43] are poles of the distribution describing the radiation probability of a photon from the lepton lines. The class of QED-Compton events features a quasi-real photon as the exchange boson and a scattered electron that is deflected into the calorimeter by radiating a hard photon. QED-Compton events can be both elastic and inelastic, i.e. in the former case the proton stays intact whereas in the latter case it splits up in the interaction.

Because of the presence of the radiated photon, the hadronic final state is ill-reconstructed, which can lead to faked events in the very high Q^2 regime. QED-Compton events belong to the potentially "dangerous" background as they contain an isolated electron which cannot be distinguished from that of an NC DIS event.

Chapter 3

Experimental setup

In this chapter, the apparatus exploited for this analysis are introduced, starting with a description of the accelerators. Then the setup and components of the ZEUS detector are explained, where instruments important for the analysis are dealt with more extensively in separate sections. The chapter ends with a brief section on the role of event generators and detector simulations in todays particle-physics experiments and their implementation in the ZEUS environment.

3.1 HERA

The Hadron-Elektron-Ringanlage (HERA), shown in Fig. 3.1, is located at the Deutsches-Elektronen-Synchrotron (DESY) in Hamburg, Germany. It was planned in the 1980s and finished in 1992, and allows the investigation of the proton structure at resolutions down to 10^{-18} m. HERA has a circumference of 6.3 km and consists of two separate rings for electrons[1] and protons, respectively. The electrons are injected from the pre-accelerator-chain, consisting of LINAC, DESY II/III and PETRA (the latter formerly the Positron-Elektron-Tandem-Ringanlage) into HERA with an energy of 12 GeV, where they are further accelerated to an energy of 27.5 GeV. The protons, coming from a separate LINAC via DESY II/III and PETRA, are injected with an energy of already 40 GeV and are accelerated to 820 GeV and, since 1998, even to 920 GeV. Protons and electrons within HERA are lumped into bunches, where a bunch consists of approximately $2 \cdot 10^{10}$ particles. Each bunch occupies one of 220 free buckets around the ring, whereas at maximum only 210 can be filled at the same time. Hence, the time between two succeeding collisions amounts to just 96 ns.

In the beginning of HERA running in 1992, e^- were filled into the ring. In June 1994 it was decided to switch to e^+ as there were severe problems with the vacuum in the case of e^-, leading to much too short lifetimes of the lepton beam. In 1998, when these problems were finally solved, HERA was operated again with e^-, however only for 9 months. In spring 1999, it was decided to switch again back to e^+ to maximize statistics, in order to further investigate effects potentially

[1] In this chapter, the term electron refers to both electron and positron if not stated otherwise.

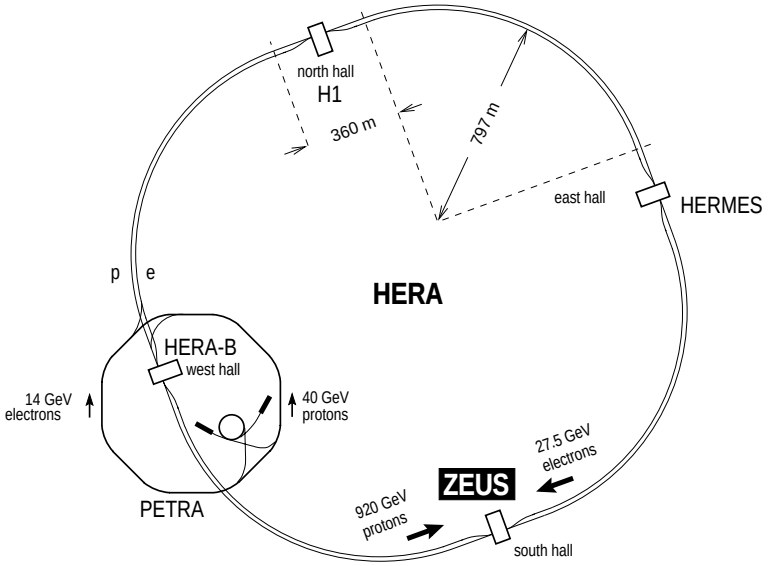


Figure 3.1. *HERA with its pre-accelerators DESY II/III and PETRA and the two collider experiments ZEUS (south hall) and H1 (north hall).*

coming from physics beyond the SM. It is planned to operate HERA with e^- after the upgrade of HERA and the collider experiments (Chap. 8), i.e. early 2002.

Due to the large mass difference between electrons and protons, the two HERA rings have quite different designs. The synchrotron-radiation energy-loss ΔE of a particle with mass m_0 and energy E during one circulation in a ring with radius R amounts to [47]

$$\Delta E \propto \frac{1}{m_0^4}\frac{E^4}{R} \ . \tag{3.1}$$

Therefore, electrons due to their small mass loose a lot of energy by synchrotron radiation in the curved sections of the ring. Hence, superconducting cavities are needed to compensate for these losses. On the contrary, the heavy protons loose only little energy by synchrotron radiation and therefore normal conducting cavities are sufficient. However, the energy of 920 GeV is large enough that superconducting magnetic coils are needed to keep the protons inside the ring.

HERA has four halls for experiments. In two of them leptons and protons are collided. In these two halls the ZEUS and H1 detectors are located. The other two halls house HERMES and HERA-B, which use only the lepton and proton beam, respectively.

3.2 ZEUS

The ZEUS collaboration consists of approximately 450 physicists from 52 different institutes located in 12 countries. Together, they have built and maintain a large, multi-purpose particle detector (Figs. 3.2 and 3.3) located in the south hall of the HERA accelerator. The main ZEUS detector [48] has a size of $15 \times 10 \times 10\,\mathrm{m}^3$ and a total weight of the order of 1000 tons. The coordinate system is a right-handed one, where the z-axis is defined along the proton-beam direction. The x-axis points to the center of HERA and the y-axis upwards, hence, in this coordinate system *large polar angles* θ correspond to *small scattering angles* of the electron.

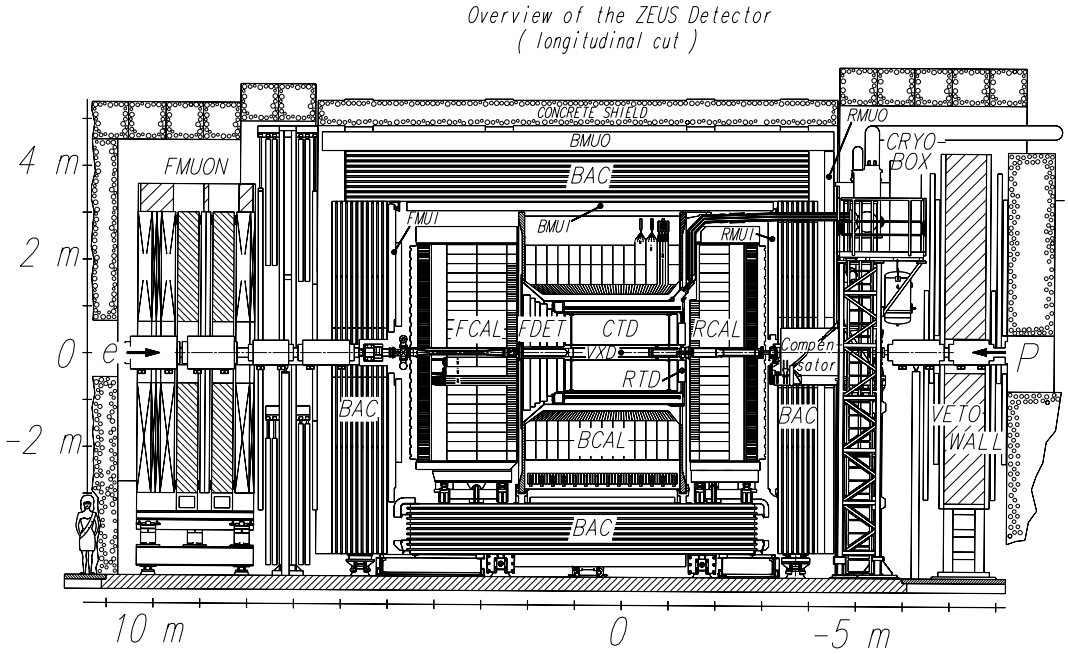


Figure 3.2: *Side-view of the ZEUS multi-purpose particle detector at HERA. Electrons enter the detector from the left and protons from the right.*

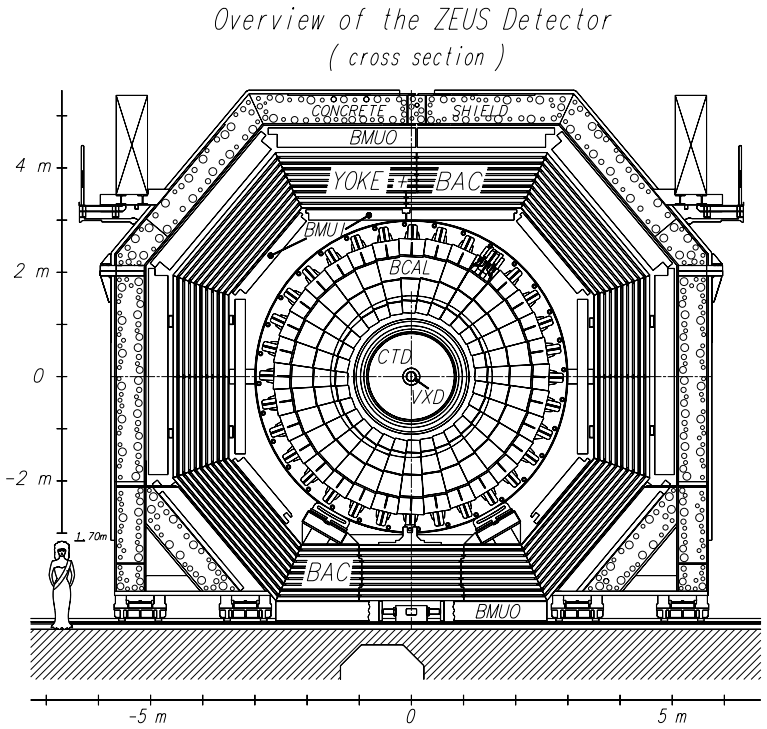


Figure 3.3: *Front-view along the z-axis of the ZEUS multi-purpose particle detector at HERA.*

The design of ZEUS is typical for recent high-energy collider experiments. Starting out from the middle of the detector with the beam pipe, tracking detectors (CTD, FDET, RTD)[2] measure the momenta and angles of charged particles coming from the interaction point (IP). Thereafter, most of the particles are stopped in the calorimeters (F/B/RCAL) and their energy is measured. However, apart from neutrinos that cannot be detected at all, muons can traverse the calorimeter and are detected in the inner and outer muon chambers (F/B/RMUI and F/B/RMUO) surrounding the calorimeter. Between the inner and outer muon chambers, the backing calorimeter is located. It consists of the instrumented iron yoke which closes the magnetic field lines of the superconducting solenoid, mounted between CTD and calorimeter and generating a magnetic field of 1.43 T.

There exist many more detector components not shown in Figs. 3.2 and 3.3, reflecting the broad physics scope covered by ZEUS analyses. In the following all those components are listed that are not used for this analysis:

FDET: The Forward-DETector is a set of 3 planar drift chambers, separated by 2 Transition Radiation Detectors (TRD). It is located in front of the FCAL and its purpose is to extend the tracking acceptance beyond the CTD towards forward angles. However, the detector suffers from high track multiplicities from the hadronic system in the forward region which renders the reconstruction of single tracks, e.g. originating from scattered electrons, highly non-trivial. Therefore, the FDET is not used for this analysis.

To improve the tracking performance of the FDET in the future, the TRDs were taken out during the upgrade period and are replaced by Straw-Tube-Trackers (STT) [50].

RTD: The Rear-Tracking-Detector consists of a single planar drift chamber and is placed in front of the RCAL. Like the FDET, its purpose is the extension of the tracking range of the detector, this time towards large polar angles corresponding to small electron scattering-angles. However, electrons in this region have a Q^2 too low for this analysis, and hence the RTD is not used here.

SRTD: The Small-angle-Rear-Tracking-Detector is placed between RTD and calorimeter and consists of two crossed layers of silicon-strip detectors, each 5 mm thick, yielding a position resolution of 3 mm [51]. However, it covers only an area of $35 \times 35\,\mathrm{cm}^2$ around the beam-pipe and can therefore, like the RTD, not be used for this analysis.

MUO: The MUOn-chamber system consists of two layers of limited streamer tubes, one between calorimeter and BAC and the second outside the BAC.

BAC: The BAcking-Calorimeter is integrated into the yoke and is mainly used to measure energy leakage from the main calorimeter.

BPC: The Beam-Pipe-Calorimeter is a tungsten-scintillation calorimeter. It is positioned $\sim 3\,\mathrm{m}$ behind the IP in the electron-beam direction and is used as an extension of the RCAL to smaller electron scattering-angles. Though it covers only a small region in ϕ it allows to extend the Q^2 range of DIS measurements down to values as low as $0.1\,\mathrm{GeV}^2$ [51].

[2] The vertex detector (VXD), plotted in Figs. 3.2 and 3.3, was removed in the winter shutdown 1995/96. The empty space is filled with a silicon micro-vertex-detector[49] (MVD) in the course of the upgrade program of the ZEUS detector.

BPT: The Beam-Pipe-Tracker [52] was installed in front of the BPC in 1997 and consists of two layers of silicon-micro-strip-detectors, separated in z by 27 cm. It allows to improve the BPC measurements and to extend them down to Q^2-values of $0.04\,\mathrm{GeV}^2$ [53].

LPS: The Leading-Proton-Spectrometer consists of several stations distributed over 65 m down the proton beam, where each of these stations contains silicon-strip detectors. The system is used to measure particles that leave the IP under small angles in forward direction, e.g. protons in deep-inelastic diffractive ep-scattering [54].

FNC: The Forward-Neutron-Calorimeter is located 106 m down the proton beam and can be used to measure the parton densities in the pion by detecting the neutron left over after the interaction of the exchanged photon with the pion inside the proton.

FPC: The Forward-Plug-Calorimeter was installed in 1998 and closes the remaining gap between beam-pipe and FCAL. With its help the energy flow in the very forward direction can be measured more accurately. However, studies [55] showed that the FPC does not improve the reconstruction of the hadronic final state in this analysis and therefore the FPC data is not included here.

PRT: The Proton-Remnant-Tagger consists of 7 pairs of scintillation counters surrounding the beam pipe at 5 m (2 pairs), 23 m (1 pair) and 24.5 m (4 pairs). It is for example used to distinguish between diffractive and non-diffractive scattered protons.

3.2.1 Uranium calorimeter

The uranium-scintillator calorimeter is one of the most important components for the investigation of ep-physics at ZEUS. Correspondingly high are the requirements:

- good hermeticity,
- good energy resolution for electrons and hadrons,
- calibration of the absolute energy scale to 1%,
- an angular resolution for jets of better than 10 mrad and good separation of jets,
- separation of electrons and hadrons both in jets and for isolated particles.

The ZEUS-calorimeter is a so-called sandwich calorimeter consisting of alternating layers of 3.3 mm steel-cladded, depleted uranium and 2.6 mm scintillator material, where a pair of uranium and scintillation layers correspond to a radiation length of $1X_0$. The thickness of the plates is chosen in such a way that the calorimeter is fully compensating, i.e. electrons and hadrons of the same energy generate the same signal-height in the detector. This is achieved within 2% for energies above 2 GeV. The setup yields a very good energy resolution of $\Delta E/E = 35\%/\sqrt{E[\,\mathrm{GeV}]} \oplus 2\%$ for hadrons and $\Delta E/E = 18\%/\sqrt{E[\,\mathrm{GeV}]} \oplus 1\%$ for electrons (obtained from test-beam measurements; $\oplus$ denotes the quadratic summation of the three terms). The calorimeter covers 99.7% of the solid angle. Only in the region of the beam pipe, particles can escape the detector undiscovered. As an exception, neutrinos cannot be detected directly and have to be reconstructed through the missing transverse momentum.

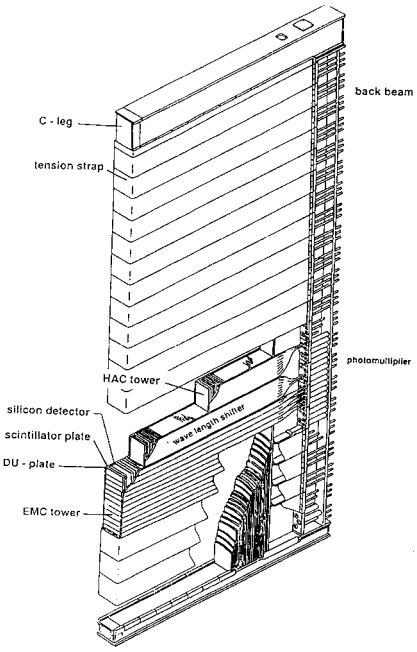


Figure 3.4. *Isometric view of an FCAL module.*

The whole calorimeter is divided into three sub-calorimeters: the forward (FCAL) and rear calorimeter (RCAL), both aligned vertically to the beam axes, and the barrel calorimeter (BCAL), which is arranged cylindrically around the beam axis. All sub-calorimeters are mechanically divided into modules, which run parallel to the y-axis for RCAL and FCAL (Fig. 3.4) and parallel to the beam axis for BCAL. Each of the 23 (32) F/RCAL (BCAL) modules is organized in read-out towers. In FCAL and RCAL these towers have a front face of $20 \times 20\,\mathrm{cm}^2$ and are longitudinally divided into a electromagnetic (EMC) and a hadronic (HAC) section. In return, the EMC section of the RCAL (FCAL) consists of 2 (4) electromagnetic cells with a cross section of $20 \times 10\,\mathrm{cm}^2$ ($20 \times 5\,\mathrm{cm}^2$) and a depth of 1λ corresponding to $25X_0$. The hadronic section in the FCAL is again subdivided longitudinally into HAC1- and HAC2-cells, whereas in the RCAL only one HAC-cell exists in each tower. Overall, the RCAL (FCAL) has a depth of 4λ (7λ) corresponding to $105X_0$ ($185X_0$). Like the FCAL, a tower in the BCAL hosts 4 EMC- and 2 HAC-cells, where a special feature is the projective orientation of the EMC-cells with respect to the nominal IP. Additionally, the BCAL modules are rotated in ϕ by 2.5° around an axis parallel to the beam axis, as otherwise, neutral particles or charged particles with a large momentum, i.e. particles with a straight track, could hit the crack between two modules, leading to a large mis-measurement of their energy[3] (Chap. A.1). The BCAL has a total depth of 5λ or $130X_0$.

Each EMC- and HAC-cell is read out from two sides via wavelength-shifters and photomultipliers (PM). This allows the determination of the shower position also inside a cell and an energy measurement independent of the impact position of the particle. Additionally, the cell can be read out and used even (with some limitations) in the case of the failure of one of the PMs. The position resolution (RMS) of electrons in the RCAL (BCAL) averages to 8.7 mrad (8.3 mrad) in ϕ and 6.3 mrad (4.1 mrad) in θ (obtained from MC studies) [56].

[3] The direction of the rotation ensures that electrons with low momenta always hit the crack under an angle. However, for positrons this is not true.

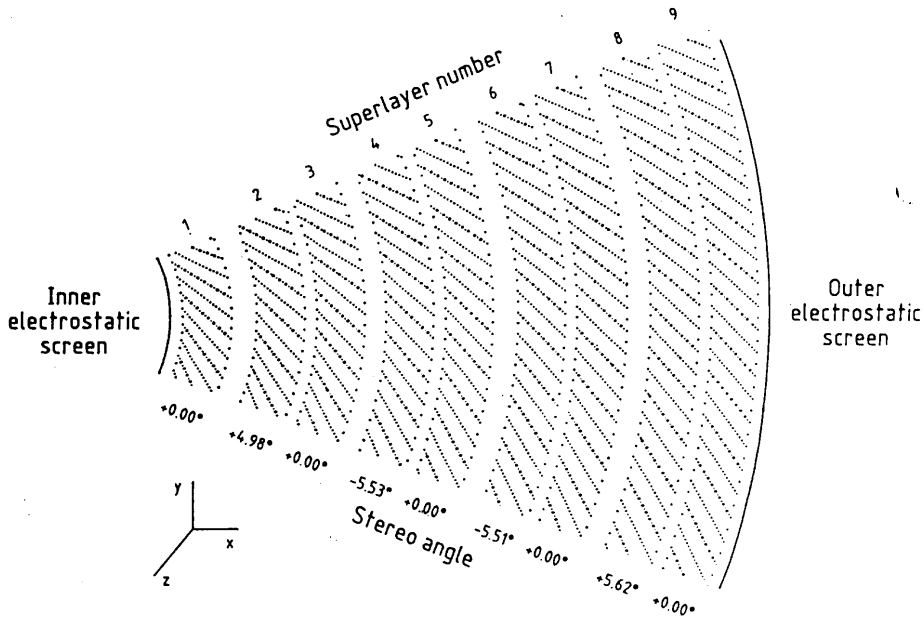


Figure 3.5: *Cross section of a CTD octant. Depicted is the structure of super-layers and drift cells, where thick dots mark signal wires and thin dots mark field-forming wires.*

3.2.2 Central tracking detector

The central tracking detector (CTD) is a cylindrical drift chamber with its wires oriented along the beam axis. Since the removal of the vertex detector it is the component closest to the beam pipe. Its active volume covers the area between $-101\,\text{cm}$ and $+105\,\text{cm}$ in z and $16.2\,\text{cm}$ and $85\,\text{cm}$ in r, where r is the radius. Internally, the CTD is subdivided into 9 super-layers, which in return consist of 8 layers of signal wires. Between the signal wires field-forming wires are strung as depicted in Fig. 3.5. Because of the orientation of the signal wires along the z-axis the CTD allows the determination of the x-y-position of a hit and accordingly the reconstruction of the associated track in this plane with high precision. In order to obtain the z-position, all odd super-layers are tilted with respect to the even super-layers by a stereo-angle of $\sim 5.5°$ (*z-by-stereo*). In a MC study the CTD yields mean resolutions (RMS) for electrons with $E > 15\,\text{GeV}$ hitting the RCAL (BCAL) of $2.8\,\text{mrad}$ ($2.5\,\text{mrad}$) in ϕ and $2.6\,\text{mrad}$ ($3.6\,\text{mrad}$) in θ [56].

The curvature of the track in the magnetic field of the solenoid serves to measure the transverse momentum p_t of the track:

$$p_t[\text{GeV}] = 0.2998 \cdot B[\text{T}] \cdot \rho[\text{m}] \ , \tag{3.2}$$

where B is the strength of the magnetic field in z-direction. Combined with the θ-information p_t can be used to calculate the momentum p of the particle generating the track:

$$p = \frac{p_t}{\sin(\theta)} \ . \tag{3.3}$$

The uncertainty on the measured momentum in the CTD is determined by the measurement error on the hit-position ($\sigma_{\text{hit}} \approx 200\,\mu\text{m}$) and by multiple scattering. Overall, the CTD yields a

momentum resolution of [57]

$$\frac{\Delta p_t}{p_t} = 0.58\% \cdot p_t[\text{GeV}] \oplus 0.65\% \oplus \frac{0.14\%}{p_t[\text{GeV}]} \tag{3.4}$$

for tracks traversing all nine super-layers. The first term accounts for the error on the hit positions and the second term for multiple scattering inside the CTD, whereas the third term takes multiple scattering between IP and CTD into account. At the moment, the last term is dominantly caused by the beam-pipe and the inner CTD wall, however after the upgrade also the micro-vertex-detector will add to this uncertainty.

As already mentioned above, (3.4) only holds for tracks passing all nine super-layers, while for tracks with less layers the error grows accordingly. If less than 3 super-layers are hit, the error on the measured z-position gets too large and the reconstruction program switches from the *z-by-stereo* to the so-called *z-by-timing* method. This method uses the runtime difference of a signal to the two ends of the wire to determine the z-position. However, this method has a resolution of the order of a few cm, yielding large errors in the θ and momentum measurement. Therefore, tracks with less than 3 super-layers are usually not considered as good tracks.

For particle tracks with $E > 10\,\text{GeV}$, relevant for this analysis, the dominant contribution to the momentum uncertainty stems from the measurement error on the hit position, as the lower Q^2-cut of 185 GeV limits the p_t of the tracks to $\gtrsim 5\,\text{GeV}$.

3.2.3 Presamplers

The presamplers consist of 5 mm thick scintillator plates (tiles) having a size of $20 \times 20\,\text{cm}^2$ [58]. They are mounted in front of the RCAL and FCAL[4] and cover an area with a radius of $\sim 190\,\text{cm}$ around the beam pipe. The presamplers are used to correct for the energy loss of particles in material on their way from the IP to the calorimeter on an event-by-event basis. The correction exploits the fact that a particle showers after a hard interaction and that the lost energy is correlated with the number of generated shower particles. Being proportional to this number of particles, the presampler signal can then be used to calculate the lost energy.

3.2.4 Hadron-Electron Separator (HES)

The Hadron-Electron Separator (HES) [59] is located inside the F/RCAL, 7 cm below the surface at the shower maximum for electrons. It is comprised of two layers of small-area silicon diodes ($29 \times 32\,\text{mm}^2$) arranged on skis that are inserted from the top of the calorimeter and cover an area with a radius of $\sim 100\,\text{cm}$ around the beam pipe. As its name suggests, HES can be used to separate electrons from hadrons due to the fact that electrons shower earlier than hadrons and therefore the shower width of the latter is smaller at the position of the HES diodes. However,

[4] In the meantime also the BCAL was equipped with a presampler, however its information is not used in this analysis. Compared to F/RCAL there is much less dead material in front of the BCAL and its variation is much weaker so that the usage of dead material maps (Chap. 4.4.4.1) is sufficient.

HES was not always fully functional during the data taking period and therefore it was decided not to include it for the identification of electrons.

Nevertheless, the HES proved to be useful in another area. Compared to the calorimeter, HES has the advantage that the z position of the deposited charge of traversing particles is well defined due to the small diode thickness of only 350 µm. Hence, the calculated x-y-position is much more precise for particles hitting the calorimeter under a large angle. Therefore, the HES was used in the calorimeter-alignment study (reported in Appendix A) to determine the position of the RCAL.

3.2.5 Luminosity-measurement system

At ZEUS the measurement of the luminosity is performed via the QED-bremsstrahlungs process (Bethe-Heitler-process)

$$ep \rightarrow ep\gamma \ . \tag{3.5}$$

Its cross section is calculable to a precision of 0.5% [60]. The luminosity system of the ZEUS detector consists of two lead-scintillator calorimeters, one positioned at $z = -35\,\mathrm{m}$ (LUMI-e) and the other at $z = -107\,\mathrm{m}$ (LUMI-γ). Photons that are emitted from the IP inside a cone of 0.5 mrad are registered with the LUMI-γ detector. Its count rate is used to determine the luminosity [61]. The error on the obtained value is 1.5% for the time period considered in this analysis. The LUMI-e detector, which is currently not used for the luminosity determination, is utilized in this thesis for electron tagging in background normalization studies of the PHP sample (Chap. 5.2.3.2).

3.3 The ZEUS trigger system

The task of a trigger system is to filter out the interesting events from a huge amount of background. At HERA for example signal events with $Q^2 > 100\,\mathrm{GeV}^2$ occur at a rate of 0.1 Hz for a typical luminosity of $2 \cdot 10^{-5}\,\mathrm{pb}^{-1}\,\mathrm{sec}^{-1}$, whereas background events like muons from cosmic radiation or collisions of beam particles with gas molecules inside the beam pipe occur at much higher rates of 5 000 Hz up to $\sim 15\,000$ Hz (for the latter the rate depends strongly on beam and vacuum conditions) [48]. As an electron-proton-beam crossing occurs every 96 ns at HERA, corresponding to a rate of 10.4 MHz, a very fast trigger system is needed. Therefore, ZEUS installed a three-level trigger system, where in each level the complexity of the detector information used increases.

The **first-level-trigger (FLT)** has to take a decision whether to except or to reject an event every 96 ns. The speed needed can only be realized in a hard-wired trigger but even then the trigger is not fast enough, on account of which each component has a pipeline attached storing information for at least 5 µs. Each component represented in the FLT takes its own first-level-trigger decision and sends the result upon request to the general first-level-trigger (GFLT), which takes the final decision. The decision of the different components are based on very crude

information like calorimeter energy above a threshold or a minimum number of hits in the CTD. The output event-rate of the FLT amounts to roughly 500 Hz.

In the **second-level-trigger (SLT)** more precise event information is used. Background suppression is mainly achieved by exploiting timing- and energy-information of the calorimeter. The SLT is realized in a transputer network. As in the case of the FLT, the information from the individual-component SLTs are collected by the general second-level-trigger (GSLT) that takes the final SLT decision. If an event passes the SLT, all detector components are read out and the information is sent to the event-builder, where the frequency must not exceed 100 Hz.

The **event-builder (EVB)** is a powerful processor farm which collects the information from all detector components and reconstructs the full event for the first time in the trigger chain. Though not an exact copy of the offline event reconstruction due to time-limitations, it provides important information of the event topology like tracks, IP, calorimeter energies, electron candidates as well as kinematic variables or global quantities of the hadronic final state. With this information, the **third-level-trigger (TLT)** is able to reduce the event rate to 3–5 Hz, which is low enough, given an average event size of ~ 100 kB, to transfer all selected events via a fast link from the detector to a storage facility on the DESY site.

Finally, all events coming from the TLT are classified by assigning DST (data summary tape) bits to them according to certain event properties.

The offline event reconstruction is performed by a PC farm. Here, the full reconstruction algorithms are used and the results are stored in the database format ADAMO [62] where they are accessible for the user via FORTRAN programs. The user then runs his analysis job on a central computer facility at DESY over the database and produces individual ntuple samples with reduced information, which have altogether a manageable size of at most a few GB. These ntuples are then further analyzed on local computers.

3.4 Event generators and detector simulation

Today's particle detectors like ZEUS have reached a complexity that makes it impossible to calculate the various detector effects. Therefore, MC simulations have become an indispensable tool in high-energy physics. They consist of two components:

- the physics generator, that generates the final state of an ep reaction including the hadronic final-state according to the differential cross section. It is based on the models of the various interactions and the knowledge about the properties of particles like their masses and lifetimes,
- the detector simulation, that tracks each particle through the detector and generates the respective detector signals. These are then fed into the standard reconstruction- and trigger-chain like the data.

Afterwards, data and MC can be compared to each other and the detector effects can be estimated from the latter, using the true quantities. With this knowledge it is now possible to correct the data and to compare them to theory and the results from other experiments.

The physics generator used in this analysis for the generation of events in deep-inelastic electron-proton scattering is called DJANGO 1.1 [63]. It consists of the programs HERACLES 4.6 [64] and LEPTO 6.5 [65]. HERACLES is responsible for the QED part of the scattering process and passes the exchange-boson afterwards on to LEPTO which simulates the hard subprocess. Internally, LEPTO is again made up of two programs. First, the final state is calculated on the parton level, where either the CDM (color dipole model), implemented in ARIADNE [66], or the MEPS (matrix element + parton shower) [65] can be used. The following fragmentation is performed by the program JETSET 7.4 [67] using the string model. To simulate diffractive events in DIS (Chap. 2.4) the program RAPGAP 2.08 [68] is used. It was modified in such a way that it produces the same ep cross sections in the x-Q^2 plane as the DJANGOH generator (Chap. 5.2.2).

The PHP samples used in this analysis were generated with the program HERWIG 5.9 [69], which in contrast to LEPTO uses a cluster fragmentation-model.

Afterwards, all MC events run through the detector simulation MOZART which is based on the program package GEANT 3.13 [70]. It describes under certain simplifying assumptions the interaction of particles with different detector components and generates signals like those provided by the real detector. These signals are then passed on to the standard reconstruction- and trigger-simulation.

3.5 Summary of Chapter 3

In this chapter the hardware used for the data taking has been described. A brief overview over the HERA and the ZEUS-detector setup was given, where components essential for the analysis were presented in more detail including the ZEUS three-level trigger system. The end of this chapter formed a brief section about the importance of event generators and detector simulations in todays high-energy experiments, followed by a list of the generators and simulations used in this analysis. In the next chapter the different detector information will be used to reconstruct events offline.

Chapter 4

Event reconstruction

In this chapter, the steps towards a fully reconstructed event are described. At first, some "pre-corrections", applied to the energy deposits in the calorimeter, and the reconstruction of the event vertex are explained. Next, the scattered electron, the most important signature of NC DIS events, is identified and its energy and direction are reconstructed. The precision on the latter is determined by the resolution of the electron position within the calorimeters (Chap. 3.2.1) and the knowledge of the absolute calorimeter positions. To ensure that the uncertainty on the absolute positions yields negligible contributions, a calorimeter-alignment study has been performed in the framework of this thesis which is discussed in detail in Appendix A. After the scattered electron has been identified, the hadronic final state can be reconstructed. Several methods exist for this task which are described and compared to each other. At the end of this chapter the methods for the reconstruction of the kinematic variables are reviewed and their corresponding resolutions are set into comparison in order to find the best method for the kinematic range covered by this analysis.

4.1 Pre-corrections

The pre-correction stage comprises the removal of faked energy deposits in the calorimeter not coming from particles, the identification and marking of noisy or dead cells and some first corrections to the energy scales of the different calorimeters.

4.1.1 Noise suppression

The radiation of the uranium in the calorimeter produces frequent but small signals in the PMs. However, due to statistical fluctuations, at times this radiation fakes energy-deposits of the order of 100 MeV. A cut on the measured energy in a cell (60 MeV for EMC and 110 MeV for HAC) has already been applied before the calorimeter information was stored in the database. However, this is a very loose cut and therefore a stricter one is applied offline, which amounts to 100 MeV for electromagnetic cells and 150 MeV for hadronic cells.

4.1.2 Dead and noisy photomultipliers

During the operation of the detector it happens from time to time that a PM, its power-supply channel or its readout electronics-channel breaks. This kind of damage can only be repaired during a shutdown and therefore its effect on the reconstruction has to be taken into account. Basically, there exist two distinct scenarios: either the channel produces no signal anymore and is then denoted as "dead" or it becomes noisy, i.e. generates random signals not associated with energy deposits in the calorimeter.

Dead channels are identified and marked by the calorimeter group in the course of the data-quality monitoring process. If a cell has only one dead channel, in general the energy of the whole cell is reconstructed by doubling the energy of the working PM[1]. However, if both channels are dead, called a "hole", the information from the cell is lost and its energy is set to zero. During the data-taking period 1998–99 there were on average 90/270/14 dead channels and 0.1/2.0/0.1 holes in the F/B/RCAL, while the calorimeter altogether consists of $\sim 12\,000$ channels.

To identify noisy cells a sophisticated algorithm is used [71]. Simplifying the procedure, noisy cells are identified by comparing the rate at which a certain cell fires to the average rate of the other cells. If the rate lies more than 5σ above the average value, the cell is marked as noisy and is removed from the cell list.

4.1.3 Energy scales of the R/BCAL

Studies of the RCAL EMC energy scale [72] showed that it is on average 2.2% too low in the data, where the distribution of the correction factors has a σ of 2.5%. Due to the large statistics, correction factors could be calculated for each individual EMC cell within a radius of about 70 cm around the beam pipe. In the remaining RCAL and the BCAL only an overall correction factor could be determined due to low electron statistics. In data the EMC and HAC energy scales of the BCAL are corrected by 5.0% up, whereas the EMC and HAC cells in the RCAL are corrected by 2.2% up. In all cases where an individual correction factor for an EMC cell could be determined, it was applied to the respective EMC cell rather than the global factor.

4.2 Calorimeter alignment

Determining the alignment of the different calorimeter parts with respect to the beam line is an important and difficult task, given that an accuracy of $\mathcal{O}(\text{mm})$ is necessary to stay below the position resolutions of the calorimeters which on average amount to ~ 8 mrad in ϕ and ~ 5 mrad in θ (Chap. 3.2.1).

Surveys can only be performed while the detector is open during a shutdown phase and they are always a major effort. Additionally, the fact that the survey measurements have a different reference system than the event reconstruction renders a transformation of these measurements

[1] For electrons in the BCAL a more sophisticated algorithm is used in this analysis that will be discussed in Chap. 4.4.4.4

into the physics coordinate-system of ZEUS quite complicated and error-prone, considering that each component has 6 degrees of freedom.

Therefore, it is desirable to independently determine the position of the calorimeters using physics events. In the framework of this thesis, an extensive study has been carried out that determines the position of the RCAL (BCAL) in x, z and ϕ (ϕ and z) by using the identified electron and its associated CTD track (Chap. 4.4). A detailed description of this study together with a list of the shifts and rotations applied in the analysis can be found in Appendix A. Unfortunately, the position of the FCAL could not be determined with these studies, as the number of electrons hitting this part of the calorimeter is much too small.

4.3 Reconstruction of the event vertex

The accurate determination of the position of the event vertex is an essential task as this position is needed for the calculation of many important quantities like the angle of the hadronic system or that of the electron if no associated track is available. The event vertex is determined from CTD tracks and the procedure, described in detail in [73], can be separated into three steps:

- First, a preliminary filter is applied to the trajectories to remove tracks that do not originate from the beam line.
- Then, a first, rather simple fit determines the weighted center of gravity for all remaining tracks. Now, all tracks with a high χ^2-distribution are discarded until the fit quality is reasonable. The obtained vertex is the starting point for the full vertex fit.
- The full vertex fit is an iterative procedure and comprises in each iteration a complete refit of the trajectories of all tracks that remain, where the vertex obtained in the previous iteration is included into the fit. The iteration is stopped when the vertex position has converged.

For this analysis a special feature of the vertex-fitting routine is exploited that allows to give a "seed" track to the algorithm, which in this case is the electron track. The seed track cannot be deleted from the track list throughout the whole procedure. This is justified, as the electron track should always come from the vertex[2]. Figure 4.1 shows for data the final vertex distributions after all selection cuts, discussed in Chaps. 5.3 and 5.4, excluding the cut on the z position of the vertex. The double-peak in Fig. 4.1 b originates from a shift of the beam position during the running period. For the reconstruction of the events only the measured z position is used, whereas the x- and y-positions are set to 0 cm, as the uncertainties on the measured positions are greater than the intersection cross-section of electron- and proton-beam of $\sim 300 \times 100\,\mu\text{m}^2$ centered at $x \approx y \approx 0\,\text{cm}$.

[2] One has to be a little bit careful here as tracks which lie outside the CTD acceptance, defined in Chap. 5.4, are not very trustworthy. Therefore, the "seed vertex" is only used if the electron track passes at least three super-layers and has a momentum $> 5\,\text{GeV}$.

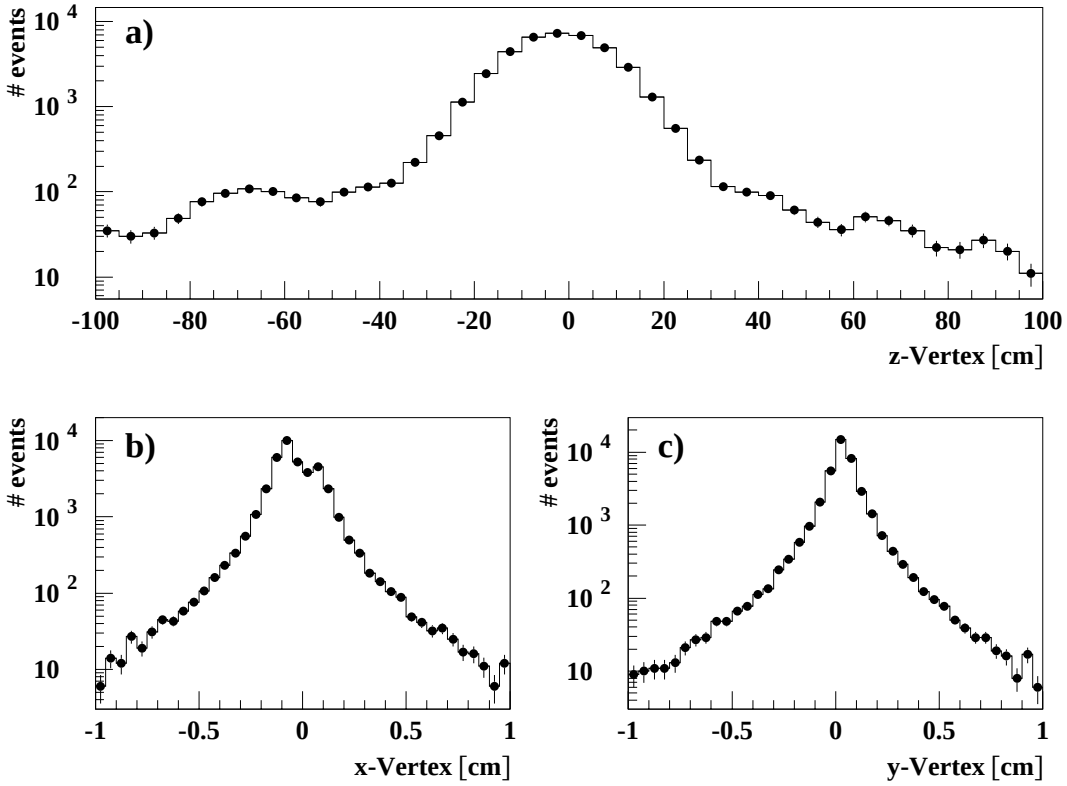


Figure 4.1: *Distribution of the reconstructed vertex in data for the (a) z-, (b) x- and (c) y-coordinate.*

4.4 Reconstruction of the scattered electron

One of the major tasks while reconstructing an NC event is the identification and reconstruction of the scattered electron. The presence of an electron is the most important feature of an NC event as it distinguishes the NC event from potential background processes, e.g. PHP where the electron escapes through the rear beam pipe. The electron scattering angle enters directly in the calculation of the kinematic variables (Chap. 4.6), but also its energy is of importance as several offline cuts are based on it(Chap. 5.4).

4.4.1 Identification

The identification of the electron with the detector is performed by so-called electron finders. Several of such programs exist within ZEUS. Each one was developed for a special purpose, e.g. finding electrons in the RCAL or finding electrons in events with high Q^2. The main electron finder, that is used for the determination of all cross sections in this analysis, is called **EM**. It was specially developed to identify electrons in the high Q^2 regime and also has a superior background suppression in the corresponding detector region [74]. For the alignment studies in the RCAL, another electron finder called **SINISTRA95** was used. It was conceived mainly for the electron identification in the RCAL and additionally provides HES information which is extensively used in these studies. Both finders select the "true" scattered electron from a list of

electron candidates. This list is sorted by probability as calculated by the electron finder. The electron candidate with the highest probability is selected as the "true" scattered electron if its probability lies above a certain threshold.

4.4.1.1 EM

To come to a decision, EM uses both calorimeter and track information and additionally takes geometrical properties of the detector into account. Overall, EM uses 7 variables to distinguish between electrons and hadrons, where 4 variables refer solely to calorimeter information, e.g. the energy spread of the electron candidate[3] in the calorimeter or its isolation with respect to other energy deposits. The remaining 3 variables describe the matching quality of calorimeter and track information like track momentum vs. calorimeter energy or the reconstructed scattering angle. The differences in the distributions for signal and background events are used to calculate a probability function for each variable [75]. The total probability of an electron candidate is then determined by multiplying the individual probabilities, obtained from the probability functions.

An EM electron candidate consists of a cell cluster in the calorimeter and, if available, a track in the CTD pointing to that cell cluster. To form the cell clusters, the following cell island algorithm is applied:

- All cells with a non-zero energy deposit are considered. Neighboring cells are defined as those which touch the main cell with at least one edge. This is also valid for cells of different types (EMC, HAC1, HAC2). A cell is defined as a local maximum, if it is surrounded only by neighboring cells with lower energy.
- The calorimeter cells are sorted by energy in descending order.
- One starts with the top cell in the list which, having the highest energy, automatically forms a local maximum. For subsequent cells j one verifies for all other cells i with $i < j$, whether cell j is a neighbor of cell i according to the definition above. If this is true, cell j is assigned to cell i. Otherwise cell j itself becomes a local maximum and therefore a seed of a new cell island.
- This procedure is repeated for all cells in the list.

A two-dimensional example for the algorithm is shown in Fig. 4.2.

The island algorithm is implemented in EM and therefore accessible via this program. To be accepted as an electron candidate, the cell island must have a calorimeter energy of at least 4 GeV and the calorimeter probability (probability product of all 4 calorimeter variables) must be greater than 10^{-5}. In addition, only 30% (50%) of the total energy of the cell island is allowed to be deposited in the hadronic section of the FCAL (BCAL).

As EM also uses track information, a track matching algorithm is applied:

- To be considered, a track has to fulfill the following requirements:
  - The transverse momentum must exceed 0.1 GeV.

[3] The exact definition of an electron candidate will be given in the next paragraph.

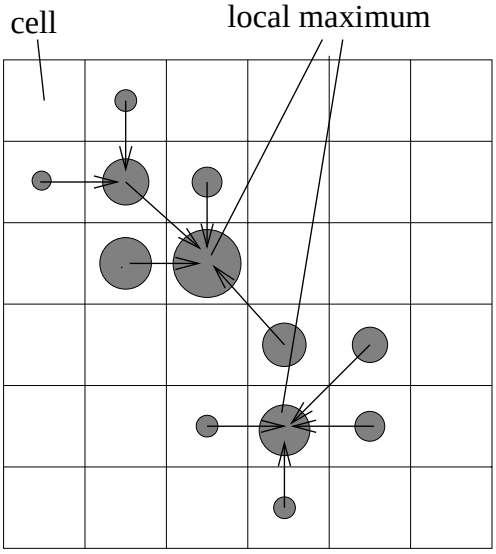


Figure 4.2. *Schematic diagram of the island algorithm in two dimensions. The arrows specify the assignment of the cells to each other. The circles symbolize energy deposits, with their areas being proportional to the energies.*

  - The distance of closest approach to the beam line must be less than 2 cm.
  - The angular distance between the shower center and the impact position of the track on the calorimeter surface must be less than 45° both for the azimuth and the polar angle.
  - The distance of closest approach of the track extrapolation into the calorimeter and the shower center must not exceed 50 cm.
- If a track survives all these cuts, EM calculates the track probability (probability product of all 3 variables containing track information) for it.
- The track with the highest track probability is assigned to the calorimeter cluster. If no track passes the cuts, the cluster is treated as a trackless electron candidate.

In addition, EM contains an algorithm which allows to merge the islands of the electron and the photon in case of FSR events. EM investigates all islands which are in an η-ϕ cone of radius 0.8 around an electron candidate, where η is the pseudo-rapidity defined as $\eta = -\ln(\tan\frac{\theta}{2})$. If the calorimeter probability of an island is greater than 0.002 and its energy exceeds 0.4 GeV but lies below the energy of the electron candidate, the electron candidate and the island are merged to a new candidate. This new candidate must have a probability greater than 0.001, otherwise it is deleted and the original electron candidates are used. If an electron candidate has more than 1 island in its η-ϕ cone of radius 0.8, then the merged candidate with the best total probability is taken. The *electron-island* in a merged candidate is the island with the highest track probability. If both islands have no track, the electron is the one with the highest energy. The position of the merged candidate and other quantities are set to those of the *electron-island*, whereas the four-momentum is the sum of the four-momenta of the electron and the photon (calculated from the energy and position of the electron in the calorimeter and the event vertex).

If an electron candidate has been identified as the scattered electron, i.e. the candidate has the greatest probability of all candidates in an event and this probability lies above a certain threshold, its properties are calculated. The total energy is the sum over the energies of the cells belonging to the electron cluster. The final position of the electron in the calorimeter is

calculated from the four-momentum and the event vertex in order to take potential FSR photons into account.

4.4.1.2 SINISTRA95

In contrast to EM, SINISTRA95 uses only calorimeter information as it was mainly developed for events with low Q^2, where the electron hits the RCAL. SINISTRA95 [76], as its predecessor SINISTRA94 [77], is based on a neural network. It analyzes the energy distribution over the calorimeter cells and evolves it in moments. 16 of these moments together with the total energy of the cluster are then used as input variables for the neural networks. The cluster-formation algorithm is comparable to that of EM, though there are small differences, e.g. in the definition of what a neighboring cell is. SINISTRA95 proved to be very efficient in finding electrons in the RCAL, however, in the higher Q^2 regime, it also accepts a lot of background, especially from PHP events [74].

4.4.2 Verification of the calorimeter electron-position

Apart from the verification of the spatial positions of the calorimeters (Appendix A), the reconstructed position of the electron within the calorimeter has to be checked. The latter is determined by an electron-position reconstruction algorithm called ELECPO [78]. This algorithm uses the imbalance information of the two photomultipliers of the most energetic cells in the electron cluster for the determination of the position perpendicular to the module boundaries (x in R/FCAL and ϕ in BCAL), whereas for the other directions (y in R/FCAL and z in BCAL) the energy spread between neighboring cells is utilized.

The verification of the calorimeter position is accomplished by comparing the reconstructed position in the calorimeter to reference positions obtained from other detector components. Like in the calorimeter-alignment studies, the extrapolated track[4] of the identified electron is used as the reference position.

For electrons in the RCAL, Fig. 4.3 displays the difference between track and calorimeter position as a function of the calorimeter position in x and y. Clearly visible are the large fluctuations of up to 1.5 cm both in data and MC, which are caused by the structure of the calorimeter. Unfortunately, the comparison of data and MC reveals significant differences. Due to the complicated structure of the curves, these fluctuations are not corrected for. However, it should be noted that the effect of these fluctuations on the cross section results are expected to be small, as in this analysis most of the electrons in the RCAL have a track assigned, so that the scattering angle can be determined from the track. A performance comparison of the "CAL" and "Track" method with respect to the reconstruction of the scattering angle can be found in the next section.

For electrons in the BCAL, Fig. 4.4 a shows the difference between the electron z-position in the calorimeter, z_{cal}, and the z position obtained from the extrapolated track, z_{trk}. The agreement

[4] The track is extrapolated to the reconstruction plane of the respective calorimeter using a so-called swim algorithm that takes the magnetic field of the ZEUS detector into account.

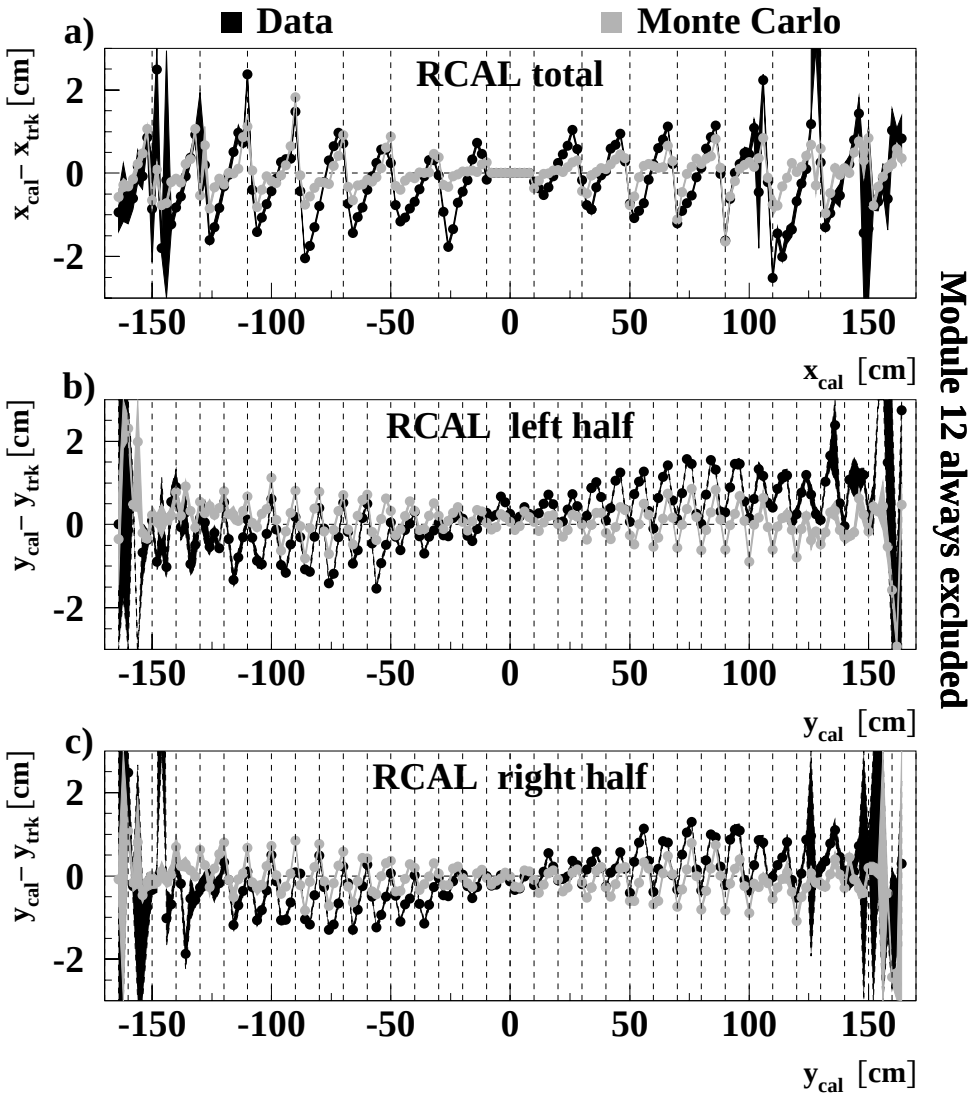


Figure 4.3: *Comparison between the calorimeter position (cal) and the position obtained from extrapolating the electron track into the calorimeter (trk).(a) Difference between* x_{cal} *and* x_{trk} *as a function of* x_{cal}*. (b,c) Difference between* y_{cal} *and* y_{trk} *as a function of* y_{cal} *for left and right RCAL half. Dark points are data and light points are MC. The thickness of the lines connecting the points represents the error on the measured differences. The vertical dashed lines indicate the cell boundaries.*

between data and MC is almost perfect and the mean values, indicating a shift of $-0.15\,\mathrm{cm}$, differ by less than 0.7%. However, the different origins of these shifts in data and MC are revealed in Fig. 4.4 b displaying the mean value of $z_{\mathrm{cal}} - z_{\mathrm{trk}}$ as a function of z_{trk}. Here, the data describes a straight line going through the origin, whereas the MC is flat with a negative offset. The slope of the data is probably an artifact of the position reconstruction algorithm caused by wrong assumptions about the depth of the shower maximum, whereas the source for the MC offset is not known. The fact that the MC does not show a slope is presumably due to an inaccurate simulation of the calorimeter. In order to correct for these effects, the position in MC is corrected to that in data with $z_{\mathrm{cor}} = z_{\mathrm{cal}} + 0.20 + 0.005 \cdot z_{\mathrm{cal}}$.

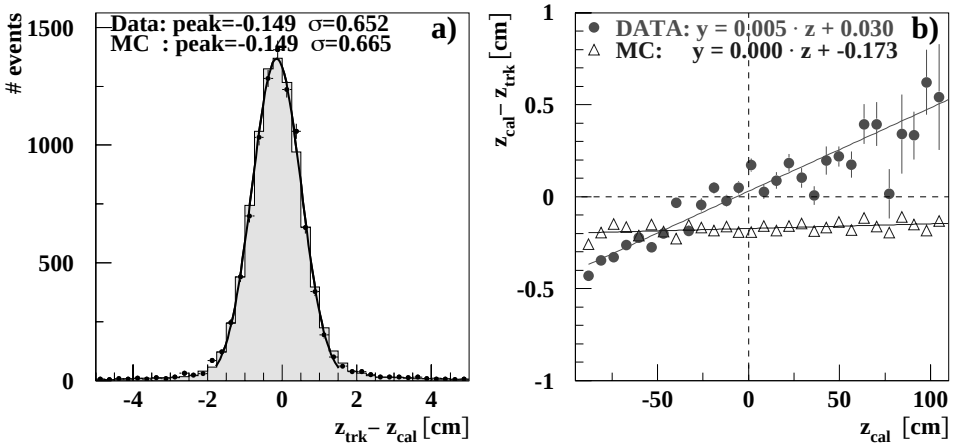


Figure 4.4: *(a) Histogram of differences between calorimeter position z_{cal} and extrapolated track position z_{trk} of identified electron in BCAL. The full dots represent the data and the shaded histogram represents the MC. (b) Mean of $z_{\mathrm{cal}} - z_{\mathrm{trk}}$ as a function of z_{trk} for data (full circles) and MC (open triangles).*

4.4.3 Determination of the electron scattering-angle

The scattering angle of the electron can be calculated in two ways. The method that is always applicable uses the position of the electron cluster in the calorimeter in combination with the reconstructed vertex. In the case of a reconstructed matching track the scattering angle is also directly available through the polar angle θ of the track at the vertex. As the latter method has a better resolution, is less biased (Fig. 4.5) and circumvents the problems with the position reconstruction in the calorimeter discussed in the previous section it is always used if a matching track with a momentum $> 5\,\mathrm{GeV}$ exists and is located inside the CTD acceptance (defined in Chap. 5.4).

4.4.4 Corrections to the electron energy

Between the IP and the calorimeter a large amount of material is located that can reduce the measured electron energy considerably. In the calorimeter itself, the support- and readout-structure or faulty PMs have a significant influence on the measured energy. In the following sections all these sources are discussed and correction methods are presented.

4.4.4.1 Dead-material correction

On its way from the vertex to the calorimeter the electron looses energy due to interaction with (mostly non-active) material in the detector. To account for these losses a so-called dead-material map, obtained from MC and detector geometry data, is available which contains the number of X_0 between the IP and the active volume of the calorimeter for all θ and ϕ angles (Fig. 4.6). As the amount of material for a given pair of θ and ϕ angles varies with the position of the vertex, the dead-material map is generated for vertex positions at $z = -45, -8, 0, 8, 45\,\mathrm{cm}$,

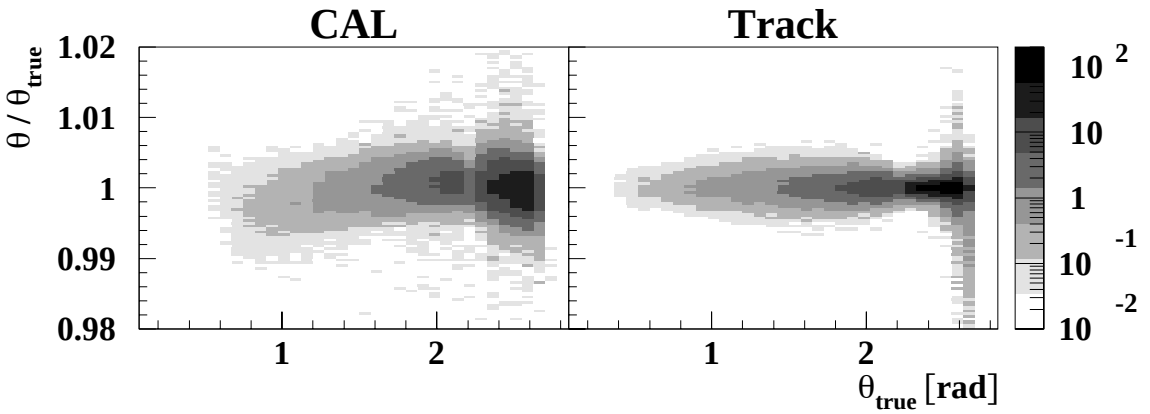


Figure 4.5: *Ratio between measured (θ) and true (θ_{true}) scattering angle of the electron as a function of θ_{true}. For "CAL" the angle is calculated from the calorimeter position using the reconstructed vertex, whereas for "Track" the polar angle is directly determined from the reconstructed track at the vertex. The density, displayed in arbitrary units, indicates the frequency.*

and for positions in between an interpolation algorithm is used. The energy correction-factor f is then determined from the number of X_0 and the measured electron energy E in the calorimeter according to the following formula:

$$f = 1 + \frac{a(iX_1)}{E^{b(iX_1)}} + (X_0 - X_1)\left(\frac{a(iX_2)}{E^{b(iX_2)}} - \frac{a(iX_1)}{E^{b(iX_1)}}\right) . \tag{4.1}$$

X_1 is either 1 or 2 and $iX_{1/2}$ are indices in the range between 1 and 3, where all these integers are derived from X_0. a and b are arrays of constants obtained from test-beam measurements [79].

Unfortunately, this method has some negative features. Though using different dead-material maps for 5 different vertex positions this is only a rough approximation of reality, especially for the R/FCAL where the amount of dead material changes very rapidly depending on the path of the electron through the detector (Fig. 4.6: $\theta \gtrsim 130°$ and $\theta \lesssim 40°$). Additionally, the maps were generated using a MC simulation of the detector which of course does not agree perfectly with the real detector. And finally the resulting correction factor is only an estimate of the mean true value whereas the individual values follow a statistical distribution.

4.4.4.2 Presampler correction

In order to be able to correct the energy of the electron on an event-by-event basis, independent from dead-material maps, this analysis uses the presampler for electrons in the RCAL whenever applicable. Unfortunately, the RCAL presampler data is not available for a small fraction of the 1998 data, which corresponds to an integrated luminosity of $\sim 0.75\,\mathrm{pb}^{-1}$. Hence, for this period the disfavored dead-material correction, described in the previous section, has to be used. The FCAL presampler is not taken into account, as the calibration method applied in this analysis needs high electron statistics which is not available in the FCAL. Also the BCAL presampler is

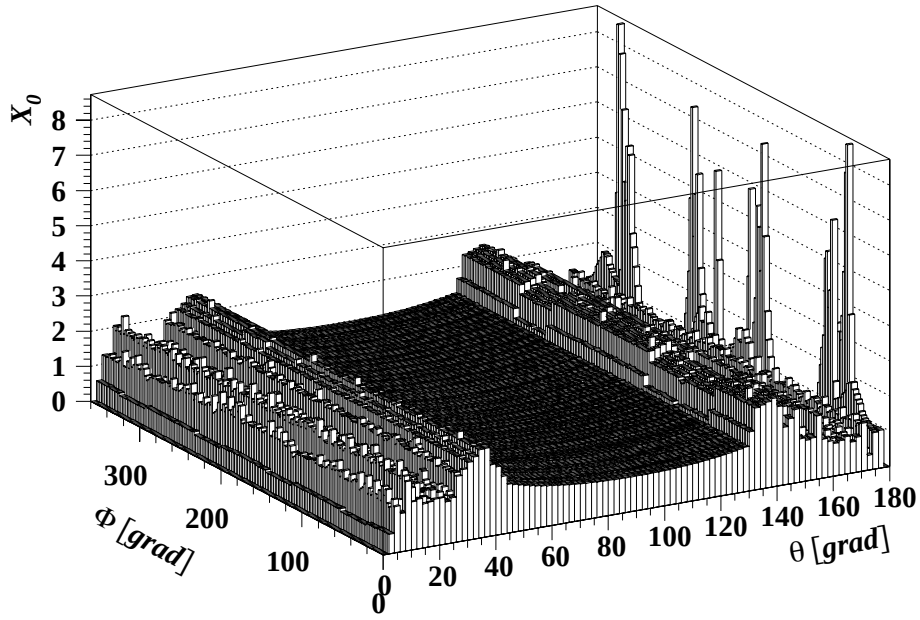


Figure 4.6: *Dead-material map for the 1998/99 running period for a z-vertex position of* 0 *cm. The amount of material in front of the calorimeter, represented by the quantity* X_0*, is plotted as a function of* θ *and* ϕ*. The size of the bins has been reduced for this plot. For the actual analysis a binning is used that is 8 (2) times finer on the* θ *(*ϕ*) axis.*

not considered here, as the dead material in the BCAL region is thin ($\sim 1X_0$) and its distribution quite uniform (Fig. 4.6: $40° \lesssim \theta \lesssim 130°$).

The presampler correction routine for RCAL electrons used here is described in more detail in [80]. As this routine was developed for a different data set, the parameters are checked by comparing the fit results from the original studies with those obtained in this analysis.

To perform these studies the electron sample is divided into three bins of the electron energy E_{cal} in the calorimeter ($21\,\mathrm{GeV} < E_{\mathrm{cal}} < 24\,\mathrm{GeV}$, $24\,\mathrm{GeV} < E_{\mathrm{cal}} < 27\,\mathrm{GeV}$ and $E_{\mathrm{cal}} = 27.5\,\mathrm{GeV}$). For each bin in E_{cal} the difference between E_{cal} and the energy predicted using the double-angle kinematics, E_{DA}, is analyzed as a function of the measured presampler signal.

Here, the double-angle energy E_{DA} is calculated via (4.9) with Q^2 obtained from the double-angle method (4.12), that does not rely on the electron energy:

$$E_{\mathrm{DA}} = \frac{1}{2E_e} \cdot \frac{Q^2_{\mathrm{DA}}}{1 + \cos\theta_e} \ . \tag{4.2}$$

E_e is the energy of the electron beam and θ_e the polar angle of the scattered electron. MC studies show (Fig. 4.7) that E_{DA} yields a good prediction of the true energy in the relevant region above $20\,\mathrm{GeV}$ where good means that the bias of the reference quantity E_{DA} (with respect to the true value) lies well below that of the quantity under investigation, E_{cal}.

As depicted in Fig. 4.8, a straight line fit of the form $g(x) = f_i \cdot x$ is performed for each bin in E_{cal}, f_i being the fit parameter, x the pulse height of the presampler signal and $g(x) = E_{\mathrm{cal}} - E_{\mathrm{DA}}$. Afterwards, the mean value f of the f_i is determined which is a priori different for MC and data. The corrected electron energy is then calculated from $E_{\mathrm{corr}} = E_{\mathrm{cal}} - f \cdot x$.

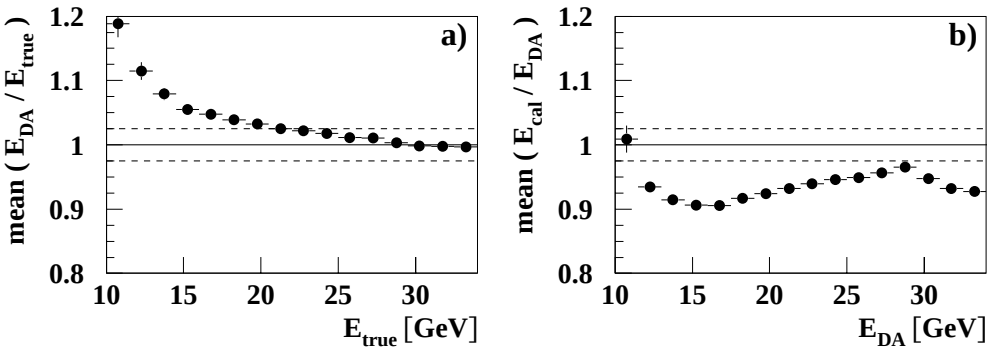


Figure 4.7: *Biases for electron-energy reconstruction methods in MC. (a) Ratio between the double-angle energy E_{DA} and the true energy E_{true} as a function of E_{true}. (b) Ratio between the calorimeter energy E_{cal} and E_{DA} as a function of E_{DA}. The dashed lines indicate a 2.5% deviation from 1.*

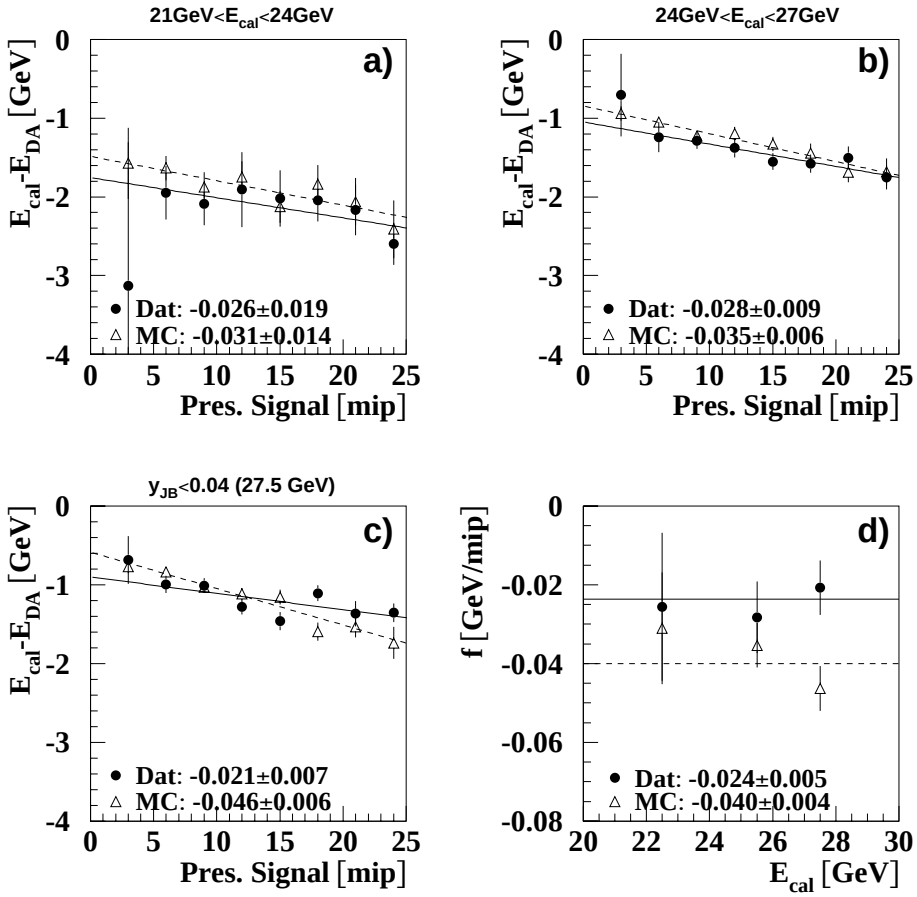


Figure 4.8: *Presampler studies to determine the correction function for the electron energy. (a–c) Difference between calorimeter electron-energy E_{cal} and double-angle energy E_{DA} for $55\ cm < R < 75\ cm$ as a function of the presampler signal in bins of E_{cal} for data and MC. (d) Slopes f of straight line fits to (a–c) as a function of the energy bin. The horizontal lines indicate the mean value for data (solid) and MC (dashed), respectively.*

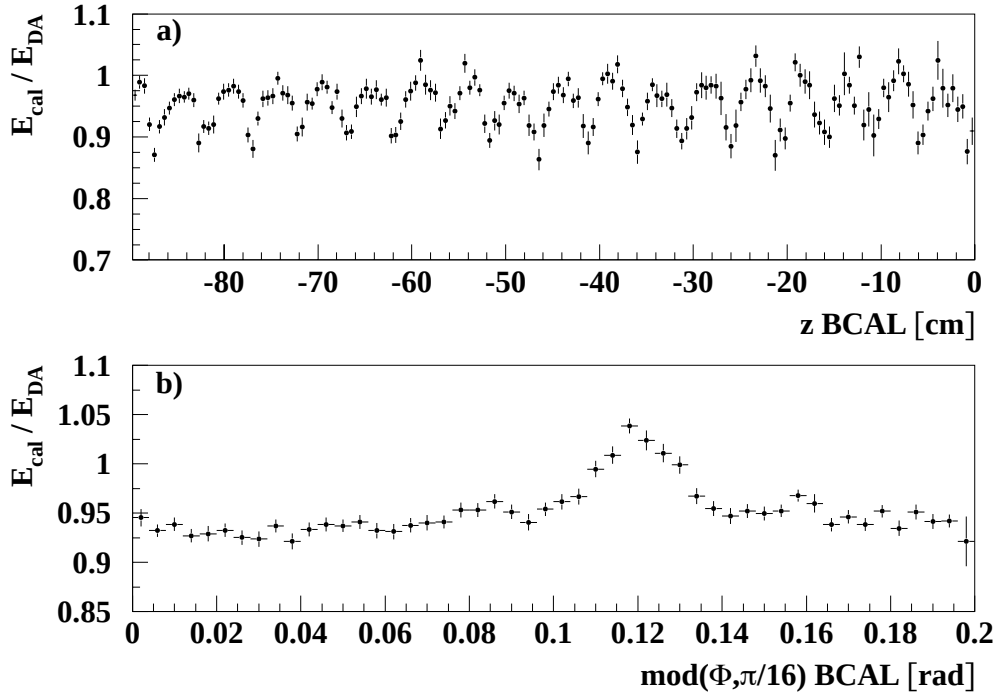


Figure 4.9: *Ratio of calorimeter energy E_{cal} and double-angle energy E_{DA} in BCAL for data as a function of (a) the z position and (b) the ϕ angle. In the latter case all 32 BCAL modules are folded on top of each other, hence only the ϕ range of one module (~ 0.2 rad) is visible.*

To account for a dependence of the fits on the radius R of the electron position in the RCAL, the studies are repeated for different R regions, where Fig. 4.8 shows the results for $55\,\mathrm{cm} < R < 75\,\mathrm{cm}$. The cut $y_{\mathrm{JB}} < 0.04$ (JB refers here to the Jacquet-Blondel method described in Chap. 4.6.3) applied in Fig. 4.8 c selects events in the "kinematic peak" at 27.5 GeV [5]. Figure 4.8 d shows the mean of the three f_i values for data and MC, which amounts to $(-0.024 \pm 0.005)\frac{\mathrm{GeV}}{\mathrm{mip}}$ and $(-0.040 \pm 0.004)\frac{\mathrm{GeV}}{\mathrm{mip}}$, respectively.

These results together with those from other R bins are in reasonable agreement with those of the original studies and therefore the correction routine is adopted in this analysis.

4.4.4.3 Non-uniformity corrections

Upon reaching the calorimeter the electron showers and deposits its energy in different cells. However, due to the shielding and support structure, the calorimeter is not homogeneous. Also the wavelength shifters of the readout, placed between modules, contribute to this inhomogeneity. If a particle hits a crack containing only inactive material, i.e. no wavelength shifters, the deposited energy is lost and the measured energy is too low (Fig. 4.9 a). However, if a particle deposits its energy partially in a wavelength shifter this leads to an increase in the measured

[5] For $y \approx 0$ the energy transfer from the electron to the proton is only very small and hence for low Q^2, i.e. small energy transfer from the proton to the electron, the energy of the scattered electron is close to that of the electron beam.

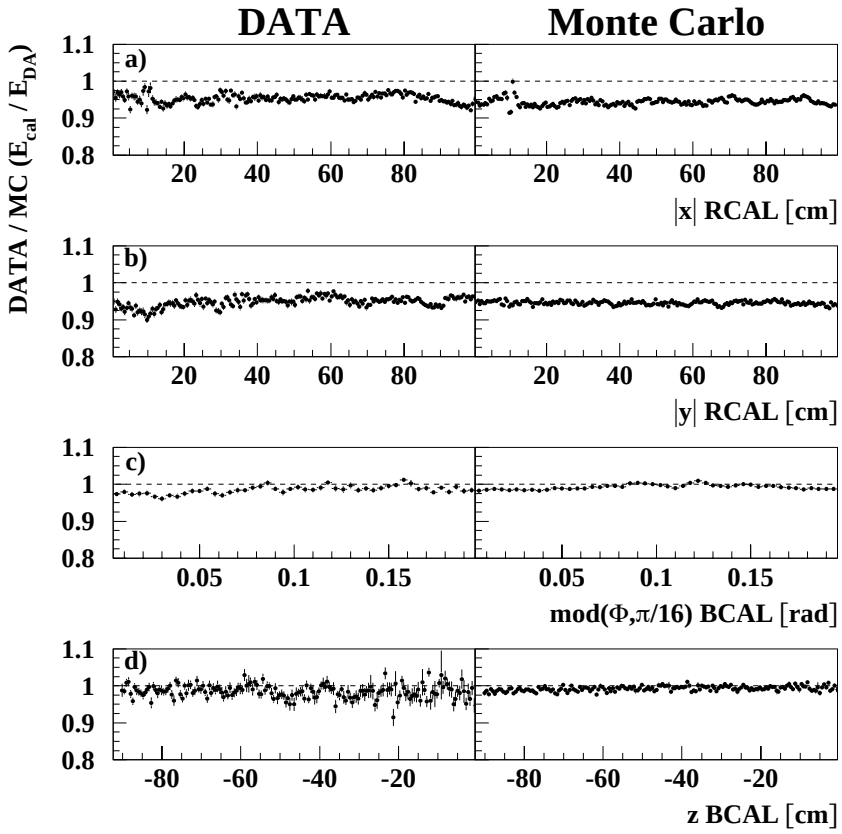


Figure 4.10: *Ratio of calorimeter energy E_{cal} and double-angle energy E_{DA} for data and MC after applying non-uniformity corrections as a function of (a) RCAL x-position, (b) RCAL y-position, (c) BCAL ϕ position and (d) BCAL z position. In (a) and (b) the negative axes are flipped onto the positive axes and in (c) all 32 BCAL modules are folded on top of each other.*

energy as the read-out electronics is calibrated for energy deposits in the active volume of the calorimeter (Fig. 4.9 b).

In order to correct for the inhomogeneities parameterizations of the energy shifts for the crack regions were determined separately for MC and data, both for RCAL [72] and BCAL. The effect of these corrections can be seen in Fig. 4.10 showing the ratio of calorimeter energy, E_{cal}, and double-angle energy, E_{DA}, as a function of x and y for the RCAL and as a function of z and ϕ for the BCAL. After corrections the ratios are almost flat and the offsets in data and MC agree quite well. However, in the data small slopes remain (Fig. 4.10 b and c) and the inhomogeneity at 10 cm for both MC and data could not be completely smoothed out (Fig. 4.10 a).

4.4.4.4 Dead-photomultiplier corrections

But even for the case that an electron hits the calorimeter near the cell center its measured energy may be distorted due to dead channels. The common procedure to correct for a single dead channel in a cell (Chap. 4.1.2) consists of doubling the energy of the working channel. However,

this method assumes that the imbalance[6] between the two PMs is not too big, otherwise the measured energy differs significantly from the true value. In the case of electrons in the BCAL a more sophisticated approach is used [81], which uses either the tracking information or the imbalance information from neighboring cells to correct the energy of the most energetic cell of the electron cluster.

The tracking method exploits the fact that a strong correlation exists between the imbalance in the most energetic cell and the distance in ϕ between the center of this cell and the impact position of the particle, obtained from the extrapolated track. The other method is based on the correlation between the imbalance of the degraded cell and that of a neighboring cell in the same module.

4.5 Reconstruction of the hadronic final state

After the identification and reconstruction of the scattered electron in the calorimeter, the hadronic final state can be reconstructed. It is defined as all energy deposits in the calorimeter and all tracks in the CTD excluding those already assigned to the electron. The angle γ_{had} of the hadronic final state is defined as

$$\cos\gamma_{\mathrm{had}} = \frac{\left(\sum_{\mathrm{had}} p_x\right)^2 + \left(\sum_{\mathrm{had}} p_y\right)^2 - \left(\sum_{\mathrm{had}} (E-p_z)\right)^2}{\left(\sum_{\mathrm{had}} p_x\right)^2 + \left(\sum_{\mathrm{had}} p_y\right)^2 + \left(\sum_{\mathrm{had}} (E-p_z)\right)^2} \tag{4.3}$$

$$= \frac{\left(p_t^{\mathrm{had}}\right)^2 - \left((E-p_z)^{\mathrm{had}}\right)^2}{\left(p_t^{\mathrm{had}}\right)^2 + \left((E-p_z)^{\mathrm{had}}\right)^2} \,. \tag{4.4}$$

The sums run over all objects of the hadronic final state with four-vector (E, p_x, p_y, p_z), where an object can be either a track, an energy cluster in the calorimeter or a single calorimeter cell depending on the algorithm applied.

Apart from imperfections of the calorimeter due to its finite energy- and spatial resolution, energy deposits, originating from particles scattered back from the FCAL surface (called back-splash), or generated by noisy cells can bias the reconstruction considerably. The same effect is caused by energy spread from collisions of particles with the beam pipe or the inner wall of the CTD in front of the calorimeter. These normally small energy deposits can have a large effect on the reconstructed hadronic angle as they also occur in regions far away from the hadronic final state and therefore have a large lever-arm. Also energy leakage from the main calorimeter is a source of mis-measurement which occurs particularly in the super-crack regions between the main calorimeter parts. The energy-reducing effect of dead material in front of the calorimeter is present everywhere, though especially prominent in the transition region between the CTD and the solenoid. In the following sections, three algorithms to reconstruct the hadronic final state are presented that are based on different quantities and have different levels of sophistication.

[6] The imbalance of a cell is defined as the difference between the heights of the two PM signals divided by their sum.

4.5.1 Cell-based approach

The simplest approach uses only energy and angular information from the cells in the calorimeter. Quantities like the total energy or the momentum vector are calculated by summing over the corresponding quantities of each cell:

$$E^{\text{had}} = \sum_i E_i \tag{4.5}$$

$$p_x^{\text{had}} = \sum_i E_i \cos\theta_i \cos\phi_i \tag{4.6}$$

$$p_y^{\text{had}} = \sum_i E_i \cos\theta_i \sin\phi_i \tag{4.7}$$

$$p_z^{\text{had}} = \sum_i E_i \sin\theta_i \,, \tag{4.8}$$

where i runs over all calorimeter cells not belonging to the electron and E_i and θ_i and ϕ_i are the corresponding energies and angles. However, this approach has the disadvantage that it does not take into account the imperfections of the detector which in many cases lead to wrong estimates of energy and hadronic angle γ_{had}.

4.5.2 ZUFOs

The *ZUFO* algorithm [80] uses apart from calorimeter- also tracking-information. The basic objects of the algorithm are energy deposits in the calorimeter grouped into clusters and CTD tracks. If possible the tracks are matched to calorimeter clusters, forming new objects. For each of the latter an algorithm decides whether the calorimeter or the track information will be used onwards. The decision depends for example on the number of tracks matched to the cluster (for more than 3 tracks the calorimeter information is always chosen) and the resolutions of the momentum and energy measurements. The *ZUFO* algorithm also corrects for back-splash and other small energy-deposits not belonging to the hadronic final state by removing all clusters with less than 3 GeV if they are far enough away from the actual hadronic system, i.e. have a polar angle greater than $\gamma_{\text{max}} = \gamma_{\text{had}} + \gamma_\delta$. γ_δ is determined from MC studies and minimizes the difference between true and reconstructed γ_{had}. This energy correction, in short called back-splash correction, was originally developed for the *CorAndCut* algorithm that will be discussed in the next section. The remaining objects are then used to calculate the different quantities of the hadronic final state according to (4.4)–(4.8), where the sums run now over all objects instead of cells.

4.5.3 CorAndCut

The *CorAndCut* algorithm [82] uses only calorimeter information. However, in contrast to the cell-based approach, it groups the energy deposits into clusters similarly to the *ZUFO* algorithm. The total energy of the cluster is the sum of the energies of the individual cells and its angle is the (energy weighted) mean of the individual angles. In contrast to the *ZUFO* algorithm,

CorAndCut corrects for energy loss caused by dead material in front of the calorimeter, where the correction is parameterized in bins of X_0 and measured energy. Additionally, differences in the measured energy of hadronic and electromagnetic deposits below 1 GeV due to the limited compensation capability of the calorimeter in this region (Chap. 3.2.1) are balanced out by parameterizing the difference as a function of the energy fraction in the EMC section.

After these corrections, low-energy deposits which are far away from the actual hadronic system are removed with the same algorithm already described in the previous section, and the hadronic energy E_{had} and angle γ_{had} are calculated. Finally, the former is corrected for energy loss in the super-crack regions located between the subcalorimeters.

4.5.4 Comparison of different algorithms

The cell-based approach has the worst performance as it lacks all corrections. A comparison between *ZUFOs* and *CorAndCut* [83] showed that each algorithm has regions where it is better than the other. However, overall *CorAndCut* shows superior performance concerning bias and resolution for the kinematic range investigated in this analysis.

For this thesis both the cell-based approach and the *CorAndCut* algorithm are used. The former has to be applied whenever information about the sub-calorimeter is needed, as *CorAndCut*, due to its corrections, only provides global calorimeter information. However, *CorAndCut* is exclusively used when cross sections are extracted.

4.6 Calculation of the kinematic variables

After the reconstruction of the scattered electron and the hadronic final state, the kinematic variables are evaluated. For this task, several algorithms exist, each based on different quantities of the electron and the hadronic final-state and hence with varying performance in the different regions of the kinematic plane.

4.6.1 Electron method (EL)

According to its name, the electron method uses only variables from the identified electron, i.e. its energy and angle:

$$Q^2_{\text{EL}} = 2E_e E'_e (1 + \cos\theta_e) \tag{4.9}$$

$$x_{\text{EL}} = \frac{E_e}{E_p} \cdot \frac{E' (1 + \cos\theta_e)}{2E_e - E'_e (1 - \cos\theta_e)} \tag{4.10}$$

$$y_{\text{EL}} = 1 - \frac{E'_e}{2E_e} (1 - \cos\theta_e) \ , \tag{4.11}$$

where E_e is the electron-beam energy, E'_e the energy of the scattered electron and θ_e its angle. This method was formerly used by fixed target DIS experiments and also frequently within ZEUS at the beginning of the data taking in 1992, as all other methods need information from the

hadronic final state whose proper reconstruction is a demanding task. However, for the electron method the electron energy has to be well known, which also needs careful investigations as discussed in Chap. 4.4.4.

4.6.2 Double-angle method (DA)

This method uses the polar angles of both the hadronic final state (γ_{had}) and the electron (θ_e).

$$Q^2_{\text{DA}} = 4E_e^2 \cdot \frac{\sin\gamma_{\text{had}}\,(1+\cos\theta_e)}{\sin\gamma_{\text{had}} + \sin\theta_e - \sin\left(\gamma_{\text{had}}+\theta_e\right)} \tag{4.12}$$

$$x_{\text{DA}} = \frac{E_e}{E_p} \cdot \frac{\sin\gamma_{\text{had}} + \sin\theta_e + \sin\left(\gamma_{\text{had}}+\theta_e\right)}{\sin\gamma_{\text{had}} + \sin\theta_e - \sin\left(\gamma_{\text{had}}+\theta_e\right)} \tag{4.13}$$

$$y_{\text{DA}} = \frac{\sin\theta_e\,(1-\cos\gamma_{\text{had}})}{\sin\gamma_{\text{had}} + \sin\theta_e - \sin\left(\gamma_{\text{had}}+\theta_e\right)}\ . \tag{4.14}$$

It has the advantage that it is independent of the calorimeter energy-scale and insensitive to photon radiation from the outgoing lepton-line. However, the angle of the hadronic final-state has to be determined with demanding accuracy.

4.6.3 Jacquet-Blondel method (JB)

The Jacquet-Blondel method [84] uses exclusively variables of the hadronic final-state:

$$Q^2_{\text{JB}} = \frac{\left(p_t^{\text{had}}\right)^2}{1-y_{\text{JB}}} \tag{4.15}$$

$$x_{\text{JB}} = \frac{Q^2_{\text{JB}}}{s\cdot y_{\text{JB}}} \tag{4.16}$$

$$y_{\text{JB}} = \frac{\left(E-p_z\right)^{\text{had}}}{2E_e}\ . \tag{4.17}$$

In general, this reconstruction method is used if no scattered electron can be identified, e.g. in CC events where the outgoing neutrino escapes the detector unseen. For NC events the Jacquet-Blondel method performs better than the other methods only in certain regions of the kinematic plane, e.g. at low Q^2 and low y.

4.6.4 Performance of reconstruction methods

The performance of the three different reconstruction methods is investigated by comparing the reconstructed kinematic variables to the true values using the generator information from the MC. Figures 4.11 and 4.12 show the distributions of the relative errors of Q^2 and x in the single-differential $d\sigma/dQ^2$ bins (Chap. 7.3.1), where the reconstructed Q^2 determines the bin in which an event is entered.

Obviously, the double-angle method provides superior performance with respect to resolution and bias over the whole kinematic range. To a large extend this holds also if the double-differential binning (Chap. 7.4.1) is used, though at very low x corresponding to very high y

the electron method is slightly better with respect to resolution. But as this difference is very small and the electron method performs much worse in other regions of the kinematic plane, the double-angle method is chosen as the standard reconstruction method for the kinematic variables in this analysis.

4.7 Summary of Chapter 4

In the course of this chapter the steps towards a fully reconstructed event have been discussed. Special attention has been given to the accurate reconstruction of the scattered electron as the most important signature of an NC event. For both the reconstruction of the hadronic final state and the kinematic variables several algorithms have been compared and those with the best performance were chosen for the future extraction of the cross sections. In the next chapter the event-reconstruction algorithms are applied to data and MC events and the reconstructed quantities are used to verify that the MC describes the data reasonably well.

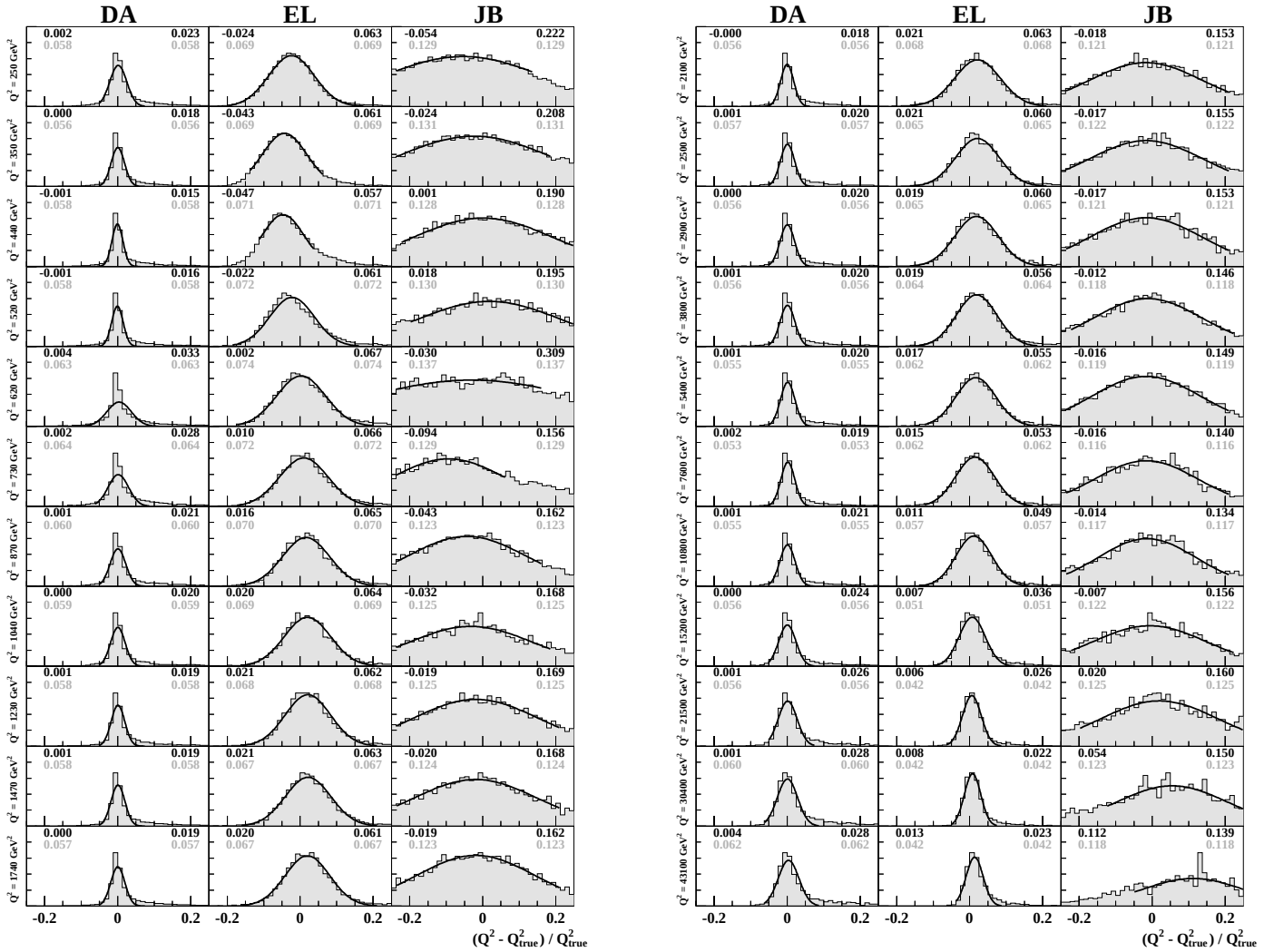


Figure 4.11: *Performance of reconstruction methods in $d\sigma/dQ^2$-bins for reconstructed Q^2. The upper numbers in each plot are the mean (left) and σ (right) from a Gaussian fit, whereas the lower numbers are the RMS (right) and the corresponding mean (left). DA = double-angle method, EL = electron method, JB = Jacquet-Blondel method.*

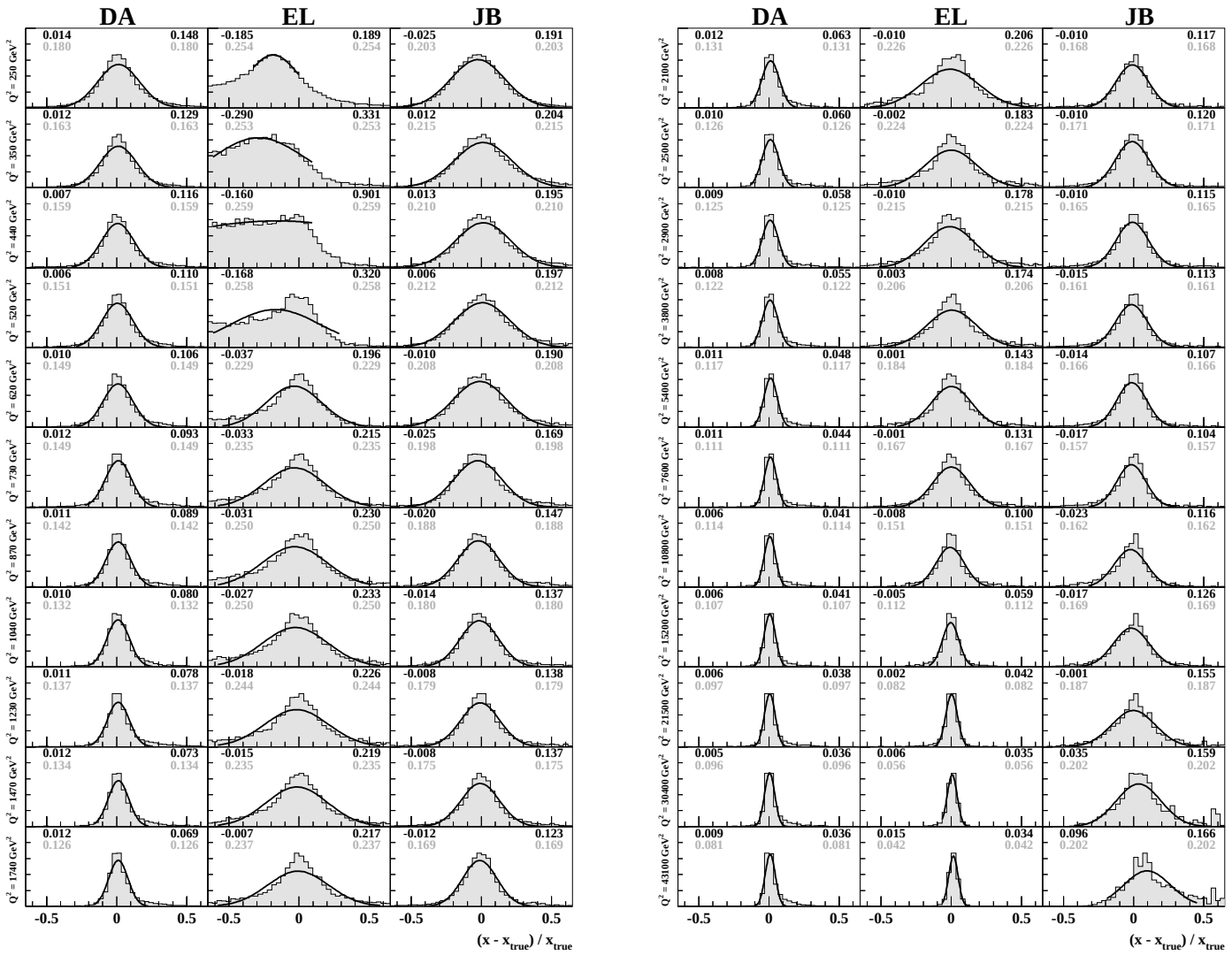


Figure 4.12: *Performance of reconstruction methods in $d\sigma/dQ^2$-bins for reconstructed x. The upper numbers in each plot are the mean (left) and σ (right) from a Gaussian fit, whereas the lower numbers are the RMS (right) and the corresponding mean (left). DA = double-angle method, EL = electron method, JB = Jacquet-Blondel method.*

Chapter 5

Data sets and event selection

Following the discussion on the reconstruction of the scattered electron and the hadronic final state in the previous chapter, this chapter starts out with a description of the data used in this analysis to extract the neutral-current e^-p cross sections. Then, the MC samples generated for signal and background processes are introduced. After presenting the online- and offline-event-selection, data and MC are compared in essential distributions and results from these comparisons are used to adjust the MC in order to achieve good agreement with the data.

5.1 Data

The data used for this analysis was recorded during the 1998/99 running period of the ZEUS detector (Fig. 5.1). In this period HERA operated with 27.5 GeV electrons in collision with 920 GeV protons, yielding a center-of-mass energy of 318 GeV. In the beginning of the running period in autumn 1998 the beam conditions were rather bad and therefore only $\sim 5\,\mathrm{pb}^{-1}$ of data were collected until Christmas. The situation improved much in the beginning of the following year and in approximately the same elapsed time more than twice the amount of data could be gathered. In May 1999 it was decided to switch back to positrons, and hence to date only $17\,\mathrm{pb}^{-1}$ of e^-p data at $\sqrt{s} = 318\,\mathrm{GeV}$ are available for analyses. The range of run numbers covered by the e^-p analysis is 30758 – 31752 for 1998 and 31784 – 32906 for 1999.

To remove runs with temporarily bad experimental conditions that could affect the physics results, all pieces of running information recorded during and after data taking are checked. Bad conditions are for example single, very noisy EMC cells that may fake electrons or problems with important detector or trigger components. Removing such runs reduces the integrated luminosity slightly to $15.75\,\mathrm{pb}^{-1}$.

To verify the quality of the remaining runs, the event yield in a run after all selection cuts (Chaps. 5.3 and 5.4) is divided by the run's integrated luminosity $\mathcal{L}$, and the result $N/\mathcal{L}$ is plotted as a function of the run number (Fig. 5.2 a). For "good" runs one expects the quantity $N/\mathcal{L}$ to scatter around a mean value, and the pull plot of this distribution should be centered at 0 and have $\sigma = 1$. As many runs have such low luminosities that only a few or even no events

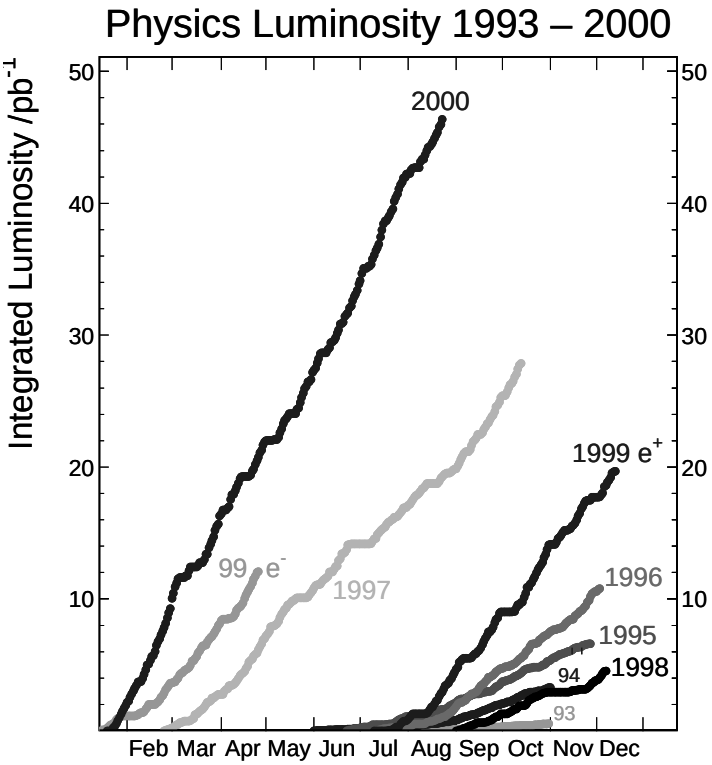


Figure 5.1. *Integrated luminosities recorded by ZEUS during the years 1993–2000. In 1993, 1998 and the first 4 months of 1999 e^- were accelerated, whereas in the other periods e^+ were used.*

are selected, all runs with $\mathcal{L} < 10\,\mathrm{nb}^{-1}$ are lumped together to a single "run" which is displayed near the left edge of Fig. 5.2 a as an open circle.

Overall, Fig. 5.2 a looks fine as it shows no distinctive features like runs with $N/\mathcal{L}$ far away from the mean value with respect to the error. The corresponding pull plot in Fig. 5.2 b has a mean value compatible with zero and the Gaussian fit yields a reasonable χ^2/ndf of 0.8. The fit width of 0.9 is $3\,\sigma$ away from the expected value of one, which is still acceptable and hence all selected runs are ready to be used in the analysis.

5.2 Monte Carlo

Due to the complexity of today's high-energy experiments MC event-simulations are inevitable for the extraction of physic results from the data (Chap. 3.4). Therefore, the best-possible description of the physics processes and of the detector is essential for the quality of the analysis.

5.2.1 Signal Monte Carlo

The prime aim of this analysis is the extraction of the NC e^-p cross section in lowest electroweak order. Therefore, strictly speaking, only DIS events without any radiation should be counted as signal events. However, most of the events with FSR or ISR can not be identified as such and as their cross section is calculable they are simulated in MC and "removed" afterwards in the extraction of the cross section (Chap. 7.1).

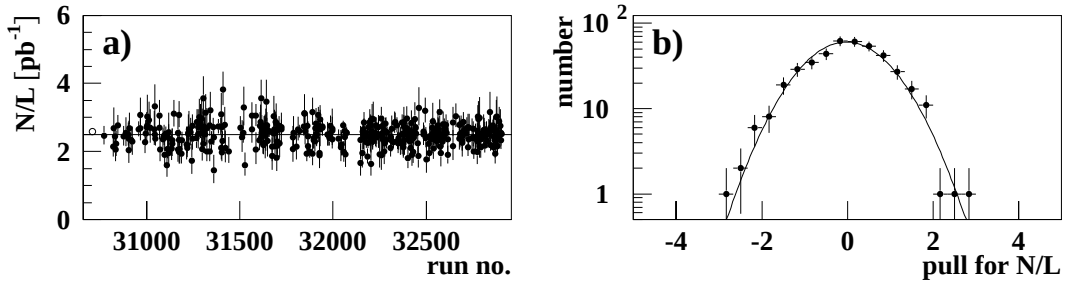


Figure 5.2: *(a) Number of events per run divided by the luminosity of that run ($N/\mathcal{L}$) as a function of run number. (b) Pull plot of $N/\mathcal{L}$. The Gaussian fit yields:* mean $= -0.028 \pm 0.048$ *and* $\sigma = 0.90 \pm 0.03$ *with* $\chi^2/ndf = 0.8$.

The signal MC is made up of samples of e^-p DIS events generated by DJANGOH 1.1, using CTEQ5D PDFs, with various lower Q^2-cuts. Table 5.1 lists the different samples where the cut is applied on the Q^2 calculated at the electron vertex. The total sample size is sufficient to ensure that the statistical errors from MC are negligible compared to those from data.

For the correct comparison of distributions of quantities between data and MC the latter has to be weighted according to the luminosities of data and MC. The number of events N and the cross section σ are related via the equation

$$N = \mathcal{L} \cdot \sigma \ , \tag{5.1}$$

where $\mathcal{L}$ is the luminosity. The weight w for MC events is then calculated with

$$w = \frac{\mathcal{L}^{\text{data}}}{\mathcal{L}^{\text{MC}}} \ , \tag{5.2}$$

where $\mathcal{L}^{\text{data}}$ and $\mathcal{L}^{\text{MC}}$ are the corresponding luminosities for data and MC.

MC version	lower Q^2-cut	cross section	# of events	Luminosity
num98v3.1	100 GeV^2	8155 pb	179 870	22.06 pb^{-1}
num98t3.1	100 GeV^2	8155 pb	174 651	21.42 pb^{-1}
num98t3.1	400 GeV^2	1196 pb	59 976	50.15 pb^{-1}
num98t3.1	1 250 GeV^2	217.0 pb	23 288	107.3 pb^{-1}
num98t3.1	2 500 GeV^2	71.77 pb	12 000	167.2 pb^{-1}
num98t3.1	5 000 GeV^2	21.65 pb	12 000	554.3 pb^{-1}
num98t3.1	10 000 GeV^2	5.363 pb	11 960	2 230 pb^{-1}
num98t3.1	20 000 GeV^2	0.8473 pb	11 997	14 160 pb^{-1}
num98t3.1	30 000 GeV^2	0.1853 pb	5 995	32 350 pb^{-1}
num98t3.1	40 000 GeV^2	0.04265 pb	5 998	140 600 pb^{-1}
num98t3.1	50 000 GeV^2	0.009190 pb	6 000	652 900 pb^{-1}

Table 5.1: *Signal MC samples with different lower Q^2-cuts. Note that there is no difference between MC version num98t3.1 and num98v3.1 except the name.*

MC version	lower Q^2-cut	cross section	# of events	Luminosity
num98v3.1	100 GeV^2	8155 pb	49 688	6.093 pb^{-1}
num98v3.1	400 GeV^2	1196 pb	30 000	25.08 pb^{-1}
num98v3.1	1250 GeV^2	217.0 pb	19 995	92.14 pb^{-1}

Table 5.2: *Diffractive MC samples with different lower Q^2-cuts.*

Moreover, in this case it has to be taken into account that the different MC samples partially overlap each other in Q^2. Therefore, the following formula is used to calculate the weight w for a MC event with $Q^2 = Q_0^2$:

$$w = \frac{\mathcal{L}^{\text{data}}}{\sum_i \mathcal{L}_i^{\text{MC}}} \ , \tag{5.3}$$

where i runs over all samples with $Q_0^2 > Q_{\text{cut}}^2$ and Q_{cut}^2 being the lower Q^2 cut on the generated sample as listed in the second column of Table 5.1.

5.2.2 Diffractive Monte Carlo

One of the rather surprising results of the HERA physics harvest is that diffractive events (Chap. 2.4) with a clear event topology contribute considerably contribution to the DIS cross section at low and medium Q^2 up to several 100 GeV^2. Though these events belong to the signal events, the topology of their hadronic final-state differs significantly from that of "normal" DIS events as there is no color flow between the hadronic jet and the proton. Therefore, using only "normal" DIS events potentially biases the result of the event selection for the MC.

To avoid this, the approach [53] chosen for this analysis is mixing a fraction f of diffractive events to the standard DIS MC, where f is determined individually in each bin of the double differential binning (Chap. 7.4.1). The aim of this approach is to modify the topology of a proper fraction of DIS events, but *not* their distribution in the kinematic variables x and Q^2 [1]. For this purpose the diffractive event generator RAPGAP 2.0806 [68] is modified such that the program continues to produce diffractive events, however to 100% in DIS mode, i.e. with the standard non-diffractive NC cross section and CTEQ5D as parton distributions.

Table 5.2 shows the available samples of diffractive events. As the fraction of diffractive events is at most 10% of the total NC DIS cross section, vanishing for $Q^2 \gtrsim 1\,000\,\text{GeV}^2$, the amount of generated events can be kept relatively small.

At first, it has to be verified that the generator indeed delivers the same x and Q^2 distributions as the standard DIS MC generator DJANGOH. Figure 5.3 shows the comparison of the distributions in all three kinematic variables Q^2, x and y for events with $Q^2 > 100$ GeV^2. Overall, good agreement between the distributions of the two generators is observed, though for $Q^2 \lesssim 30\,\text{GeV}^2$ RAPGAP produces fewer events than DJANGOH. However, this lies far below

[1] Note that this method does not claim to measure the diffractive cross section but rather takes care of the difference in the hadronic final state between "normal" DIS and diffractive events.

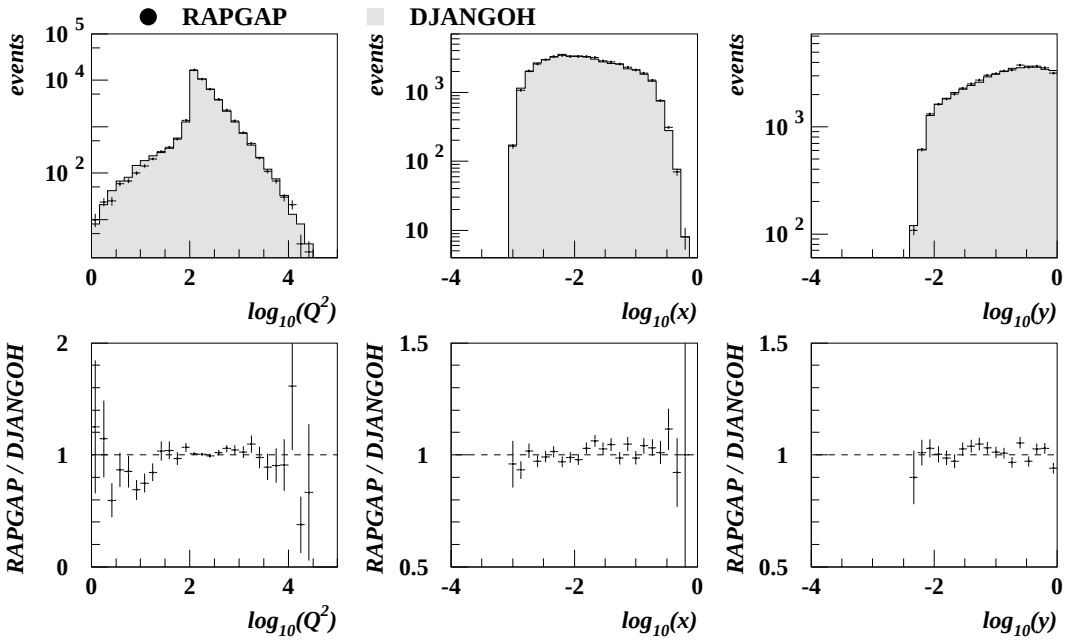


Figure 5.3: *Comparison of distributions of kinematic variables between* RAPGAP *and* DJANGOH *for events with* $Q^2 > 100\ GeV^2$*. In the upper row the crosses represent the* RAPGAP *sample whereas the histograms mark the* DJANGOH *sample. In the lower row the ratios of the distributions are plotted.*

the region of $Q^2 > 185\,\mathrm{GeV}^2$ covered by this analysis and can therefore affect the results of the analysis only indirectly through migrations to higher Q^2 values.

The fraction f of diffractive events is a function of x and Q^2 and is determined by fitting $f \cdot V_{\mathrm{DIF}} + (1-f) \cdot V_{\mathrm{DIS}}$ to V_{data}, where V is the distribution of a variable in which diffractive and non-diffractive events are well distinguishable. In the kinematic range considered in this analysis the so-called η_{max} variable, the pseudo-rapidity of the most forward calorimeter object, proved to be suitable for this task. Figure 5.4 compares η_{max} from data and standard DIS MC (DJANGOH) for events with $Q^2 > 185\,\mathrm{GeV}^2$. The region of missing diffractive events at low η_{max} is clearly visible. As the right flank also shows deviations between data and MC, a known weakness of the simulation, only the left flank up to $\eta_{\mathrm{max}} = 4$ is used for fitting. In order to be independent of the nominal normalizations of the DJANGOH and RAPGAP samples, their η_{max} distributions are normalized to the data histogram in the range $\eta_{\mathrm{max}} < 4$.

Fits of η_{max} distributions in the individual x-Q^2 bins yield fractions f that are plotted in Fig. 5.5 a – f as a function of Q^2 in different bins of x. As the fractions are almost constant within a single x bin, each x bin gets assigned a mean fraction $\bar{f}$. The $\bar{f}$ values are plotted in Fig. 5.5 g as a function of $\log(x)$ together with a straight line fit to the points. The function $\bar{f}(x)$ obtained from the fit is then used to weight the standard DIS and diffractive MC according to

$$w'_{\mathrm{DIS}} = w_{\mathrm{DIS}} \cdot \big(1 - \bar{f}(x)\big) \qquad w'_{\mathrm{DIF}} = w_{\mathrm{DIF}} \cdot \bar{f}(x)\ , \tag{5.4}$$

where w_{DIS} and w_{DIF} are the nominal MC weights. The reweighted MC is then used to generate new η_{max} distributions and the whole procedure is iterated until it converges. After 2 cycles

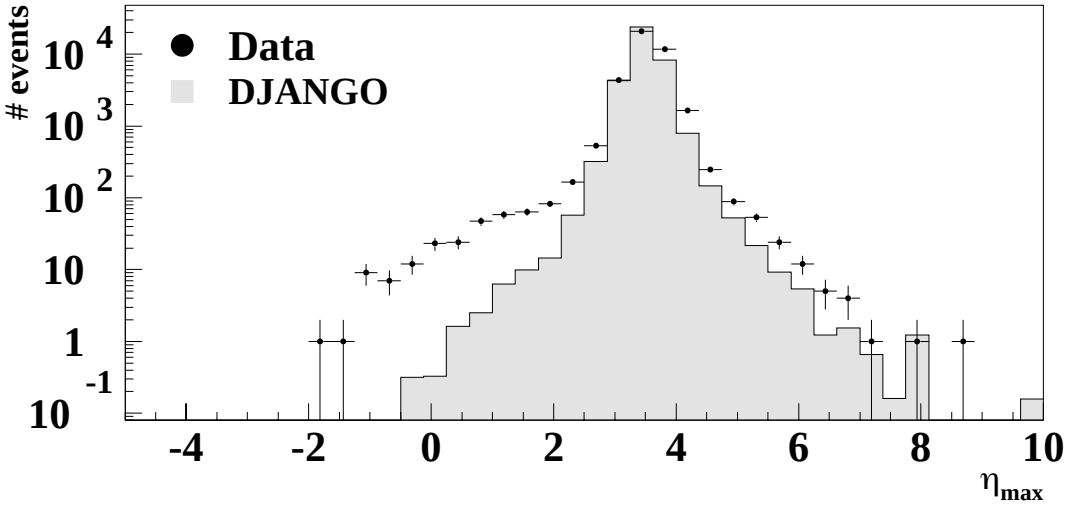


Figure 5.4: *Comparison between data and standard DIS MC (*DJANGOH*) in* $\eta_{\max}$.

the variation of the function is down to the sub-per-mill level and therefore the procedure is stopped.

The final function obtained from the studies and used for reweighting the MC in the future is

$$\bar{f}(x) = -(0.064 \pm 0.036) - (0.071 \pm 0.017) \cdot \log x \ . \tag{5.5}$$

The resulting $\eta_{\max}$ distribution is plotted in Fig. 5.6 showing that the left flank is now reasonably well described by MC.

5.2.3 Background Monte Carlo

5.2.3.1 Photoproduction

Studies [9] showed that the main background for NC events originates from PHP. Though the probability that PHP events can fake a high-Q^2 NC event is very small, the enormous cross section of the PHP process and the absence of a true scattered electron in the final state make up for the low probability.

The event generator for PHP events is HERWIG 5.9 [69]. Table 5.3 lists the PHP samples used for this analysis. A combination of different MC versions with only minor differences is used to maximize statistics. In order to limit the number of PHP events that have to be generated, several cuts are imposed on the generator level. First, the transverse momentum of the hard subprocess has to be above 4 GeV. Then, either the total transverse-energy E_T has to be greater than 30 GeV or the total transverse-momentum has to exceed 6 GeV, where for these quantities only particles hitting the main calorimeter are considered. Studies showed [55] that these generator-level cuts, originally developed for CC analyses, also remove some events with $E_T < 30$ GeV that would otherwise pass the NC selection cuts. However, this affects less than 10% of the events and studies are underway with the aim to optimize the generator-level cuts for NC. However, these studies together with the planned generation of several millions additional

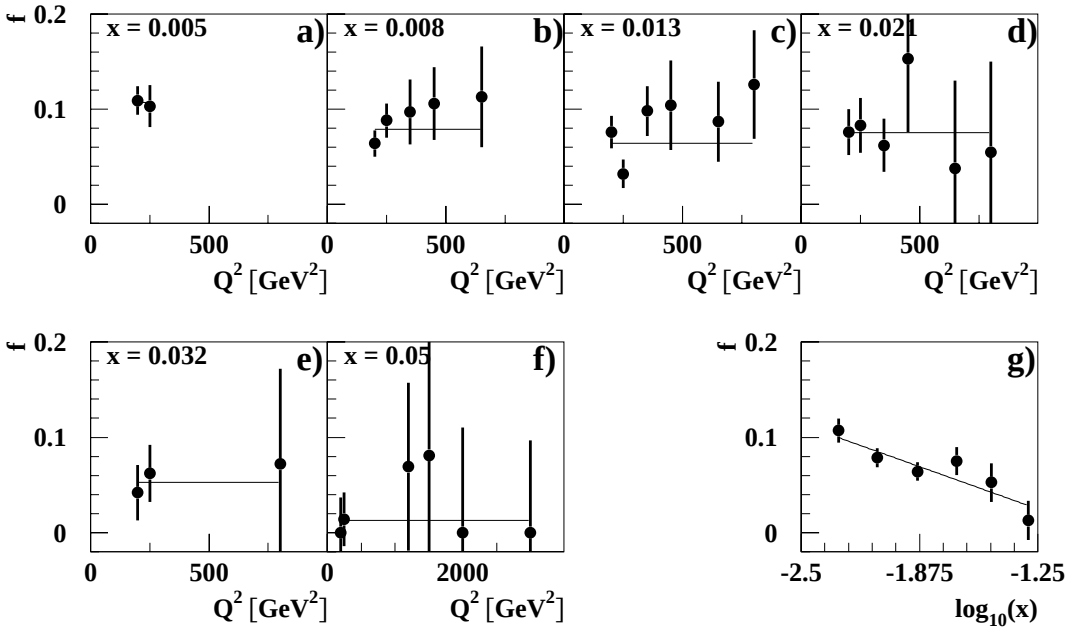


Figure 5.5: *Determination of the fraction of diffractive events for MC. (a–f) show the fraction f of diffractive events as a function of Q^2 in several x bins. In (g) the mean values of f for each x bin is plotted as a function of x, where the solid line represents a straight-line fit.*

PHP events will take at least 1–2 months and hence, these samples are not available for this thesis.

The need for more PHP events is also apparent when considering that after all selection cuts (Chaps. 5.3 and 5.4) only 68 direct and 69 resolved PHP events remain out of 570 000 respectively 1 490 000 generated events.

5.2.3.2 Photoproduction normalization

The absolute normalization of the PHP MC is not very well known. Therefore, a study has been performed to verify the normalization in MC by comparing the number of tagged PHP events in data and MC. The tagging detector used here is the LUMI-e detector (Chap. 3.2.5). The basic

HERWIG	MC versions	upper Q^2-cut	# of events	Luminosity
direct	num98t3.1 num98v3.1	$4\,\mathrm{GeV}^2$	570 000	$183.2\,\mathrm{pb}^{-1}$
resolved	num98t3.0 num98t3.1 num98t3.3 num98v3.1	$4\,\mathrm{GeV}^2$	1 490 000	$108.4\,\mathrm{pb}^{-1}$

Table 5.3: *Direct and resolved PHP MC samples used in the analysis.*

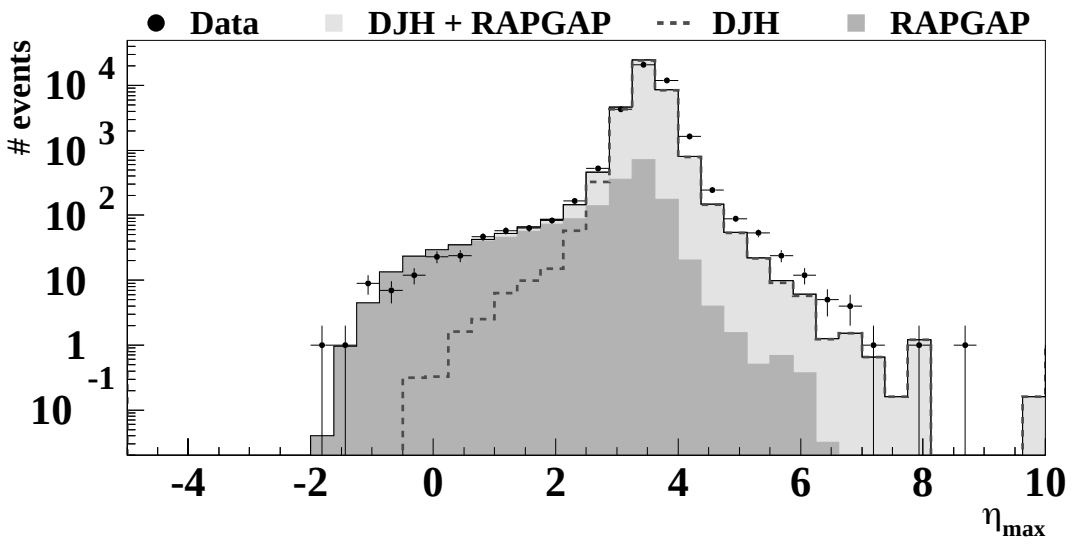


Figure 5.6: *Comparison of $\eta_{\max}$ distribution between data and MC after including an x-dependent fraction of diffractive events. Points are data, the dark histogram and the dashed line represent the* RAPGAP *and standard DIS MC (DJH), whereas the light histogram is the sum of DJH and* RAPGAP*.*

idea is to normalize the MC sample to the data sample for the tagged events and then scale the MC with its tagging efficiency to obtain the total number of PHP events. This scaled sample together with the DIS sample should then yield a good description of the data.

Unfortunately, only a small fraction of the available PHP MC events can be used for this study as the rest of the events was generated with a faulty simulation of the LUMI-e tagger, rendering it unusable for these kind of investigations. In order to obtain reasonable statistics the selection cuts are loosened with respect to the rest of the analysis: the Q^2 cut is lowered to $120\,\mathrm{GeV}^2$ and the electron selection is completely omitted except for the electron energy cut[2] at 10 GeV and the cut on the z position of the reconstructed vertex. Additionally, a total transverse energy of at least 30 GeV is required which simulates a similar cut already performed on the generator level for the PHP MC. Tagged events are also required to have less than 2 GeV in the LUMI-γ detector. This cut largely suppresses *ep*-bremsstrahlung events as the radiated photon is detected with 97% efficiency [85] in the LUMI-γ detector. Also note, that for electrons interacting with beam- or gas-protons within $\sim 6\,\mathrm{m}$ of the nominal IP[3] the tagger acceptance is limited to 5–18 GeV due to the beam optics.

In an event where all particles are detected and perfectly measured the quantity $E - p_z = \sum_i (E_i - p_{z_i})$ yields exactly 55 GeV due to the conservation of energy and longitudinal momentum. The sum includes both the hadronic final state and the electron. If particles escape through the forward beam-pipe the effect on $E - p_z$ is negligible. However, if particles are lost through the rear beam-pipe as in the case of PHP events, $E - p_z$ is reduced by roughly twice their energy.

[2] Removing also this cut leads to an excess in data over MC around $E - p_z \approx 40\,\mathrm{GeV}$. The reason for this has not been found yet.

[3] The first bending/focusing magnets with respect to the IP are located at $\sim \pm 6\,\mathrm{m}$. In between the electrons follow a straight flight path.

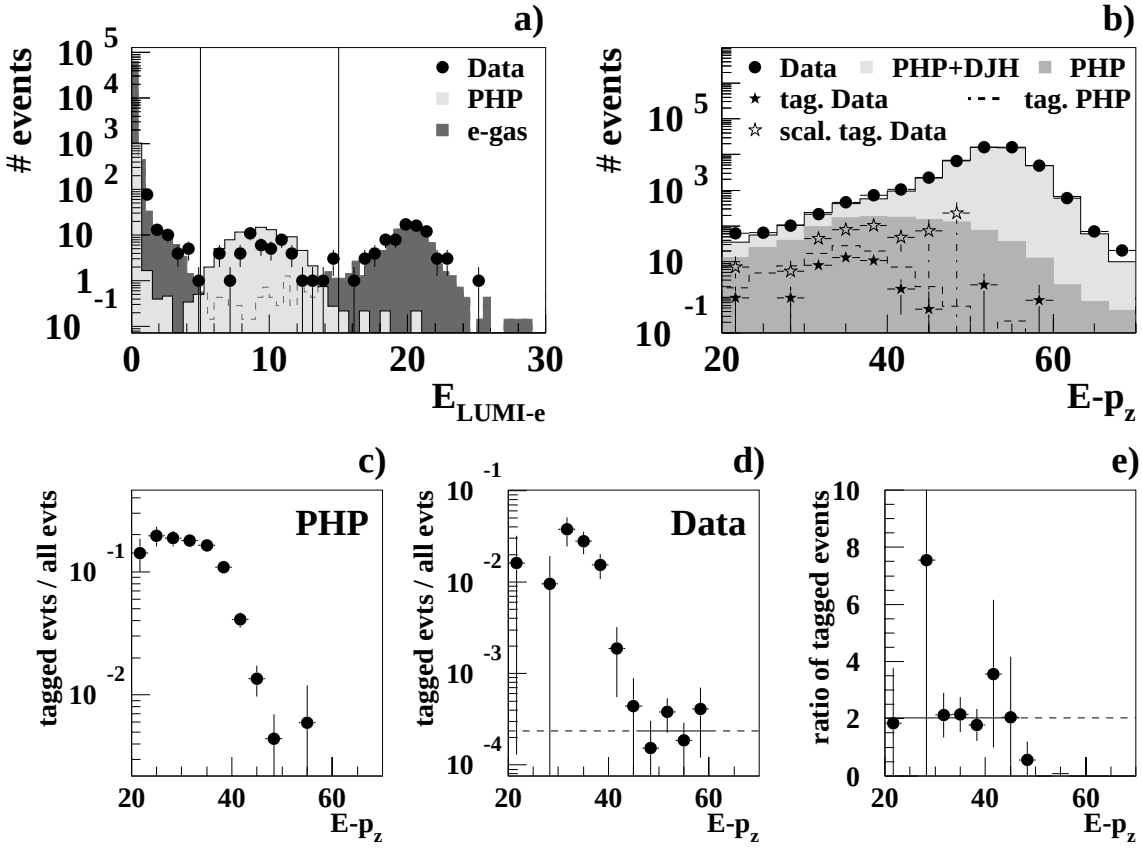


Figure 5.7: *Normalization of the PHP sample: (a) comparison of energy distribution in the LUMI-e tagger for data and MC; (b) $E - p_z$ distribution for data and MC (for individual contributions see text); (c,d) tagging efficiency for data and MC; (e) ratio between the number of tagged events in data and MC. The vertical lines in (d,e) are fits to the distributions, where the solid parts mark the fit ranges. Note that for these plots the electron energy-smearing discussed in Chap. 5.7 has already been applied.*

Therefore, the comparison of data and MC in this study is performed in this variable. The normalization is determined according to the number of generated events and the cross section used in the generator (nominal luminosity). In order to verify that the LUMI-e tagger is well simulated, first, the energy distributions of PHP MC and data in this tagger are compared (Fig. 5.7 a). Apart from an obvious normalization difference in the region between 5 and 15 GeV, the second peak as well as the low-energy behavior of the data is not at all reproduced by the PHP MC. A closer inspection of these events shows that they look like normal DIS events with an $E-p_z$ around the nominal value of 55 GeV, but with an additional energy cluster of about 20 GeV in the LUMI-e tagger. The dark shaded histogram (normalized to the second peak), matching both the high and low $E_{\mathrm{LUMI-e}}$ region of the data almost perfectly, represents electron-gas bremsstrahlung (e-gas) events, taken with a random trigger. In these events an electron interacts with atoms of the residual gas inside the beam pipe via emitting a bremsstrahlung photon. Due to its reduced energy the electron is afterwards deflected from the nominal electron-beam path and can hit the LUMI-e tagger. Such e-gas events occurring between the two inner magnets of ZEUS at ± 6 m are suppressed by the LUMI-γ cut and the limited acceptance of the tagger,

forming the minimum in Fig. 5.7 a (dashed line). MC studies [85] showed that the outer dark shaded regions in Fig. 5.7 a are not caused by PHP, *ep*-bremsstrahlung or *e*-gas events near the nominal IP. A reasonable explanation would be *e*-gas reactions occurring upstream the inner magnet of ZEUS as for these electrons the tagger acceptance is different due to the additional magnet(s) the electron has to pass before reaching the tagger.

The *e*-gas events occur in all samples of the $e^- p$ 1998/99 data as so-called overlay events, where a "normal" DIS event and for example an *e*-gas event occur at the same time. Limiting the energy range for tagged events to 5–15 GeV removes most of the overlay events, which is sufficient for this study.

The tagging efficiencies for data and MC are displayed in Fig. 5.7 c,d as functions of $E - p_z$. Note that the two distributions are not directly comparable as the PHP MC, in contrast to the data, does not contain any DIS events. Though both distributions have similar shapes, that of the data has a flat part around the nominal $E - p_z$ value of 55 GeV. This is caused by overlay events containing a QED-bremsstrahlung process (Chap. 3.2.5), which deposits some energy in the LUMI-e tagger. These events do not only appear around $E - p_z = 55$ GeV but are expected to be uniformly distributed over the whole $E - p_z$ range. Therefore, a constant is fitted to the region between 45 and 60 GeV (horizontal line in Fig. 5.7 d) and the resulting offset value is used to subtract the overlay events statistically from the tagged $E - p_z$ distribution in data.

The normalization of the PHP MC to the data is performed with the $E - p_z$ distribution of the tagged events, displayed in Fig. 5.7 b as full stars (data) and dashed line (PHP MC). Obviously, the number of tagged PHP MC events comes out too high as already observed in Fig. 5.7 a. A fit to the ratio of the two distributions (Fig. 5.7 e) yields a correction factor of 2.0 ± 0.4. Accordingly, also the expected total number of PHP events, obtained by multiplying the tagged events in data with the MC efficiency and represented by the open stars in Fig. 5.7 b, is too low compared to the corresponding distribution of the PHP MC (dark histogram, labeled PHP). Astonishingly enough, the total $E - p_z$ distributions of data (points) and MC (light histogram, labeled PHP+DIS) show very good agreement, except for the lowest bin, where the data is higher than the MC. However, this is far away from the lower $E - p_z$ cut at 38 GeV and can therefore be ignored.

The reason for the discrepancy in the PHP normalization between "all events" and "tagged events" is still unknown. The data is described quite well by the MC, using the nominal PHP luminosity. Therefore, this luminosity will be used for the analysis. The uncertainty on the PHP normalization is then estimated according to the fitted ratio in Fig. 5.7 e, i.e. as a systematic check the contribution of the PHP MC will be halved.

5.2.3.3 QED-Comptons

QED-Compton events, though containing a scattered electron, are not considered as signal events, as they comprise an additional vertex and are therefore, similar to FSR and ISR events, no longer of lowest electromagnetic order. QED-Compton events can be either elastic or inelastic. The former case has a very clear signature (empty calorimeter apart from two electromagnetic clusters, one with a track pointing towards it if inside the CTD acceptance) and the events

are therefore easy to reject. The latter case is already simulated within DJANGOH. However, the cross section is not well known and therefore, though the overall cross section is very small compared to the standard DIS cross section, migrations from low to high Q^2 can lead to a significant uncertainty in the predicted number of events in the highest Q^2-bins.

A detailed study [86] shows good agreement between data and MC in distributions of various variables like calorimeter energy and θ angles of photon and electron for tagged inelastic QED-Compton events. Additionally, all events in data with $Q^2 > 20\,000\,\mathrm{GeV}^2$ have been scanned visually but no elastic QED-Compton event was found. Hence, no effort is made to reject this kind of events.

5.3 Online event-selection

The basic setup of the ZEUS three-level trigger-system has been described in Chap. 3.3. This section is intended to give a more detailed description of the trigger chain used to select the events for this analysis.

The event selection starts with the requirement of **DST-bit 12**. This DST bit consists of a TLT part, i.e. a logical OR of TLT bits, a calorimeter-timing part that ensures that the selected events are compatible with an *ep* interaction, and a part that imposes cuts on quantities available after the event builder. In this case an electron with an energy of at least 5 GeV, a resulting $Q^2_{\mathrm{DA}} > 80\,\mathrm{GeV}^2$ and $y_{\mathrm{el}} < 0.95$ is demanded, where the electron has to be found by at least one of five different electron finders including SINISTRA95. Additionally, $E - p_z + 2 \cdot E_{\mathrm{LUMI}-\gamma}$ has to be greater than 30 GeV, where $E_{\mathrm{LUMI}-\gamma}$ is the energy deposited in the LUMI-γ tagger (Chap. 3.2.5).

The **TLT** part of DST bit 12 contains an OR of 13 different TLT bits. Apart from various SLT and FLT trigger bits which are each combined by a logical OR, the main requirement is an electron found by one of the electron finders, where the selection criteria are lower with respect to those of the DST bit. In addition, cuts are imposed for example on the radius of the electron in the RCAL or on the $E - p_z$ value and the calorimeter timing.

The **SLT** bits are a reduced form of the TLT bits with a combination of FLT bits, relaxed timing requirements but without electron finder algorithms.

On the **FLT** level only basic information from the detector components is available. For the selection of NC DIS events mainly calorimeter information is used. There exist several trigger bits each requiring certain conditions in order to fire the bit. Table 5.4 lists the main FLT bits used in this analysis together with a brief explanation of their main features.

The online event-selection has the task to reduce the amount of data, coming from the detector, to a manageable (storable) size and hence, rather loose cuts are applied in order not to sacrifice interesting events. The following offline event-selection uses much stricter cuts to reduce the background as far as possible while preserving most of the signal events.

trigger bit	selects events
GFLT 28	with high total transverse momentum E_T
GFLT 30	with isolated energy in the RCAL and at least 2 GeV in the EMC section of the RCAL
GFLT 40	where the total energy in the EMC cells is greater than 15 GeV
GFLT 41	where the transverse energy is greater than 21 GeV
GFLT 43	with $E_T > 12$ GeV and a "good track"
GFLT 44	where the BEMC energy is greater than 4.8 GeV and that have a good track or the REMC energy is greater than 3.4 GeV
GFLT 46	similar to GFLT 28
GFLT 47	similar to GFLT 46 and 28
GFLT 50	with energy in REMC and BPC
GFLT 62	with an isolated energy deposit in the calorimeter (QED-Compton bit)

Table 5.4: *FLT bits used for this analysis and a short explanation of the main features of the selected events. Each of the GFLT bits can also have one or more vetoes, i.e. if certain conditions are met the FLT bit does not fire even if the requirement described above is fulfilled. Tracks are reconstructed by a simple algorithm using z-by-timing information of the CTD signals and crude pattern recognition. R/BEMC refers to the energy in the EMC sections of the R/BCAL and E_T is the total transverse energy excluding the two rings of calorimeter towers nearest to the forward beam pipe.*

5.4 Offline event-selection

The main signature of a DIS event is an isolated electron in the main calorimeter. Therefore, an electron identification with high efficiency and purity is an essential requirement for the offline event-selection. A major feature that distinguishes an electron cluster from other electromagnetic clusters, e.g. photons, is an associated track with matching momentum. However, the CTD has a limited acceptance that depends on the minimum number of super-layers a track is required to pass. In general, for a well reconstructed track, three super-layers, i.e. 2 normal- and 1 stereo-layer, are demanded. This translates, with a safety-margin added, into a minimum radius of $r = 45$ cm at the CTD end-caps. Assuming for example an interaction at the nominal IP, i.e. at $z = 0$ cm, the corresponding polar angles are 23° and 169°, as depicted in Fig. 5.8. In general, the angles are calculated for each event individually, depending on the respective reconstructed z-vertex position.

5.4.1 Electron identification

- **Electron finder:** The electron finder EM is described in Chap. 4.4.1. The electron candidate with the highest probability is selected. For the event to be accepted, the probability of this electron has to exceed $1 \cdot 10^{-3}$. For electrons with angles outside the acceptance of the CTD in the forward region of the detector the probability cut is increased to $1 \cdot 10^{-2}$.

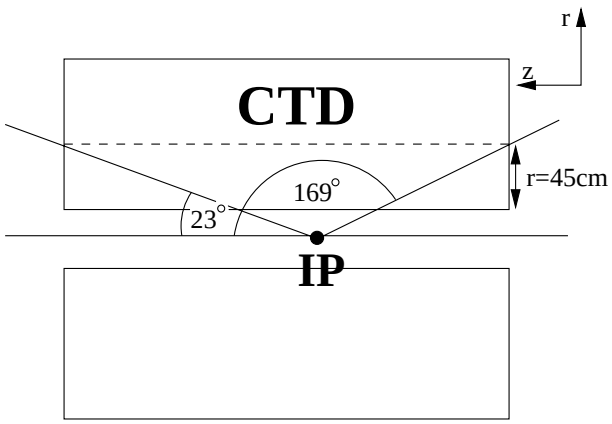


Figure 5.8: *Definition of the CTD acceptance. The picture is not drawn to scale.*

- **Isolation:** In contrast to electrons within jets or electromagnetic clusters from other particles, the energy deposits of scattered electrons are in most cases isolated in the calorimeter. Therefore, it is required that at most 5 GeV of the calorimeter energy inside an η-ϕ-cone (ϕ measured in radians) with a radius of 0.8 around the electron-candidate position does not belong to the electron cluster.
- **Electron energy:** Electromagnetic clusters from neutral pions with energies $\gtrsim 2\,\text{GeV}$ decaying into two photons are indistinguishable from those of electrons[4]. However, their energy spectrum falls steeply with increasing energy and therefore, a cut of 10 GeV on the electron energy is imposed which suppresses most pions. For electrons outside the forward CTD-acceptance this cut is replaced by a cut of 30 GeV on the transverse energy E_t of the electron. This much harder cut is necessary to account for the high background originating from the proton remnant.
- **Track matching:** Inside the CTD acceptance a track associated with the electromagnetic cluster with at least 5 GeV is required. The selection criteria for the track are described in Chap. 4.4.1.1. Additionally, the distance of closest approach between track and cluster center has to be less than 10 cm.

5.4.2 Background suppression

- **Cut on $E-p_z$:** A lower cut on $E-p_z$ reduces background from PHP events, as the electron escapes through the rear beam-pipe and hence, reduces $E-p_z$ by roughly twice its energy. Also ISR events with a high-energy radiated photon are excluded. An upper cut on $E-p_z$ on the other hand removes so-called overlay events, where for example a Bethe-Heitler or beam-gas event occurs at the same time as the DIS event, leading to additional energy in the rear part of the calorimeter. However, also the measured $E-p_z$ of true DIS events can vary considerably due to the limited energy resolution of the calorimeter and energy leakage out of the calorimeter. For this analysis only events with $38\,\text{GeV} < E-p_z < 60\,\text{GeV}$ are

[4] For small energies $\lesssim 2\,\text{GeV}$ and hence small boosts the opening angle of the two photons is large enough to separate them in the EMC section of the calorimeter.

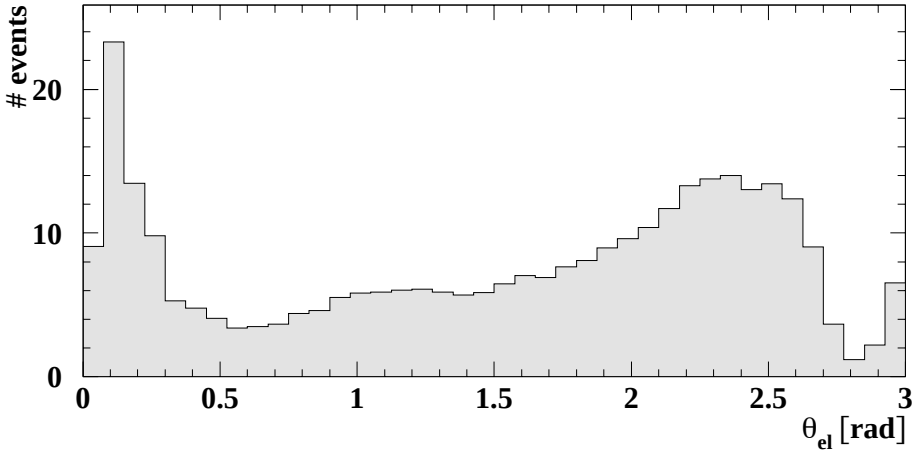


Figure 5.9: *Distribution of the polar angle θ_{el} for fake electrons in PHP MC events. Pions, contained in the proton remnant and decaying into two photons, accumulate at low θ_{el}. All events were selected with DST bit 12.*

accepted. For events with the electron outside the CTD-acceptance in the forward region of the detector, the lower cut is raised to 44 GeV.

- **Cut on $P_T/\sqrt{E_T}$:** If a DIS event is fully reconstructed the total transverse momentum P_T should be 0 GeV. P_T is calculated from the individual energy deposits in the calorimeter using the *CorAndCut* method (Chap. 4.5.3). As the error on the energy measurement in the calorimeter is proportional to $\sqrt{E}$ (Chap. 3.2.1) an upper cut is imposed on the quantity $P_T/\sqrt{E_T}$. The cut value amounts to $4\sqrt{\mathrm{GeV}}$. This reduces for example the background from cosmic radiation, where muons pass through the calorimeters and in most cases produce a net-P_T[5].
- **Cut on y_{el}:** PHP events faking a DIS event accumulate at high y-values. A π^0, often occurring in the proton remnant (Fig. 5.9) and decaying into two photons, can easily be misidentified as an electron as the event comprises no real scattered electron. Such fake electrons have rather low energies and hence a small $(E-p_z)_{\mathrm{el}}$ value leading to high $y_{\mathrm{el}} = 1 - \frac{(E-p_z)_{\mathrm{el}}}{2E_e}$ (see (4.11) with $E'_e(1-\cos\theta_e) = (E-p_z)_{\mathrm{el}}$). Therefore, an upper cut on y_{el} of 0.95 is imposed which suppresses a large fraction of PHP events but preserves most of the signal events.
- **Elastic-QED-Compton rejection:** Elastic QED-Compton events (Chap. 2.5.2) are relatively easy to reject as they contain only two electromagnetic clusters in the calorimeter plus an associated track if the scattered electron lies within the CTD-acceptance.

5.4.3 Geometrical cuts on electron position

Certain regions in the detector complicate the identification of the scattered electron. These are also regions where the detector simulation is found to perform not as well as in other areas. Therefore, events with electrons in these regions are excluded.

[5] Cosmics that traverse the detector in a short distance to the beam line produce only a small net P_T and can therefore not be removed with this cut.

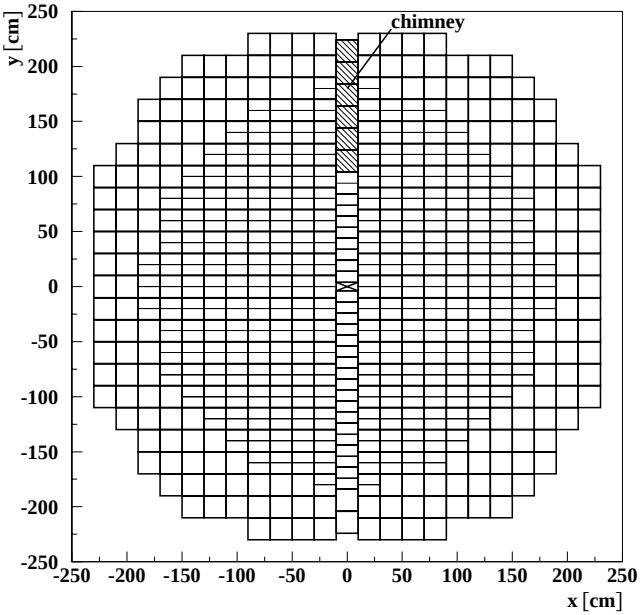


Figure 5.10. *RCAL structure as seen from the IP. The thicker lines mark the towers, whereas the thinner lines display the cell boundaries. The hatched area in the upper RCAL half indicates the chimney region and the cross at (0,0) marks the hole for the beam pipe.*

- **Chimney-cut:** The so-called chimney is located in the middle of the upper RCAL half (Fig. 5.10). In its area the EMC cells are left out to make room for the cryogenic pipes for the super-conducting coil of the ZEUS detector. The EMC-less area has a width of 20 cm, is centered at $x = 0\,\text{cm}$, and starts at $y = 104\,\text{cm}$ for the calorimeter configuration of the 1998–99 data-taking period. As the finer segmentation of the EMC cells in the y direction is crucial for the electron identification, electrons hitting this region are excluded from the analysis. However, the limits set are not equal to the module boundaries as the electron has a shower width (Molière-radius) of $\sim 2\,\text{cm}$ in the EMC section of the ZEUS calorimeter [48] and therefore the x-cut is extended to $|x| > 12\,\text{cm}$. The y-cut is set 14 cm below the edge of the chimney, i.e. $y > 90\,\text{cm}$, as the electrons hit the calorimeter surface in this region under a large angle of $\sim 35°$. Due to the missing EMC cells in the chimney an electron hitting the calorimeter a few centimeters below the chimney deposits only a fraction of its energy in the EMC cells and enters the chimney region afterwards (Fig. 5.11). All cuts are not only applied on the reconstructed calorimeter position of the electron (exactly this position suffers from the missing EMC cells) but also on the "calorimeter position" of the extrapolated electron track.

- **Super-crack cut:** The regions between the main parts of the calorimeter are called super-cracks. A particle hitting the BCAL at its edge facing the R/FCAL, would leak into the R/FCAL after traveling a short distance through the BCAL EMC cells where it is very likely to shower. This scenario is not very well modeled in the detector simulation and therefore electrons in the BCAL within $z < -90\,\text{cm}$ (RCAL) and $z > 162.5\,\text{cm}$ (FCAL) are excluded, where the cut is applied both on the calorimeter and the track position. Investigations also revealed that the MC does not describe the number of electrons in the RCAL with a radius $r > 175\,\text{cm}$ (corresponding to the region where the RCAL is shadowed by the BCAL) very well. Therefore, these electrons are also excluded. Due to statistical limitations these studies are so far not possible in the FCAL.

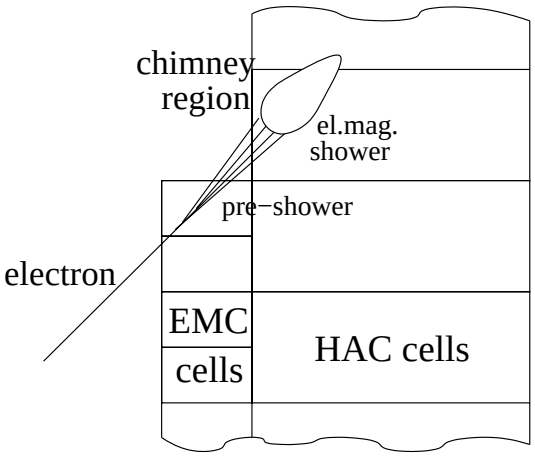


Figure 5.11. *y-z cross section of the RCAL for $x = 0$ cm. The fate of an electron, hitting the EMC section of the RCAL near the chimney region under a large angle, is displayed.*

5.4.4 Other cuts

- **z-vertex cut:** A cut on the reconstructed z-position of the vertex of $|z| < 50\,\text{cm}$ ensures that the outer regions, which are not well modeled by the MC, are discarded. Furthermore, events with vertices near the end-caps of the CTD are rejected. These events are difficult to reconstruct, as the CTD acceptance is highly asymmetric and the particles hit the calorimeter under a large angle, leading to complications in the position reconstruction of the calorimeter clusters.
- **Validity range of MC:** The DJANGOH manual states [87][6] that in the high-x region for low Q^2 values the calculated cross section becomes negative, due to missing higher order corrections and is therefore artifically set to 0. The manual sets a boundary of $y \cdot (1-x)^2 = 0.004$ which is also used as an exclusion limit in this analysis, i.e. events with $y_{\text{JB}} \cdot (1-x_{\text{DA}})^2 < 0.004$ are not considered in this analysis.
- **Bunch-crossing type:** A priori events from all bunch crossings are selected by DST bit 12. This also includes those where for example an electron from a regular bunch collides with a proton out of an "empty" bunch, as even empty bunches contain a small amount of protons or electrons, respectively. However, these collisions are not included in the calculated luminosity and therefore have to be removed from the data sample.
- **Q^2-cut:** As this analysis deals with events of medium and high Q^2 only, a lower Q^2_{DA} cut of $185\,\text{GeV}^2$ is applied.

5.5 Selected events in data

After applying all offline-selection cuts 39 988 out of $\sim$ 210 000 events (after the online selection) remain. Their distributions in the x-y and x-Q^2 planes are displayed in Fig. 5.12 a,b. The super-crack cut, described in Chap. 5.4.3, leads to a suppression of events in the corresponding

[6] Though this is the manual for an older DJANGO version the validity range of the MC is the same for the version employed.

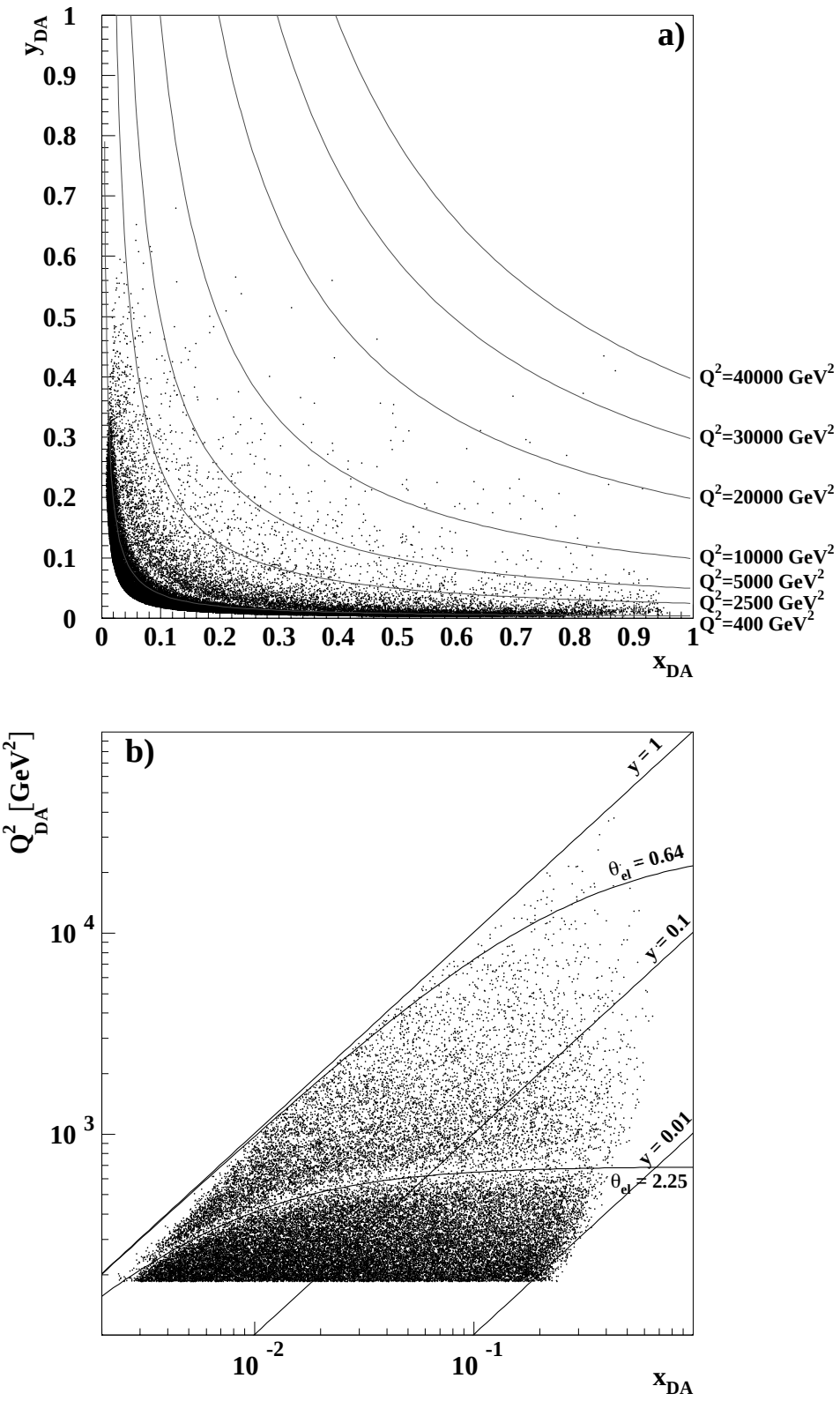


Figure 5.12: *Distribution of reconstructed events after all selection cuts for data in the (a) x-y and (b) x-Q^2 plane. Each dot represents a single event. Also shown are lines of constant Q^2, y and θ_{el}.*

detector region that is clearly visible along the $\theta_{\mathrm{el}} = 2.25\,\mathrm{rad}$ line in Fig. 5.12 b. Examples for NC DIS events can be found in Appendix B.

5.6 Comparison of data and Monte Carlo

To ensure that the data is well described by both the event generator and the detector simulation, data and MC are compared to each other in various distributions of essential variables (Figs. 5.13 – 5.15), where the luminosities are used to normalize the MC to the data. Overall, good agreement is observed in almost all plots. The dip in θ_e (Fig. 5.13 c,d) is an artifact of the super-crack cut. Larger differences between data and MC distributions appear for the track momentum P_{trk} (Fig. 5.13 e and f), the distance of closest approach (DCA) of calorimeter object and matched track (Fig. 5.13 g and h), and for the electron energy (Fig. 5.13 a), where the latter discrepancy is also apparent in the $E - p_z$ distribution (Fig. 5.14 e).

The difference in the track-momentum distributions of data and MC is not yet understood. Though the excess of the data around 25 GeV is more obvious, the more critical discrepancies are located at 8 and 2 GeV as the selection cut of 5 GeV lies between them. To account for this uncertainty, a systematic check is introduced in which the track-momentum cut is varied by $\pm 100\%$ covering the problematic regions.

For the DCA, the distribution of the data is apparently broader than that of the MC. This effect is probably caused by the differences in the electron position reconstruction in the calorimeter between data and MC already discussed in Chap. 4.4.2. A correction routine is not yet available and therefore the resulting uncertainty is estimated with a systematic check, where the cut is tightened from 10 cm to 8 cm.

A similar but somewhat smaller difference between data and MC can be observed in the electron-energy distribution, where again the one for data is noticeably broader than the one for MC. In this case, the energy resolution of the calorimeter is obviously not simulated correctly in the MC. Though the double-angle method, chosen as the reconstruction method for the kinematic variables, does not rely on the electron energy (in contrast to the electron method), the latter enters the analysis like other quantities via cuts and hence should be simulated as well as possible. A correction method for it will be presented in the next section.

Another detector region that proved to be difficult to simulate is located around the forward beam pipe. This refers not so much to the detector simulation but rather to the simulation of the hadronic final state. In this region the hadronic final state is massively influenced by the proton remnant, a complicated object that interacts with the scattered quark via color exchange. Figure 5.15 shows the energy distribution in the FCAL, which is not very well simulated in the MC. Note that this plot uses the cell-based approach to reconstruct the hadronic final state (Chap. 4.5.1). However, other studies using corrected calorimeter information from *CorAndCut* see similar deviations between data and MC in this detector region. Understanding these discrepancies and accounting for them needs careful investigations which are underway at the moment. A preliminary systematic check to cover the difference in the distribution will be presented in Chap. 6.2.

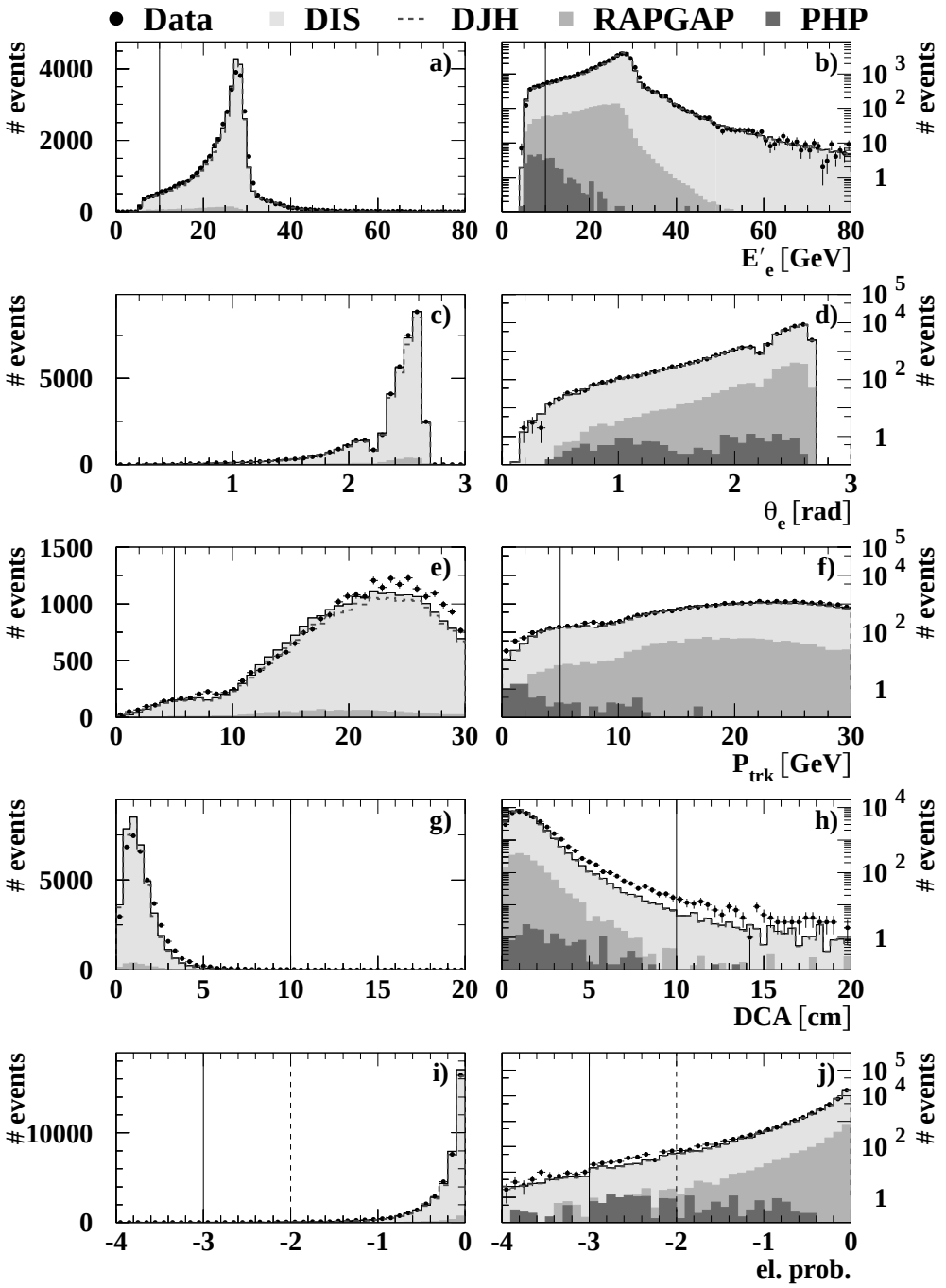


Figure 5.13: *Comparison of distributions of various quantities between data and MC, where the left column shows a linear scale on the y axis, whereas the plots in the right column have logarithmic y axes. (a,b) electron energy; (c,d) polar angle of the electron; (e,f) momentum of the electron track; (g,h) distance of closest approach between track and calorimeter cluster; (i,j) probability of the electron. The MC distributions are normalized to those of the data according to the luminosities. DIS is the sum of DJH- and* RAPGAP*-MC and PHP represents the PHP MC. The solid vertical lines indicate the standard cut boundaries, whereas the dashed lines represent the tightened cuts for electrons outside the CTD acceptance in the forward region.*

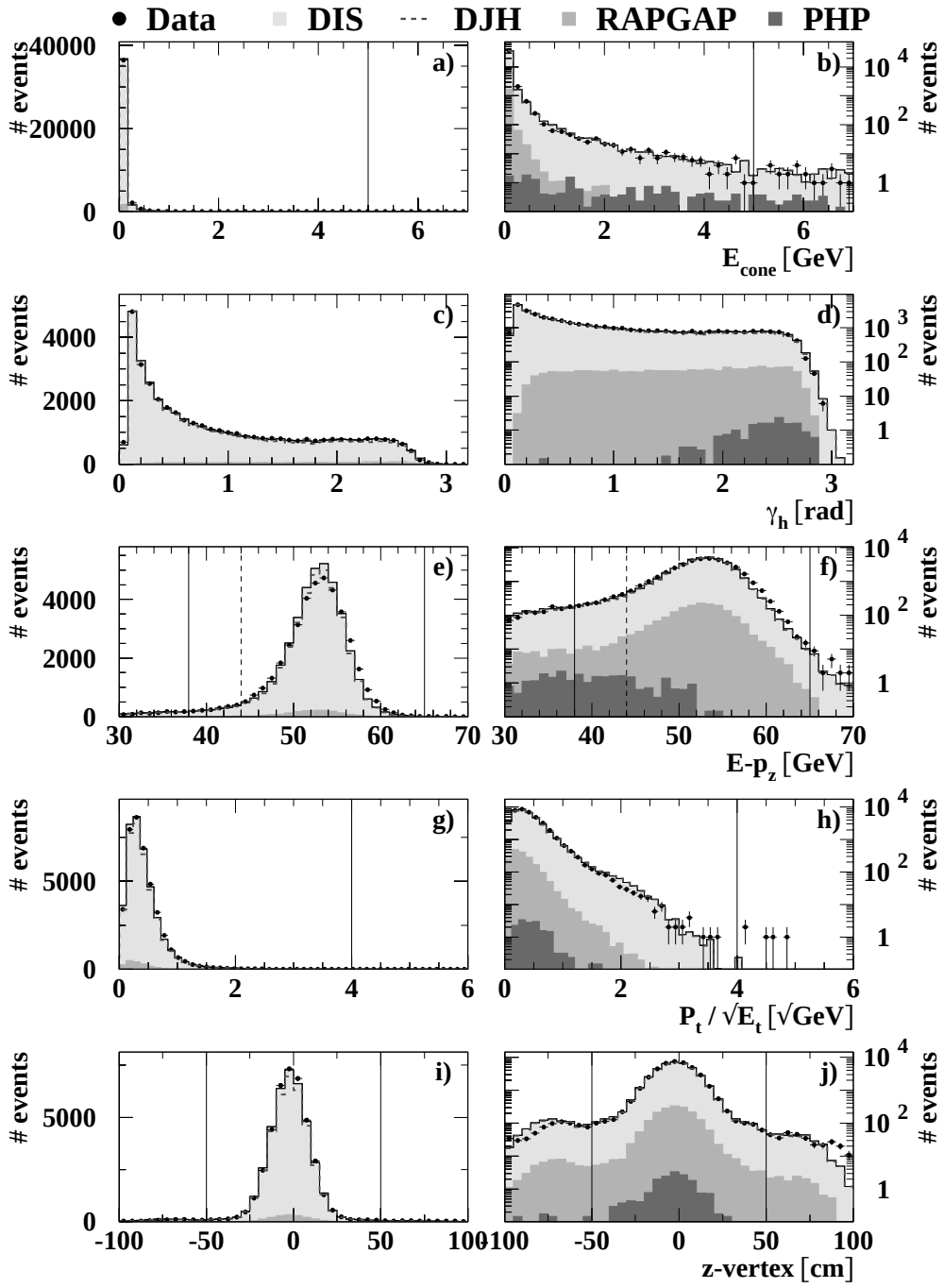


Figure 5.14: *Comparison of distributions of various quantities between data and MC, where the left column shows a linear scale on the y axis, whereas the plots in the right column have logarithmic y axes. (a,b) energy in η-ϕ-cone with radius 0.8 not associated with the electron cluster; (c,d) hadronic angle γ_{had}; (e,f) $E - p_z$; (g,h) $P_T/\sqrt{E_T}$; (i,j) z position of the vertex. The MC distributions are normalized to those of the data according to the luminosities. DIS is the sum of DJH- and* RAPGAP*-MC and PHP represents the PHP MC. The solid vertical lines indicate the standard cut boundaries, whereas the dashed lines represent the tightened cuts for electrons outside the CTD acceptance in the forward region.*

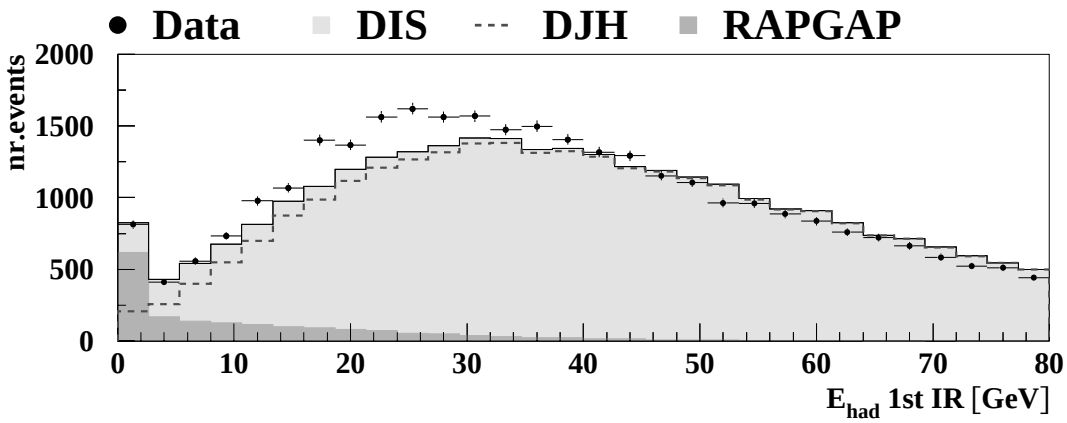


Figure 5.15: *Comparison of the hadronic energy distribution in the first inner ring of cells in the FCAL for data and MC. The energy is calculated with the cell-based approach. The MC distribution is normalized to that of the data according to the luminosities. DIS is the sum of DJH- and* RAPGAP*-MC. Note that the diffractive MC is essential to make data and MC agree in the lowest energy bin which is dominated by the* RAPGAP *contribution.*

Figure 5.16 finally displays the comparison between data and MC in the kinematic variables obtained from the different reconstruction methods used in this analysis, where the correction to the electron, discussed in the next section, has already been applied. All plots show very good agreement between data and MC. As expected, the PHP background is located mainly in the high-y region, whereas the distributions of the diffractive sample are similar to those of the "standard" DIS sample (DJH).

5.7 Electron-energy smearing in Monte Carlo

In order to correct for the differences in the energy resolution between data and MC, observed in Fig. 5.13 a and Fig. 5.14 b, the double-angle energy E_{DA} from (4.2) is used as an estimate of the true energy. Then, the relative difference between E_{DA} and the calorimeter energy E_{cal} is plotted and a Gaussian is fitted to the distribution yielding the energy resolution. The correction "factor" for MC is then obtained by comparing the resolutions for data and MC.

In order not to depend strongly on the accurate simulation of the resolution of E_{DA}, it should be significantly better than that of E_{cal}. Figures 5.17 a,b show for MC the resolutions of E_{DA} with respect to the true energy E_{true} as a function E_{true}, whereas in Figs. 5.17 c,d the resolutions of E_{cal} with respect to E_{DA} are displayed. Obviously, for small E_{cal}, E_{DA} is not a good reference energy as its resolution (with respect to E_{true}) is equal to or greater than the resolution of E_{cal}. However for $E_{\mathrm{cal}} \gtrsim 25\,\mathrm{GeV}$, $\sigma(E_{\mathrm{DA}}/E_{\mathrm{true}})$ becomes very small compared to $\sigma(E_{\mathrm{cal}}/E_{\mathrm{DA}})$ and can therefore be neglected. In the following, only electrons with a double-angle energy $E_{\mathrm{DA}} > 25\,\mathrm{GeV}$ are used for the study.

First investigations showed that the difference in the energy resolution between data and MC is present both in BCAL and RCAL. However, they are not restricted to the module cracks as one might suspect due to the complicated structure of these regions but are rather distributed

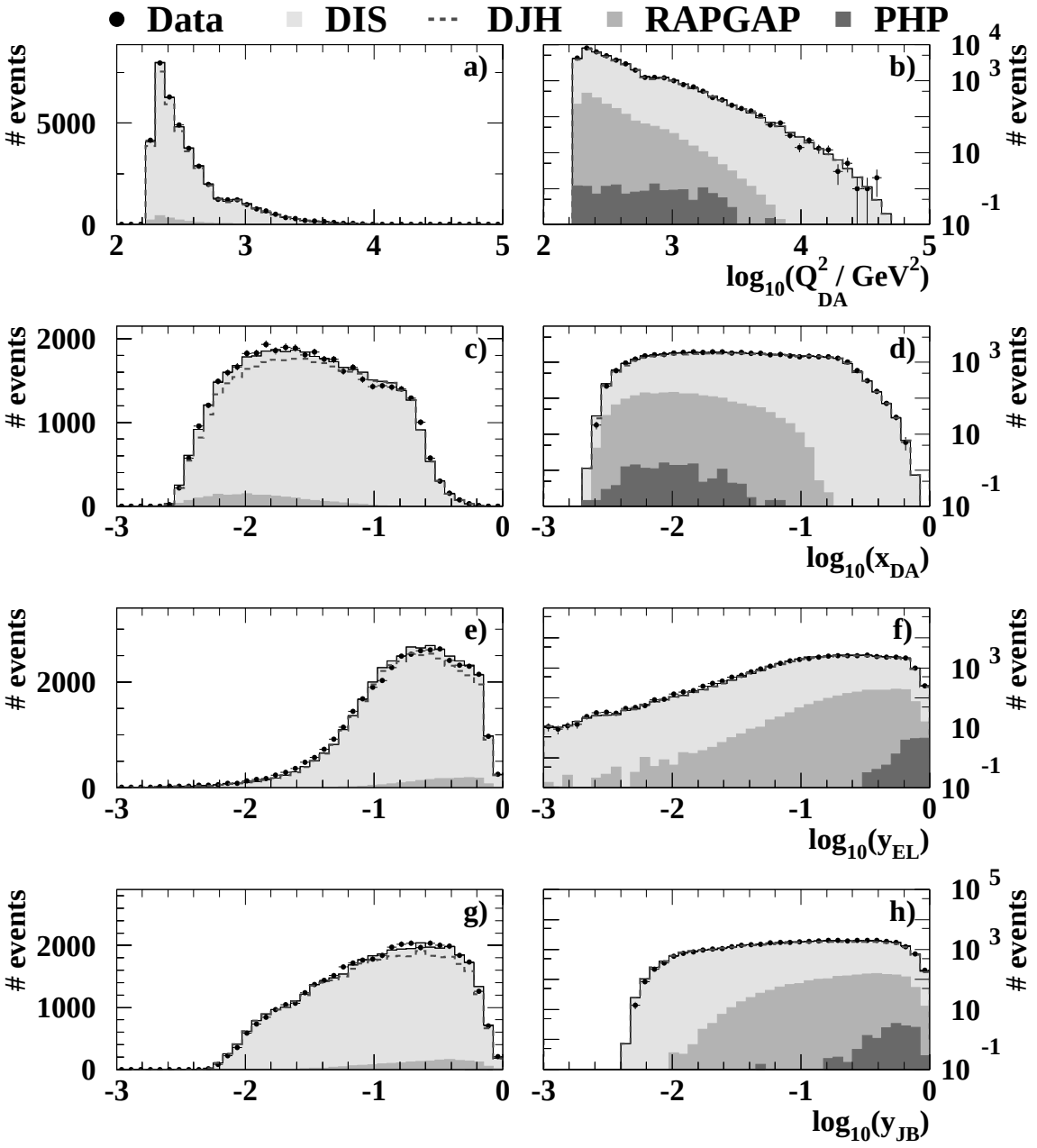


Figure 5.16: *Comparison of distributions of kinematic variables between data and MC, where the left column shows a linear scale on the y axis, whereas the plots in the right column have logarithmic y axes. (a,b) Q^2 reconstructed with double-angle method; (c,d) x reconstructed with double-angle method; (e,f) y reconstructed with electron method; (g,h) y reconstructed with Jacquet-Blondel method. The MC distributions are normalized to those of the data according to the luminosities. DIS is the sum of DJH- and* RAPGAP*-MC and PHP represents the PHP MC.*

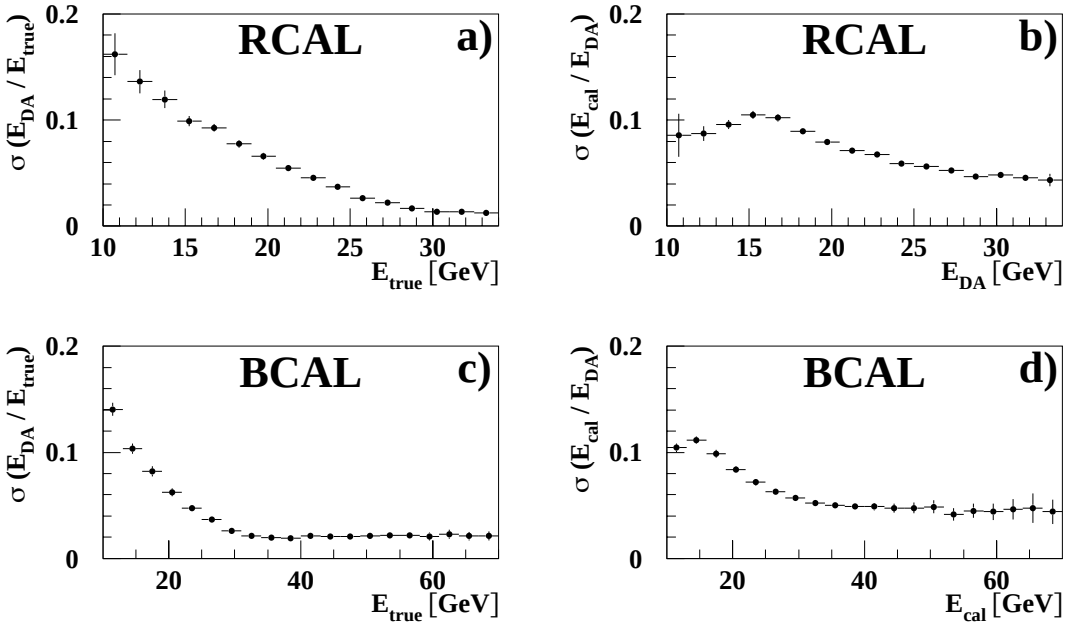


Figure 5.17: *MC energy resolution in R/BCAL for: (a,c) double-angle energy E_{DA} with respect to the true energy E_{true} as a function of E_{true}; (b,d) calorimeter energy E_{cal} with respect to E_{DA} as a function of E_{DA}.*

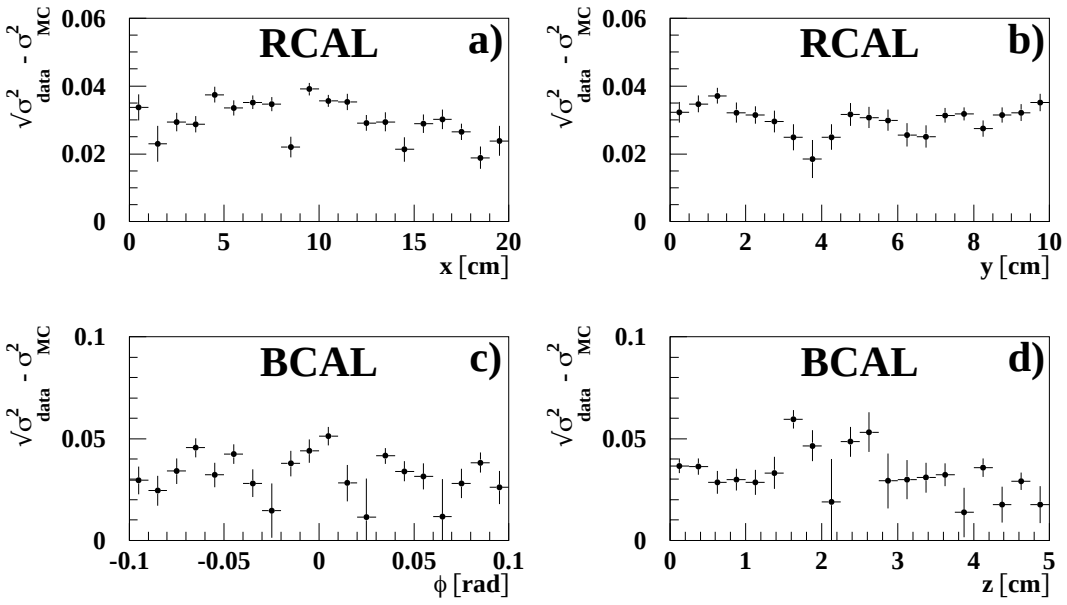


Figure 5.18: *Quadratic difference between the energy resolutions in data and MC within a cell for (a,b) RCAL and (c,d) BCAL. For these plots all cells are projected on top of each other.*

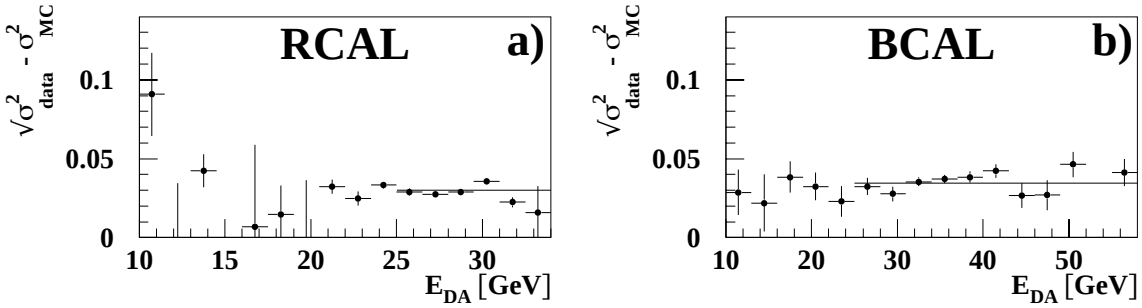


Figure 5.19: *Quadratic difference between the energy resolutions in data and MC as a function of the double-angle energy E_{DA} for (a) RCAL and (b) BCAL. The fits yield offsets of (a) 0.030 and (b) 0.035.*

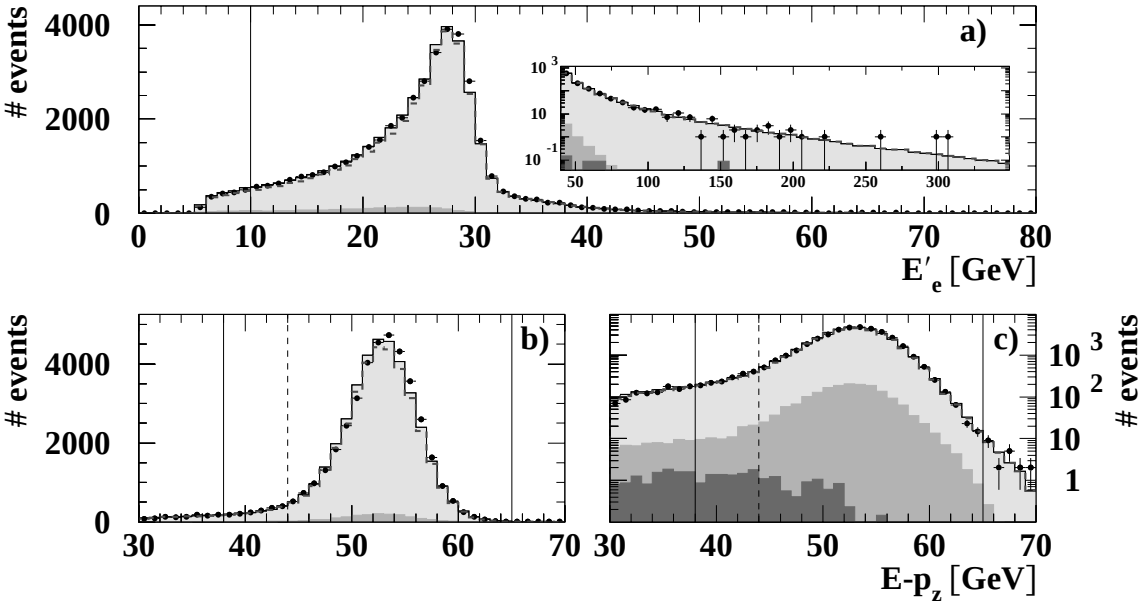


Figure 5.20: *(a) Electron energy and (b,c) $E - p_z$ distribution after smearing the electron energy in MC for electrons in the RCAL and BCAL. The inset in (a) displays the high-energy tail of the energy distribution. DIS is the sum of DJH and* RAPGAP *MC and PHP represents the PHP MC. The vertical lines indicate the cut boundaries where the dashed lines represent the tightened cuts for electrons outside the CTD acceptance in the forward region.*

uniformly over the whole cell (Fig. 5.18). Another reason for the difference could be a false intrinsic energy resolution in the MC. To verify this Fig. 5.19 displays the quadratic difference between the energy resolution of data, σ_{data}, and MC, σ_{MC}, as a function of the double-angle energy E_{DA}. In the relevant region above 25 GeV the difference is basically flat which translates into an additional constant term for the energy resolution $\Delta E/E$, where the default calorimeter resolution in the MC is

$$\frac{\Delta E}{E} = \frac{18\%}{\sqrt{E[\,\text{GeV}]}} \oplus 1\% \ . \tag{5.6}$$

Straight-line fits to Figs. 5.19 a,b yield offsets of 0.03 and 0.035 which are then used to "smear" the electron energy in MC with a Gaussian distribution of the respective width. The following formula shows the final energy resolution in MC for RCAL (BCAL):

$$\frac{\Delta E}{E} = \frac{18\%}{\sqrt{E[\,\text{GeV}]}} \oplus 1\% \oplus 3\% \ (\oplus 3.5\%) \ . \tag{5.7}$$

Figure 5.20 finally displays the resulting total electron energy (a) and $E - p_z$ (b) distributions which show now good agreement between data and MC, with only a small remaining shift to higher energies for the distribution of the data events.

5.8 Summary of Chapter 5

In this chapter the data and MC samples needed for the analysis were introduced and the selection criteria for NC DIS events were presented and discussed. The following comparison between data and MC revealed differences in some distributions. One of the larger differences were removed by modifying the MC, whereas for the other cases systematic checks have been introduced that cover the differences. These checks together with other systematic studies will be discussed in the next chapter, the introduction of which forms a general discussion about different approaches and general problems with the implementation of systematic uncertainties.

Chapter 6

Studies of systematic effects

A reasonable estimate of systematic measurement-uncertainties is a highly non-trivial task. Todays high energy experiments have reached a level of complexity that gives origin to numerous sources of systematic uncertainties. Unlike statistical errors that in most cases follow a Gaussian or at least a Poisson distribution, systematic uncertainties show a wide variety of distributions that are only in rare cases Gaussian-like or even known.

The first step in the determination of the total systematic uncertainty is the identification of the most important sources. Afterwards, methods have to be developed that yield reasonable estimates of the individual contributions and a way of combining them has to be found. In general, the sources of systematic uncertainties can be organized into four classes: uncertainties on input parameters for the MC (e.g. PDFs), uncertainties on the parton level (e.g. missing higher orders), uncertainties on the hadron level (e.g. fragmentation model) and deficiencies in the detector simulation.

A common source for systematic uncertainties are imperfections in the simulation of the detector, e.g. simplified description of the calorimeter structure or ignorance about the true energy scale of the calorimeter. The latter for example would lead to differences in the measured electron energy in data and MC and therefore to different acceptances when applying the electron-energy cut. A common method to cope with this kind of uncertainties is the variation of the cut value followed by the recalculation of the measured cross section. The difference between the nominal and the modified-cut cross section is then taken as the systematic uncertainty.

Though very simple and easy to implement this method has conceptional weaknesses. At first, the amount by which the cut is varied is often rather subjective. Furthermore, in the case of the cut on the electron energy for example, the underlying uncertainty on the energy scale of the calorimeter also effects the hadronic final state. For the latter, however, the situation is more complicated as no direct cut is applied on the hadronic energy but on $E - p_z$ which itself comprises both the hadronic and the electron energy. Hence, the variation of the $E - p_z$ cut is correlated (in an unknown way) with that of the electron energy making the combination of the systematic uncertainties problematic. In addition, the cut-variation method completely neglects effects of the energy scale uncertainty on the hadronic angle γ as, a priori, the energy scale is different for the various calorimeters.

An obvious conclusion from these considerations is that cut variations should be avoided whenever possible. Instead, the energy scale of the calorimeter can be varied directly in the example above, where this is done separately for the three different sub-calorimeters.

However, the actual source of a systematic uncertainty is not always that easy to identify. Often one is left with an apparent discrepancy between data and MC that could originate from either of the four classes of systematic uncertainties. The difference in the energy distributions of data and MC in Fig. 5.15 for example could either be caused by a wrong simulation of the energy leakage from the FPC into the FCAL (detector simulation), by inaccuracies in the fragmentation model (hadron level), or by deficiencies in the simulation of the interaction between struck quark and proton remnant (parton level). An exact localization of the problem in this case is at least difficult if not impossible.

If the true source for the discrepancy cannot be located the only possibility left is to reweight the distribution in the MC to that of the data, with the hope that it has the same effect on the kinematics and cut variables as a correction of the true reason. This "effective correction" can then be used either for the nominal analysis or as a systematic check.

6.1 First-level-trigger studies

As described in Chap. 5.3, the FLT bits form the basics for the event selection and all other bits from higher trigger-levels rely heavily on them. If an event fails to fire one of the FLT bits it is irrevocably lost for any analysis, and as only rough detector information is available at this stage of the reconstruction, the reduction of the event rate by the FLT without discarding interesting physics events is a demanding task. But even more important is the need for an accurate simulation of this first stage in the event-selection chain in the MC. Data and MC events are passed through the same selection chain, and hence, the simulated signals from the detector components must match their real-world counterparts very accurately. In order to verify this, the performances of the FLT bits most relevant for this analysis are investigated and the results from MC and data are finally compared to each other.

To find out what the relevant FLT bits are, Figure 6.1 shows for each FLT bit used in this analysis (for a description of the different bits see Table 5.4) the relative number of events (after DST bit 12 selection) for which it was fired. Obviously, the TLT bits 40, 44 and 46 (TLT bit 30 is similar to bit 46) are the most important bits for DST bit 12 and hence, these bits are subject to further investigation.

6.1.1 Trigger efficiencies

The basic idea for the study of trigger efficiencies is the selection of events via a trigger bit which is orthogonal to the one under study. The resulting event sample is the basic sample. Afterwards, the trigger bit under study is applied and the efficiency with respect to the basic sample is computed and plotted as a function of relevant quantities. Finally, the results from data and MC are checked for the degree of agreement. All events are preselected with the

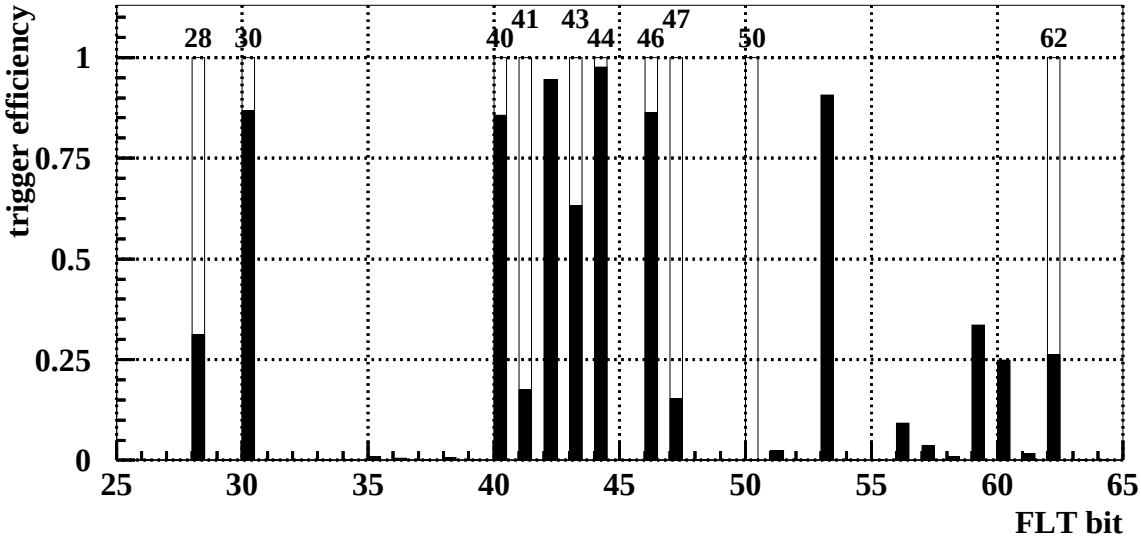


Figure 6.1: *Trigger efficiencies in data for DST bit 12 selection as a function of the FLT bit number. The hollow bars together with the number on top mark the trigger bits contained in DST bit 12. The solid bars indicate the efficiency of a bit, i.e. how often it has fired with respect to the number of events that have passed DST bit 12.*

standard NC DIS cuts, excluding those having a direct impact on the trigger investigated (e.g. the electron-energy cut is ignored for the triggers that use the EMC energy). In the following investigations the chimney region (Chap. 5.4.3) is always excluded.

The investigations start off with bit 44 ($E_{\mathrm{REMC}} > 4.8\,\mathrm{GeV}$ + good track or $E_{\mathrm{REMC}} > 3.4\,\mathrm{GeV}$) which fires most frequently. As the bit consists of two parts, one for RCAL and one for BCAL, it is split up and the individual trigger requirements are simulated offline[1]. The BEMC part is chosen as the independent trigger for the REMC part with the additional requirement that EM must have found an electron in the RCAL. Figure 6.2 shows the REMC-trigger efficiency as a function of the energy and the radius of the electron in the RCAL. The overall efficiency is almost always 1 and data and MC agree to better than 0.5%. Similarly good agreement is observed for the BEMC part (Fig. 6.3), although with a somewhat lower efficiency. This time the basic sample has been selected with FLT bit 41 (high E_T), where the additional requirement has been imposed that an electron in the BCAL has been found by EM.

The next bit investigated is bit 46 (RCAL ISOe). The selection of the basic sample is again done with FLT bit 41, where this time additionally an electron in the RCAL is required. Figure 6.4 shows the trigger efficiency as a function of energy and radius of the electron in the RCAL. The energy plot shows an overall high efficiency and deviations between data and MC below 0.5% though for $E > 20\,\mathrm{GeV}$ the data tends to lie below the MC. On the other hand, the radius plot displays a significant dip of $\sim 1\%$ in the data efficiency around $r \approx 110\,\mathrm{cm}$. The origin of this effect is revealed when looking at the efficiency distribution in the x-y plane of the

[1] The original values of the various quantities that enter the trigger bit are also available offline and are contained in the ntuple. However, vetoes are not simulated but the effects from this inaccuracy are expected to be small as all events are already preselected via DST bit 12 containing these vetoes.

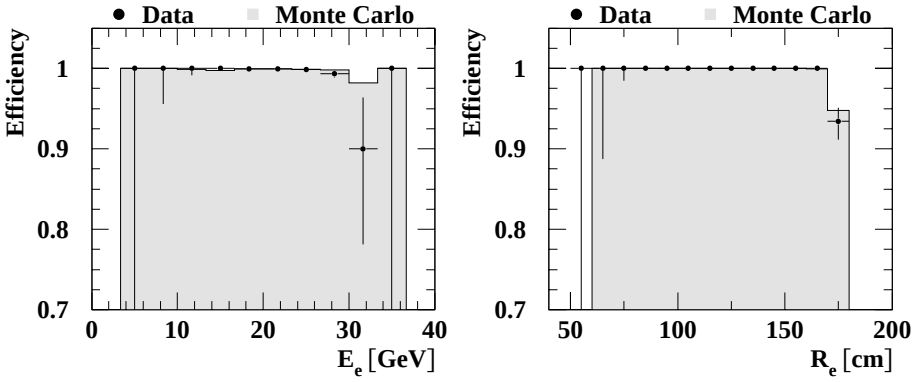


Figure 6.2: *Efficiency of the REMC part of FLT bit 44 as a function of (a) energy and (b) radius of the electron in the RCAL. The basic sample is selected with the BEMC part of FLT bit 44.*

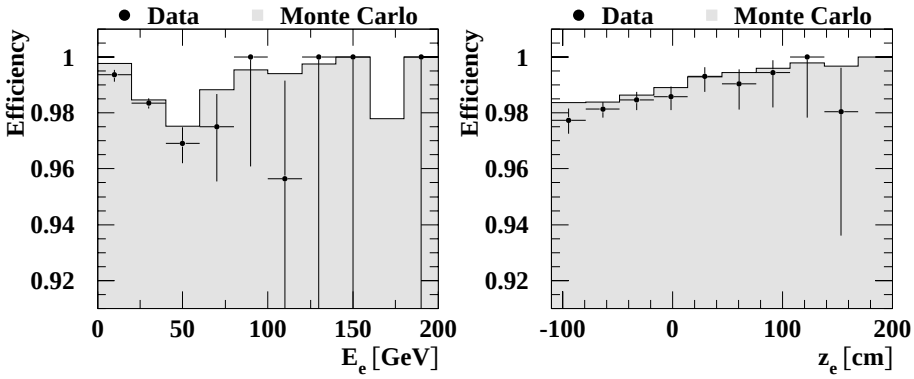


Figure 6.3: *Efficiency of the BEMC part of FLT bit 44 as a function of (a) energy and (b) z-position of the electron in the BCAL. The basic sample is selected with FLT bit 41 (high E_T).*

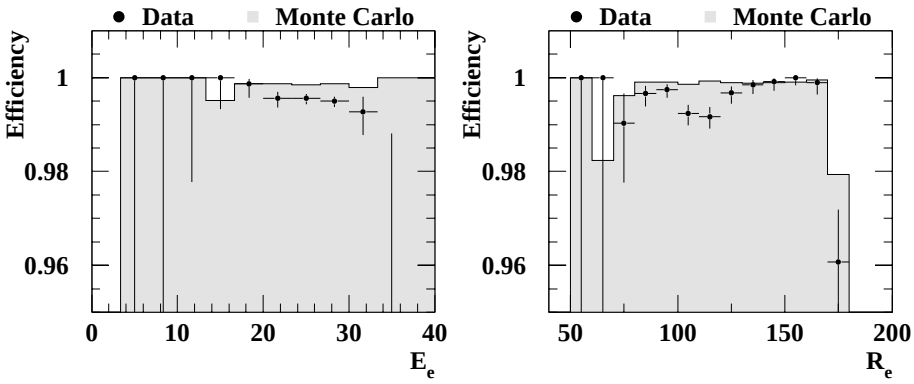


Figure 6.4: *Efficiency of FLT bit 46 (RCAL ISOe) as a function of (a) energy and (b) radius of the electron in the RCAL. The basic sample is selected with FLT bit 41 (high E_T).*

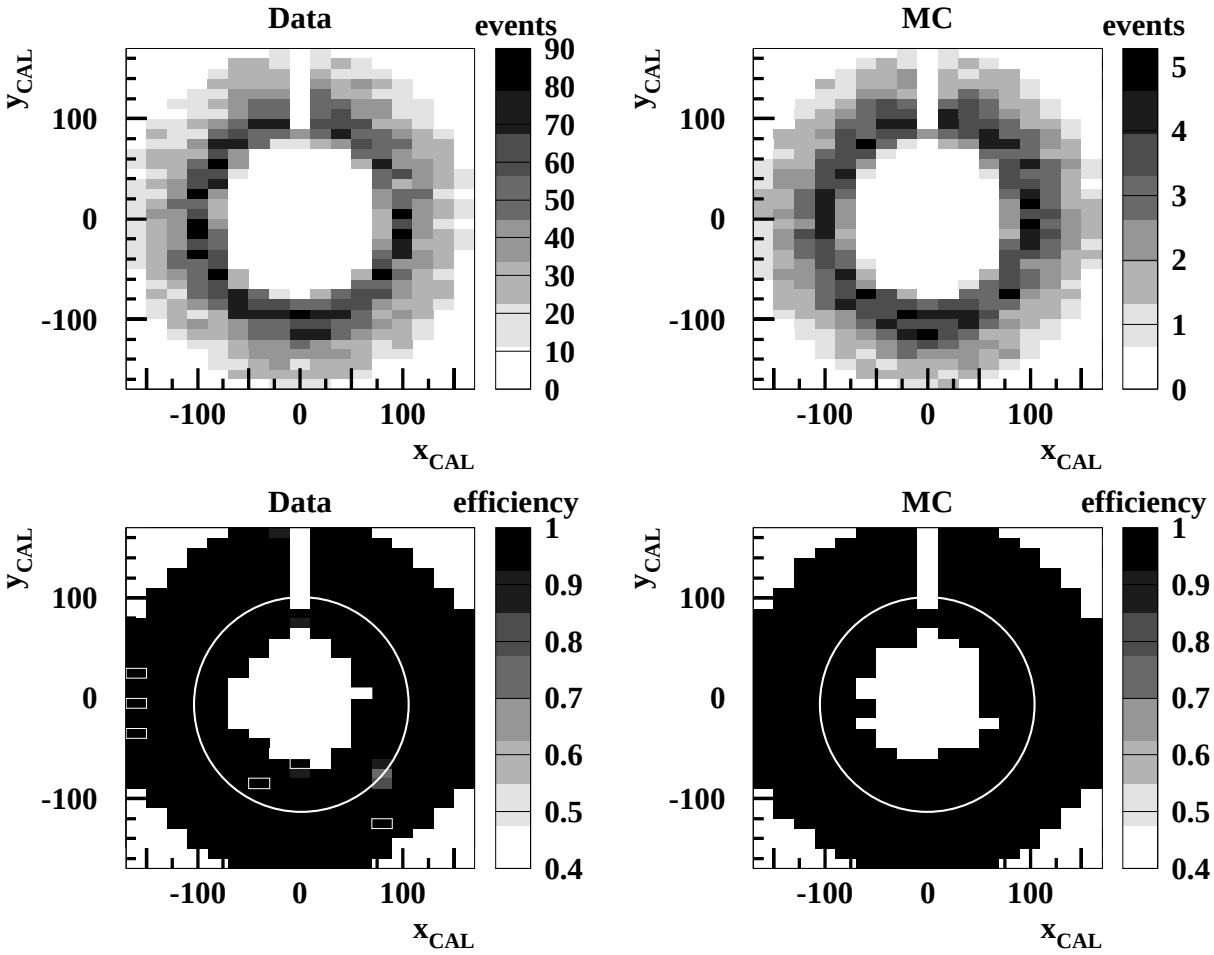


Figure 6.5: *Number of events (upper row) and FLT bit 46 efficiency (lower row) for data (left column) and MC (right column) in x-y bins for electrons in the RCAL. The basic sample was selected with FLT bit 41 (high E_T). Each bin corresponds to an EMC cell and the white circles in the lower plots mark a radius of* 110 *cm. The rectangular gap in the upper center of all plots is caused by the chimney cut and the white-framed cells had a damaged PM for a certain period of the data taking.*

RCAL, displayed in the lower plots of Fig. 6.5. Obviously, there is a region around $x = 80\,\text{cm}$, $y = -80\,\text{cm}$ (the white circles indicate a radius of $\sim 110\,\text{cm}$) in the real RCAL where the efficiency is significantly lower than in the rest of the calorimeter. Further inspection revealed that this is not caused by damaged PMs (Chap. 4.4.4.4) as the PM-affected cells, marked with a white frame, are far away from the inefficient region. A logical explanation for the observed effect is provided by a water leak that occurred during the shutdown phase in March 1995 in the HES cooling circuit in RCAL module 18, i.e. in the column of cells right of the cells with the low efficiency. The influence of this efficiency dip on the results is expected to be negligible as FLT bit 44 is almost 100% efficient and all FLT bits are always combined by a logical OR.

FLT bit 40 (total EMC) is not investigated separately as the efficiency of the EMC triggers which are highly correlated with FLT bit 40 were already discussed extensively for REMC and BEMC.

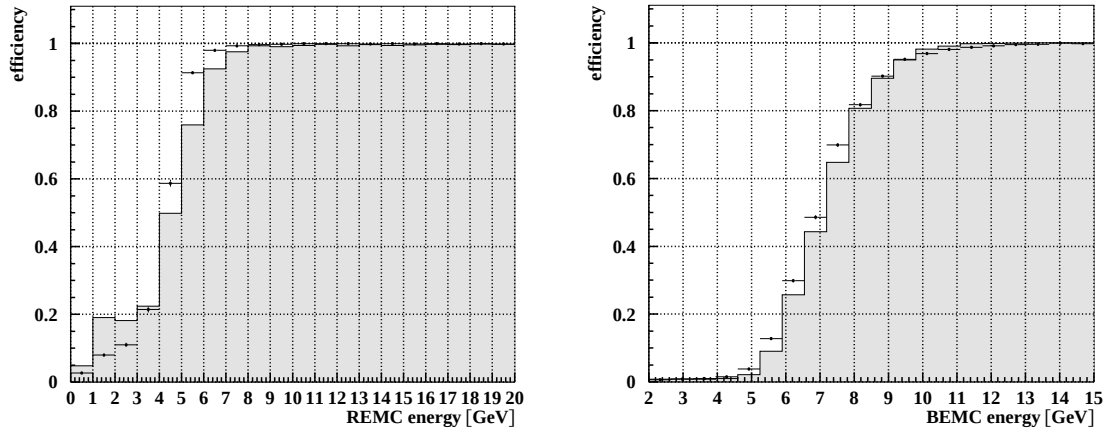


Figure 6.6: *Turn-on curves for (a) REMC trigger ($E_{\mathrm{REMC}} > 3.4\,GeV$) and (b) BEMC trigger ($E_{\mathrm{BEMC}} > 4.8\,GeV$ + good track). The histograms are MC and the dots are data. The basic samples are selected with DST bit 12.*

6.1.2 Turn-on curves for FLT

The idea of turn-on curves is to measure the trigger efficiencies, based on FLT quantities, as a function of the corresponding offline variables. On the one hand, this shows the level of agreement between the information from different reconstruction levels together with the point beyond which the trigger is 100% efficient (it would be dangerous to cut right in the rising flank of the curve). On the other hand, it again allows a comparison of data and MC to ensure that the simulation of this important stage in the event selection is simulated correctly.

The basic samples for the following studies are solely selected with DST bit 12. Figure 6.6 a,b show the turn-on curves for the REMC and BEMC triggers[2]. Though differences in the rise of the curve are present, good agreement between data and MC is observed for $E > 8\,\mathrm{GeV}$. The offline selection-cut on the energy of the identified electron at 10 GeV is sufficiently far away from the region of discrepancy so that the effect on the event selection is negligible.

Overall, the investigations of the last two sections showed that the MC yields a reasonable description of the data on the FLT level.

6.2 Variation of parameters

This and the following section contain descriptions of all systematic checks used to estimate the systematic uncertainty on the measured cross sections. The term "variation of parameters" refers here to those systematic checks that vary specific parameters of the simulation within their errors. In contrast to this, the section "variation of cuts" contains all systematic checks

[2] A similar plot for the RCAL ISOe trigger is not possible as the isolation information on the FLT level is reduced to 4 bins, which cannot be compared to the isolation criterion discussed in Chap. 5.4.1.

calorimeter	CAL section	
	EMC	HAC
FCAL	$(0.0 \pm 1.5)\%$	$(0.0 \pm 1.0)\%$
BCAL	$(2.0 \pm 1.5)\%$	$(5.0 \pm 1.0)\%$
RCAL	$(0.0 \pm 2.0)\%$	$(0.0 \pm 2.0)\%$

Table 6.1: *Hadronic energy-scale differences between data and MC together with their uncertainties for all three main calorimeters. A positive number indicates that the energy has to be increased in data.*

that perform the disfavored variation on one of the offline-selection cuts. Unfortunately, this kind of systematic checks cannot be completely omitted yet.

Calorimeter energy-scale: The EMC and HAC energy scales of the three calorimeter parts together with their uncertainties were subject to intense investigations [88]. The basic idea of these studies is to use the measured energy of the electron (after all corrections) together with its polar angle to predict the hadronic transverse momentum. The difficulty lies in the fact that the hadronic final state is in most cases spread out over a large solid angle which makes it very difficult to determine the energy scales of all three calorimeters independently. These problems can be overcome by using diffractive events as here no energy flow between the hadronic jet and the outgoing proton disturbs the measurement. The separation of the EMC and HAC energy scales is achieved by plotting the ratio between predicted and measured hadronic transverse momentum as a function of the energy fraction in the EMC cells. The extrapolated values at 0 and 1 then give the corresponding EMC and HAC energy scales.

Table 6.1 lists the differences in the energy scales between data and MC together with their uncertainties [88]. The MC has already been corrected for these differences. The systematic uncertainties are obtained by varying the energy scale of all six calorimeter sections in the MC separately up and down by the listed errors yielding 12 individual checks.

First inner ring: The discrepancy in the energy distribution between data and MC in the innermost ring of FCAL cells (Fig. 5.15) is accounted for by multiplying all cell energies in the innermost FCAL ring in MC by a factor 0.9. The resulting energy distribution, plotted in Fig. 6.7, shows now reasonable agreement between data and MC. Note that this is only an "effective" correction as discussed in the last paragraph of the introduction to this chapter. Hopefully, investigations currently under way will yield a better handle on this uncertainty in the future.

No electron smearing: To estimate the influence of the electron-energy smearing in MC (Chap. 5.7) on the measured cross sections the smearing is turned off.

Vertex distribution: The vertex distribution in the MC has to be adjusted to that of the data for each data taking period [89]. The error on the fit is used to determine the corresponding systematic uncertainty.

Variation of $\gamma_{\max}$ in CorAndCut: To delete energy deposits in the main calorimeter originating from back-splash or noisy cells, *CorAndCut* removes all energy deposits with less than

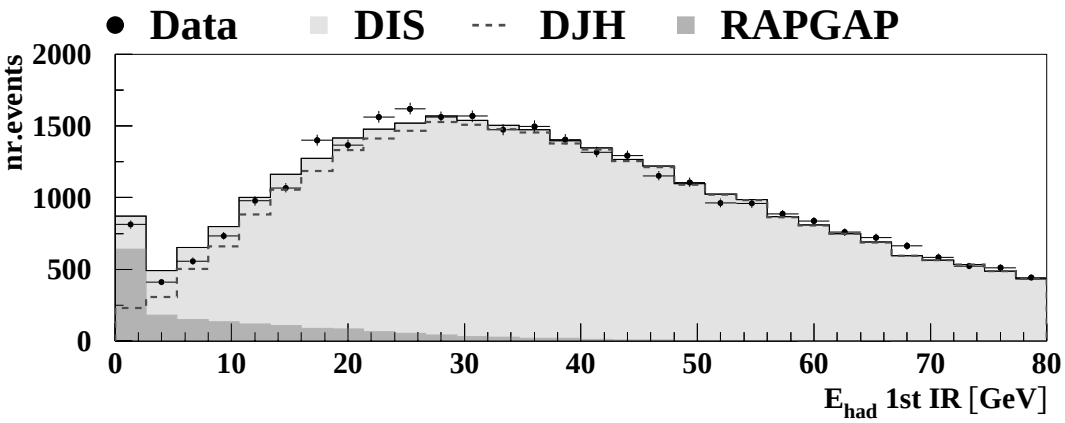


Figure 6.7: *Comparison of the hadronic energy distribution in the innermost ring of cells in the FCAL for data and MC after multiplying the energies in MC with a factor 0.9. The energy is calculated with the cell-based approach. The MC distribution is normalized to that of the data according to the luminosities. DIS is the sum of DJH and* RAPGAP *MC.*

3 GeV if they have a polar angle larger than $\gamma_{\max} = \gamma_{\text{had}} + \gamma_\delta$ (Chap. 4.5.2). The sensitivity of the measured cross sections on the exact choice of γ_δ is estimated by using the built-in systematic check of the back-splash-correction routine.

RAPGAP fraction: The fraction of diffractive events was determined in Chap. 5.2.2 by fitting a mixture of standard DIS and diffractive MC to the $\eta_{\max}$ distribution in data. The fractions f_i obtained were supposed to be constant in Q^2 and a straight line was fitted to the x dependence yielding (5.5):

$$\bar{f}(x) = -(0.064 \pm 0.036) - (0.071 \pm 0.017) \cdot \log x \ . \tag{6.1}$$

The worst-case scenarios are therefore

$$\bar{f} = -0.028 - 0.090 \cdot \log x \qquad \text{and} \tag{6.2}$$

$$\bar{f} = -0.100 - 0.053 \cdot \log x \ , \tag{6.3}$$

which are used to determine the systematic uncertainty.

PHP fraction: To account for the uncertainty in the amount of PHP events used, their fraction is varied down by 50% as discussed at the very end of Chap. 5.2.3.2.

6.3 Variation of cuts

$\boldsymbol{E - p_z}$**:** The $E - p_z$ distribution is sensitive to the amount of PHP events at its lower end and to overlay events at its upper end. However, as the uncertainty on the PHP fraction has already been covered, only the upper $E - p_z$ cut is varied by ± 2 GeV yielding an estimate of the uncertainty arising from the missing simulation of overlay events.

$\boldsymbol{y_{\text{el}}}$**:** The y_{el} cut is supposed to suppress fake electrons originating from pions often contained in the remnant (Chap. 5.4.2). Lowering the cut from 0.95 to 0.9 yields an estimate of the uncertainty of the simulation of these kind of events.

Isolation of the DIS electron in the calorimeter: Isolation in the calorimeter is a characteristic feature of most of the scattered electrons. However, due to noisy cells, which are not simulated at all, or problems in the simulation of "indirect" energy deposits like back-splash (Chap. 4.5), the energy distribution in the calorimeter in the near vicinity of the electron could be different between data and MC. Therefore, the cut on the isolation of 5 GeV (Chap. 5.4.1) is varied by ±2 GeV.

$p_T/\sqrt{E_T}$: There exists no reliable simulation of cosmics, hence the cut value $p_T/\sqrt{E_T} < 4\sqrt{\text{GeV}}$ is varied by $\pm 1\sqrt{\text{GeV}}$ to obtain an estimate of the uncertainty of the measured cross sections due to the missing simulation.

Lower track momentum cut: To account for the differences in the P_{trk} distribution between data and MC at low P_{trk} (Fig. 5.13 f) the corresponding cut is varied by ±100%, i.e. $P_{\text{trk}} > (5 \pm 5)\,\text{GeV}$.

DCA between calorimeter object and track: Due to differences in the electron-position reconstruction the distance of closest approach between electron-cluster center in the calorimeter and associated track is not very well described by the MC (Fig. 5.13 h). The influence on the measured cross section is taken into account by lowering the cut from 10 cm to 8 cm.

6.4 Summary of Chapter 6

After the discussion of the individual systematic checks that are used to estimate the systematic uncertainties of the measured cross sections, it is clear that these checks are not completely uncorrelated as they should be. For example, the variation of the $E - p_z$ cut is still correlated with the energy-scale uncertainty. Unfortunately such overlaps cannot always be disentangled completely. In the absence of a better alternative to cope with the unknown correlations all systematic uncertainties are added in quadrature to obtain the total error.

At the end of this chapter finally everything is set up for the extraction of the cross sections. MC samples for different reactions were compiled to match the processes present in the data. NC events with $Q^2 > 185\,\text{GeV}^2$ have been selected by the offline- and online selection and were reconstructed, where large efforts have been made to correct for the various imperfections of the detector. Then, MC and data have been compared in essential distributions with mostly satisfying results, where for the less satisfying cases and other sources of systematic uncertainties checks have been developed. The next chapter deals with the extraction of the differential $e^- p$ cross sections and then compares them to those from $e^+ p$. Finally, both cross sections are used to extract the mass of the Z boson as well as the parity violating structure function xF_3.

Chapter 7

Results

This chapter presents the results of the DIS NC analysis of the e^-p data recorded by ZEUS in 1998/99. The chapter begins with a presentation of the unfolding method employed and the definition and discussion of efficiency and purity. The subsequent sections deal with the extraction of the three single- and the double-differential cross sections. Each section contains detailed information on binning, efficiency, purity and systematic checks. Finally, the e^-p cross sections obtained in this analysis are used in combination with e^+p cross sections [9, 10] to extract the Z-boson mass and, in particular, the parity violating structure function xF_3. The essential results of this analysis are highlighted with a frame.

7.1 Unfolding method

The aim of this analysis is the extraction of the differential cross sections at the Born level, i.e. in lowest order of the electroweak interaction. Therefore, the radiative corrections δ_r, introduced in (2.32) and included both in MC (Chap. 5.2) and (naturally) data, have to be removed by an unfolding procedure. The method used here is the so-called one-step bin-by-bin unfolding: First, the number of events in the data, N_{obs}, and in MC, N_{MC}, are counted in each bin. Then, the expected number of background events in this bin, N_{bg}, is subtracted from N_{obs}. Finally, the ratio between $N_{\text{obs}} - N_{\text{bg}}$ and N_{MC} is multiplied by the theoretical Born-level cross section. Equation (7.1) gives an example for the double-differential cross section at (x_0, Q_0^2) inside a bin:

$$\left(\frac{d^2\sigma^{\text{NC}}}{dx\,dQ^2}\right)^{\text{Born}}_{\text{meas}}(x_0, Q_0^2) = \frac{N_{\text{obs}} - N_{\text{bg}}}{N_{\text{MC}}} \cdot \left(\frac{d^2\sigma^{\text{NC}}}{dx\,dQ^2}\right)^{\text{Born}}_{\text{theo}}(x_0, Q_0^2)\ . \tag{7.1}$$

This unfolding method has the advantage that bin centering is automatically taken care of and that the cross section obtained from the bin can in principal be quoted at any position (x_0, Q_0^2) inside the bin.

However, strictly speaking, this method is only valid if measured and theoretical cross sections agree exactly. Otherwise, even with equal numbers of measured and expected events in a bin, the measured cross section would be wrong if the distributions of the events in data and MC inside the bin differ. Agreement between measured and theoretical cross sections can be achieved in

principal by an iterative procedure of reweighting the MC and extracting the cross section. This requires a good parameterization of the Q^2 and x dependences of the cross section, commonly achieved by parameterizing the parton density functions and evolving them with the DGLAP equations (Chaps. 2.1.3.3 and 2.3.2), indeed a major task. As the next sections will show, the agreement between measured and theoretical cross sections, evaluated with CTEQ5D PDFs, is very good from the start. Hence, omitting MC reweighting has a very small effect on the measured cross section values. Therefore, the iterative procedure is waived in this analysis.

Though a cross section can in principle be quoted at any (x, Q^2) within a bin, in most cases this is done near the logarithmic bin-center, i.e. the mean value of the logarithmic bin boundaries. Only for the single-differential cross section in y the linear bin center is quoted.

7.2 Definition of efficiency and purity

The bin-dependent quantities efficiency and purity are used to judge, whether the chosen bin size is appropriate, or whether a bin should be used at all for the extraction of a cross section. As the true kinematic quantities are needed, these studies can only be performed with MC events. Efficiency and purity in a bin are calculated from the total number of events that are generated in this bin, N_{gen}, the number of events that are generated and reconstructed in that same bin, $N_{\mathrm{gen}}^{\mathrm{rec}}$, and the number of events that are generated anywhere and reconstructed in this bin, N^{rec}. All three numbers of events are calculated with cuts on the true kinematic variables that are equivalent to the corresponding offline cuts on the reconstructed quantities, i.e. $y_{\mathrm{true}} < 0.95$ and $Q^2_{\mathrm{true}} > 185\,\mathrm{GeV}^2$ (or for the sets of $d\sigma/dx$ and $d\sigma/dy$ cross sections the corresponding lower Q^2 cut). For the reconstructed events, in addition the standard on- and offline selection cuts are applied.

Efficiency and purity are defined as follows:

$$\mathit{eff} = \frac{\mathrm{N}_{\mathrm{gen}}^{\mathrm{rec}}}{\mathrm{N}_{\mathrm{gen}}} \quad , \quad \mathit{pur} = \frac{\mathrm{N}_{\mathrm{gen}}^{\mathrm{rec}}}{\mathrm{N}^{\mathrm{rec}}} \; . \tag{7.2}$$

The acceptance, $\mathcal{A}$, listed in the tables in Appendix D, is defined as the ratio between efficiency and purity. Note that this quantity can be larger than 1, in contrast to purity and efficiency.

An efficiency far below 50% indicates either that the selected bin size is too small with respect to the detector resolution or that problems in the reconstruction lead to large migrations into other bins. In principle, the latter would be tolerable if the source for the migrations was simulated very accurately in the MC. However, migrations are often caused by imperfections of the detector which are always hard to simulate. Hence, even small inaccuracies in the simulation of the source of the migrations in conjunction with a large number of migrating events may lead to large alterations of the extracted cross sections.

Similar standards hold for the purity. If the purity is below 50%, the bin is dominated by events that were generated anywhere in the kinematic plane, including regions that are not well modeled by the MC. Again, this would lead to large and hardly controllable uncertainties in the extracted cross sections.

Therefore, in general bins with efficiency or purity below 50% are omitted from the analysis.

7.3 Extraction of single-differential cross sections

In the following, the extraction of the three single-differential cross sections $\boldsymbol{d\sigma/dQ^2}$, $\boldsymbol{d\sigma/dx}$ and $\boldsymbol{d\sigma/dy}$ is treated, where for the last two cases three scenarios with lower Q^2 cuts of 200, 2 500 and 10 000 GeV^2 are considered. In the first subsection, the binnings are described and the efficiencies and purities are determined. Then, the effects of the systematic checks on the extracted cross section are investigated. Finally, the cross sections obtained in this analysis are presented and compared to theory and to those from e^+p scattering [9].

7.3.1 Binning

The guiding principles for the choice of the binning for the various cross sections are efficiency and purity above 50%, as well as approximately equally spaced bin centers. Both $\boldsymbol{d\sigma/dQ^2}$ and $\boldsymbol{d\sigma/dx}$ cross sections are plotted logarithmically in Q^2 and x, respectively, whereas $\boldsymbol{d\sigma/dy}$ is plotted linearly in y. Hence, for the first two cases the linear bin width grows continously with increasing Q^2 and x, while the "logarithmic bin-width", i.e. the difference between the logarithms of the bin boundaries, stays constant. For $d\sigma/dy$ the difference between bin boundaries is constant. Details on the bin boundaries can be found in tab. D.1 – D.4, here only general features will be discussed.

The lowest Q^2 bin for $\boldsymbol{d\sigma/dQ^2}$ starts at 200 GeV^2 and the highest bin ends at 51200 GeV^2. As the ep cross section drops like $1/Q^4$, the logarithmic bin-width is doubled beyond $Q^2 = 3\,200$ GeV^2 to increase bin statistics.

For the lowest Q^2 cut the $\boldsymbol{d\sigma/dx}$ bins start at 0.0063 and end at 0.4. With increasing Q^2 values the covered x range is shifted towards higher x. The high-x regions for $Q^2 > 200$ GeV^2 and 2 500 GeV^2 cannot be considered, as here the validity range of the MC is exceeded (Chap. 5.4.4). For $Q^2 > 10\,000$ GeV^2 the highest bin ($0.63 < x < 1.0$) is discarded as it contains no events in the data.

For $\boldsymbol{d\sigma/dy}$ and $Q^2 > 200,\ 2\,500$ GeV^2 the bins start at $y = 0.05$ and end at $y = 0.8$. All bins have a width Δy of 0.05. For $Q^2 > 10\,000$ GeV^2 the bin width is doubled to $\Delta y = 0.1$.

7.3.2 Efficiencies and purities

Efficiency and purity of the different bins are calculated according to (7.2). Figures 7.1 and 7.2 display the results.

The efficiencies for most $\boldsymbol{d\sigma/dQ^2}$ bins in Fig. 7.1 lie well above 70%. The large dip around 600 GeV^2 is "artificial" and is caused by the super-crack cut, that has not been accounted for in the cuts for N_{gen}. The increase in efficiency for $Q^2 > 3200$ GeV^2 is caused naturally by the wider binning in this region. Also the purity always exceeds 70%. The step-like structure again originates from the larger bin widths. All Q^2 bins are accepted for extraction of the cross section $d\sigma/dQ^2$.

In the left column of Fig. 7.2 efficiencies and purities for $\boldsymbol{d\sigma/dx}$ are displayed for the three sets of lower Q^2 cuts. In all bins but the highest bin for $Q^2 > 200$ GeV^2 efficiency and purity are

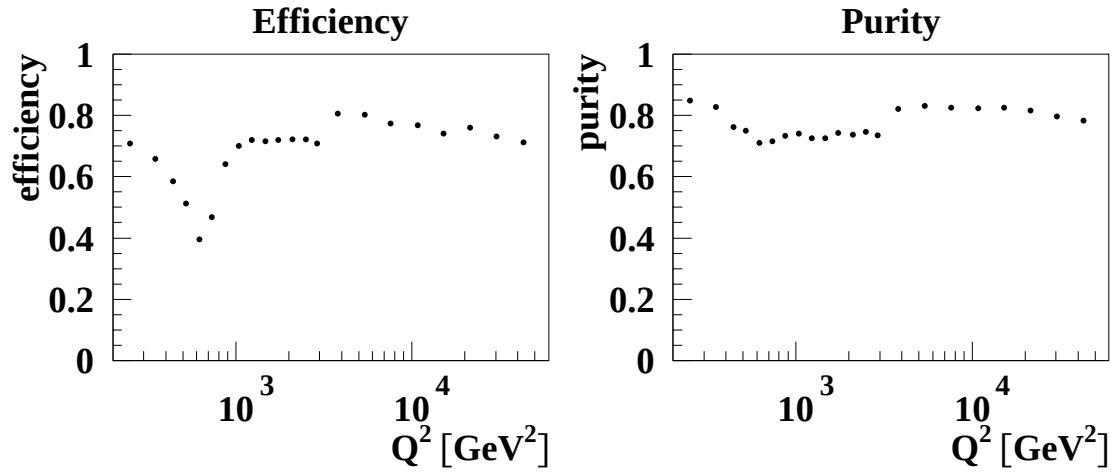


Figure 7.1: *Efficiencies and purities in the $d\sigma/dQ^2$ bins.*

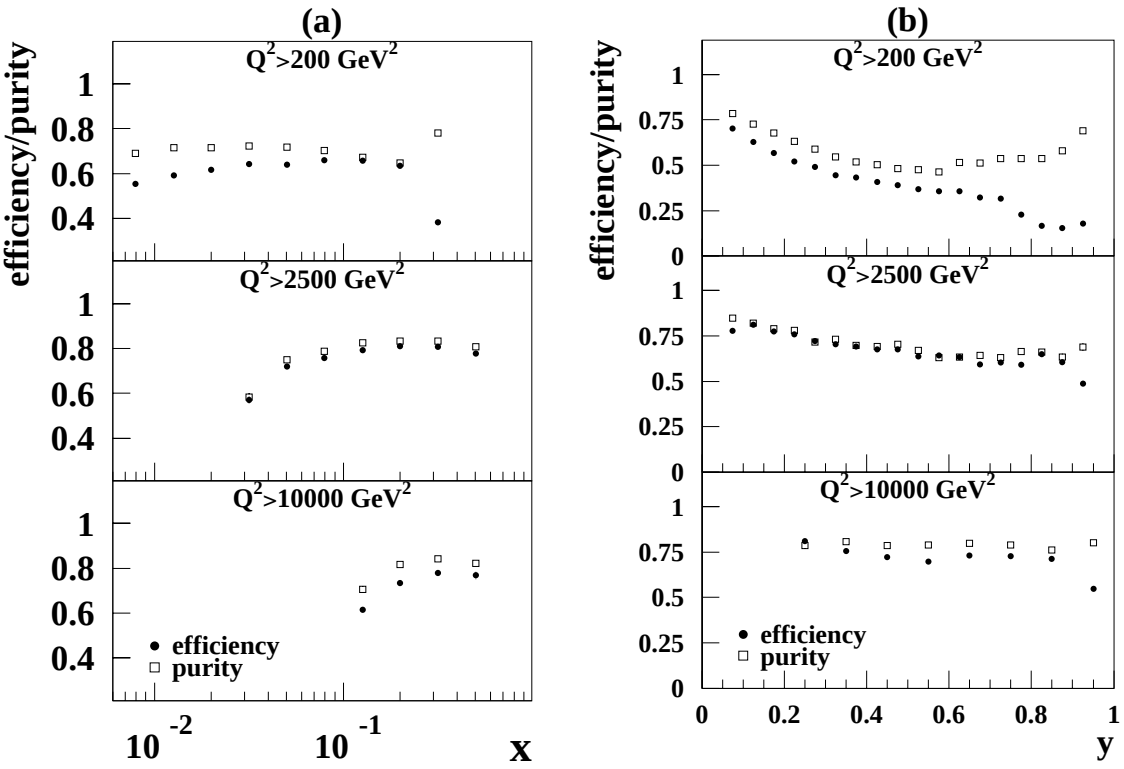


Figure 7.2: *Efficiencies and purities in the (a) $d\sigma/dx$ and (b) $d\sigma/dy$ bins for the three lower Q^2 cuts.*

above 50% which is sufficient to accept all bins. For the highest Q^2 bin the efficiency is 40% which is partially caused by the super-crack cut and hence also this bin is used for the extraction of the cross sections.

Finally, the right column of Fig. 7.2 shows the efficiencies and purities for $\boldsymbol{d\sigma/dy}$. Except for the efficiencies for $Q^2 > 200\,\mathrm{GeV}^2$ both quantities lie above 50% in all bins. The efficiency for $Q^2 > 200\,\mathrm{GeV}^2$ decreases continously with increasing y where at $y = 0.75$ a steep fall-off can be observed which extends down to 15%. The low efficiencies are again partially caused by the super-crack cut and removing it from the event selection raises the efficiency to at least 40% if the last 4 bins are excluded. This together with the good purity is sufficient to accept all but the last four bins for the extraction of the cross section.

7.3.3 Systematic checks

In this section the effects of the systematic checks on the single-differential cross sections are evaluated. A detailed description of the various checks has been given in Chap. 6.

Figure 7.3 shows a summary of the resulting variation of the cross section $d\sigma/dQ^2$ as a function of the Q^2 bins (the individual contribution of each check can be found in Appendix C). The systematic uncertainty y'_{sys} is plotted on the ordinate according to the following formula:

$$y'_{\mathrm{sys}} = \log_{10}(1 + 100 \cdot |y_{\mathrm{sys}}|) \cdot \frac{y_{\mathrm{sys}}}{|y_{\mathrm{sys}}|} \,, \tag{7.3}$$

where y_{sys} is the relative systematic uncertainty plotted in Fig. C.1. The areas of different gray scales in Fig. 7.3 represent the uncertainties obtained from the various checks, that are listed on top of the picture. Inspecting the error bands from the inside outwards, the innermost originates from the variation of the REMC energy scale and the outermost from the energy scaling of the innermost ring of FCAL cells ($E_{1\mathrm{st}}$). The contribution of each source of uncertainty considered is added in quadrature to the sum of the previous uncertainties, thus yielding in the end the total systematic uncertainty. The individual contributions of the checks are proportional to the corresponding shaded regions.

The largest contributions for $\boldsymbol{d\sigma/dQ^2}$ come from lowering the y_{el} cut and from the variation of the isolation and P_{trk} cuts (Fig. C.1), where the latter yields the largest contribution in the low-Q^2 region. The other variables identified as problematic in Chap. 5.6, i.e. DCA and $E_{1\mathrm{st}}$, give only small contributions below 2%.

The variations of the cross sections in case of the isolation and y_{el} cuts show two typical "patterns". For y_{el} the variation in the high Q^2 regions seems to be more or less erratic whereas decreasing (increasing) the P_{trk} cut leads to a steady increase (decrease) of the cross section. The former finding indicates that this systematic uncertainty is dominated by statistical fluctuations in the accepted fraction of a small number of events. The second case indicates a problem with the simulation, in this case with the hadronic energy distribution in the forward region of the calorimeter.

Despite some large systematic uncertainties, the measured $d\sigma/dQ^2$ cross section is dominated by the statistical errors except for the first two bins where statistical and systematic errors are of approximately the same size.

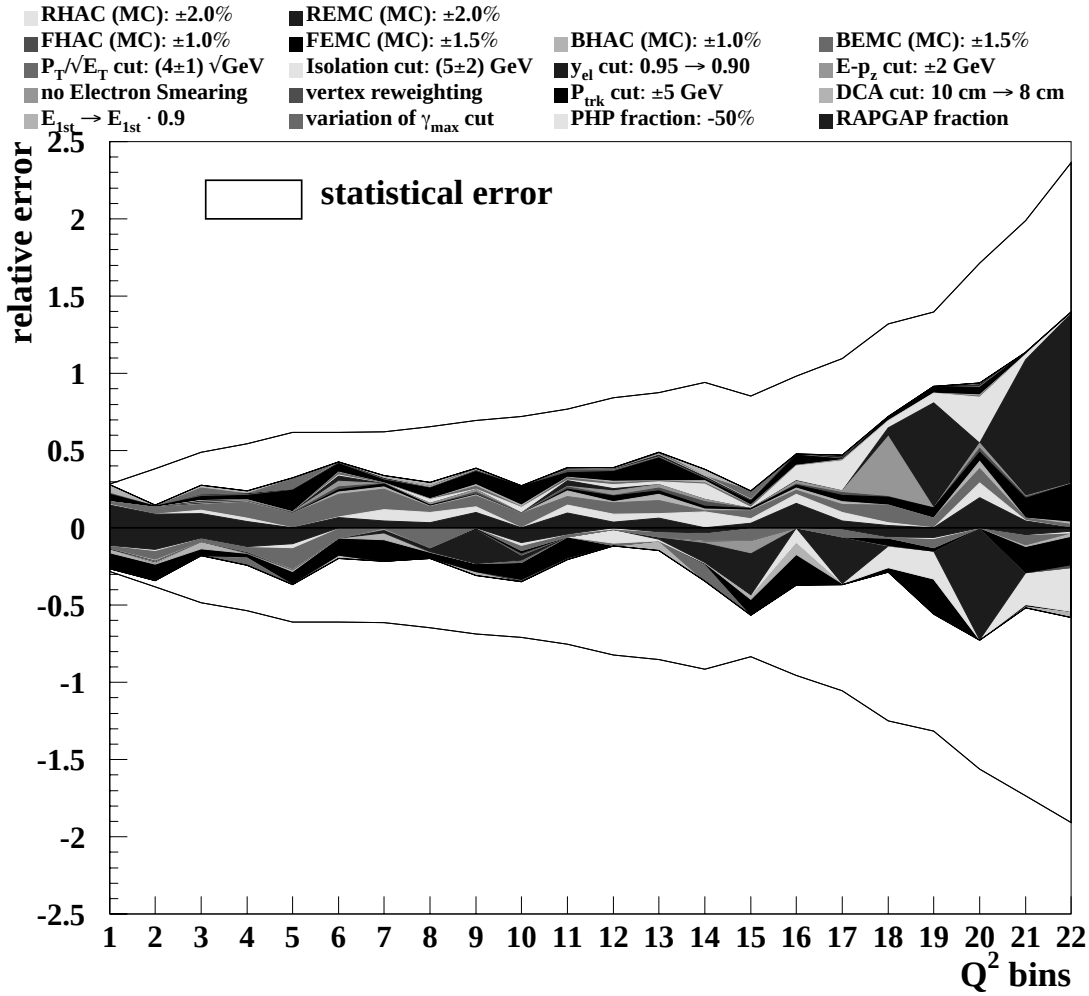


Figure 7.3: *Relative systematic uncertainties for $d\sigma/dQ^2$, obtained from systematic checks, as a function of the Q^2 bins. The error value on the ordinate is calculated according to the formula $y'_{\mathrm{sys}} = \log_{10}(1 + 100 \cdot |y_{\mathrm{sys}}|) \cdot \frac{y_{\mathrm{sys}}}{|y_{\mathrm{sys}}|}$, where y_{sys} is the relative systematic uncertainty. The shaded region represents the total systematic uncertainty, whereas the solid lines indicate the statistical error. For further explanations see text.*

Figure 7.4 summarizes the systematic uncertainties for $\boldsymbol{d\sigma/dx}$ in all three Q^2 bins. The individual contributions are displayed in Figs. C.2 – C.4. The measurement is dominated by statistics in all but the two highest bins for $Q^2 > 200\,\mathrm{GeV}^2$, where the rescaling of the energy in the innermost FCAL ring of cells yields a contribution of up to 10%. This large effect is located at low y rather than high x (visible on its decrease with increasing Q^2) as here the hadronic angle is very small and hence a large fraction of the hadronic final state is located in the innermost ring of cells. For $Q^2 > 200\,\mathrm{GeV}^2$ ($2500\,\mathrm{GeV}^2$) the largest contributions at low x originate from the variation of the P_{trk} and γ_{max} (P_{trk}, isolation and γ_{max}) cuts. For $Q^2 > 10\,000\,\mathrm{GeV}^2$ also the variation of the $E - p_z$ and y_{el} cuts yield large systematic uncertainties at high and low x, respectively, where the former is probably a statistical fluctuation. A clearly non-statistical contribution to the systematic uncertainty comes from the variation of the isolation cut starting at high x with $\sim 2\%$ and rising to 5% towards low x.

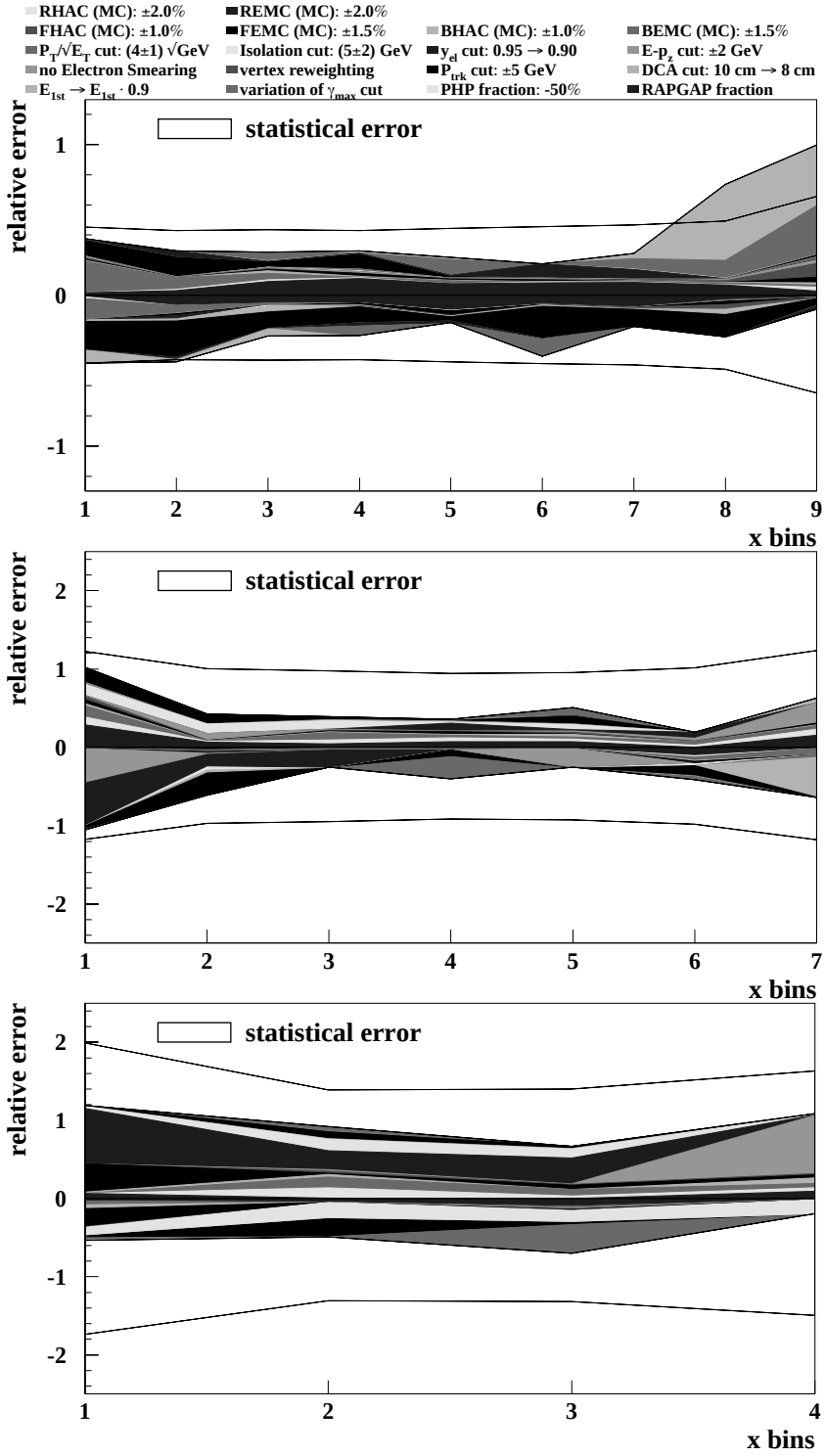


Figure 7.4: *Relative systematic uncertainties for $d\sigma/dx$, obtained from systematic checks, as a function of the x bin number, for (a) $Q^2 > 200\,GeV^2$, (b) $Q^2 > 2\,500\,GeV^2$ and (c) $Q^2 > 10\,000\,GeV^2$. The error value on the ordinate is calculated according to the formula $y'_{\mathrm{sys}} = \log_{10}(1 + 100 \cdot |y_{\mathrm{sys}}|) \cdot \frac{y_{\mathrm{sys}}}{|y_{\mathrm{sys}}|}$, where y_{sys} is the relative systematic uncertainty. The shaded region represents the total systematic uncertainty, whereas the solid lines indicate the statistical error. For further explanations see text.*

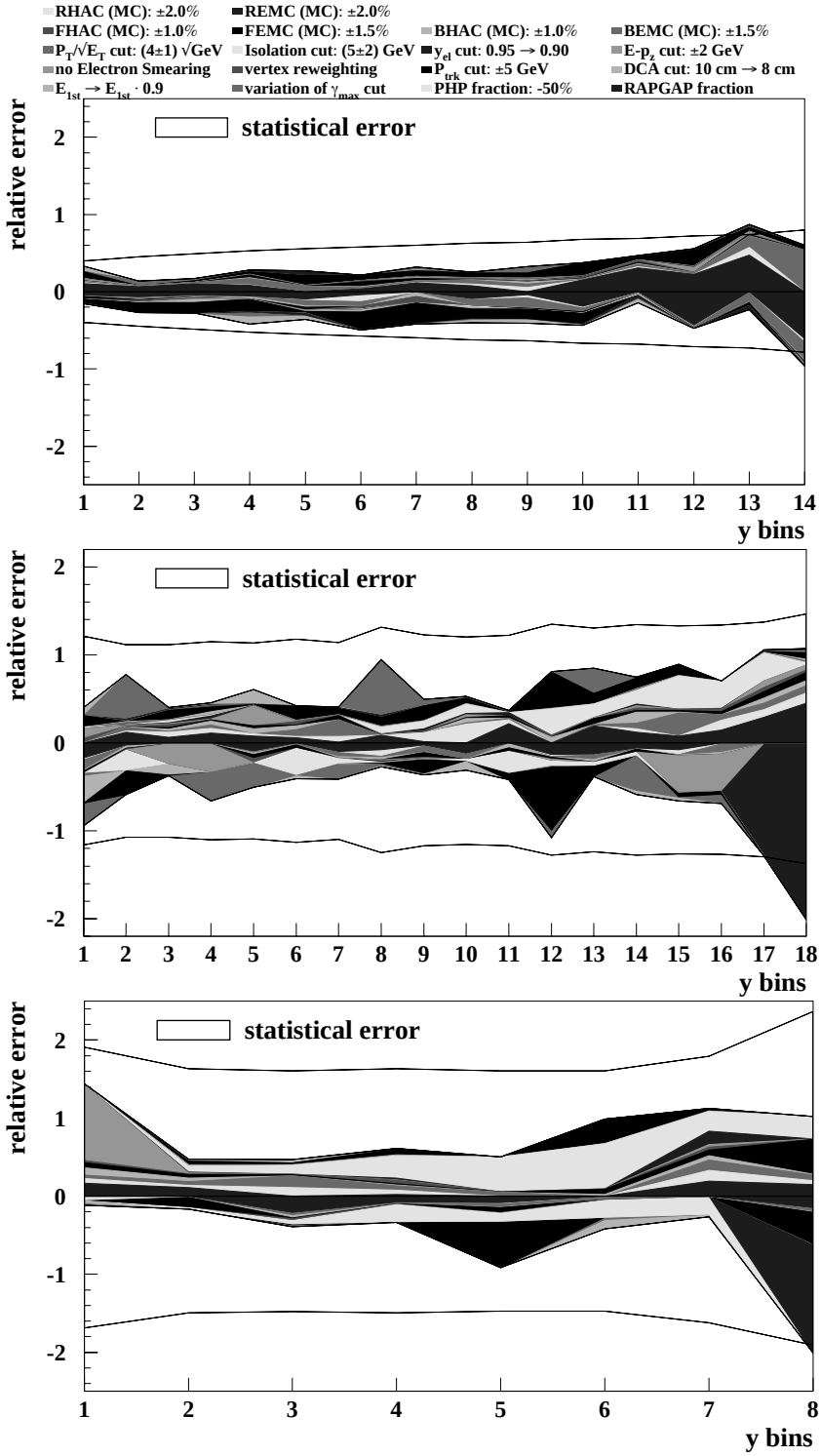


Figure 7.5: *Relative systematic uncertainties for $d\sigma/dy$, obtained from systematic checks, as a function of the y bin numbers, for (a) $Q^2 > 200\ GeV^2$, (b) $Q^2 > 2\,500\ GeV^2$ and (c) $Q^2 > 10\,000\ GeV^2$. The error value on the ordinate is calculated according to the formula $y'_{\mathrm{sys}} = \log_{10}(1 + 100 \cdot |y_{\mathrm{sys}}|) \cdot \frac{y_{\mathrm{sys}}}{|y_{\mathrm{sys}}|}$, where y_{sys} is the relative systematic uncertainty. The shaded region represents the total systematic uncertainty, where the solid lines indicate the statistical error. For further explanations see text.*

The systematic checks for the $\boldsymbol{d\sigma/dy}$ cross sections are summarized in Fig. 7.5. Details can be found in Figs. C.5 – C.7. The largest systematic uncertainties for $Q^2 > 200\,\mathrm{GeV}^2$ (Fig. C.2) are located at high y and stem from the variation of the REMC and BEMC energy scales, as well as from the variation of the P_{trk} cut. For $Q^2 > 2\,500\,\mathrm{GeV}^2$ (Fig. C.3) the variations of the isolation, P_{trk} and γ_{max} cuts yield the largest uncertainties. For $Q^2 > 10\,000\,\mathrm{GeV}^2$ (Fig. C.4) the variation of the isolation cut is the dominant uncertainty at high y, while the variation of the P_{trk} cut yields some larger contributions at medium y. Except for the highest y bin(s) the errors on $d\sigma/dy$ are dominated by statistics.

7.3.4 Single-differential cross sections

In this section the results on the three single-differential cross sections are presented.

Figure 7.6 shows the measured cross sections for $\boldsymbol{d\sigma/dQ^2}$, where (a) displays the absolute cross section and (b) the ratio of the measured to the theoretical cross section calculated from the ZEUS NLO parameterization[1]. Also shown are theoretical calculations based on CTEQ5D and MRST(99) PDF parameterizations. The shaded band indicates the uncertainty on the ZEUS NLO fit.

The measured cross sections show very good agreement with the theoretical ones and the normalization agrees within 0.1% compared to CTEQ5D. No excess is observed in the high-Q^2 region. All theoretical cross sections agree within the PDF uncertainty and the precision of the data is not sufficient to distinguish between the parameterizations bearing in mind that the normalization uncertainty is 1.5%.

Figures 7.7 a – c show the measured $\boldsymbol{d\sigma/dx}$ cross sections for $Q^2 > 200\,\mathrm{GeV}^2$, $2\,500\,\mathrm{GeV}^2$ and $10\,000\,\mathrm{GeV}^2$. Figure 7.7 a seems to show a tendency of increasing cross sections with decreasing x for the middle x region. An opposite effect can be observed in Fig. 7.7 b. However, all three sets of measured cross sections are in good agreement with the predictions from CTEQ5D and ZEUS NLO considering the errors on the measured cross sections and the PDF uncertainty. On the other hand, for $Q^2 > 200\,\mathrm{GeV}^2$ and $x \lesssim 0.01$ the MRST(99) parameterization is considerably lower than CTEQ5D and ZEUS NLO. In this region the data favors the latter two parameterizations. For the higher Q^2 cuts the differences between the parameterizations are negligible in the region covered by the measured cross sections.

Figures 7.7 a – c also clearly show the rapidly increasing PDF uncertainty towards $x = 1$. In this region data is scarce and hence the differences between the PDF sets are large. Also this analysis cannot improve this uncertainty as statistics is still too low in this region.

Finally, Figs. 7.8 a – c display the measured $\boldsymbol{d\sigma/dy}$ cross sections together with the theoretical predictions. Overall good agreement between measured and calculated cross sections is observed. For $Q^2 > 200\,\mathrm{GeV}^2$ the data seems to favor the CTEQ5D and ZEUS NLO parameterizations as the cross section for MRST(99) is 2–3% lower than the other parameterizations over almost the whole y range. On the other hand, the PDF uncertainty together with the normalization error

[1] Note that the ZEUS NLO fit utilized in this thesis (Chap. 2.3.2.3) does not use any of the e^-p data analyzed in this thesis.

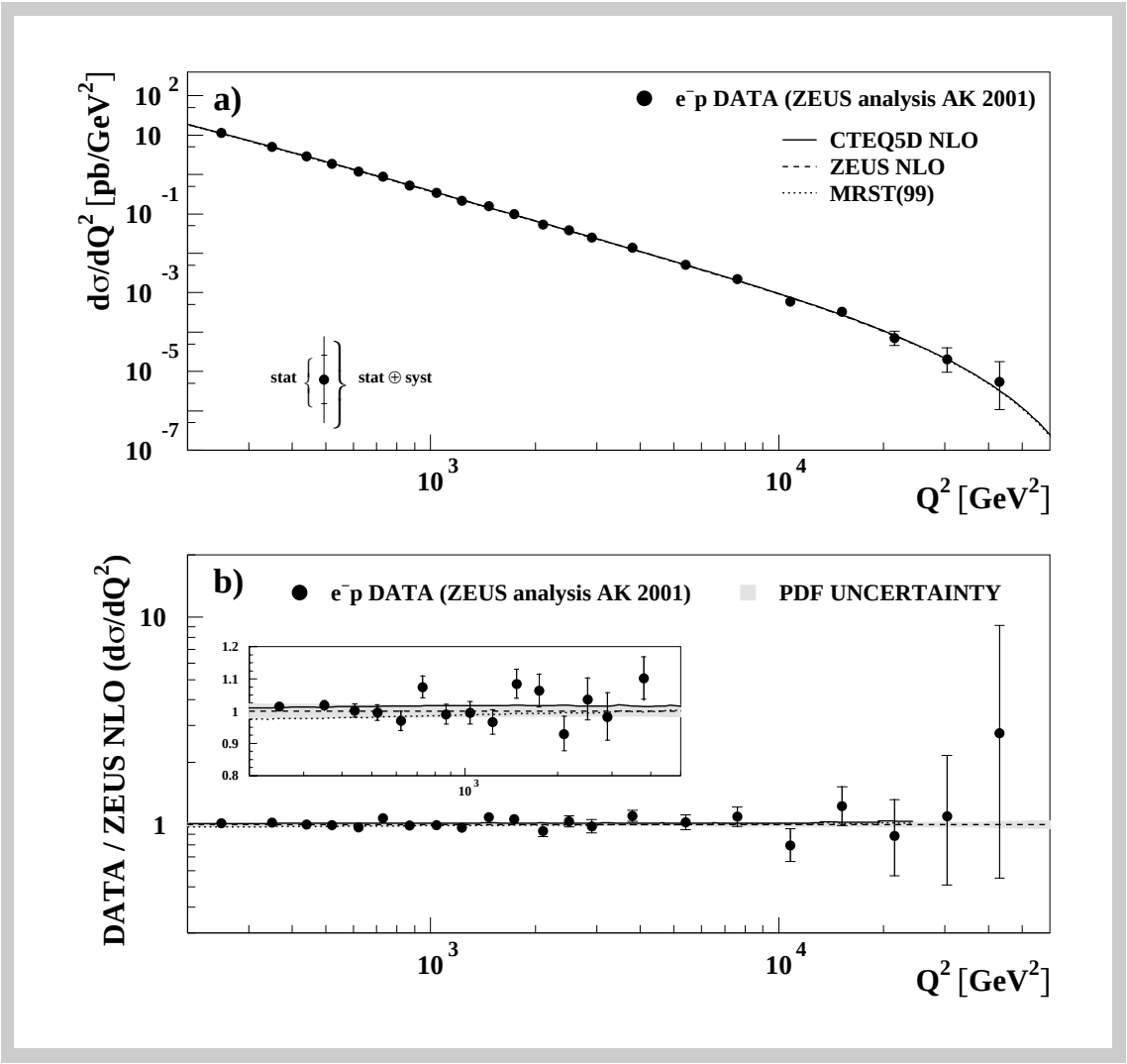


Figure 7.6: *(a) Single-differential cross section $d\sigma/dQ^2$ as a function of Q^2 compared to theoretical calculations based on CTEQ5D, ZEUS NLO and MRST(99) PDFs. (b) Ratio of measured cross sections to ZEUS NLO calculations. Also shown are the ratios of the other two theoretical calculations to the ZEUS NLO fit. The inset is a blow-up of the range $200\,GeV^2 < Q^2 < 5\,000\,GeV^2$, with the ordinate zoomed in to 0.8–1.2. The shaded band indicates the uncertainty on the calculated cross section.*

of 1.5% is of similar size. The difference between the parameterizations vanishes for the higher Q^2 cuts. For $Q^2 > 10\,000\,\mathrm{GeV}^2$ and $y < 0.2$ both CTEQ5D and MRST(99) are considerably higher than ZEUS NLO, corresponding to the high x regions in Figs. 7.7 a – c where data is not available yet.

7.3.5 Comparison to e^+p cross sections

In this section, the measured e^-p cross sections are compared to those of e^+p scattering, obtained in the years 1994–97 at a center-of-mass energy of 300 GeV [9]. Figure 7.9 displays the e^-p cross

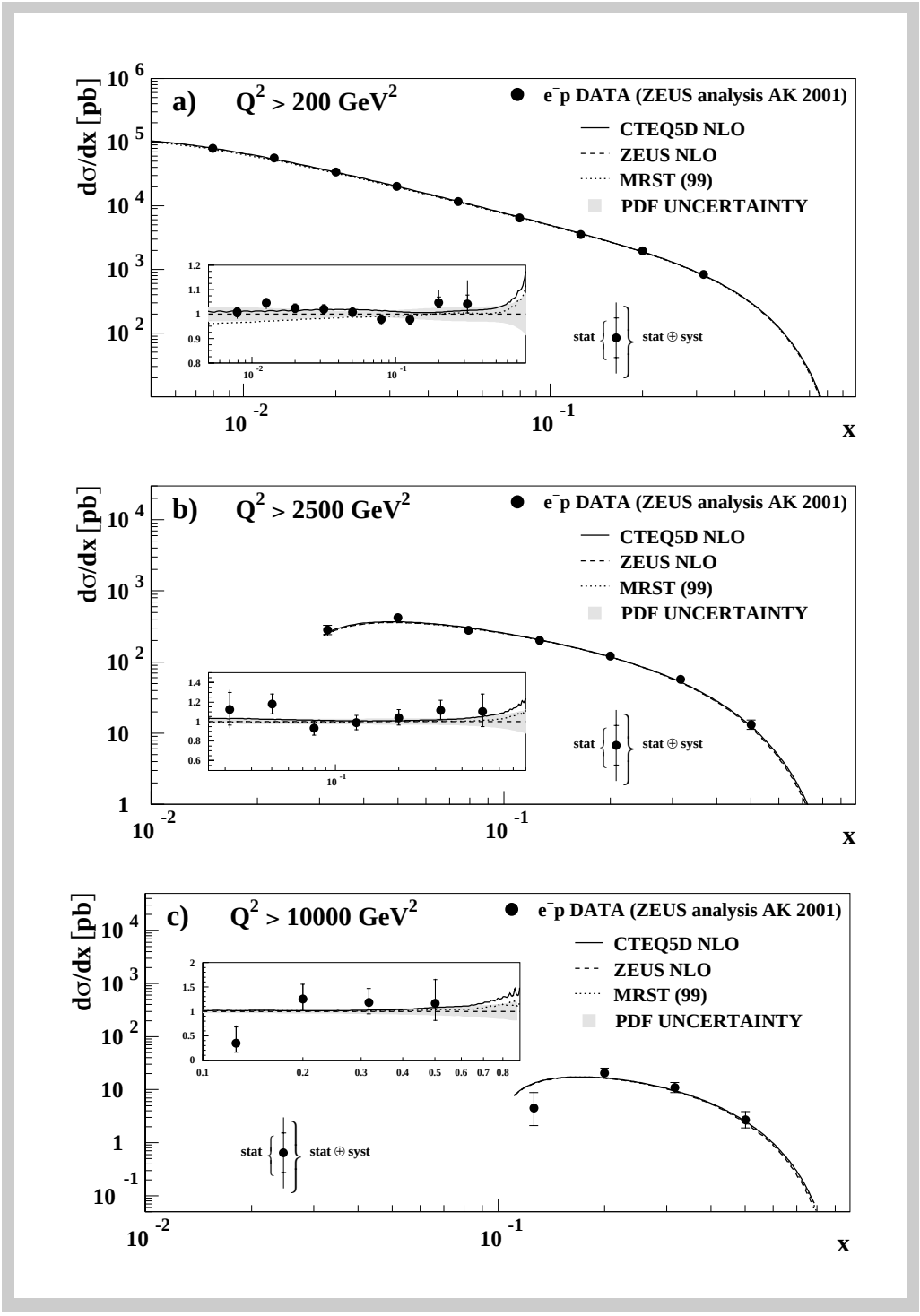


Figure 7.7: *Single-differential cross sections $d\sigma/dx$ as a function of x together with theoretical calculations based on CTEQ5D, ZEUS NLO and MRST(99) PDFs for (a) $Q^2 > 200\,GeV^2$, (b) $Q^2 > 2\,500\,GeV^2$ and (c) $Q^2 > 10\,000\,GeV^2$. The insets show the ratios between the measured cross sections and the theoretical calculations based on the ZEUS NLO fit. Also shown are the ratios of the other two theoretical calculations to the ZEUS NLO fit. The shaded band indicates the uncertainty on the ZEUS NLO fit.*

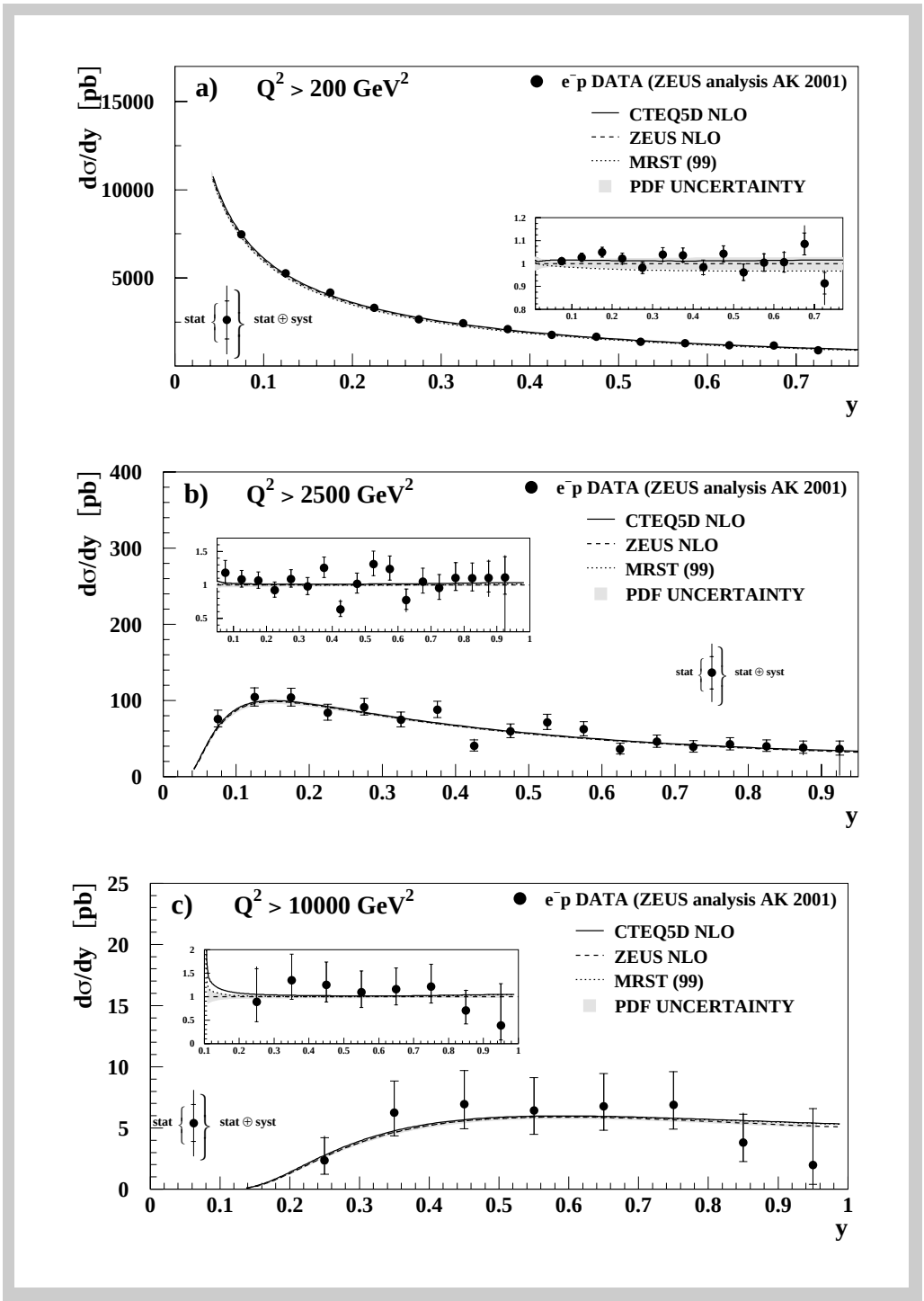


Figure 7.8: *Single-differential cross sections $d\sigma/dy$ as a function of y compared to theoretical calculations based on CTEQ5D, ZEUS NLO and MRST(99) PDFs for (a) $Q^2 > 200\,GeV^2$, (b) $Q^2 > 2\,500\,GeV^2$ and (c) $Q^2 > 10\,000\,GeV^2$. The insets show the ratios between the measured cross sections and the ZEUS NLO fit. Also shown are the ratios of the other two theoretical calculations to the ZEUS NLO fit. The shaded band indicates the uncertainty on the ZEUS NLO fit.*

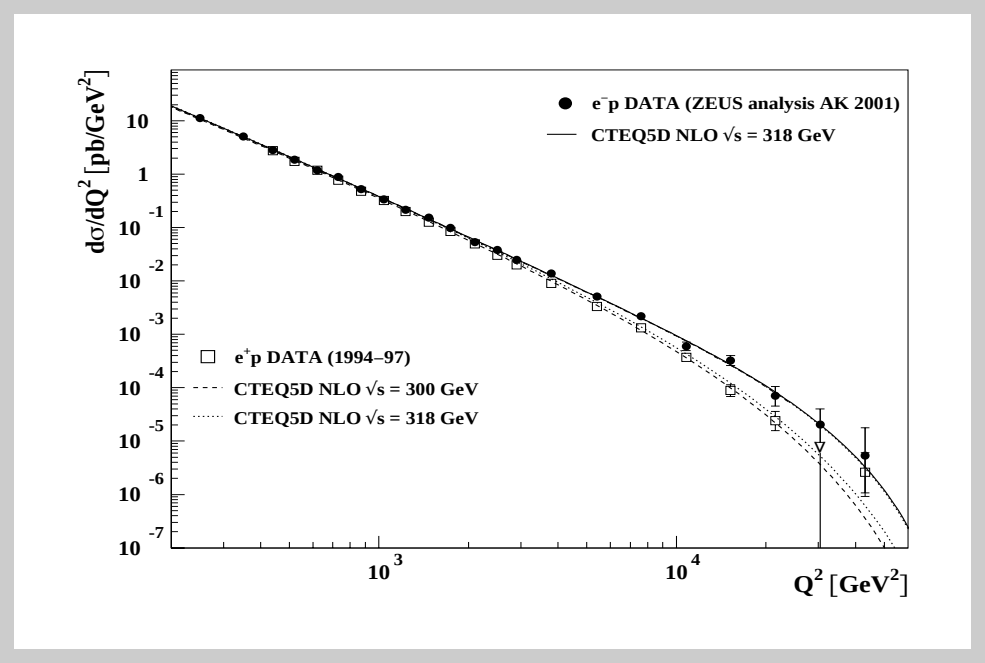


Figure 7.9: *Comparison of measured cross sections $d\sigma/dQ^2$ for e^-p (recorded at $\sqrt{s} = 318\,GeV$) and e^+p (recorded at $\sqrt{s} = 300\,GeV$) scattering as a function of Q^2. The lines represent theoretical predictions using CTEQ5D for e^-p at $\sqrt{s} = 318\,GeV$ (solid), e^+p at $\sqrt{s} = 300\,GeV$ (dashed) and e^+p at $\sqrt{s} = 318\,GeV$ (dotted).*

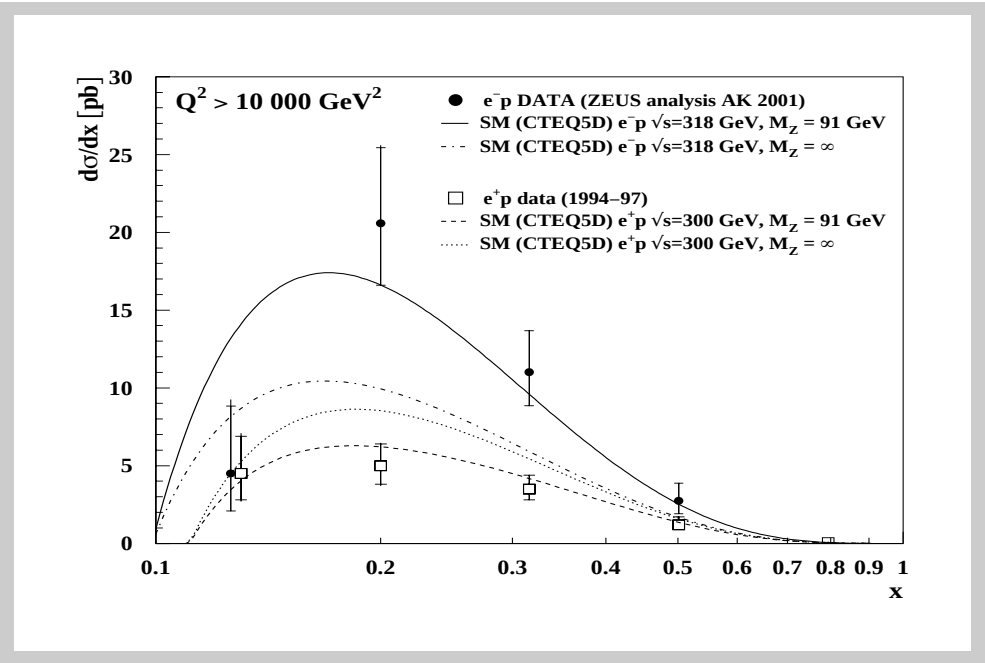


Figure 7.10: *Comparison of measured cross sections $d\sigma/dx$ for e^-p (recorded at $\sqrt{s} = 300\,GeV$) and e^+p (recorded at $\sqrt{s} = 300\,GeV$) scattering as a function of x for $Q^2 > 10\,000\,GeV^2$. In addition to the "nominal" calculated cross sections, denoted by solid/dashed lines, the plot shows those for an infinite Z mass as dash-dotted/dotted lines. The lowest e^-p point in x is moved slightly to the left as otherwise e^-p and e^+p would fall on top of each other.*

sections as full circles and the e^+p cross sections as open squares. The lines represent the theoretical predictions. The enhancing (decreasing) effect of the structure function xF_3 on the cross section for e^-p (e^+p) (see (2.19)) is clearly visible. Though the lower center-of-mass energy for e^+p also reduces the cross section, as can be seen from the difference between the dashed and dotted line, the effect of the Z exchange is clearly dominating.

Figure 7.10 compares the measured $d\sigma/dx$ cross sections from e^+p and e^-p scattering for $Q^2 > 10\,000\,\mathrm{GeV}^2$. Again the enhancing (decreasing) effect of the Z exchange for e^-p (e^+p) is visible. In addition to the nominal theoretical cross sections those for an infinite Z mass (weak interaction turned off) are plotted as dash-dotted and dotted lines. The measured cross sections clearly reject such a model and support the presence of the weak force with its nominal SM strength.

In Chaps. 7.5 and 7.6 the difference between the cross section for e^+p and e^-p will be exploited to extract the mass of the Z boson and the structure function xF_3.

7.4 Extraction of double-differential cross sections

In this section the extraction of the double-differential cross section is described. As the area of the kinematic plane covered by the individual bins is much smaller than in the case of the single-differential cross sections, the extracted cross sections are much more sensitive to inaccuracies in the simulation of the underlying physics process and the detector. Therefore, more attention has to be payed to systematic effects.

7.4.1 Binning

The bins used for the extraction of the double-differential cross section, displayed in Fig. 7.11, cover the kinematic plain between $185\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $0.0037 < x < 1$. The bins are almost equally spaced in $\log(Q^2)$ where the higher bins are a little bit broader due to the rapidly decreasing cross section. Also the width in x increases towards low y to compensate for larger migrations. Bins beyond the $y(1-x)^2 = 0.004$ limit were omitted as they would span a large region of the kinematic plane where the MC is assumed to be invalid (Chap. 5.4.4). Bins at low Q^2 and low x were dropped due to their bad efficiencies and purities of the order of 30% or below. The bins equal those of the 1996–97 e^+p analysis [9], as the results from the latter are later needed for the extraction of xF_3.

7.4.2 Efficiencies, purities and migrations

The efficiencies and purities are calculated according to (7.2). Figure 7.12 displays for each bin the number of reconstructed data events in large font in the middle of the bin and the efficiencies (purities) in small font in the lower left (right) corner. The efficiencies lie almost always well above 50% and in the high Q^2 regime even above 70%. Only in the region of the R/BCAL super-crack indicated by the line of constant $\theta = 2.25\,\mathrm{rad}$ the efficiencies drop to 40%. The purity in all bins exceeds 50% and lies above 70% in the high Q^2 region. No effect of the

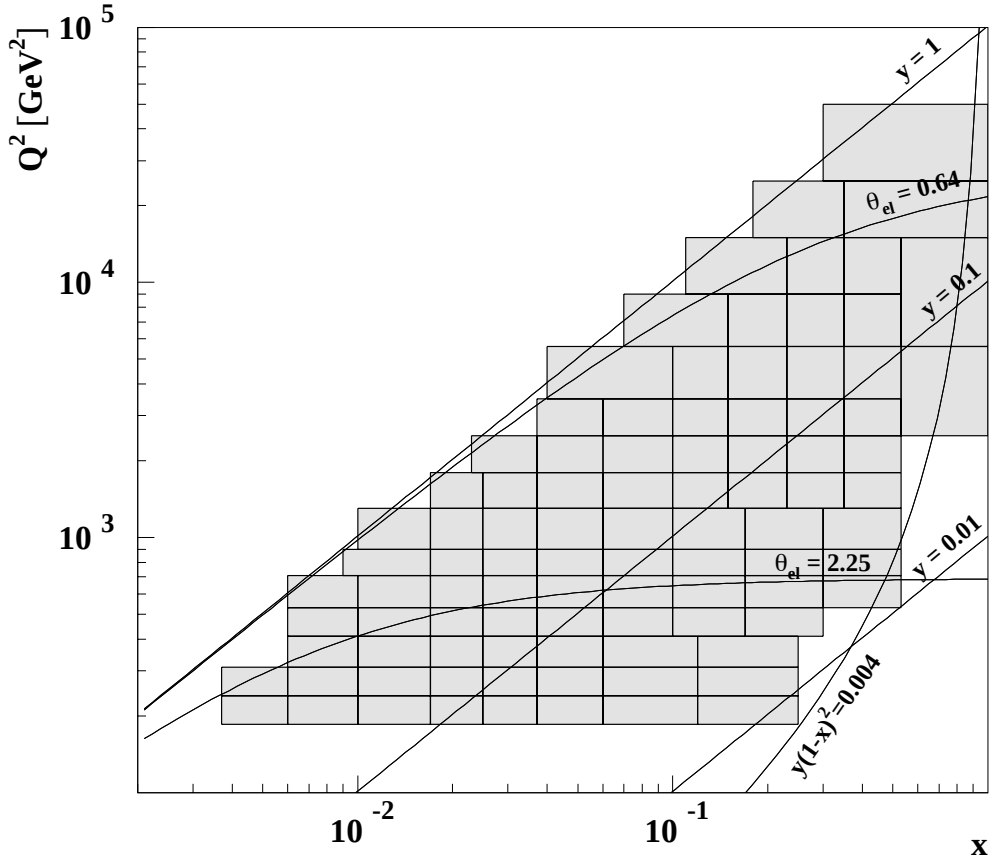


Figure 7.11: *Double-differential binning in the x-Q^2 plane. Also shown are lines of constant y and θ_{el}, where the latter mark the super-crack regions between the calorimeter parts. The line labeled $y(1-x)^2 = 0.004$ indicates the validity limit of the MC. All displayed bins are accepted for the extraction of the cross sections.*

super-crack cut on the purity is visible. All bins are accepted for the extraction of the cross sections.

Apart from the bare efficiency and purity numbers it is interesting and helpful to investigate the event migrations in the kinematic plane, displayed in Fig. 7.13 and obtained from MC studies. In Figure 7.13 a the beginning of an arrow is positioned at the logarithmic center of the bin the events were generated in and the arrowhead marks the mean reconstructed position. In Figure 7.13 b the arrowhead points to the logarithmic bin center where the events where reconstructed, and the beginning of the arrow marks the mean generated position. As almost all arrows remain in a single bin, these plots support the conclusions drawn from Fig. 7.12, that both efficiency and purity are sufficiently high to accept all bins for the extraction of the cross sections.

7.4.3 Systematic checks

The systematic checks equal those for the single-differential cross sections. Figure 7.14 displays a summary of the individual systematic checks as a function of the x-Q^2 bins, where a detailed

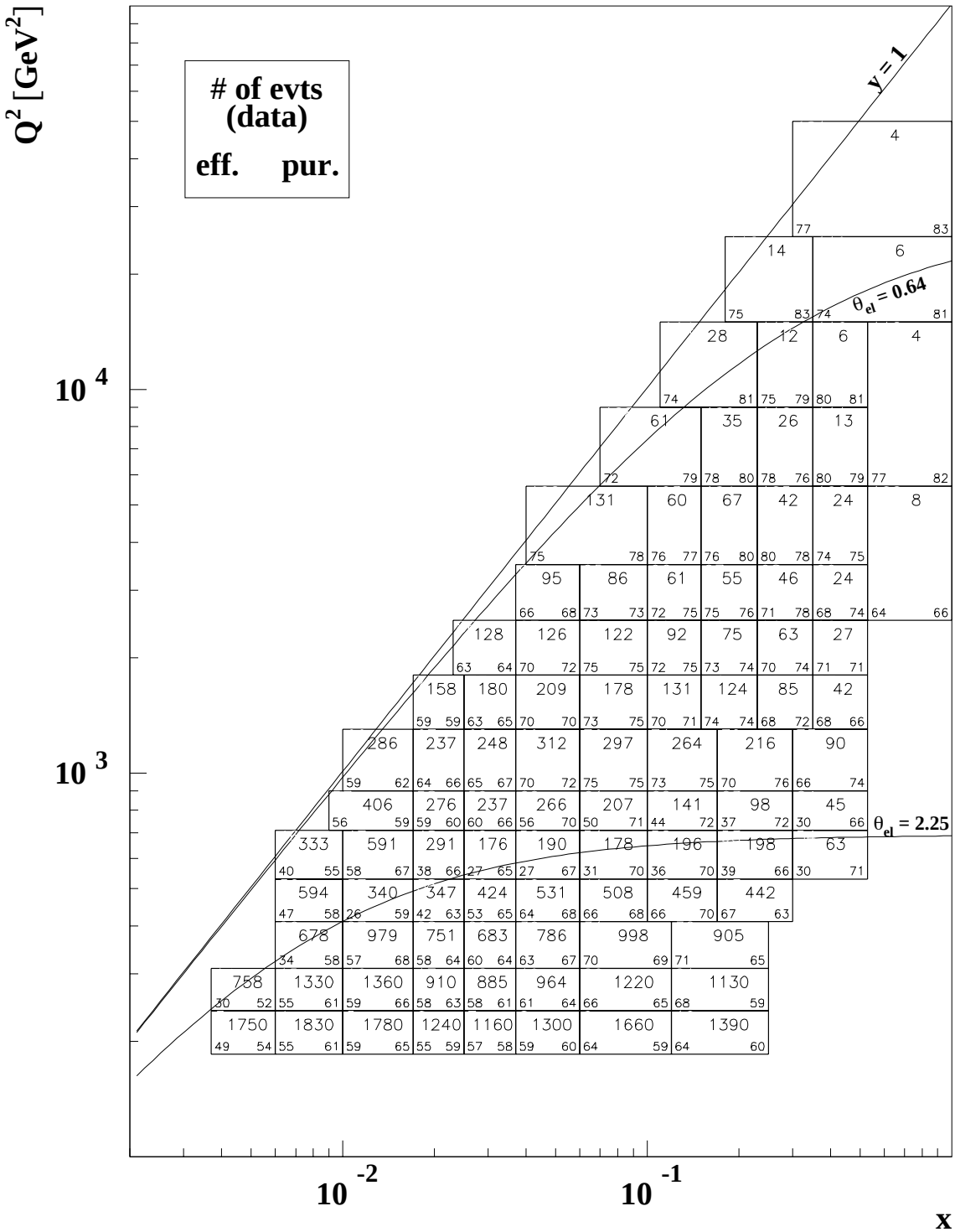


Figure 7.12: *Double-differential binning with number of events (data), efficiency and purity (both in %) for each bin. Also shown are lines of constant* θ_{el} *marking the super-crack regions between the calorimeter parts.*

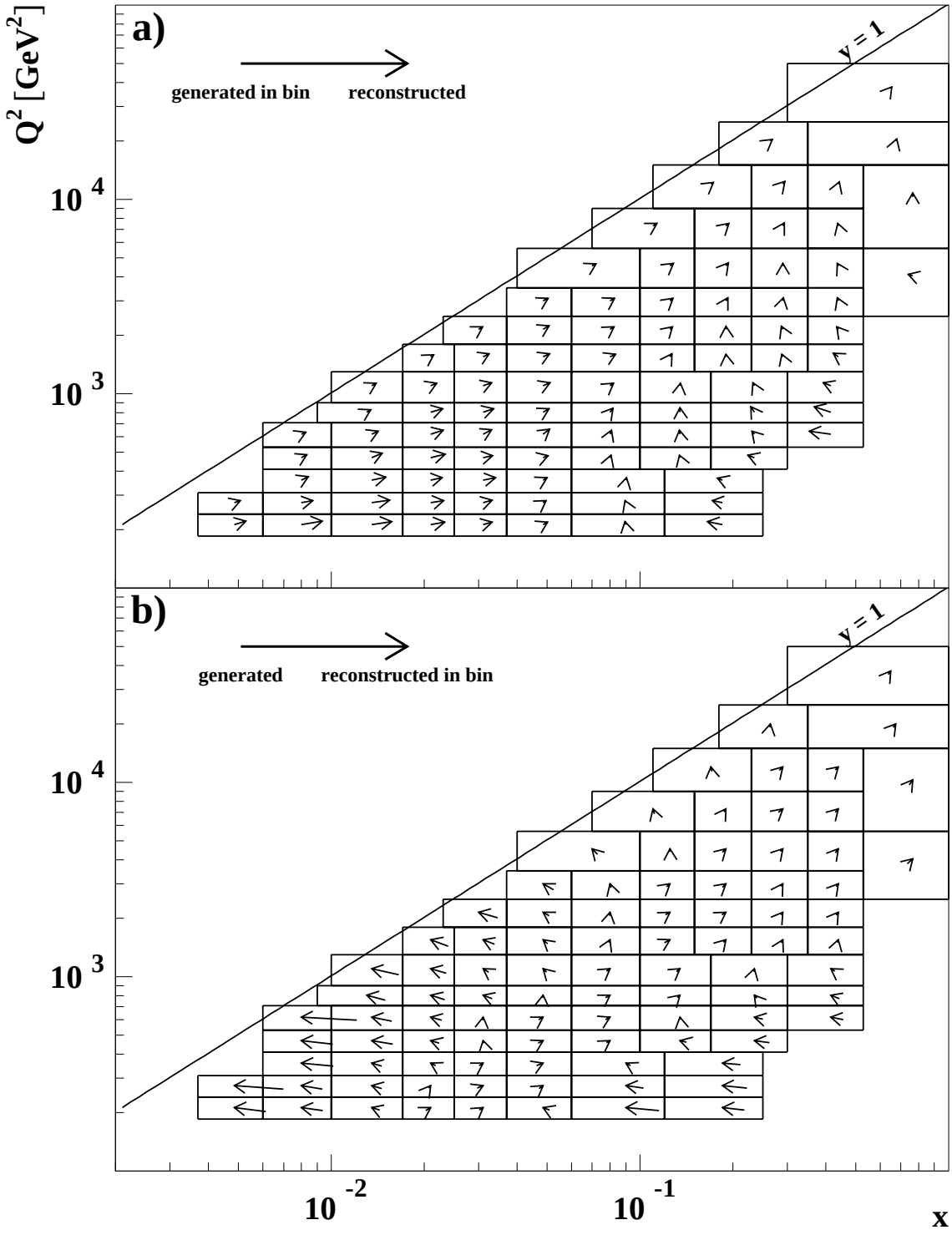


Figure 7.13: *Migrations of events in the kinematic x-Q^2 plane: (a) arrow from the center of the bin the events were generated in to the mean reconstructed position and (b) arrow from the mean true position to the center of the bin the events are reconstructed in.*

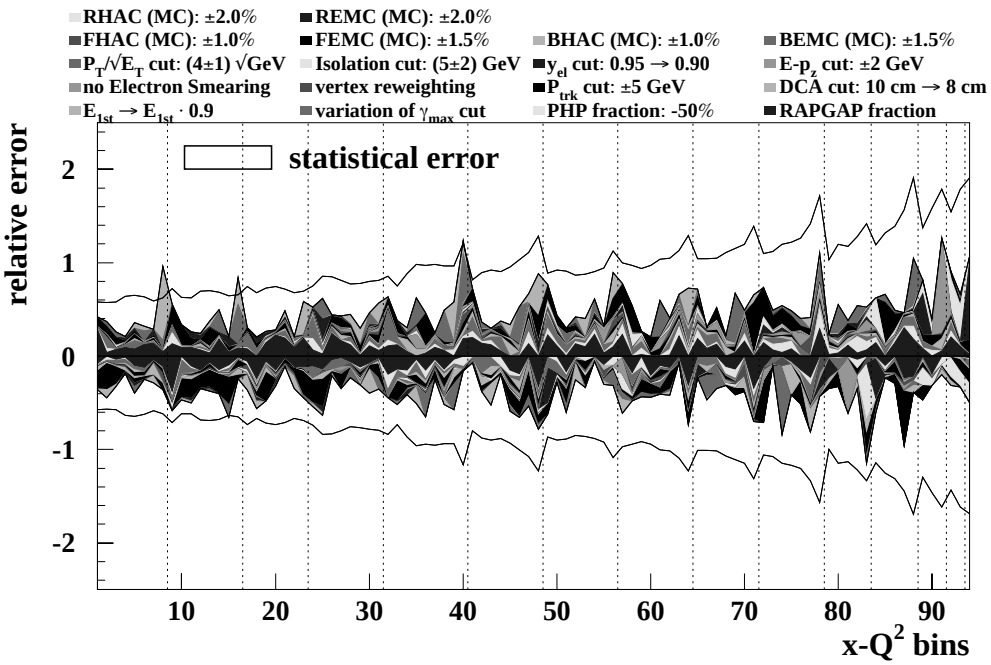


Figure 7.14: *Systematic uncertainties for $d^2\sigma/dxdQ^2$ as a function of the x-Q^2 bin. The bins are ordered by increasing x in each Q^2 bin where the vertical dashed lines mark the transition from one Q^2 bin to the next. The value on the ordinate is calculated according to the formula $y'_{\mathrm{sys}} = \log_{10}(1 + 100 \cdot |y_{\mathrm{sys}}|) \cdot \frac{y_{\mathrm{sys}}}{|y_{\mathrm{sys}}|}$, where y_{sys} is the relative systematic uncertainty. The inner shaded region represents the total systematic uncertainty, where the solid lines indicate the statistical error. For further explanations see Chap. 7.3.3.*

list of the individual contributions can be found in Appendix C.4.

The total systematic uncertainty, represented by the inner shaded band, is almost always lower than the statistical error indicated by a solid line. Hence, the uncertainties on the measured cross sections are dominated by statistics. Larger contributions to the systematic uncertainty of up to 8% stem from the variation of the isolation, DCA, y_{el} and P_{trk} cuts and are mainly located at high Q^2 and low or high x. Larger systematic uncertainties in the medium and low Q^2 region originate from the variation of the γ_{max} cut and the rescaling of the energies in the innermost ring of FCAL cells. The majority of the systematic checks contribute less than 2% to the systematic uncertainty.

7.4.4 Double-differential cross sections

This section presents the measured double-differential cross sections. Figures 7.15 and 7.16 show the cross sections as a function of x in different Q^2 bins, where on the ordinate the so-called reduced cross section $\tilde{\sigma}$ is plotted, defined as (neglecting $y^2 F_L^{\mathrm{NC}}$)

$$\tilde{\sigma} := \frac{x\,Q^4}{2\pi\alpha^2} \cdot \frac{1}{Y_+} \cdot \frac{d^2\sigma^{\mathrm{NC}}}{dx\,dQ^2} \approx F_2^{\mathrm{NC}} \mp \frac{Y_-}{Y_+} x F_3^{\mathrm{NC}} \,. \tag{7.4}$$

Neglecting $y^2 F_L$, $\tilde{\sigma}$ can be interpreted as an "effective" F_2^{NC} structure function, modified by xF_3^{NC}. Using this quantity instead of the absolute cross section has the advantage that the

"trivial" $1/Q^4$ dependence of the cross section is eliminated leading to a better presentability. In addition to the measured cross sections those from theoretical predictions are plotted using the CTEQ5D, ZEUS NLO and MRST(99) PDF parameterizations (Chap. 2.3.2).

The precision of the measured cross sections is high enough to distinguish between the PDFs for $Q^2 \lesssim 450\,\mathrm{GeV}^2$ and $x \lesssim 3 \cdot 10^{-2}$. In this region the data favors the CTEQ5D and ZEUS NLO parameterizations. However, for $Q^2 = 1200\,\mathrm{GeV}^2$ the measured cross sections between $x = 0.02$ and 0.08 seem to agree better with MRST(99). This behavior is not yet understood and needs further investigation although a statistical fluctuation cannot be excluded either.

For a better comparability of measured and theoretical cross section Figs. 7.17 and 7.18 display the ratio between both, where the latter is based on the ZEUS NLO parameterization. Again the large PDF uncertainties at high x for all Q^2 bins are visible. For $Q^2 < 1500\,\mathrm{GeV}^2$ no distinctive features are visible if one abstains from the somewhat lower MRST(99) parameterization and the $Q^2 = 1200\,\mathrm{GeV}^2$ bin already visible in the reduced cross section plot. For $Q^2 \geq 1500\,\mathrm{GeV}^2$ CTEQ5D tends to lie above the ZEUS NLO fit where the difference becomes larger towards lower x. However, the error on the measured cross section is too large to distinguish between both PDFs. Again, the large uncertainties on the PDFs at high x and the even larger differences between the theoretical predictions in this region stand out.

7.4.5 Comparison to e^+p data

Figure 7.19 displays the comparison between the e^-p and e^+p double-differential cross sections, where the latter is taken from the analysis of the 1996/97 data [9], recorded at a center-of-mass energy of 300 GeV. Both reduced cross sections are plotted as function of Q^2 in different bins in x, ranging from $x = 0.08$ to 0.65. Note that the measured cross sections are multiplied by a factor, indicated in brackets, which allows displaying all cross sections in the same plot. Furthermore, the plot does not show all measured points as the interesting area, where both cross sections can be distinguished due to opposite effects of the Z exchange, is located at high Q^2. A comparison of the data to the theoretical predictions cannot separate between the two scenarios as statistics in this high Q^2 region is still too poor.

7.5 Extraction of M_Z

The mass of the Z boson has been measured with very high precision by the LEP experiments over the last decade (see [90] and references therein). This was made possible by precisely scanning the Z resonance via the variation of the center-of-mass energy. It is obvious that this precision will never be reached by HERA experiments as the Z is not produced (in resonance) in the s channel (as at LEP), but exchanged in the t channel. On the other hand, the different production mechanism and the completely different initial state with strong interactions present constitutes a worthwhile test of the electroweak theory complementarily to the LEP experiments.

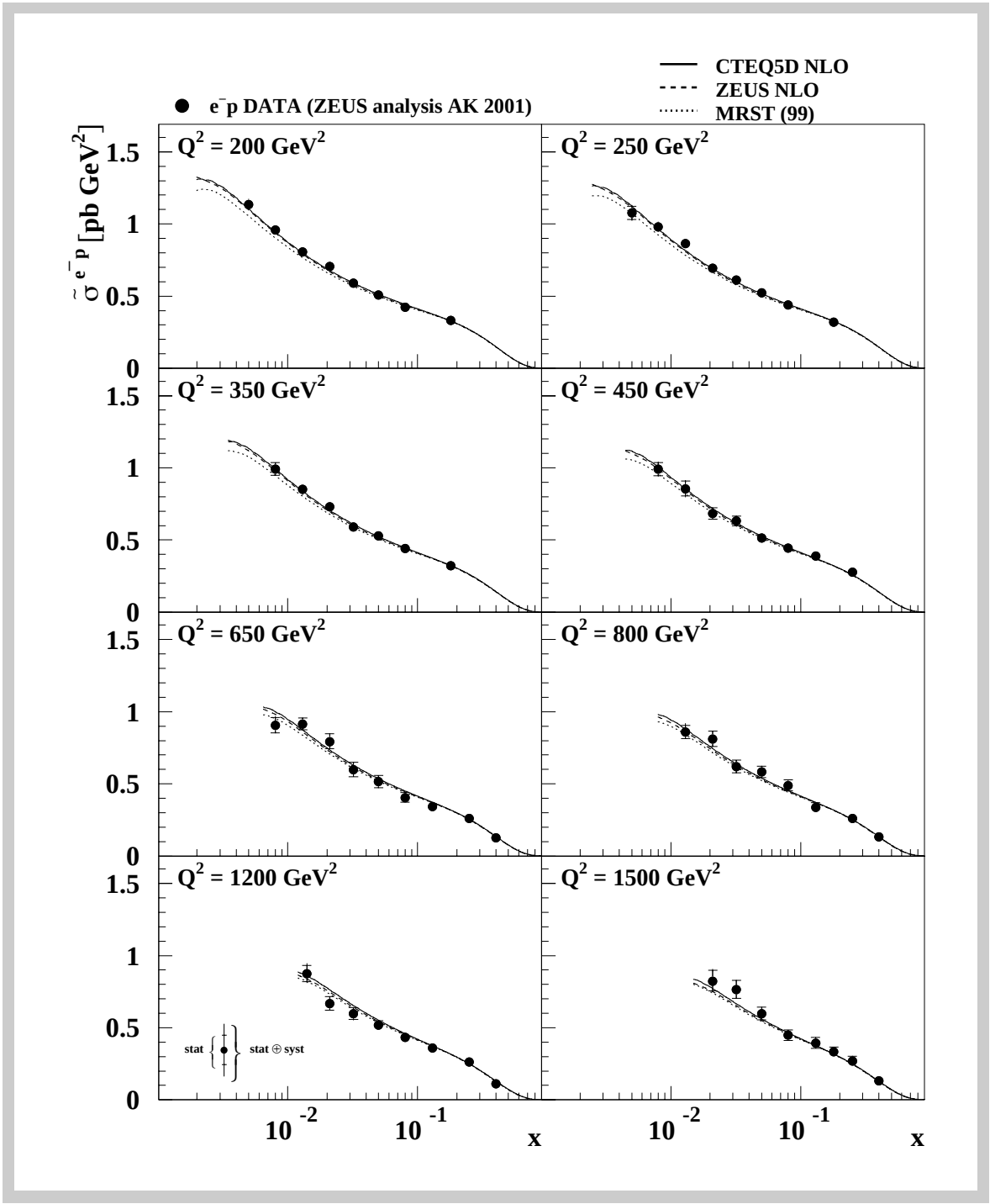


Figure 7.15: *Double-differential reduced cross sections $\tilde{\sigma}^{e^-p}$ as a function of x in bins of Q^2 between 200 GeV^2 and 1500 GeV^2. Theoretical predictions are drawn for the CTEQ5D (solid line), ZEUS NLO (dashed line) and MRST(99) (dotted line) PDF sets.*

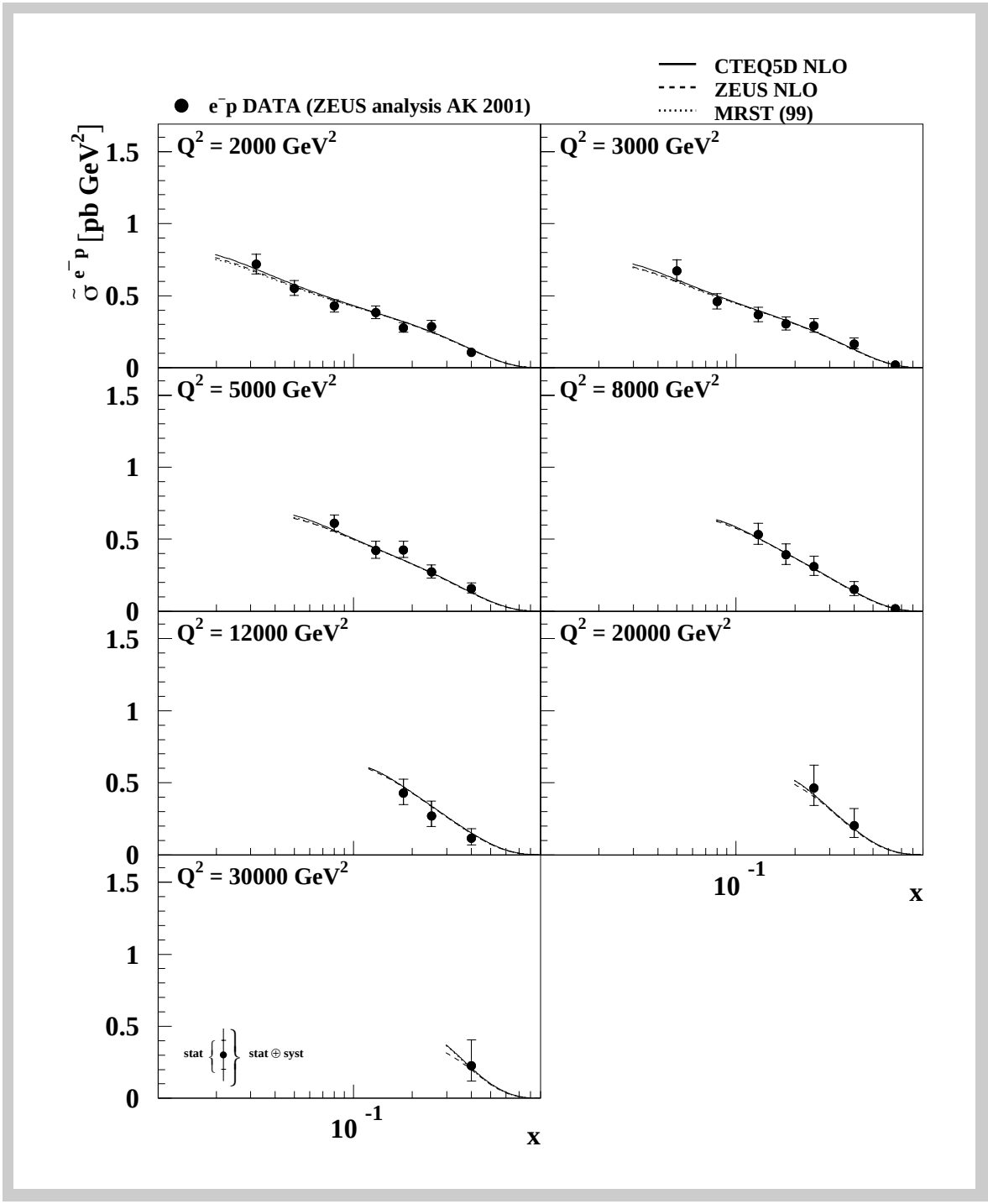


Figure 7.16: *Double-differential reduced cross sections $\tilde{\sigma}^{e^-p}$ as a function of x in bins of Q^2 between 2 000 GeV^2 and 30 000 GeV^2. Theoretical predictions are drawn for the CTEQ5D (solid line), ZEUS NLO (dashed line) and MRST(99) (dotted line) PDF sets.*

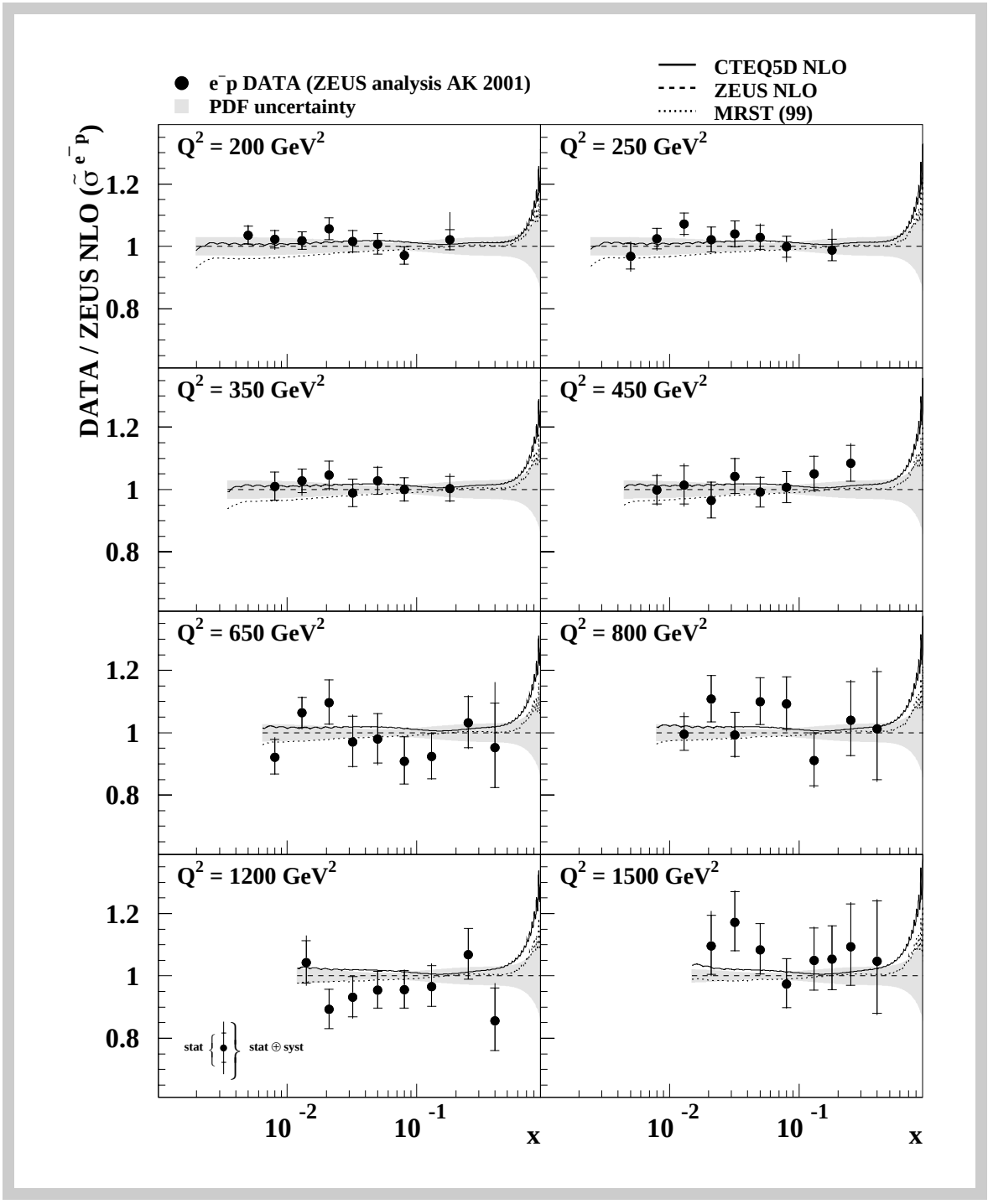


Figure 7.17: *Ratio between measured cross sections and calculations based on the ZEUS NLO fit as a function of x in different Q^2 bins in the range $200\ GeV^2 < Q^2 < 1500\ GeV^2$. Also drawn are the ratios of CTEQ5D and MRST(99) calculations to the ZEUS NLO fit. The shaded band indicates the error on the ZEUS NLO fit.*

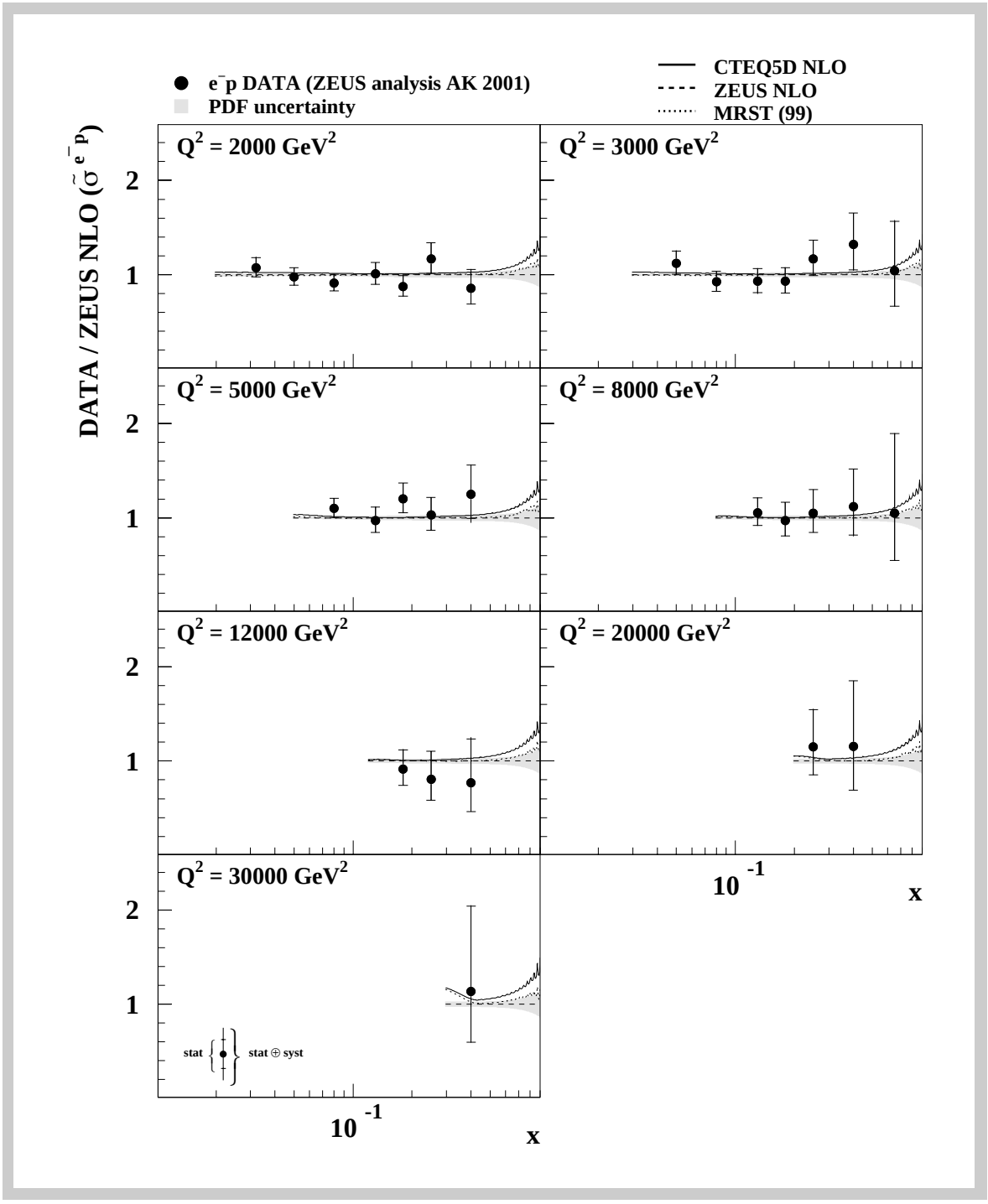


Figure 7.18: *Ratio between measured cross sections and calculations based on the ZEUS NLO fit as a function of x in different Q^2 bins in the range $2\,000\,GeV^2 < Q^2 < 30\,000\,GeV^2$. Also drawn are the ratios of CTEQ5D and MRST(99) calculations to the ZEUS NLO fit. The shaded band indicates the error on the ZEUS NLO fit.*

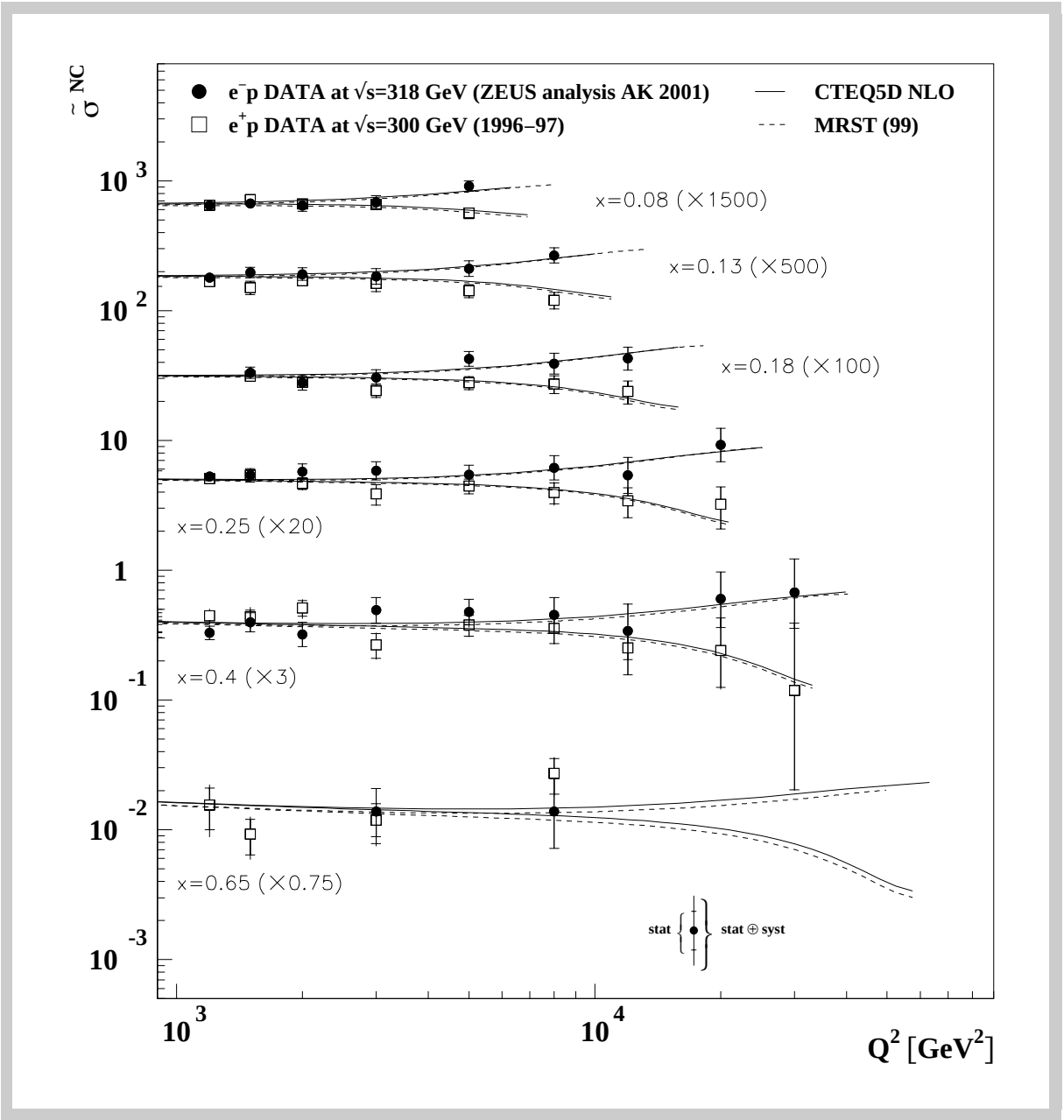


Figure 7.19: *Reduced cross sections $\tilde{\sigma}^{\mathrm{NC}}$ for e^-p and e^+p scattering as a function of Q^2 in six different bins of x. The measured values are compared to theoretical predictions using the CTEQ5D and MRST(99) PDF sets. In order to be able to plot all cross sections in a single plot they are multiplied by an x-dependent factor indicated in brackets behind the respective x value.*

The mass of the Z boson enters the neutral current ep cross section (2.19) in various places. For better reference the relevant equations from Chap. 2.3.1 are listed here again:

$$\frac{d^2\sigma^{\mathrm{NC}}(e^{\pm}p)}{dx\,dQ^2} = \frac{2\pi\alpha^2}{x\,Q^4}\left[Y_+F_2^{\mathrm{NC}}(x,Q^2) \mp Y_-xF_3^{\mathrm{NC}}(x,Q^2) - y^2F_L^{\mathrm{NC}}(x,Q^2)\right] \tag{7.5}$$

$$F_2(x,Q^2) = x\sum_{\mathrm{q=quarks}} A_f(Q^2)\left[q(x,Q^2)+\bar{q}(x,Q^2)\right] \tag{7.6}$$

$$xF_3(x,Q^2) = x\sum_{\mathrm{q=quarks}} B_f(Q^2)\left[q(x,Q^2)-\bar{q}(x,Q^2)\right] \tag{7.7}$$

$$A_f(Q^2) = q_f^2 - 2v_ev_fq_fP_z + (v_e^2+a_e^2)(v_f^2+a_f^2)P_z^2 \tag{7.8}$$

$$B_f(Q^2) = -2a_ea_fq_fP_z + 4v_ea_ev_fa_fP_z^2 \tag{7.9}$$

$$P_z = \frac{1}{4\sin^2\theta_W\cos^2\theta_W}\cdot\frac{Q^2}{Q^2+M_Z^2}\,. \tag{7.10}$$

The most obvious M_Z dependence occurs in the propagator term $\frac{Q^2}{Q^2+M_Z^2}$ in (7.10). However, M_Z is also contained in the vector coupling constants v_e and v_f in the prefactors A_f (7.8) and B_f (7.9). M_Z is connected through the electroweak theory to other electroweak parameters such as the Fermi constant G_F and the mass of the W boson M_W. At the Born level, i.e. lowest electroweak order, the relations can be written as

$$\cos\theta_W = \frac{M_W}{M_Z} \tag{7.11}$$

$$G_F = \frac{\pi\alpha}{\sqrt{2}\sin^2\theta_W\cdot M_W^2}\,, \tag{7.12}$$

θ_W being the weak mixing angle.

Moreover, the Z mass appears in various loop corrections such as those mentioned in Chap. 2.3.3. This leads to much more involved relations between the various electroweak parameters.

The contributions of the γ and Z exchanges can be separated into a purely electromagnetic part, a purely weak part and an interference part. Equation (7.8), i.e. the prefactor in F_2, contains all three parts, represented by terms without P_Z (electromagnetic part), with P_Z (interference part) and with P_Z^2 (weak part). On the other hand, (7.9) and hence the structure function xF_3 contains the latter two terms only. The contributions of these terms to the single-differential cross section $d\sigma/dQ^2$ are displayed in Fig. 7.20. Note that the cross sections of the purely weak and the interference parts are multiplied by 10. The pure γ term in lowest order[2] is the same for e^-p and e^+p and represents the largest contribution. The interference term yields the second largest contribution. It is positive for e^-p and negative for e^+p scattering[3].

In this thesis two methods to extract M_Z are investigated which use the e^-p data from this analysis as well as the e^+p data [10] from the 1999/2000 run period. The latter data set was

[2] Higher order photonic corrections involving box diagrams (photons couple to both the lepton and the scattered quark) and interference terms of leptonic and quarkonic radiation do depend on the charge of the lepton.

[3] This sign must not be confused with the sign in front of the xF_3 term in (7.5). Both the γZ and the Z term contain parity conserving and violating contributions which can be separated and then form F_2 and xF_3, respectively.

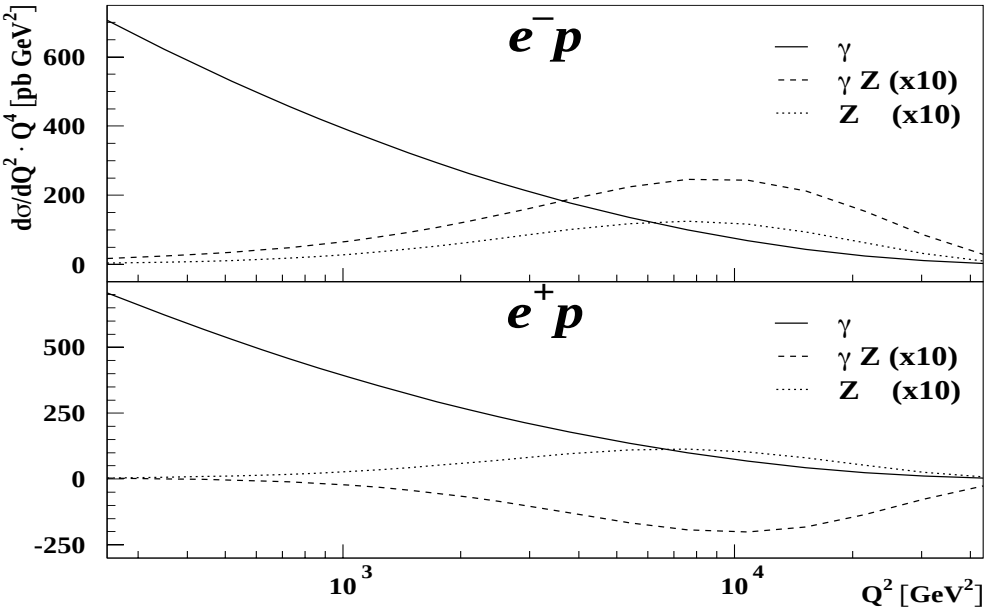


Figure 7.20: *Individual contributions of the γ, γZ and Z terms to the total single-differential $d\sigma/dQ^2$ cross section. All cross sections are multiplied by Q^4 where those for the γZ and Z terms are additionally multiplied by a factor 10.*

collected at the same center-of-mass energy as the e^-p data and corresponds to an integrated luminosity of $\sim 60\,\mathrm{pb}^{-1}$.

7.5.1 Fit without electroweak corrections

The idea of this method is to factorize out all parts of the cross section that contain an M_Z dependence and then fit the shape of this dependence by varying M_Z. This method can only be applied to the Born level cross sections as otherwise the M_Z dependence of the cross section becomes too complicated. The only kinematic variable that appears in (7.10) together with M_Z is Q^2. Hence, it seems sensible to fit M_Z in terms of $d\sigma/dQ^2$. However, (7.10) only holds for the lowest electroweak order. The measured cross sections are corrected to this order using a MC with electroweak corrections[4] and the unfolding method described in Chap. 7.1. The MC is based on HERACLES 4.6 [5] and utilizes the CTEQ5D PDF parameterization.

In the following, $\sigma^{\pm}$ is defined as the integral of (7.5) over x:

$$\sigma^{\pm} = \frac{d\sigma^{\mathrm{NC}}(e^{\pm}p)}{dQ^2} = \int dx\, \frac{d^2\sigma^{\mathrm{NC}}(e^{\pm}p)}{dx\, dQ^2}\ . \tag{7.13}$$

[4] Note that at the moment the MC does not contain any photonic box diagrams or interference terms of leptonic and quarkonic photon-radiation which depend on the charge of the lepton. However, the effect of this on the measured M_Z value is expected to be small with respect to the error on the mass.

[5] The following parameters have been used: $M_Z = 91.1880\,\mathrm{GeV}$, $\Gamma_Z = 2.351\,\mathrm{GeV}$, $M_W = 80.3738\,\mathrm{GeV}$, $\Gamma_W = 1,9710\,\mathrm{GeV}$, $\sin^2\theta_W = 0.2231$, $M_u = 0.062\,\mathrm{GeV}$, $M_d = 0.083\,\mathrm{GeV}$, $M_s = 0.215\,\mathrm{GeV}$, $M_c = 1.5\,\mathrm{GeV}$, $M_b = 4.5\,\mathrm{GeV}$, $M_t = 174.3\,\mathrm{GeV}$ and $M_h = 150.0\,\mathrm{GeV}$.

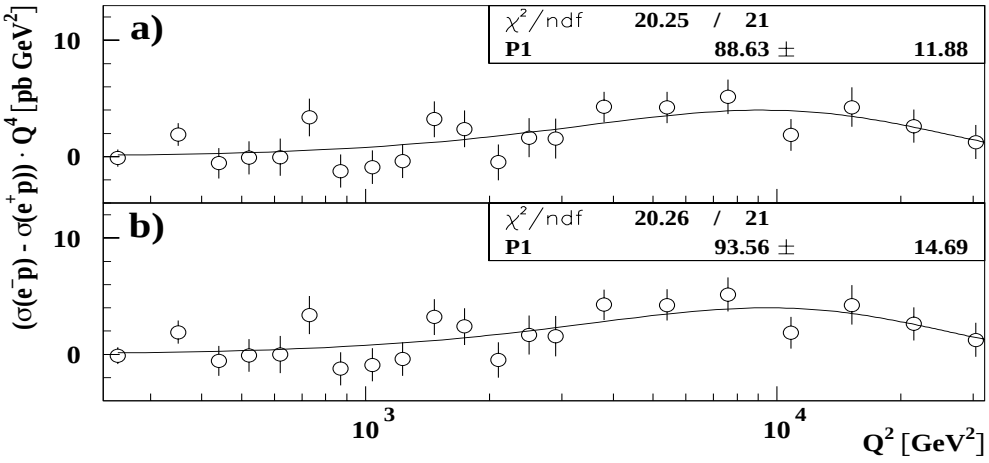


Figure 7.21: *Fit of the Z boson mass M_Z to $(\sigma^{e^-p} - \sigma^{e^+p}) \cdot Q^4$ as a function of Q^2 assuming different M_Z dependencies: (a) $Q^2/(Q^2 + M_Z^2) \cdot C_{\text{theo}}$ and (b) $M_Z^2 \cdot Q^2/(Q^2 + M_Z^2) \cdot C_{\text{theo}}$. C_{theo} is constant for all Q^2 values and is calculated from the theoretical $\sigma^{e^\pm p}$ cross sections (see text). The ordinate is scaled down by a factor* 100 000.

In order to get rid of the charge term q_f^2 in $A_f(Q^2)$ in (7.8) σ^+ is subtracted from σ^- which at the same time removes F_L. Afterwards, one is left with:

$$\sigma^- - \sigma^+ = \int dx \, \frac{4\pi\alpha^2}{x\,Q^4} \cdot Y_- x F_3(x, Q^2) \; . \tag{7.14}$$

The contribution of the P_Z^2 term in (7.9) to the cross section can be neglected as it only amounts roughly to 1/140 of the P_Z term due to the smallness of v_e[6]. Using (7.7)–(7.10) with (7.14) yields

$$\sigma^- - \sigma^+ = \frac{1}{\sin^2\theta_W \cos^2\theta_W} \cdot \frac{Q^2}{Q^2 + M_Z^2} \cdot \int dx \, (\dots) \; . \tag{7.15}$$

Here, $\int dx(\dots)$ represents the remaining integral determined by multiplying the theoretically calculated difference $\sigma^- - \sigma^+$ by the inverse of the two factors in front of the integral in (7.15) using the nominal values of M_Z and $\sin^2\theta_W$ ($M_Z = 91.1180\,\text{GeV}$ and $\sin^2\theta_W = 0.2231$).

Two possibilities exist for the determination of M_Z. One can either fit just the $Q^2/(Q^2 + M_Z^2)$ dependence leaving θ_W constant or one can also exploit the relations between G_F, M_W and M_Z given by (7.11). In both cases the right side of (7.15) is fitted to the difference of the measured cross sections σ^-_{meas} and σ^+_{meas}.

Figure 7.21 a shows $\sigma^-_{\text{meas}} - \sigma^+_{\text{meas}}$ multiplied by Q^4 as a function of Q^2. The factor Q^4 is applied to get rid of the steep fall-off of the cross section. The Z mass is fitted to the $Q^2/(Q^2 + M_Z^2)$ dependence, i.e. $\sin^2\theta_W = 0.2231$ is assumed to be constant. For Fig. 7.21 b $\sin^2\theta_W$ has been replaced with (7.12) ($G_F = 1.125 \cdot 10^{-5}\,\text{GeV}^{-2}$) yielding $M_Z^2 \cdot \frac{Q^2}{Q^2+M_Z^2}$ as the M_Z dependent part. The systematic uncertainties are determined by varying the e^-p and e^+p cross sections

[6] This Z term must not be confused with the one depicted in Fig. 7.20 which also contains contributions from (7.8).

		No.	e^+p	e^-p	ΔM_Z [GeV]
$\frac{Q^2}{(Q^2+M_Z^2)}$	no correlation	1	—	↑	−10.9
		2	—	↓	+11.4
		3	↑	—	+10.2
		4	↓	—	−9.9
	positive correlation	5	↑	↑	−1.1
		6	↓	↓	+1.2
	negative correlation	7	↑	↓	+22.3
		8	↓	↑	−20.7
$M_Z^2 \cdot \frac{Q^2}{Q^2+M_Z^2}$	no correlation	9	—	↑	+5.0
		10	—	↓	−5.2
		11	↑	—	−3.9
		12	↓	—	+3.6
	positive correlation	13	↑	↑	+1.3
		14	↓	↓	−1.4
	negative correlation	15	↑	↓	−9.2
		16	↓	↑	+8.5

Table 7.1: *Influence of the luminosity uncertainty of* 1.5% *on the extracted M_Z value for different levels of correlation between the e^-p and e^+p luminosities. "—" represents no variation and ↑ (↓) denotes a positive (negative) change of the cross sections by* 1.5%. *The column No. is added for better reference.*

simultaneously with the same individual systematic uncertainty[7]. The results are then added in quadrature in order to obtain the total systematic error. The uncertainty on the integrated luminosity of the data sets, i.e. their normalization, is not included and will be discussed below.

The two fits yield $M_{Z,1} = (88.6 \pm 11.9^{+5.4}_{-9.0})\,\text{GeV}$ and $M_{Z,2} = (93.6 \pm 14.7^{+6.5}_{-2.1})\,\text{GeV}$. Both fits have a very good χ^2/ndf of 0.96 and rather large errors, dominated by statistics. Obviously, accounting for the M_Z dependence of θ_W does not help to better constrain the fit.

The variation of the normalizations by their uncertainties of 1.5%[8] is an important systematic check as it influences all data points at the same time and in the same direction. Unfortunately, the exact correlation between the luminosity measurements of 1998/99 and 1999/2000 is unknown. Therefore, all possibilities listed in Table 7.1 are inspected. For better reference, the second column contains a running number. For the uncorrelated cases the two positive and

[7] Not all systematic checks applied for the determination of the e^-p cross sections are also available for e^+p. The missing checks are: variation of γ_{max} cut, variation of PHP fraction and rescaling of energies in the innermost ring of FCAL cells. However, the contribution of these checks to the total systematic uncertainty is small (Appendix C.1).

[8] For the 1999/2000 e^+p data the preliminary normalization uncertainty currently amounts to 2.25%. However, in order to get an estimate of the smallest possible normalization error on M_Z, for these studies an uncertainty of 1.5% is assumed, representing the best value achieved so far at ZEUS.

negative ΔM_Z values have to be added in quadrature in order to obtain the total normalization uncertainty, yielding $(+15.3, -14.7)$ for $M_{Z,1}$ and $(+6.2, -6.5)$ for $M_{Z,2}$.

Except for the positively correlated luminosities (Nos 5,6 and 13,14 in Table 7.1), all ΔM_Z values are at least of the same size as the total systematic uncertainty, or even much higher as in the case of the negative correlation. Obviously, this kind of fit is very sensitive to a global change of the normalization. Apart from this common feature, the first fit has the advantage that it arises only from the existence of a propagator term of the form $Q^2/(Q^2 + M_Z^2)$. It does not rely on detailed information about the relations between the different parameters, i.e. on electroweak theory, and therefore its sensitivity to model uncertainties is potentially smaller. On the other hand, the first fit completely neglects the dependence of θ_W on M_Z. The second fit builds on the validity of (7.11) and (7.12) and hence on the electroweak theory. As G_F is very precisely known from muon decay it is sensible to take G_F as a constant leaving M_W and M_Z as free but correlated parameters.

Both fits have in common that they completely neglect higher order electroweak corrections which also depend on the mass of the Z boson. A method that takes this kind of effects into account to a certain level is presented in the next section.

7.5.2 Fit with electroweak corrections

Considering higher order corrections leads to a modification of (7.5), where several schemes for the implementation exist. In this context, only the crude basics will be explained, whereas a much more detailed discussion can be found for example in [91, 92]. The scheme used in this analysis is called "on shell" scheme in which (7.11) holds exactly to all orders. However, this has the consequence that G_F has to be modified to

$$G_F = \frac{\pi\alpha}{\sqrt{2}s_W^2 M_W^2}\left[1 + \Delta r(\alpha, M_W, M_Z, M_H, m_t)\right] \ , \tag{7.16}$$

where s_W symbolizes the "fixed" $\sin\theta_W$ relation and Δr contains the loop corrections. Additionally, all vector and axial-vector couplings in (7.8) and (7.9) have to be modified in a non-trivial way (see for example [91]), rendering them x and Q^2 dependent. As a consequence the e^+p and e^-p cross sections have to be recalculated for each quoted point and each Z mass value, rendering the normal fit procedure with a continuous variation of the Z mass inappropriate. Therefore, the cross sections are calculated beforehand in a Z mass range between 77 GeV and 110 GeV with a step width of 0.1 GeV and fixed $G_F = 1.1664 \cdot 10^{-5}$ using the program eprc[9] [93] with the CTEQ5L[10] PDFs. The step size chosen is small enough considering that the errors are expected to be of the order of several GeV. Then, the discrepancy between measured (σ^{theo}) and predicted (σ^{data}) cross sections for different M_Z values is calculated:

$$\chi^2(M_Z) = \sum_i \left(\frac{\sigma_i^{\text{data}} - \sigma_i^{\text{theo}}(M_Z)}{\Delta_i^{\text{data}}} \right)^2 \ , \tag{7.17}$$

[9] This program has electroweak corrections included but does not take any photonic corrections into account. However, these corrections depend only weakly on M_Z and hence the effect on the measured M_Z is expected to be small compared to M_Z errors of several GeV.

[10] eprc calculates only to LO QCD.

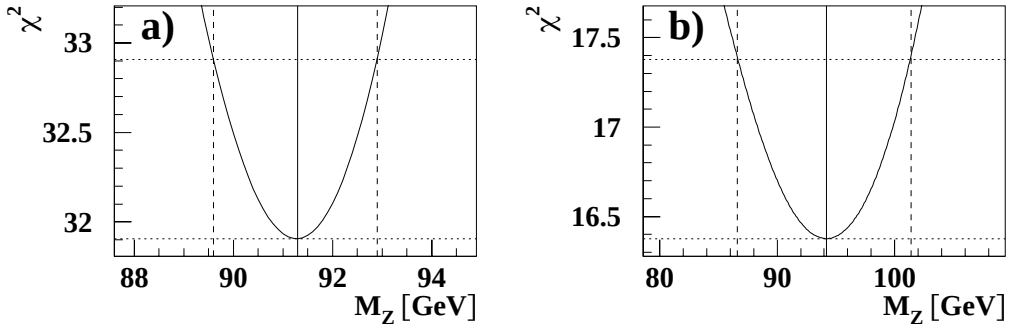


Figure 7.22: *χ^2 distribution for (a) sum of individual χ^2 of σ^{e^-p} and σ^{e^+p} and (b) for $\sigma^{e^-p} - \sigma^{e^+p}$. The solid vertical line indicates the central fit value, whereas the dashed lines represent the 1σ-statistical error. The horizontal lines mark the corresponding χ^2 values.*

where i runs over all measured points and Δ^{data} is the error on the measured cross sections. The M_Z value with the smallest χ^2 , χ^2_{min}, is taken to be the central value for M_Z and the upper and lower statistical errors are defined as the M_Z values where $\chi^2 = \chi^2_{\text{min}} + 1$. The systematic uncertainties are obtained by altering the cross sections and re-performing the χ^2 minimization, where the total systematic uncertainty is the quadratic sum of the individual contributions.

In general two ways exist to combine the e^-p and e^+p cross sections. Both methods will be investigated. For the first, χ^2 is calculated separately for both cross sections and added afterwards. The sum is then used to find the M_Z value with the lowest χ^2. In the second method the two cross sections are first combined into a function and then χ^2 is calculated for this function. The function taken here is the difference between the two cross sections, which supposingly yields a high sensitivity to xF_3 and hence to M_Z. The five lowest cross section points in Q^2 are discarded as the χ^2 contribution of the preliminary e^+p data is very high in this region.

The χ^2 minimization yields a Z mass of $M_{Z,3} = (91.3^{+1.6+2.7}_{-1.7-2.0})\,\text{GeV}$ for the first method and of $M_{Z,4} = (94.2^{+7.2+3.1}_{-7.6-1.0})\,\text{GeV}$ for the second. Figure 7.22 displays the corresponding χ^2 curves. The two curves have a very good minimum χ^2/ndf of 1.03 and 1.09, respectively. The two mean values are compatible with each other within errors. Surprisingly, $M_{Z,4}$ has a much higher statistical error than $M_{Z,3}$, whereas the systematic uncertainties are roughly the same (the uncertainty on the luminosity is not yet included at this stage).

The basic difference between the two methods is that the first exploits the whole Q^2 range, whereas the second, like the methods utilized for the determination of $M_{Z,1}$ and $M_{Z,2}$, is limited to the region where xF_3 is large, i.e. the high Q^2 regime. At $Q^2 \approx M_Z^2$, both the shape and the normalization of the distribution are sensitive to M_Z, whereas for $Q^2 \lesssim M_Z^2$ only the normalization varies with M_Z. The upshot is that in the first method M_Z is mainly determined from the normalization, which however is known to only 1.5%. Therefore, the variation of the normalization is an important systematic check (Table 7.2). For the uncorrelated cases the two positive and negative ΔM_Z values are added in quadrature in order to obtain the total luminosity uncertainty, yielding $(+2.5, -2.8)\,\text{GeV}$ for $\chi^2(\sigma^{e^-p}) + \chi^2(\sigma^{e^+p})$ and $(+3.1, -3.2)\,\text{GeV}$ for $\chi^2(\sigma^{e^-p} - \sigma^{e^+p})$.

		No.	e^+p	e^-p	ΔM_Z [GeV]
$\chi^2(\sigma^{e^-p}) + \chi^2(\sigma^{e^+p})$	no correlation	1	—	↑	+0.7
		2	—	↓	−0.9
		3	↑	—	+2.4
		4	↓	—	−2.7
	positive correlation	5	↑	↑	+3.2
		6	↓	↓	−3.6
	negative correlation	7	↑	↓	+1.6
		8	↓	↑	−1.9
$\chi^2(\sigma^{e^-p} - \sigma^{e^+p})$	no correlation	9	—	↑	+2.5
		10	—	↓	−2.6
		11	↑	—	−1.9
		12	↓	—	+1.9
	positive correlation	13	↑	↑	+0.7
		14	↓	↓	−0.7
	negative correlation	15	↑	↓	−4.6
		16	↓	↑	+4.3

Table 7.2: *Influence of the luminosity uncertainty of* 1.5% *on the extracted M_Z value for different levels of correlation between the e^-p and e^+p luminosity. "—" represents no variation and ↑ (↓) denotes a positive (negative) change of the cross sections by* 1.5%. *The column No. is added for better reference.*

As expected the effect on M_Z depends strongly on the level of correlation between the two luminosity errors. Both the smallest and largest ΔM_Z are achieved for $\chi^2(\sigma^{e^-p} - \sigma^{e^+p})$ with positively and negatively correlated luminosity errors. The much smaller errors of checks Nos. 1 and 2 compared to Nos. 3 and 4 are explained by the four times higher luminosity of the e^+p data which dominates the determination of M_Z if χ^2 is calculated for both cross sections individually.

Though the exact correlation between the two luminosity measurements is unknown the studies show that the uncertainty is in the range of 2–3 GeV for $\chi^2(\sigma^{e^-p}) + \chi^2(\sigma^{e^+p})$ and of 1–4 GeV for $\chi^2(\sigma^{e^-p} - \sigma^{e^+p})$. Consequently, the error on $M_{Z,3}$ is dominated by systematics, whereas that for $M_{Z,4}$ is dominated by statistics.

7.5.3 Conclusions

The mass of the Z boson has been determined in four different ways. The resulting M_Z values are summarized in Table 7.3. All four values are in agreement with the world average of $M_Z = (91.882 \pm 0.0022)$ GeV [94]. The overall smallest error is obtained by minimizing the sum of the χ^2 values for e^+p and e^-p ($M_{Z,3}$).

Apart from the very high sensitivity on the normalization, the method employed for the first two M_Z values completely neglects the dependence of higher order corrections on M_Z, whereas

central value	errors on M_Z [GeV]		
	statistical	systematic	normalization
$M_{Z,1} = 88.6$	± 11.9	$^{+5.2}_{-9.0}$	1–20
$M_{Z,2} = 93.6$	± 14.7	$^{+6.5}_{-2.1}$	2–9
$\boldsymbol{M_{Z,3} = 91.3}$	$\boldsymbol{^{+1.6}_{-1.7}}$	$\boldsymbol{^{+2.7}_{-2.0}}$	**2–3**
$M_{Z,4} = 94.2$	$^{+7.2}_{-7.6}$	$^{+3.1}_{-1.0}$	1–4

Table 7.3: *The four determined M_Z values with their statistical, systematic and normalization errors. For the normalization error only an approximate range is given as the exact correlation between the luminosity measurements of the e^+p and e^-p cross sections is unknown. The result with the overall smallest error ($M_{Z,3}$) is printed in boldface.*

for $M_{Z,3}$ and $M_{Z,4}$ these are taken into account to a certain level. However, compared to the errors on the measured M_Z values, the effect of ignoring higher electroweak corrections on $M_{Z,1}$ and $M_{Z,2}$ is small. The best result is obtained by minimizing the sum of χ^2 for e^+p and e^-p ($M_{Z,3}$), where this method profits with respect to the statistical errors both from combining the integrated luminosities of the two data sets (obviously, one does not gain much exploiting the difference in the cross sections of e^-p and e^+p scattering due to the alternating sign of xF_3) and from the smaller errors on the measured cross sections for $Q^2 < M_Z^2$.

For $M_{Z,3}$ the systematic and normalization uncertainties are already larger than the statistical errors, and hence higher statistics will not reduce the total error considerably. Especially the normalization error, i.e. the error on the luminosity measurement, which by itself has roughly the same size as all other systematic errors together, has to be reduced. During recent years the measurement of NC $d\sigma/dQ^2$ cross sections in the region $65\,\mathrm{GeV}^2 < Q^2 < 140\,\mathrm{GeV}^2$ yielded very precise normalization values. Hence, in the future these can be used to determine the relative normalization between data sets of different run periods to better than 1% [95] providing an independent cross check of the luminosity measurements. The insights gained can then be used to reduce the uncertainty on the luminosity measurements.

Another possibility to determine electroweak parameters is the combination of NC and CC cross sections [91] which is beyond the scope of this thesis.

7.6 Extraction of xF_3

Former measurements of the structure function xF_3 have been performed by the CCFR [96] and IHEP-JINR [97] collaborations in neutrino-proton scattering. However, these measurements were limited in Q^2 to a few $10\,\mathrm{GeV}^2$. Together with similar results from H1 [98] the measurements presented in this thesis are the first extractions of xF_3 at high Q^2 in lepton-proton scattering.

The parity violating nature of the weak force leads, depending on the charge of the incoming lepton, to different signs in parts of the interference and the Z-only contributions to the NC DIS cross section in (2.19). The size of these parity-violating contributions is predicted by the SM electroweak theory and hence the latter can be tested by extracting the structure function xF_3 from the data. In principle, this could already be performed with either e^+p or e^-p data. However, in this case one would have to calculate the F_2 contribution[11] to the cross section from theory, where both electromagnetic and weak forces are involved as can be seen from (2.22), or, to put it in a different way, the electroweak theory would be used to test itself. In order to avoid this, both e^-p and e^+p data have to be available[12] as then the double-differential cross sections can be subtracted leading to:

$$xF_3^{\mathrm{NC}} = \frac{xQ^4}{2\pi\alpha^2} \cdot \frac{1}{2Y_-} \cdot \left(\sigma^- - \sigma^+\right) \tag{7.18}$$

with

$$\sigma^{\pm} = \frac{d^2\sigma^{\mathrm{NC}}(e^{\pm}p)}{dx\, dQ^2} \tag{7.19}$$

from (2.19). Note that this formula only holds if both data sets were recorded at the same center-of-mass energy, whereas for different ones (300 GeV and 318 GeV in the case of HERA) (7.18) transforms into:

$$xF_3 = \frac{xQ^4}{2\pi\alpha^2} \cdot \left(\frac{Y_-^{300}}{Y_+^{300}} + \frac{Y_-^{318}}{Y_+^{318}}\right)^{-1} \cdot \left(\frac{1}{Y_+^{318}} \cdot \sigma^- - \frac{1}{Y_+^{300}} \cdot \sigma^+\right) + \Delta F_L \ , \tag{7.20}$$

where $Y_{\pm}^{300,318}$ denote $Y_{\pm}$ calculated for center-of-mass energies of 300 GeV and 318 GeV, respectively. The term ΔF_L is a remnant of the difference between the $y^2 F_L^{\mathrm{NC}}$ terms for e^-p and e^+p in (2.19) which do not exactly cancel due to the y^2 factor and the different center-of-mass energies. The contribution of ΔF_L (calculated from QCD) to xF_3 in (7.20) is of the order of 7% at $Q^2 = 1500\,\mathrm{GeV}^2$ and $x = 0.04$ and decreases rapidly with increasing Q^2 and x. For $x \gtrsim 0.1$ it is always below 1%.

In addition to the e^-p cross sections measured in this thesis, those from the published e^+p data set [7] taken during the 1996/97 run period at a center-off-mass energy of 300 GeV are used to extract xF_3[13]. The latter are based on a data set corresponding to an integrated luminosity of $30\,\mathrm{pb}^{-1}$.

7.6.1 Binning

The double-differential binning (Fig. 7.11) cannot be adopted unmodified for the extraction of xF_3 as statistics is very low in the high Q^2 region. This in combination with the subtraction

[11] The longitudinal structure function F_L is neglected here due to its small contribution at high Q^2.

[12] This not only holds for the extraction of xF_3 but also for that of F_2 in the high Q^2 regime $\gtrsim M_Z^2$. Here, the xF_3 contribution to the cross section is no longer negligible and the extraction of F_2 relies heavily on the correct prediction of the size of xF_3. This difficulty can be solved by adding the measured double-differential cross sections for e^-p and e^+p. However, as the statistical errors are very small, systematic uncertainties play a significant role and have to be investigated very carefully, which is a major task when combining two different data sets.

[13] The 1999/2000 e^+p cross sections, recorded at $\sqrt{s} = 318\,\mathrm{GeV}$, have not been published yet.

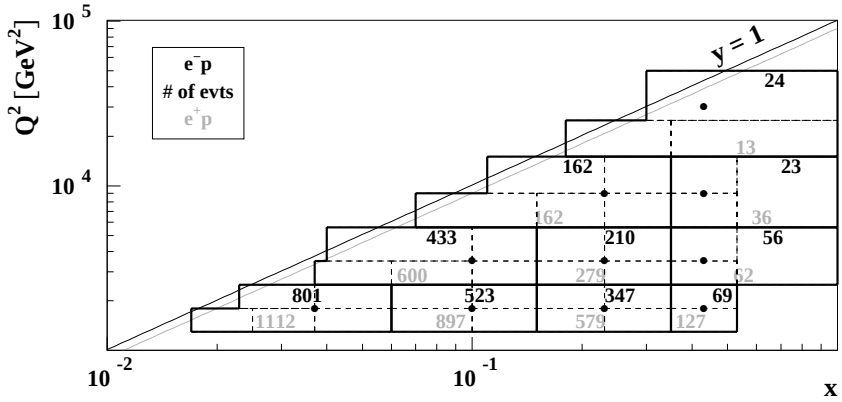


Figure 7.23: *xF_3 binning in the x-Q^2 plane. Bins with a thick frame are xF_3 bins, whereas the original double-differential bins are marked with dashed lines. The numbers in the xF_3 bins denote the number of events for e^-p and e^+p in that bin. The full circles indicate the quoted points for the measurement.*

of the two cross sections leads to huge statistical errors. Therefore, several bins are merged yielding the binning displayed in Fig. 7.23. The "old" double-differential binning is also shown for reference. The points in the kinematic plane where the xF_3 will be quoted are marked with a full circle.

7.6.2 Systematic checks

Systematic investigations are of major importance when combining two data sets as the experimental conditions like detector setup or beam conditions can be quite different over the years. In the special case of subtracting two cross sections the uncertainty of the integrated luminosity plays also a significant role. The major problem in the present case is that not all systematic checks are made in the same way and that some checks are completely missing for one of the two measured cross sections. Thus, for these calculations the systematic uncertainties for e^-p and e^+p are assumed to be uncorrelated. Later it will become clear that the systematic uncertainties are much smaller than the statistical errors and hence small inaccuracies in the treatment of the systematic uncertainties do not affect much the total error of the measured xF_3 values.

For e^-p the systematic uncertainties incorporated into the extraction of xF_3 have been discussed in Chaps. 6.2 – 6.3 and Chap. 7.4.3, whereas for e^+p a detailed description can be found in [9]. The error on the measured structure-function values due to the 1.5% uncertainty on the normalization of both data sets is not contained in the total systematic error but is calculated separately. As the correlation between the normalization errors is unknown for now a worst-case scenario is assumed, i.e. that the normalization uncertainties are negatively correlated.

7.6.3 xF_3 results

Using (7.20) with the measured e^-p ($16\,\mathrm{pb}^{-1}$) and e^+p ($30\,\mathrm{pb}^{-1}$) cross sections yields the xF_3 values depicted in Figs. 7.24 a,b (the corresponding table can be found in Appendix D.5). Note

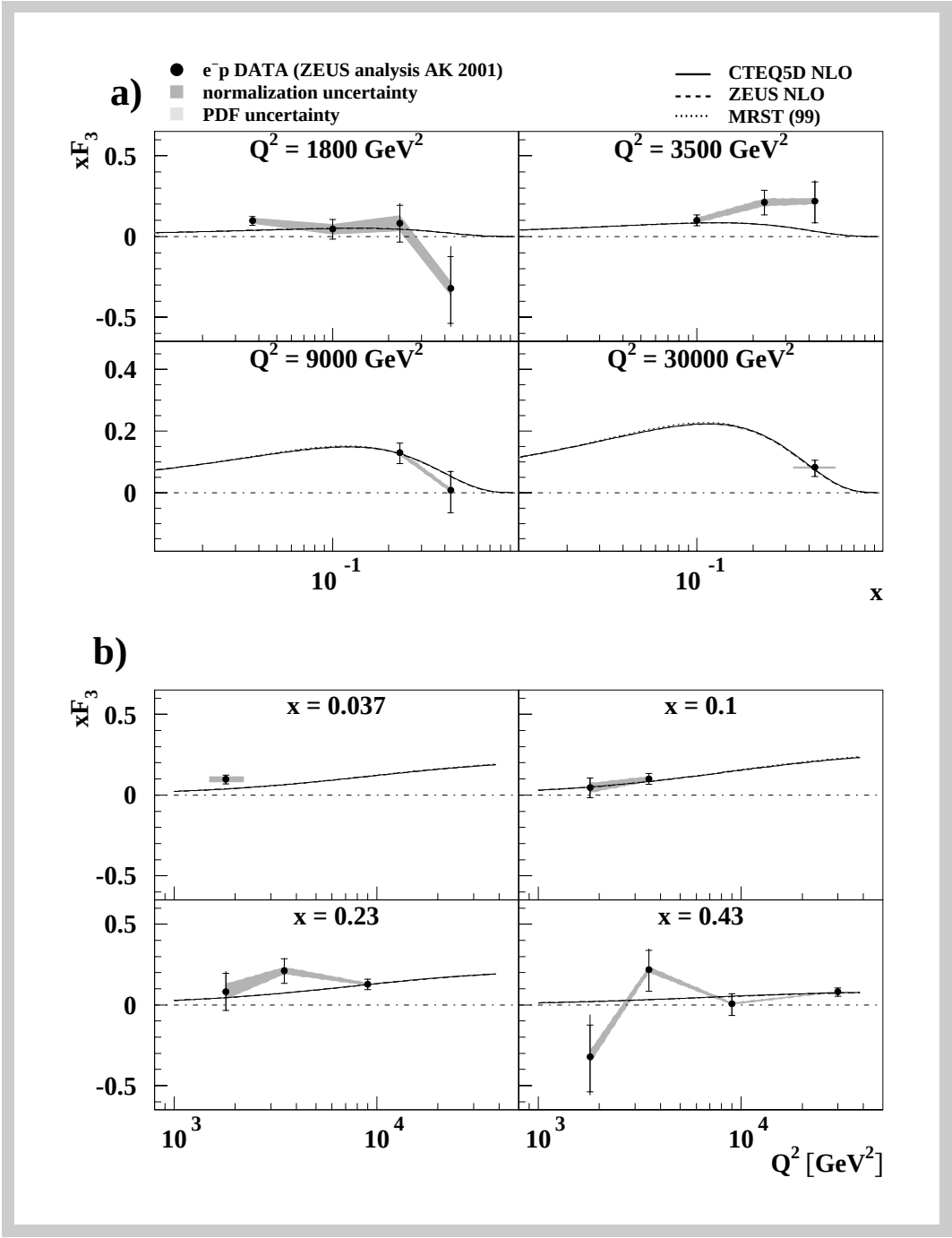


Figure 7.24: *Extracted structure function xF_3 as (a) a function of x in bins of Q^2 and (b) a function of Q^2 in bins of x. In addition calculations based on CTEQ5D, ZEUS NLO and MRST(99) PDFs are shown together with a light shaded band (hardly visible) representing the uncertainty on the ZEUS NLO calculations. The dark shaded band indicates the uncertainty on the measured structure functions due to the limited precision of the normalization of both data sets of* 1.5%*, where a worst-case scenario, i.e. negatively correlated errors, is assumed.*

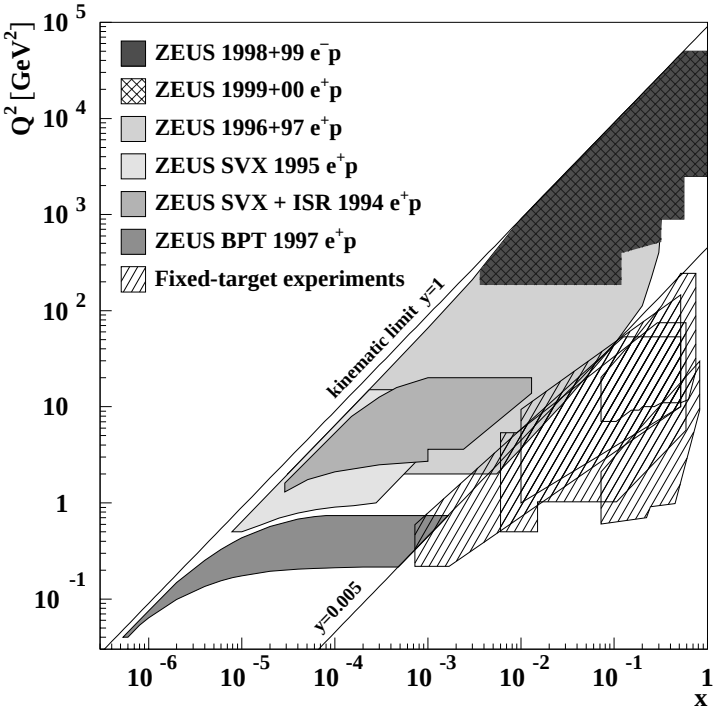


Figure 7.25. *Kinematic x-Q^2 plane with areas covered by ZEUS analyses and fixed target-experiments. The area labeled "ZEUS 1998+99 e^-p" is the subject of this analysis.*

that the contribution of ΔF_L to xF_3 has been neglected in these plots due to its smallness compared to the errors. In addition, theoretical calculations based on CTEQ5D, ZEUS NLO and MRST(99) PDF parameterizations are shown together with the uncertainty on the ZEUS NLO fit as a light shaded band (hardly visible). The dark shaded band marks the normalization uncertainty. In Figs. 7.24 b the rise of xF_3 with increasing Q^2, already observed in the difference between the e^-p and e^+p cross sections (Figs. 7.9 and 7.19) for $Q^2 \gtrsim M_Z^2$, is apparent. It should be noted that the fall-off of xF_3 towards low x is not caused by the actual structure function F_3 but solely by the x factor.

A comparison between the calculations and the measured values yields good agreement. The largest deviation for a single point occurs at $Q^2 = 1800\,\mathrm{GeV}^2$ and $x = 0.037$ and amounts to 2.2σ. The errors are too large to allow discriminating the different PDFs. The error on the measured structure function points is completely dominated by the statistical errors, even if the worst-case normalization uncertainty is considered.

The e^-p data set with its much lower luminosity contributes by far the most to the statistical error. Consequently, gathering more e^+p data does not further reduce the error on xF_3 but high-luminosity data sets of similar size are needed for improved precision.

7.7 Summary of Chapter 7

In this chapter the results of the analysis have been presented. Both single- and double-differential cross sections have been measured in the region $185\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $3.7 \cdot 10^{-3} < x < 1$, displayed in Fig. 7.25 as dark area together with other ZEUS measurements and fixed-target results.

To obtain a reasonable estimate of the systematic uncertainty on the cross section, 18 different systematic checks were carried out. The uncertainty on the measured cross section is almost everywhere dominated by the statistical errors, only at high x for $d\sigma/dx$ ($x > 0.2$, $Q^2 > 200\,\mathrm{GeV}^2$) and in the highest y bins for $d\sigma/dy$ the systematics are equal to or larger than the statistics.

The measured cross sections have been compared to theoretical calculations using the PDF sets CTEQ5D, ZEUS NLO and MRST(99). Also the error band from the ZEUS NLO parameterization has been shown. Within this error band CTEQ5D and ZEUS NLO yield very similar results in the kinematic range covered by the analysis. Only for $Q^2 \geq 1500\,\mathrm{GeV}^2$ CTEQ5D has the tendency to lie above the ZEUS NLO parameterization in the lowest x bins of the double-differential cross section. On the other hand, the MRST(99) parameterization falls below the ZEUS NLO error band for $x \lesssim 2 \cdot 10^{-2}$. The largest difference between CTEQ5D/MRST(99) and ZEUS NLO is observed at high x where also the error band becomes very broad due to the lack of precise measurements in this region.

Overall good agreement between measured and calculated cross sections using CTEQ5D and ZEUS NLO is observed. An excess of data events at high Q^2 is not visible. For $x \lesssim 2 \cdot 10^{-2}$ the data disfavors the MRST(99) parameterization which is too low in this region. With the kinematic range covered by the analysis and the available statistics it is not possible to constrain the PDFs for $x > 0.6$ where their uncertainties are still very large.

Afterwards, the measured e^-p cross sections have been compared to those from e^+p scattering. For $Q^2 \gtrsim M_Z^2$ the positive (negative) influence of the parity violating structure function xF_3 on the e^-p (e^+p) cross sections is clearly visible, requiring the presence of the weak force with its nominal strength.

The sensitivity of the NC cross section to the strength of the weak force has then been used to measure the mass of the Z boson with different methods. The result with the smallest errors yields $M_Z = (91.3^{+1.6}_{-1.7}[\mathrm{stat.}] \oplus {}^{+2.7}_{-2.0}[\mathrm{sys.}] \oplus 2\text{–}3[\mathrm{norm.}])\,\mathrm{GeV}$.

Finally, the e^-p and e^+p cross sections have been utilized to extract the structure function xF_3 for the first time in NC lepton-proton scattering at high Q^2. The kinematic region covered is $1300\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $0.017 < x < 1.0$. The extracted xF_3 values are in good agreement with theoretical predictions, however the statistical errors are very large due to the small e^-p data set of $16\,\mathrm{pb}^{-1}$ and do not allow to discriminate between different PDF parameterizations.

Chapter 8

Upgrade and outlook

After nearly 10 years of successful running, HERA was shutdown in summer 2000 for one year to execute a substantial upgrade program which also includes several modifications and enhancements of the collider experiments. The main aims of the upgrade are a significant enlargement of the delivered luminosity and the ability to collide polarized lepton beams with the protons. In the imminent HERA II phase it is planned to gather an integrated luminosity of $1\,\mathrm{fb}^{-1}$ until 2006.

The upgrades enable the experiments to investigate the high Q^2 region with much higher precision and will therefore permit to test the predictions of the SM also in regions where today statistical limitations dominate, e.g. in the electroweak sector with $Q^2 \gtrsim M_Z^2$. It will also lead to better constraints of the PDFs at very high $x > 0.6$.

On the other hand, the possibility to scatter polarized leptons off protons will open the window to a completely new field of HERA physics. The neutral current cross section may serve as an example. In the unpolarized case, the Z contribution either increases ($e^- p$) or decreases ($e^+ p$)

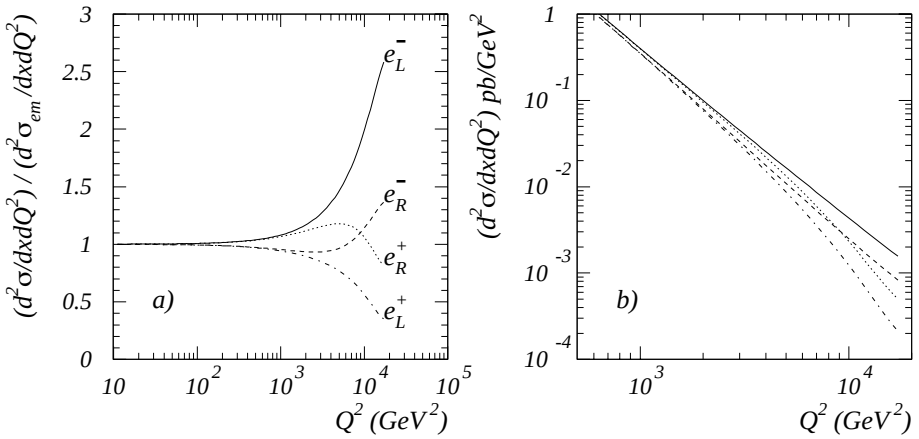


Figure 8.1: *Sensitivity to weak contributions of double-differential NC cross sections as a function of Q^2 at $x = 0.2$. (a) Ratio of Born level cross sections including weak terms with purely electromagnetic contributions. (b) Absolute values of Born level NC cross sections. (taken from [99])*

the cross section continously with increasing Q^2. However, in the polarized case the situation is much more complex as can be seen in Fig. 8.1. Here, only the cross section for the left-handed electron e_L^- and positron e_L^+ behave similarly to the unpolarized case, whereas that of the right-handed electron e_R^- for example first decreases and only becomes larger than the purely electromagnetic cross section for $Q^2 \gtrsim 10\,000\,\mathrm{GeV}^2$.

Apart from HERA also the ZEUS detector will change. Several detector components like the BPC/BPT or the LPS had to be removed to make space available for new HERA magnets. They are needed to focus the beam more strongly in order to achieve the target luminosity. Additionally, new components that will enhance the detector performance considerably are installed. The installation of the silicon micro vertex-detector (MVD) will among other things help to improve the neutral-current analysis as it permits to measure tracks in the very forward direction down to polar angles of 8°, which will facilitate the electron identification and background suppression in this difficult region.

Hence, even after 10 years of HERA I running, there will be many interesting things to study in the upcoming years and maybe some surprises await to be revealed.

Chapter 9

Summary

In this thesis the e^-p data set recorded with the ZEUS detector in 1998/99 has been used to measure the single- and double-differential neutral current (NC) cross sections in deep inelastic ep scattering (DIS) in the range $185\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $3.7 \cdot 10^{-3} < x < 1$. The results have been compared to theoretical calculations using different PDF sets and to cross sections from e^+p scattering. Then, the e^-p and e^+p cross sections have been used to determine the mass of the Z boson and to extract the parity violating structure function xF_3.

In order to reconstruct the kinematic variables of an event as accurately as possible various corrections to the energy and position (polar angle) of both the scattered electron and the hadronic final state have been applied. In this context detailed studies of the R/BCAL alignment with respect to the CTD have been carried out in the framework of this thesis. These studies show that the z and ϕ (x, y and ϕ) positions of the BCAL (RCAL halves) in the real detector are known to a very good precision of 0.5 mm and 1 mrad, respectively. For the reconstruction of the kinematic variables several algorithms using different combinations of information from the scattered electron and the hadronic final state have been tested and their performance has been compared and quantified.

As the extraction of the cross sections relies extensively on MC simulation an accurate description of both the physics event and the detector is essential. The agreement between data and MC has been carefully verified. Apparent differences have been investigated and solved whenever possible. For example, for $Q^2 \lesssim 1\,000\,\mathrm{GeV}^2$ the data clearly shows the presence of events with a gap in the energy flow between the hadronic final state and the proton remnant (diffractive events) that are not simulated in the standard NC DIS MC. The fraction of these diffractive events has been determined in each of the double-differential bins and such events have been properly admixed to the standard DIS MC. Also, the resolution of the electron-energy distribution in MC proved to be considerably narrower than in data. This has been corrected for by smearing the energy in MC by 3% (3.5%) in RCAL (BCAL).

Due to the complexity of today's high energy experiments and the extensive use of MC simulation, systematic studies are at premium. In this context, an accurate simulation of the early stages of the event selection is essential as all following steps rely on them. Therefore, detailed studies of the First-Level-Trigger have been performed yielding good agreement between reality

and simulation. Furthermore, the combination of different checks to obtain a reasonable estimate of the total systematic uncertainty on the measured cross sections is highly non-trivial due to the mostly unknown correlation between different systematics. Therefore, in this analysis the commonly used cut-variation method is disfavored and replaced by more fundamental checks wherever possible. Altogether, 18 different checks have been applied to obtain the total systematic uncertainty on the measured cross sections.

The measurements of the single-differential cross sections have been carried out for $d\sigma/dQ^2$ as a function of Q^2 ($200\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$), $d\sigma/dx$ as a function of x ($6.3 \cdot 10^{-3} < x < 6.3 \cdot 10^{-1}$) and $d\sigma/dy$ as a function of y ($0.05 < y < 1.00$), where for the latter two cross sections three different lower Q^2 cuts of 200, 2 500 and 10 000 GeV^2 have been applied. Double-differential cross sections have been measured in a range $185\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $3.7 \cdot 10^{-3} < x < 1.0$. The results obtained have been compared to theoretical calculations using the CTEQ5D, ZEUS NLO and MRST(99) PDF parameterizations. Overall, good agreement between the measured cross sections and the CTEQ5D and ZEUS NLO PDFs is observed. No excess of data events at high Q^2 is visible. For $x \lesssim 2 \cdot 10^{-2}$ the measured cross sections disfavor the MRST(99) parameterization, that lies considerably below the data in this kinematic region. A comparison of the measured e^-p cross sections to those obtained from e^+p scattering for $Q^2 \gtrsim M_Z^2$ clearly shows the positive and negative influence of the parity violating structure function xF_3, respectively, requiring the presence of the weak force with its nominal strength.

The sensitivity of the NC cross sections to the strength of the weak force has been exploited to measure the mass of the Z boson. For this two different methods have been investigated. The first one neglects electroweak corrections and extracts M_Z from the difference between the $d\sigma/dQ^2$ cross sections of e^-p and e^+p scattering, exploiting the fact that in the NC DIS cross section formula the sign of the xF_3 term depends on the charge of the lepton. The second method takes electroweak corrections into account and minimizes a χ^2 from a combination of e^-p and e^+p cross sections. All fits show a large sensitivity to the absolute normalization of the cross sections, which is only known to a precision of 1.5%. Altogether four M_Z values have been obtained, where the one with the smallest total error amounts to $M_Z = (91.3^{+1.6}_{-1.7}[\mathrm{sta.}] \oplus {}^{+2.7}_{-2.0}[\mathrm{sys.}] \oplus 2\text{–}3[\mathrm{norm.}])\,\mathrm{GeV}$, compatible with the world average value of $M_Z = (91.882 \pm 0.0022)\,\mathrm{GeV}$. The systematic and normalization uncertainties are larger than the statistical errors, and hence higher statistics will not reduce the size of the errors considerably. Especially the normalization error which by itself is of roughly the same size as all other systematic errors together has to be reduced in the future.

The different sign of the xF_3 term for e^-p and e^+p scattering has been used to extract the parity violating structure function xF_3 for the first time in lepton-proton scattering at high Q^2. The covered kinematic region extends over $1300\,\mathrm{GeV}^2 < Q^2 < 50\,000\,\mathrm{GeV}^2$ and $0.017 < x < 1.0$. The extracted xF_3 values are in agreement with theoretical predictions. However, the statistical errors are very large due to the small e^-p data set of $16\,\mathrm{pb}^{-1}$ (compared to $30\,\mathrm{pb}^{-1}$ for e^+p data) and hence no differentiation between the predictions is possible yet. For an optimal exploitation of the data with respect to the extraction of xF_3 in the future e^-p and e^+p data sets of similar size are mandatory.

Appendix A

Calorimeter alignment (R/BCAL)

Accurate reconstruction of the kinematic variables relies on accurate alignment of the three calorimeter parts (F/B/RCAL). In order to achieve the latter alignment measurements are performed by exploiting the fact that the measured energy of a particle (electron) deviates characteristically from its predicted "true" energy if the particle hits a crack between the calorimeter modules. This provides a method to obtain position information without using the position reconstruction algorithm of the calorimeter which itself could be biased. Unfortunately, this method is only applicable to the BCAL as the crack structure is "suppressed" if particles hit the cracks at an angle (R/FCAL). Hence, for the RCAL alignment a different method has to be developed. Furthermore, for this study the alignment of the FCAL has to be completely omitted as electron statistics is much too low in the forward detector region.

By ZEUS convention the CTD serves as the reference system and therefore the extrapolated track of the reconstructed electron is used as the reference position. The extrapolation is performed using a so called swim-algorithm that accounts for the deflection of the electron in the magnetic field. For the RCAL (BCAL) the track is extrapolated to the x-y (ϕ-z) plane used by the electron-position reconstruction-algorithm in the calorimeter. For the RCAL part of the alignment study SINISTRA95 instead of EM is used as it was specifically written to find electrons in the RCAL and also directly provides HES position information, extensively used in this part of the study. For all studies data and MC events are selected with the standard NC DIS cuts described in Chaps. 5.4 and 5.3.

A.1 Barrel calorimeter

Two of six degrees of freedom have been checked in this study, i.e. movement along the z-axis and a rotation about the z axis. In contrast to the other 4 degrees of freedom, both the z movement as well as the rotation about the z axis enter directly in the reconstructed angular position of a calorimeter cluster, where this position is not only important for the reconstruction of the kinematic variables but also for example for the track matching of the EM electron finder. Figures A.1 a,b show the ratio between the calorimeter energy E_{cal} and the double-angle energy E_{DA} (4.2) as a function of the z position of the electron (obtained from the extrapolated track),

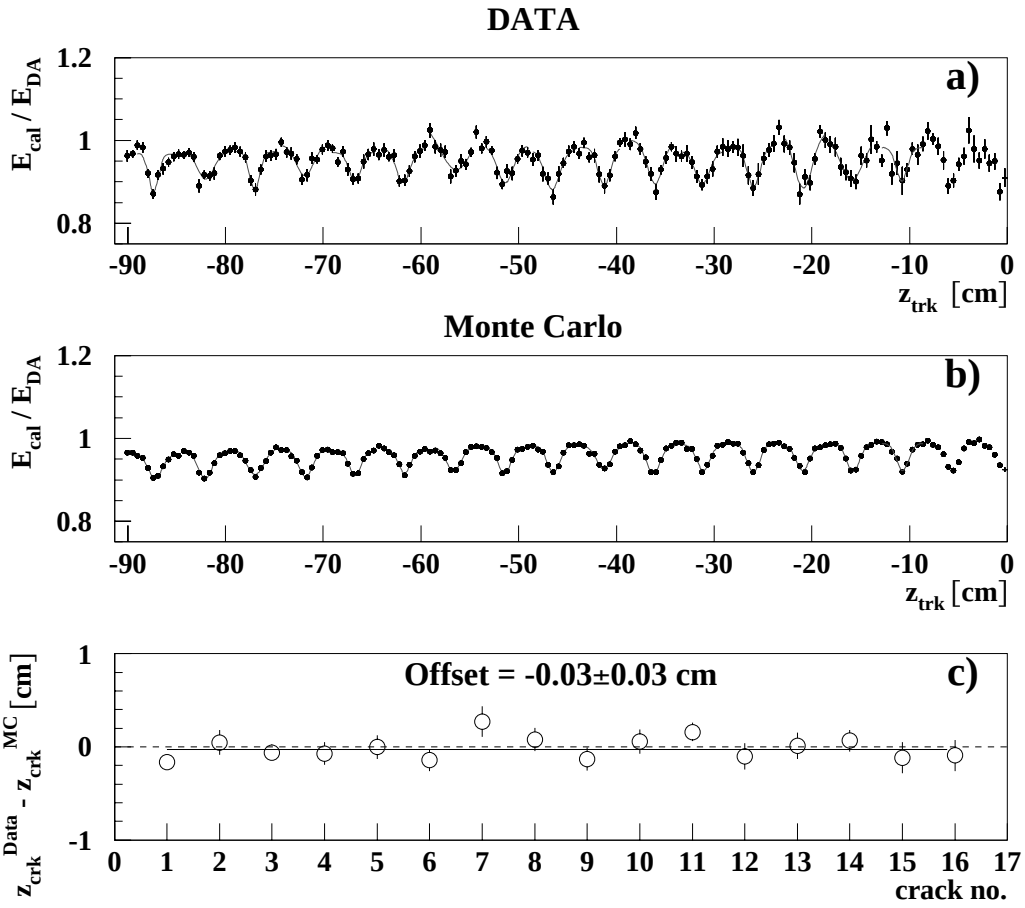


Figure A.1: *Ratio between calorimeter energy E_{cal} and double-angle energy E_{DA} as a function of the z position of the electron z_{trk} obtained from the extrapolated track for (a) data and (b) MC. (c) Difference between data and MC in the center values of the dips, z_{crk}, as a function of the crack number.*

z_{trk}, for data and NC DIS MC. The crack positions are clearly visible as dips in the ratio. Each dip has been fitted to an inverted Gaussian, and the differences between the mean values z_{crk} of data and MC are plotted in Fig. A.1 c as a function of the crack number. A fit to the points yields an offset of (-0.03 ± 0.03) cm which is compatible with zero within the error.

Figure A.2 shows the ratio of calorimeter and double-angle energies as a function of ϕ. All 32 BCAL modules have been projected on top of each other in order to enhance statistics. The ϕ position of the cracks between the modules is visible as a broad peak in the ratio. Fitting Gaussians to the crack regions for both data and MC yields peak values of (0.1199 ± 0.0002) rad and (0.1199 ± 0.0006) rad, respectively, indicating very good agreement between data and MC. On the other hand, the cracks in data are obviously much broader than in MC which could be caused either by an improper treatment of this region in the detector simulation or by small irregularities in the distances between two modules in the real detector.

From these investigations it can be concluded that the agreement between data and MC concerning the z and ϕ positions of the cracks is very good and that the BCAL is within 0.3 mm (0.6 mrad) at the expected z (ϕ) position.

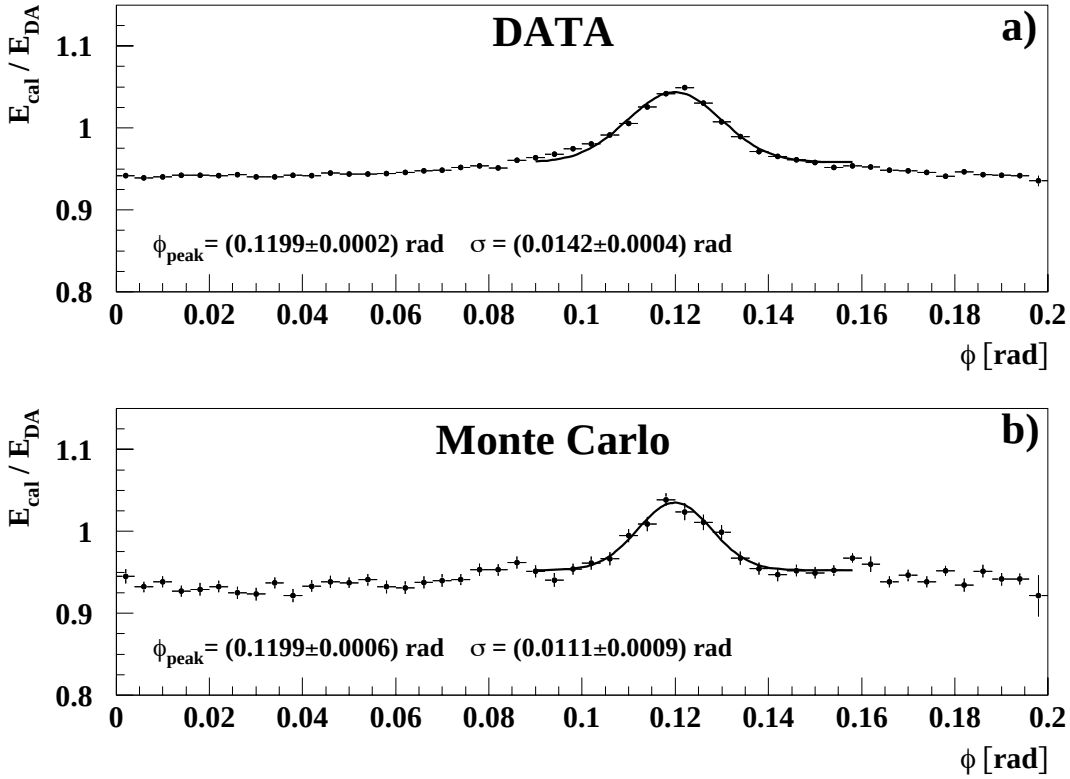


Figure A.2: *Ratio between calorimeter energy E_{cal} and double-angle energy E_{DA} as a function of ϕ for (a) data and (b) MC. In order to gain better statistics all 32 BCAL modules have been projected on top of each other.*

A.2 Rear calorimeter

The determination of the RCAL position is somewhat more difficult as it is not possible to use the module cracks due the non-projectiveness of the RCAL. As a result only a few of the cracks are visible. In addition, those which are visible are not very pronounced, leading to large uncertainties in the calculation of the calorimeter position. Therefore a different approach has to be chosen where the most obvious one uses the calorimeter position together with the position information from a CTD track associated with the electron. However, this approach has the disadvantage that it relies on the calorimeter-position reconstruction which is heavily biased as we shall see later. This method also implies, due to the limited CTD acceptance, that only electrons hitting the RCAL at a radius $r \gtrsim 70\,\mathrm{cm}$ can be used for this analysis, resulting in limited statistics. Here, the CTD acceptance is defined by tracks that pass at least 4 superlayers of the CTD, corresponding to a minimum radius of 45 cm at the CTD end-plate. Extrapolating the line going through this point and the nominal IP up to the electron-position reconstruction plane of the RCAL at $z = -153.03\,\mathrm{cm}$, yields the mentioned minimum radius of $r \approx 70\,\mathrm{cm}$.

Figures A.3 a–c, already depicted in Fig. 4.3, show the mean difference between calorimeter and track position in x (y) as a function of the x (y) position in the calorimeter, where in the case of y, the RCAL is split into its left and right half. Module 12 is always excluded as it can be moved in y independently from the rest of the RCAL. All plots show strong fluctuations in the

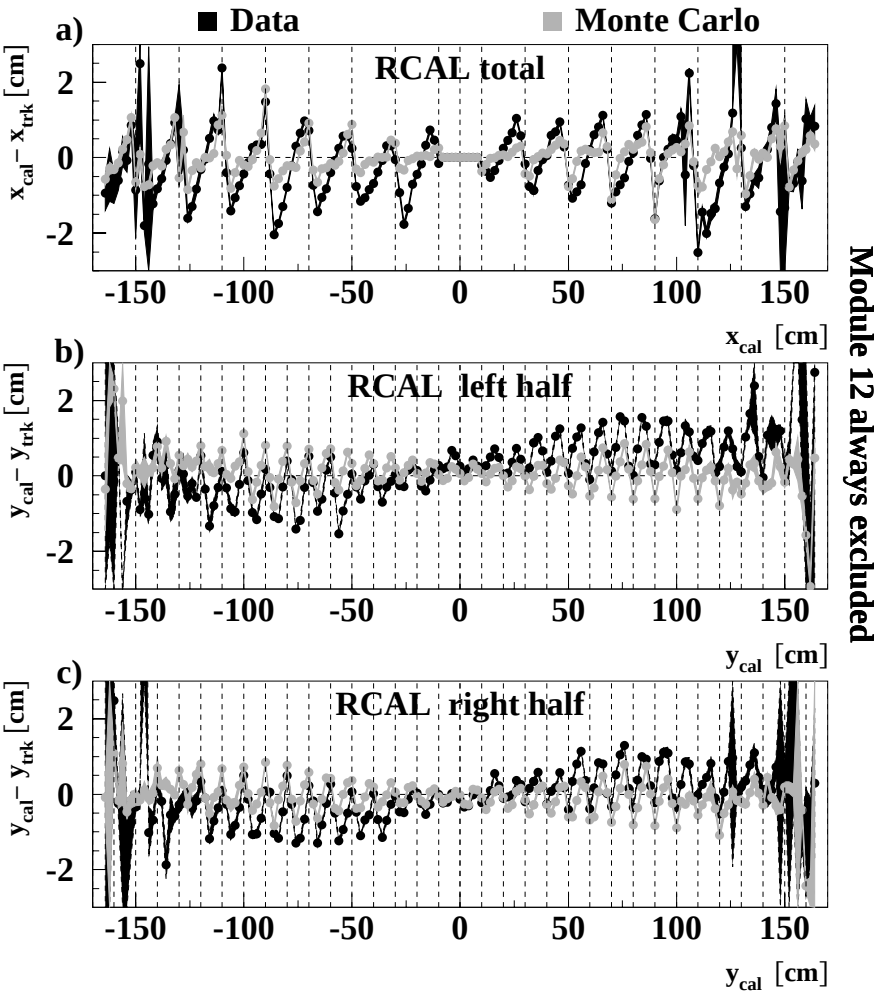


Figure A.3: *(a) Difference in x between calorimeter position x_{cal} and track position x_{trk} as a function of x_{cal}. (b) Difference in y between calorimeter position y_{cal} and track position y_{trk} as a function of y_{cal} for left RCAL half. (c) Difference in y between calorimeter position y_{cal} and track position y_{trk} as a function of y_{cal} for right RCAL half. Dark points are data and light points are MC. The vertical lines indicate the cell boundaries. The thickness of the lines connecting the points represent the error on the measured differences.*

differences between the calorimeter and track positions, which are correlated with the physical structure of the calorimeter and appear both in MC and data. In the x direction (Fig. A.3 a) this fluctuation has a periodicity of 20 cm, the width of a module, and an amplitude of up to 2 cm. In y (Fig. A.3 b and c) the fluctuations have a periodicity of only 10 cm, reflecting the subdivision of the EMC section within a tower into two cells. Due to the lack of a flat area which could be used as an unbiased position information, the calorimeter position is not suitable for aligning the RCAL.

As an alternative the position information of the RHES is used. This has the advantage that, in contrast to the calorimeter, the signals are already produced in a plane, which simplifies the spatial position reconstruction. The RHES plane runs parallel to the RCAL surface and

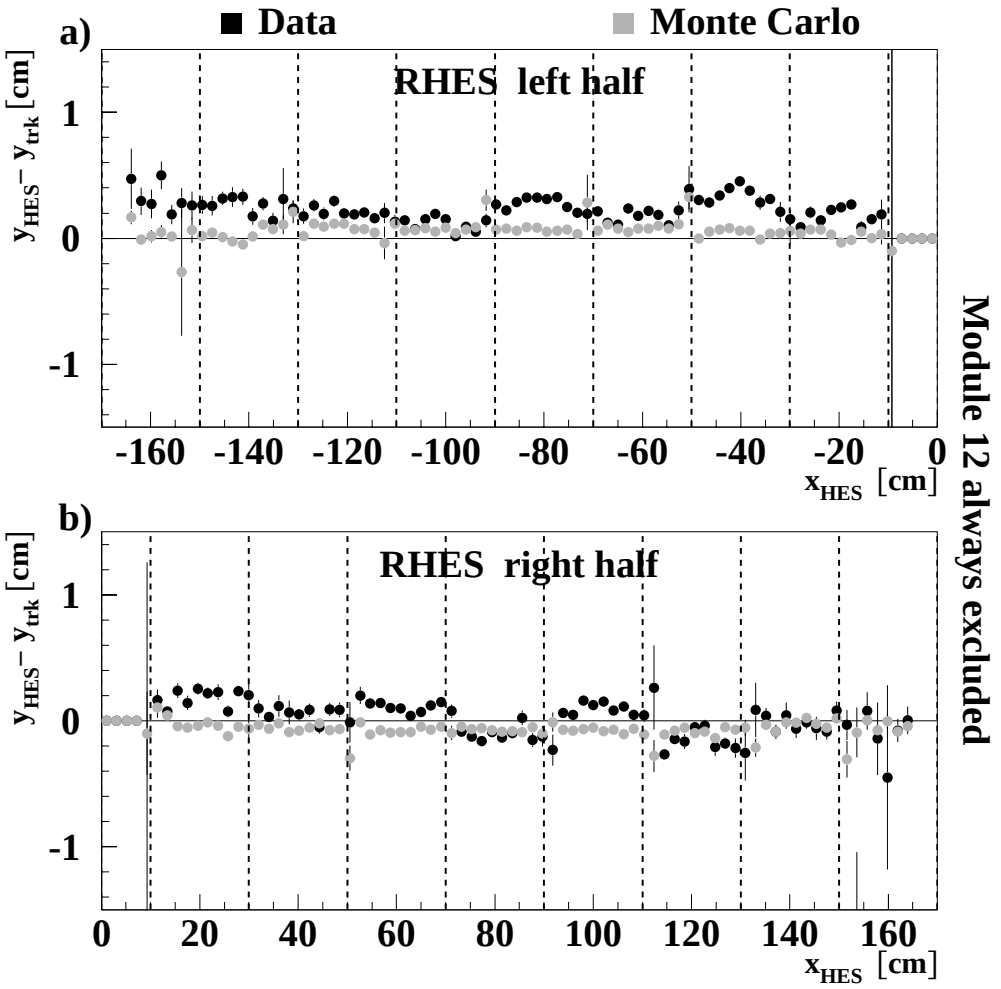


Figure A.4: *Difference in y between HES position y_{HES} and track position y_{trk} as a function of x_{HES} for left (a) and right (b) RHES half. Common offsets of skis in individual modules show displacements in y. Dark points are data and light points are MC.*

is located 7.1 cm inside the RCAL. The x-y position for a track in this plane is obtained by a linear extrapolation of the track from the calorimeter reconstruction plane to the RHES plane ($\Delta z = 2.0\,\mathrm{cm}$). The following parameters are determined for each half of the RHES: shift in y for the skis in each module, rotation about the z axis, shift in z and shift in x. As both the upper and the lower half of module 12 can be moved independently relative to the corresponding RCAL half, additional shifts in x and y relative to the corresponding RCAL half have been determined.

Figure A.4 shows the difference between HES and track position in y as a function of x for the left (a) and right (b) RHES half. The dashed vertical lines indicate the module boundaries. This kind of plot is supposed to show shifts in y of the skis in one module. These common shifts are possible as the skis of a module are inserted from the top of the calorimeter independently from the other modules. For data the different module offsets are clearly visible in the plots, whereas the MC points have a rather flat distribution. The offsets for data are obtained by fitting a constant for each module. The resulting values are listed in Fig. A.13 a.

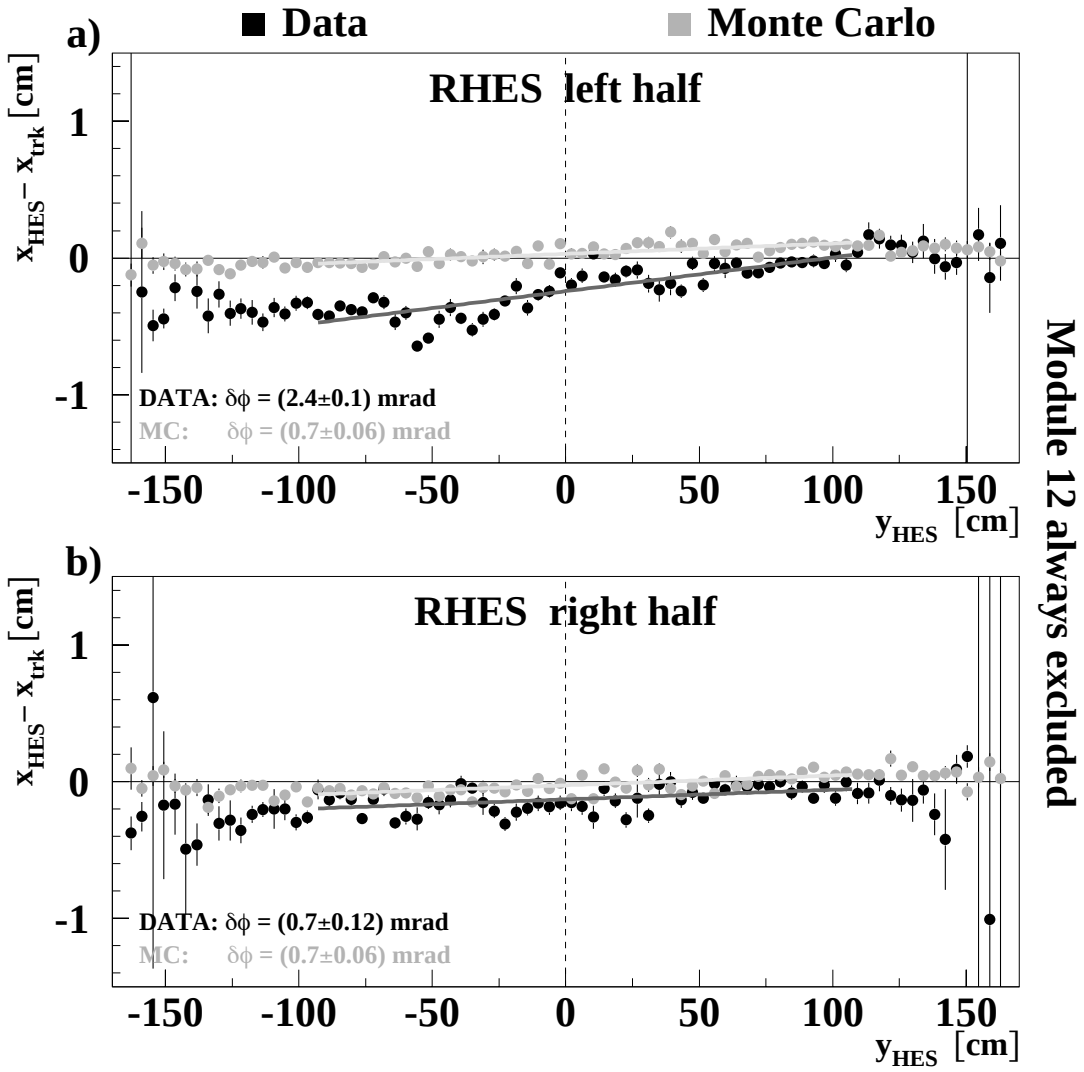


Figure A.5: *Difference in x between HES position x_{HES} and track position x_{trk} as a function of y_{HES} for left (a) and right (b) RHES half. Slopes in the distributions indicate a rotation about the z axis. Dark points are data and light points are MC.*

Figure A.5 displays the difference in x between RHES and track position as a function of y for the left (a) and right (b) RHES half. To first order a rotation about the z axis results in a slope in these diagrams. As can be seen from the plots all slopes correspond to rotations below 1 mrad, except for that of the left RHES half in data which yields a rotation of 2.4 mrad. Note that the errors given are only statistical, whereas the final errors will also contain systematic errors, which will be discussed in the next section. The fact that even the MC has a non vanishing slope suggests that the reconstruction method has a slight bias which would imply that for data the right RHES half is not rotated about the z axis. Therefore, the final shifts/rotations for data will be calculated relative to those of the MC.

The following plots already contain the previous corrections for data, i.e. a rotation about the z axis by 1.7 mrad for the left RHES half and the y shifts of the individual modules. No rotation for the right half is applied as the difference between data and MC is 0.0 mrad.

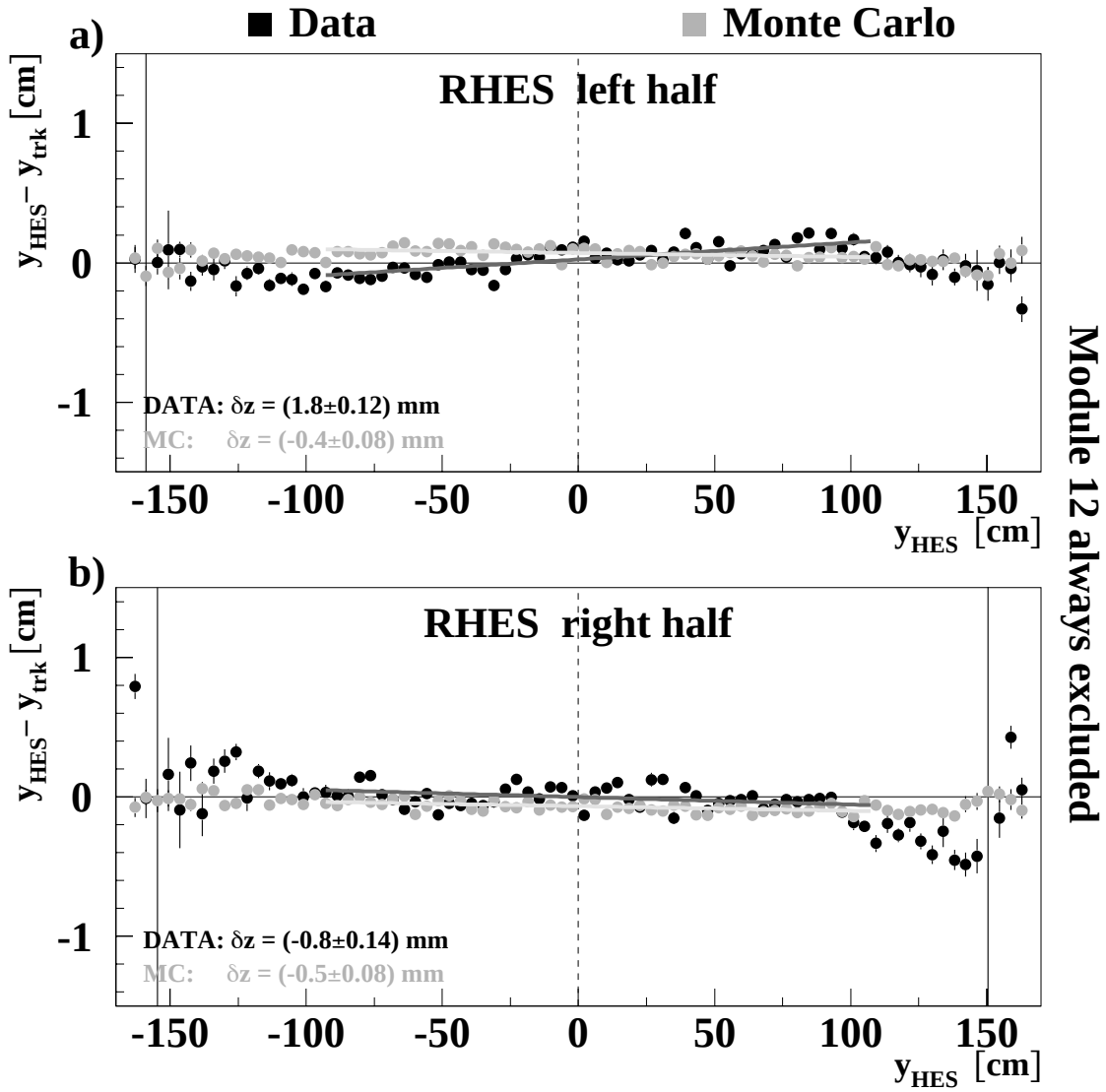


Figure A.6: *Difference in y between HES position y_{HES} and track position y_{trk} as a function of y_{HES} for left (a) and right (b) RHES half. Slopes in the distributions indicate a shift in z. Dark points are data and light points are MC.*

Figure A.6 shows the difference in y between HES and track position as a function of y. A slope in these plots indicates a shift in z. All shifts are below 1 mm except for the data of the left RHES half, where the fit yields a z shift of 1.8 mm away from the IP. The fit was restricted to ± 100 cm as the data shows a non-linear behavior in the outer parts of the calorimeter. This feature is not yet understood and needs further investigations.

For the following plots z shifts of -2.2 mm for the left and 0.3 mm for the right RHES half in data are included.

Figure A.7 a displays the difference in x as a function of x indicating shifts of the RHES halves in x. As in the case of the calorimeter (Fig. A.3) the structure of the HES is clearly visible. However, in contrast to the RCAL the RHES has a flat region between two module gaps which

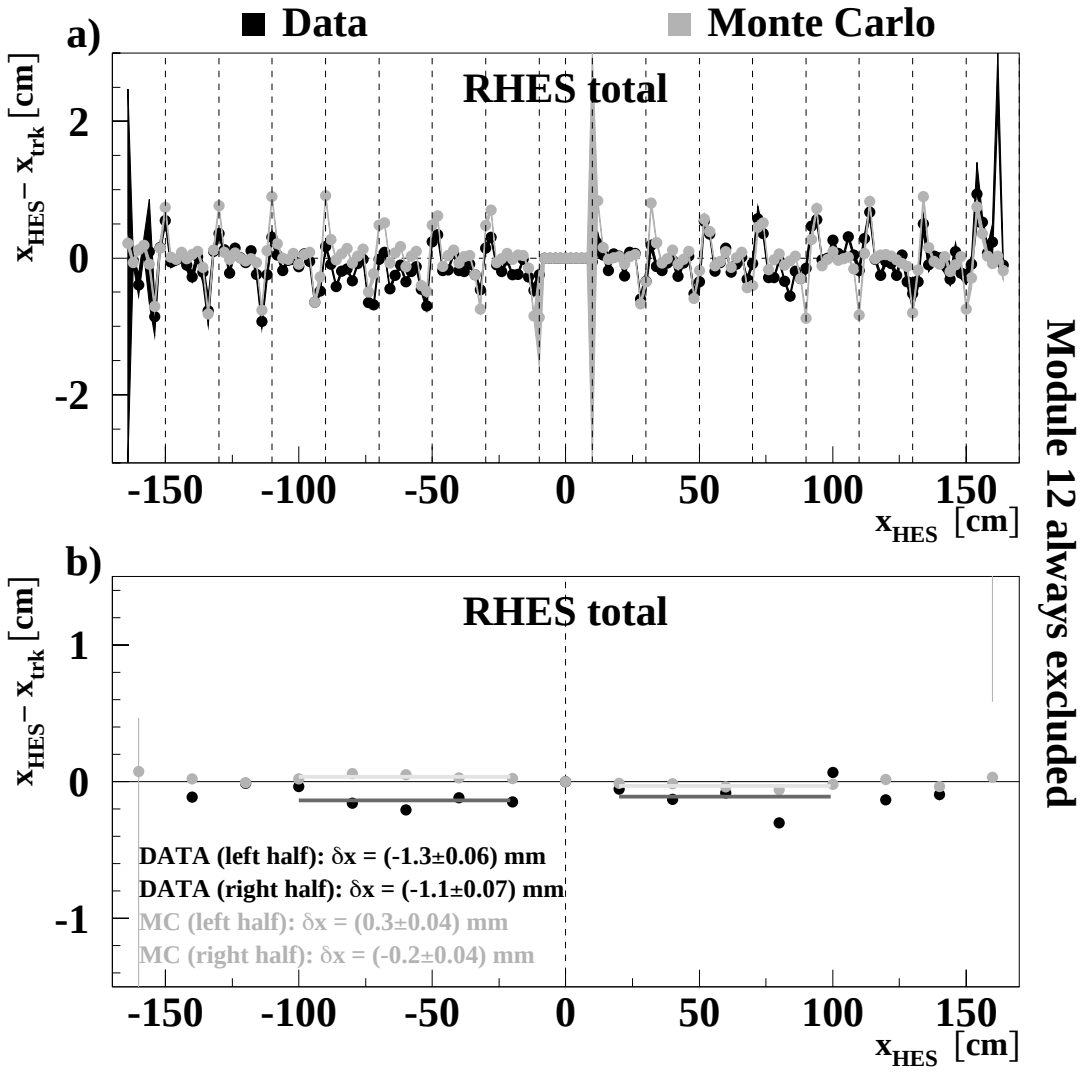


Figure A.7: *Difference in x between HES position x_{HES} and track position x_{trk} as a function of x_{HES}. (a) shows all x positions whereas in (b) the mean value for each module is shown, obtained from fitting a constant to the central region of the corresponding module. Offsets are caused by shifts in x. The left and right halves were fitted separately. Dark points are data and light points are MC. The thickness of the lines connecting the points repesents the error on the measured differences.*

can be used to fit a constant and take the offset as the x position in this module (the fit excludes 5 cm to either edge of the module). The results of these fits are displayed in Fig. A.7 b. The MC distribution is flat and has offsets less than 0.5 mm, whereas the data show offsets of more than 1 mm for both RHES halves. The offset values are determined by fitting a constant to the five inner modules of each RHES half. The offsets in data are 1.3 mm (left half) and 1.1 mm (right half) which leads to corrections in x for the data of 1.6 mm and 0.9 mm, respectively.

Figures A.8–A.11 show the previous plots after applying all corrections to the data. Ignoring the non-linear behavior in the outer regions of Fig. A.10, all data distributions agree very well with those of the MC.

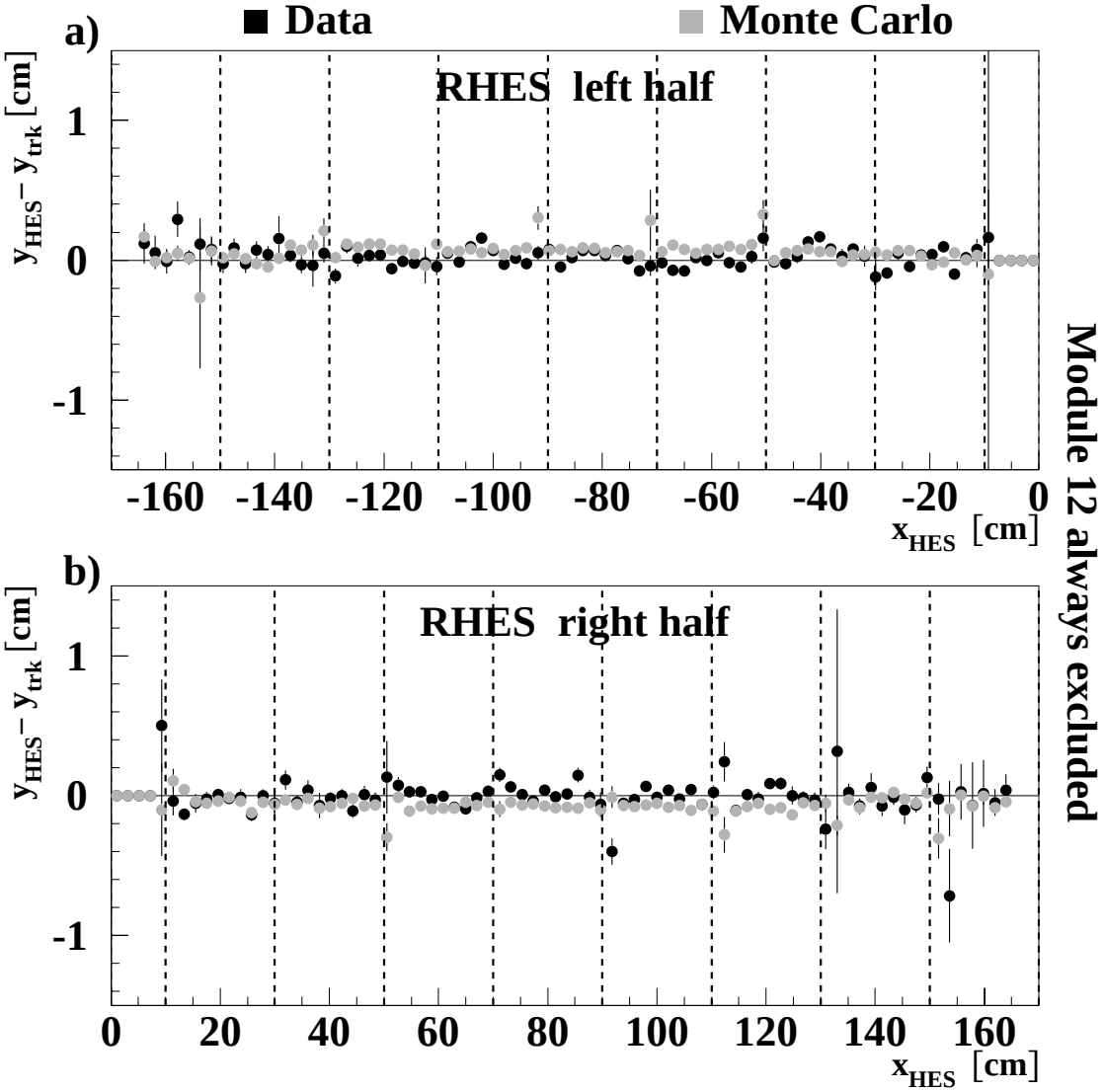


Figure A.8: *Difference in x between HES position x_{HES} and track position x_{trk} as a function of y_{HES} for left (a) and right (b) RHES half after all corrections. Dark points are data and light points are MC.*

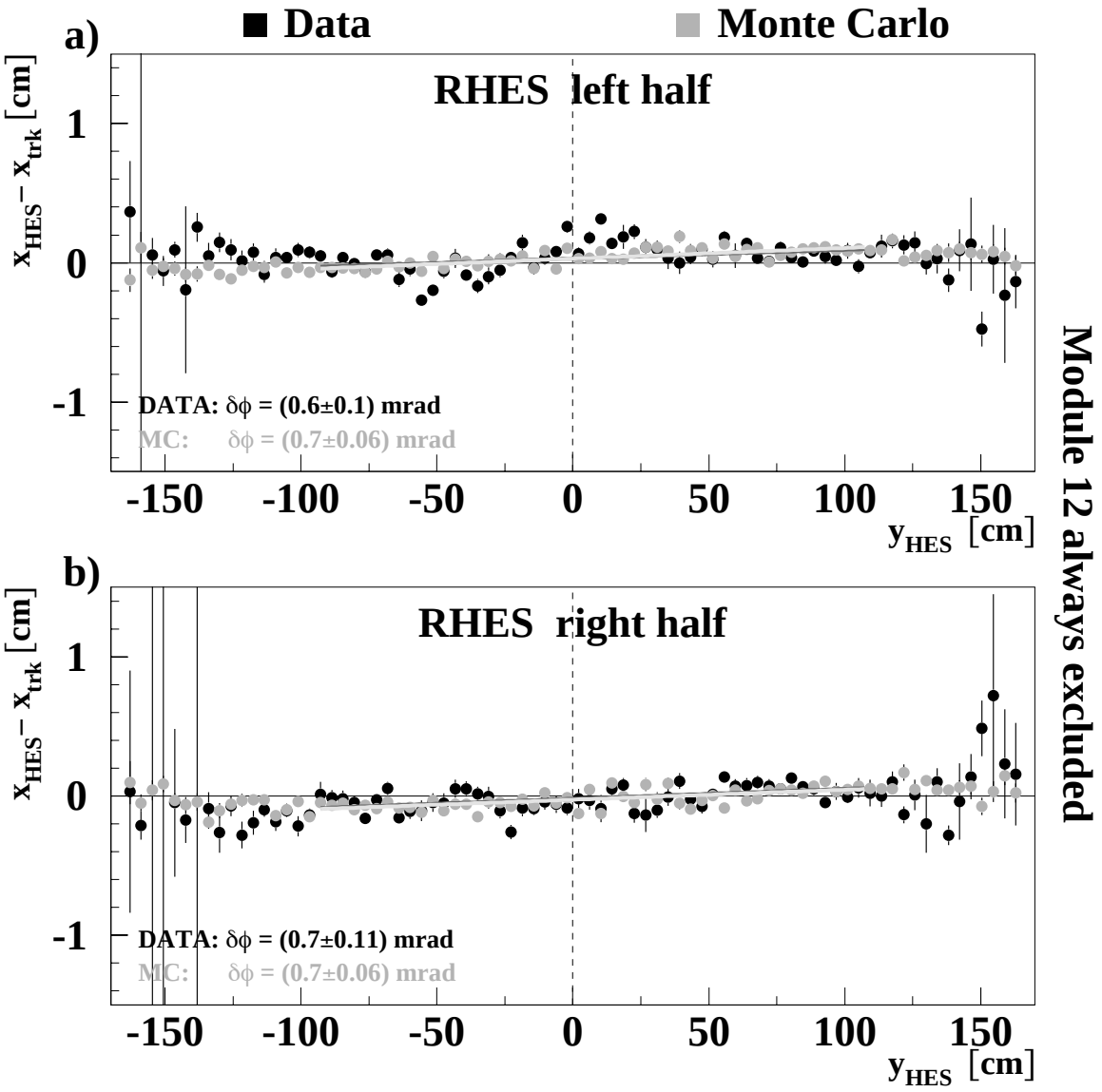


Figure A.9: *Difference in y between HES position y_{HES} and track position y_{trk} as a function of x_{HES} for left (***a***) and right (***b***) RHES half after all corrections. Dark points are data and light points are MC.*

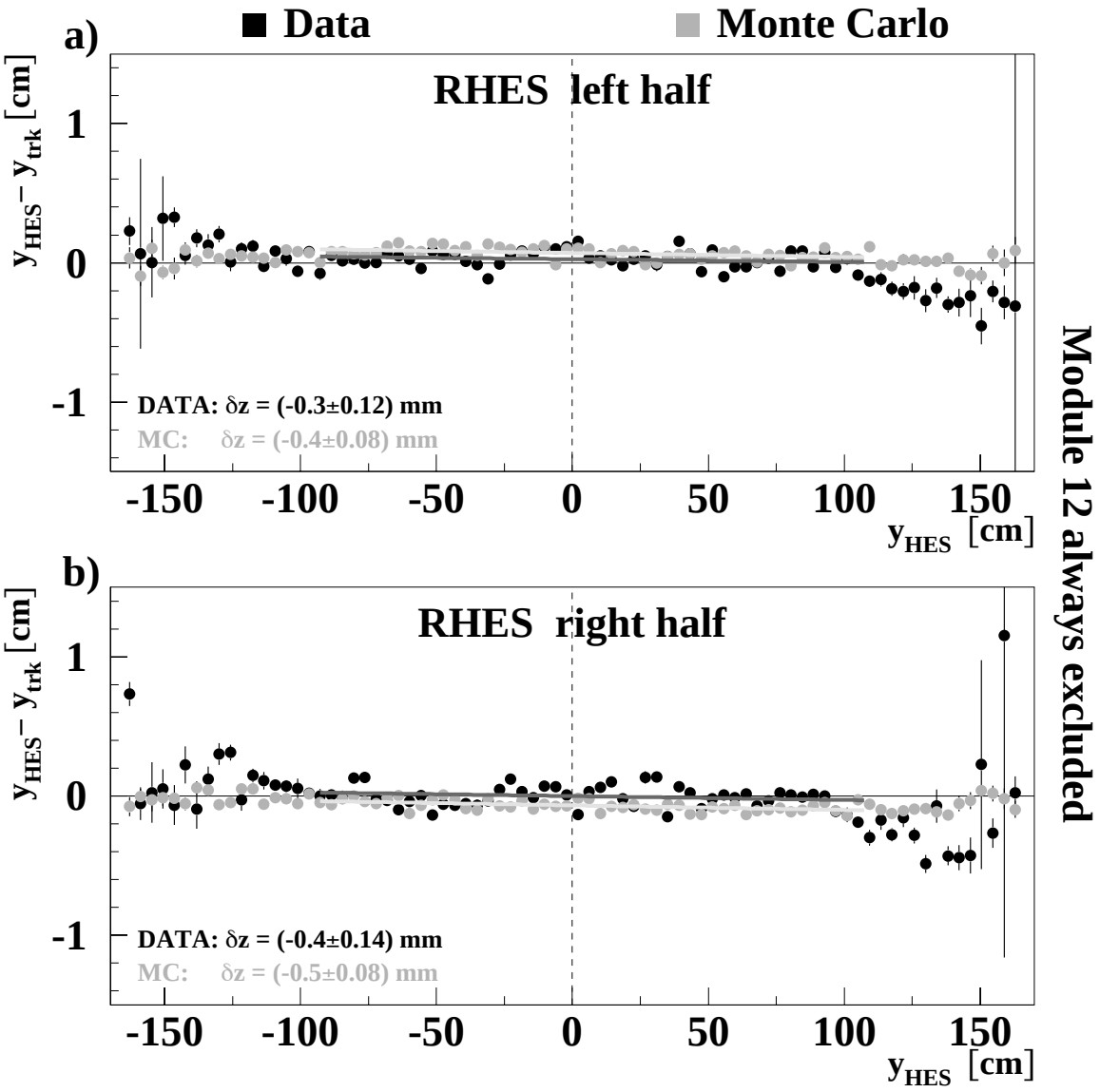


Figure A.10: *Difference in y between HES position y_{HES} and track position y_{trk} as a function of y_{HES} for left (***a***) and right (***b***) RHES half after all corrections. Dark points are data and light points are MC.*

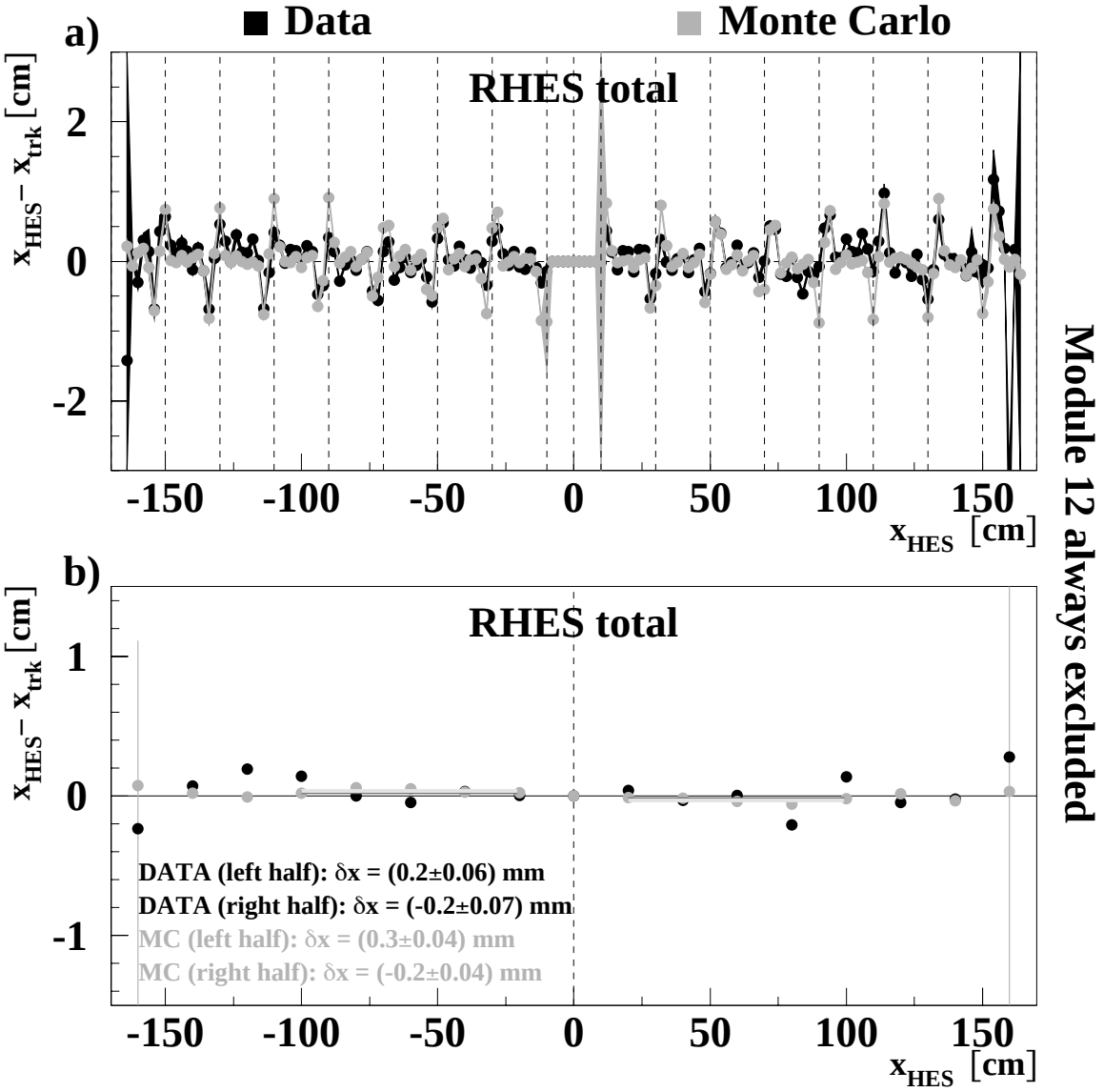


Figure A.11: *Difference in x between HES position x_{HES} and track position x_{trk} as a function of x_{HES} after all corrections. (a) shows all x positions whereas in (b) the mean value for each module is shown, which was obtained from fitting a constant to the central region of the corresponding module. Dark points are data and light points are MC.*

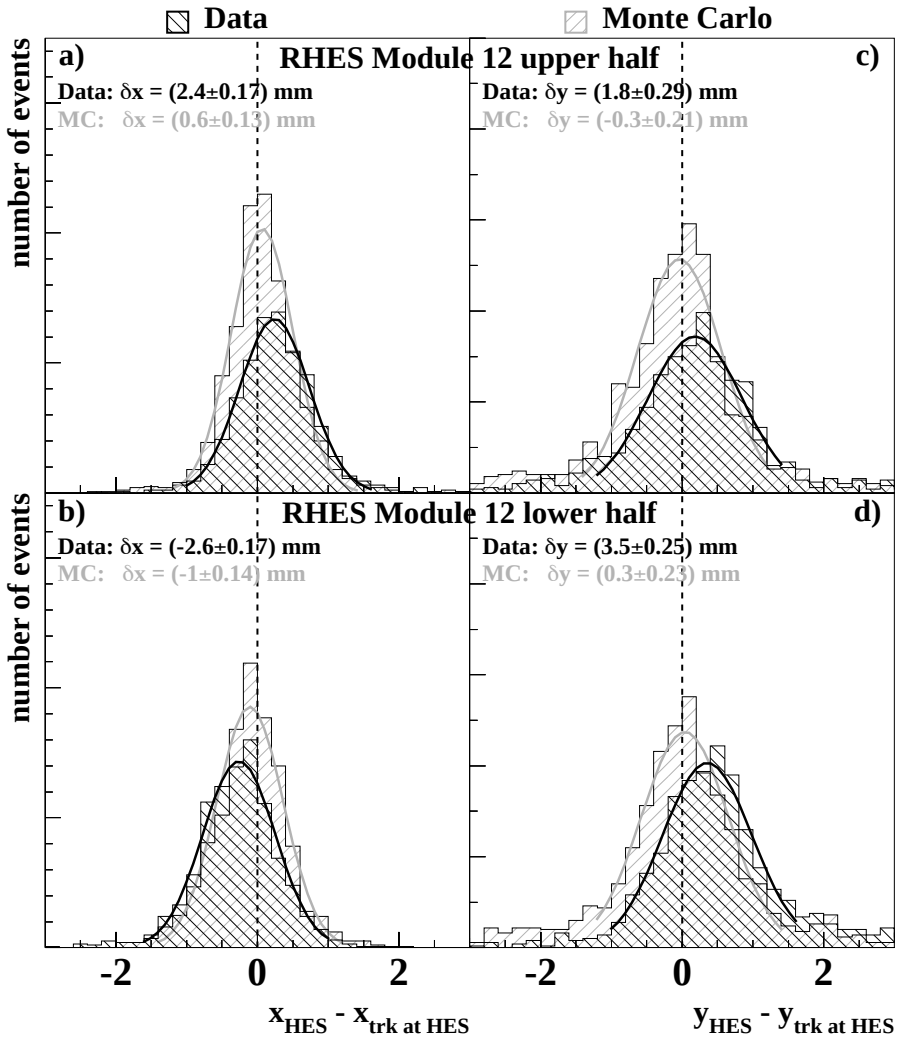


Figure A.12: *Difference in x and y between HES and track position for RCAL module 12 for upper half (a,b) and lower half (c,d). Data and MC are not normalized to each other.*

Both halves of module 12 can be moved in y with respect to the RCAL half they are attached to. Therefore independent alignment studies in x and y for these two module halves are performed. As the number of events is much lower than in the previous studies, the difference between HES and track position is plotted as a histogram. In Fig. A.12 the upper and lower two plots show δx and δy histograms for the upper and lower half of module 12, respectively. As neither MC nor data events are weighted the histograms are not normalized to each other. All shifts for MC are below 1 mm, whereas for data they range from 1.8 mm (upper left plot) to 3.5 mm (lower right plot). Astonishingly enough, for data the fits suggest a shift in x of over 2 mm towards the corresponding RCAL half. The fact that these shifts are also present in MC, though below 1 mm, indicates that the reconstruction method is biased. As a movement of a module by more than 1 mm towards the RCAL half it is attached to seems mechanically impossible, these shifts are ignored. As a result from this study, the upper (lower) half of module 12 is shifted down by 2.1 mm (3.2 mm) in data.

left RCAL half	
δx	(1.6 ± 0.2) mm
δz	(-2.2 ± 0.4) mm
$\delta\phi$	(1.7 ± 0.9) mrad

right RCAL half	
δx	(0.9 ± 0.3) mm
δz	(0.3 ± 0.7) mm
$\delta\phi$	(0.0 ± 0.8) mrad

Table A.1: *Table with shifts and rotations for both RCAL halves, where $\delta\phi$ is the rotation about the z axis.*

A.3 Systematic checks

In order to estimate the influence of systematics on the results obtained, various systematic checks are performed for each of the fits. These checks comprise variations of the fit range of the Gaussians to the δx and δy distributions in each x and y bin[1] and of the straight line fits to the $\delta x, \delta y$ vs. x, y distributions (Figs. A.4 and A.7), a variation of the minimum track radius at the CTD end-plate and an additional requirement of at least 40 units for the HES signal. Statistical and systematic errors are added in quadrature. The total error is dominated by the systematic uncertainties. A summary of all alignment results together with the corresponding total errors for data and MC can be found in Fig. A.13. The arrows indicate the direction of the "shifts" and "rotations". A subdivision of the data into two samples of equal integrated luminosity (first and second half of the run period) yields values that are compatible within errors with the numbers from the total sample. Hence no time dependence of the measured parameters is observed.

A.4 Final shifts and rotations

Up to now only the position of the HES plane relative to the CTD has been measured, whereas the original task was the alignment of the RCAL. However all results except for the shifts in y should be applicable to the RCAL as the relative position of the HES in x and z with respect to the RCAL is fixed. Hence, all RCAL shifts listed in Table A.1 are taken into account, whereas y shifts, including those of module 12, are omitted.

A.5 Summary

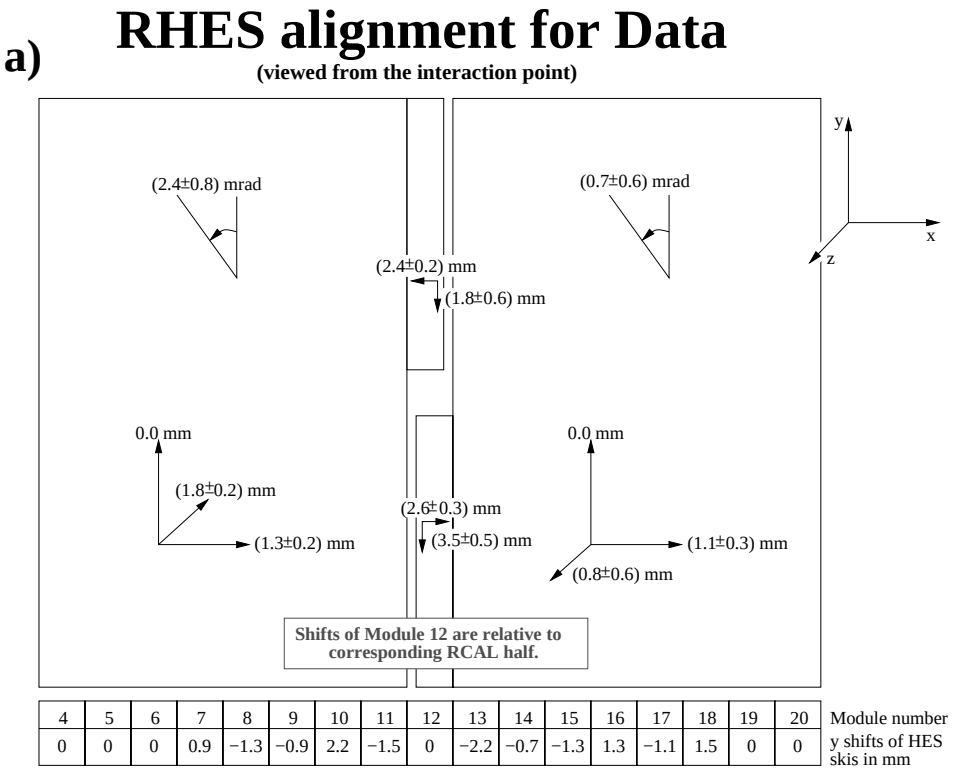
The positions of the BCAL and RCAL have been determined relative to the CTD using extrapolated tracks of identified electrons. For the BCAL a method insensitive to the calorimeter electron-position reconstruction algorithm has been used, showing that the z (ϕ) position in data and MC agree within 0.3 mm (0.6 mrad), where the mean positions obtained are identical in data and MC. Hence, no shifts are applied to the BCAL.

[1] The mean deviations of the HES positions from the track positions in a variable b, plotted on the ordinates, are obtained by filling 2-dimensional histograms (b vs. a) and fitting a Gaussian to the b distribution in each bin of a.

For the determination of the RCAL position the HES has been used as the calorimeter electron-position reconstruction shows large biases. The obtained shifts are listed in Table A.1 and are applied to the data.

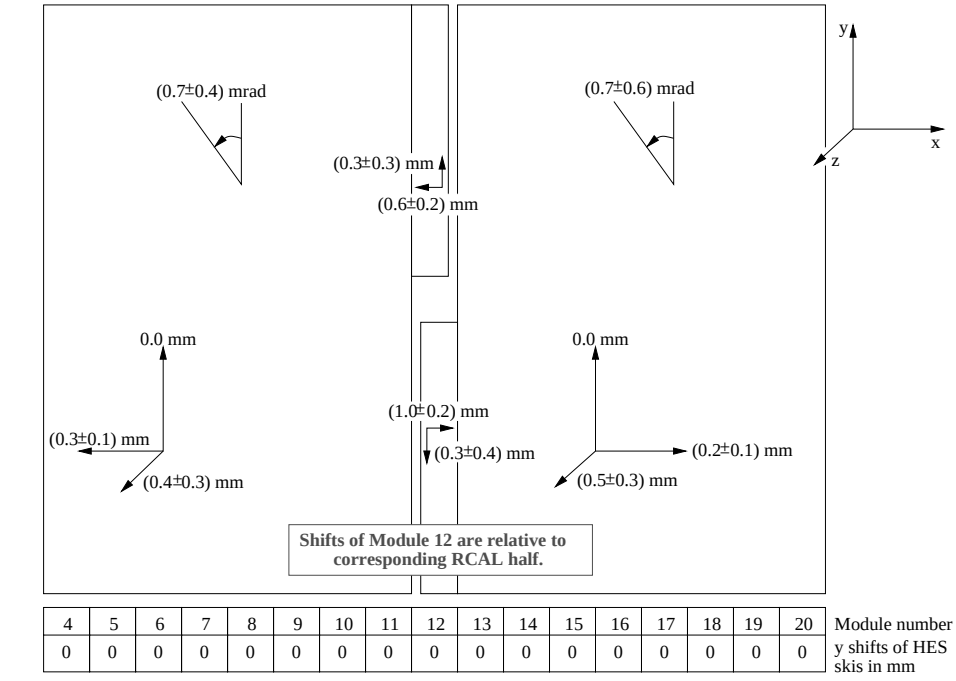
All examined calorimeter positions are known to a precision of better than 0.7 mm and 1 mrad, respectively, which is far better than the overall position resolution for electrons in the calorimeter.

a)



4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Module number
0	0	0	0.9	−1.3	−0.9	2.2	−1.5	0	−2.2	−0.7	−1.3	1.3	−1.1	1.5	0	0	y shifts of HES skis in mm

b)



4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Module number
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	y shifts of HES skis in mm

Figure A.13: *Alignment results of the RHES for (a) data and (b) MC. Statistical and systematic errors are added in quadrature.*

Appendix B

Examples for NC events

The following event displays show different views of NC DIS events in the ZEUS detector. The large picture on the left side is a z-r projection. Electrons enter the detector from the left and protons from the right. The lower right picture is an x-y projection of the detector, where only CTD and BCAL are shown. The upper right plot displays the transverse-energy distribution in the calorimeter in the η-ϕ plane, where black (white) bars represent EMC (HAC) energies.

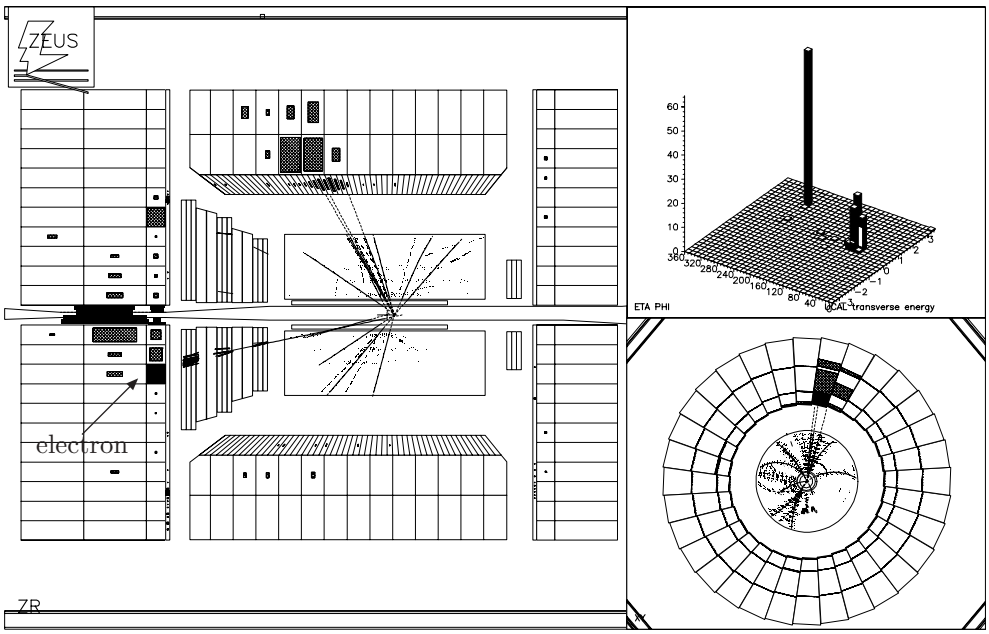


Figure B.1: *Event with the highest Q^2 in the 1998/99 e^-p data sample: Run number:* 32821*, event number:* 42090*, $Q^2 = 37400\,GeV^2$, $x = 0.43$, $y = 0.85$. The scattered electron with a corrected energy of* 310 *GeV hits the FCAL under an angle of* 12.6°*, whereas the hadronic final state is mainly contained in the BCAL. The upper right plot shows the very good isolation of the electron and the collimation of the jet.*

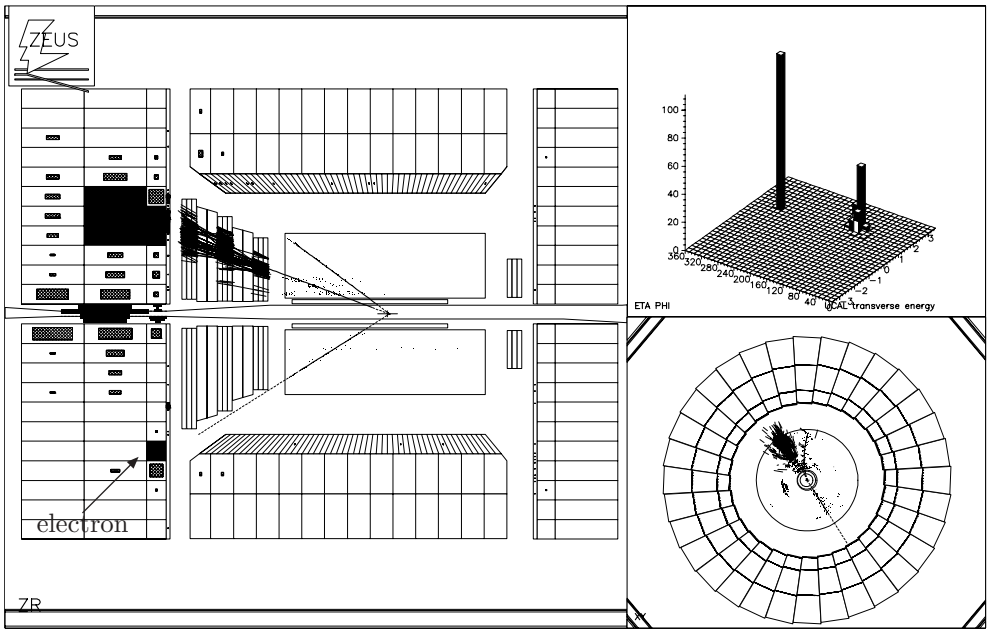


Figure B.2: *Run number:* 31244, *Event number:* 31806, $Q^2 = 22100\,GeV^2$, $x = 0.56$, $y = 0.39$. *In this event the hadronic final state is completely contained in the FCAL together with the electron which has a matched track. The x-y projection of the tracks in the lower right plot shows that the hadronic final state (upper left quarter) is balanced by the scattered electron (lower right quarter).*

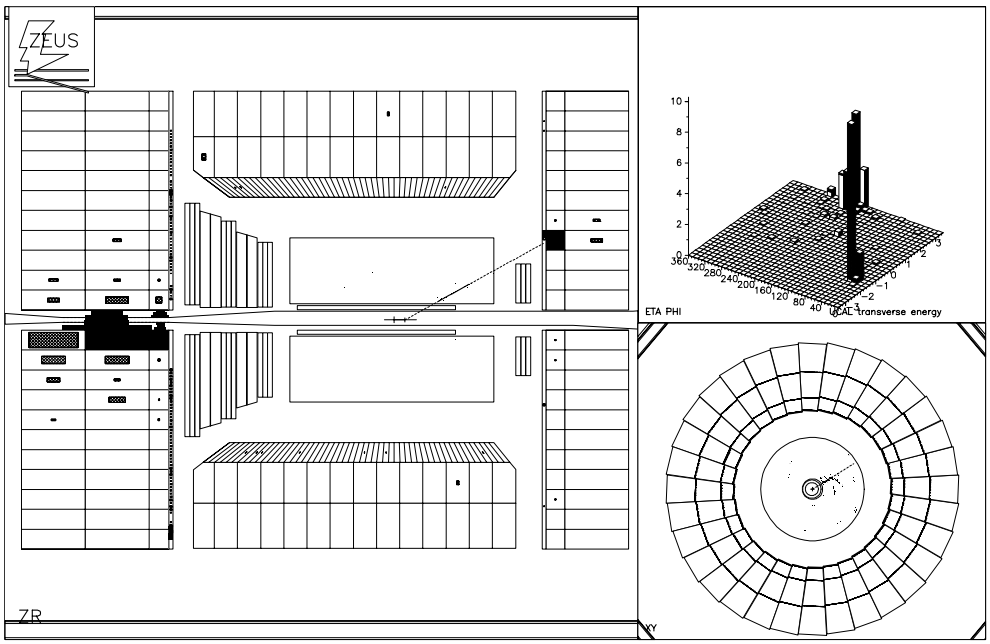


Figure B.3: *Run number:* 31316, *event number:* 19682, $Q^2 = 200\,GeV^2$, $x = 0.20$, $y = 0.01$. *The electron in the RCAL has an energy of* $28\,GeV$ *close to the beam energy of* $27.5\,GeV$ *and hence y is very small. The hadronic final state deposits its energy mainly in the first inner ring of FCAL cells and leaves no tracks in the CTD.*

Appendix C

Systematic checks

C.1 Single-differential cross sections: $d\sigma/dQ^2$

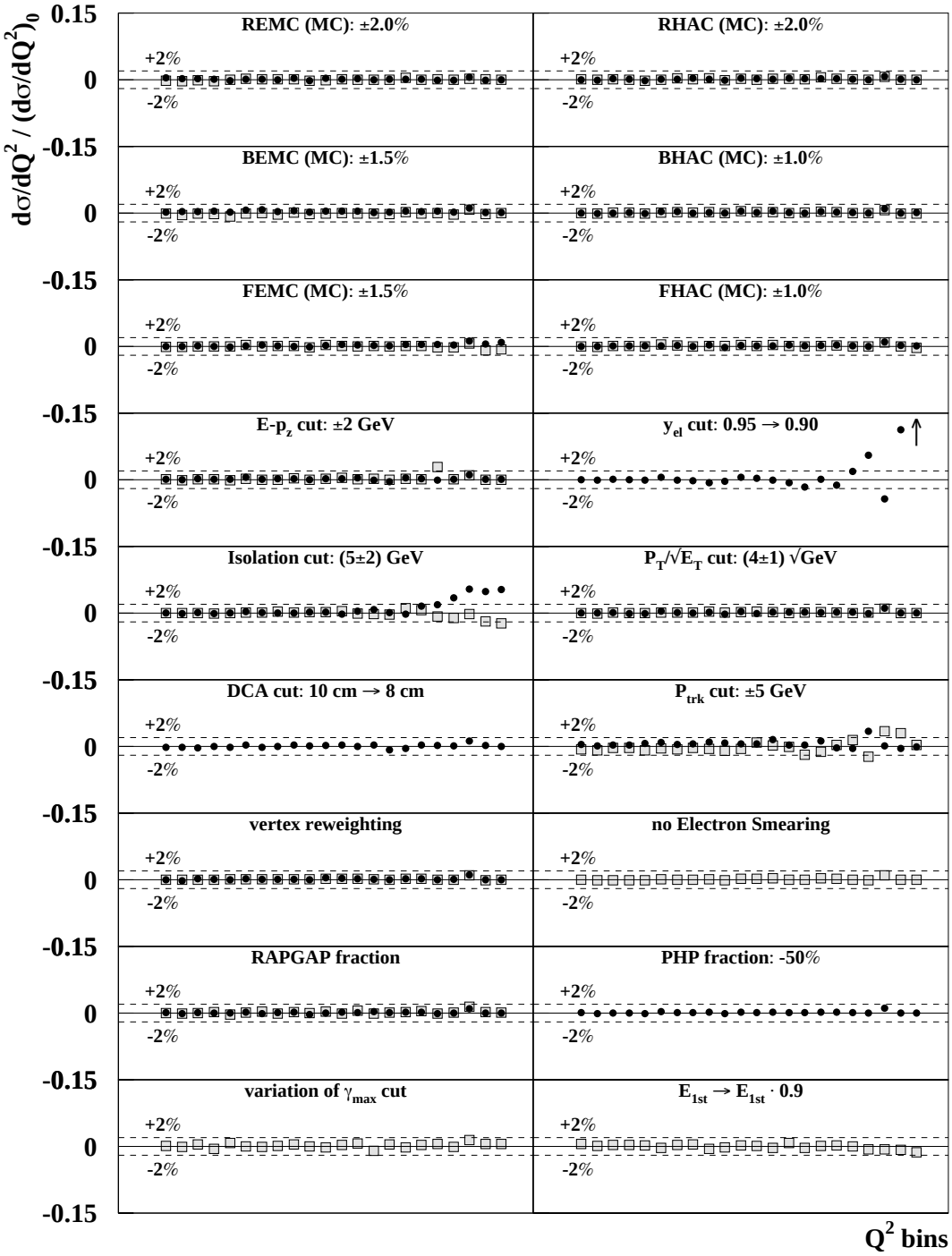


Figure C.1: *Ratio between the nominal $d\sigma/dQ^2$ cross section, $(d\sigma/dQ^2)_0$, and the same cross section after applying the systematic check, as a function of the Q^2 bin. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. Variations beyond the ordinate range are marked with an arrow. The dashed lines indicate a variation of ±2%.*

C.2 Single-differential cross sections: $d\sigma/dx$

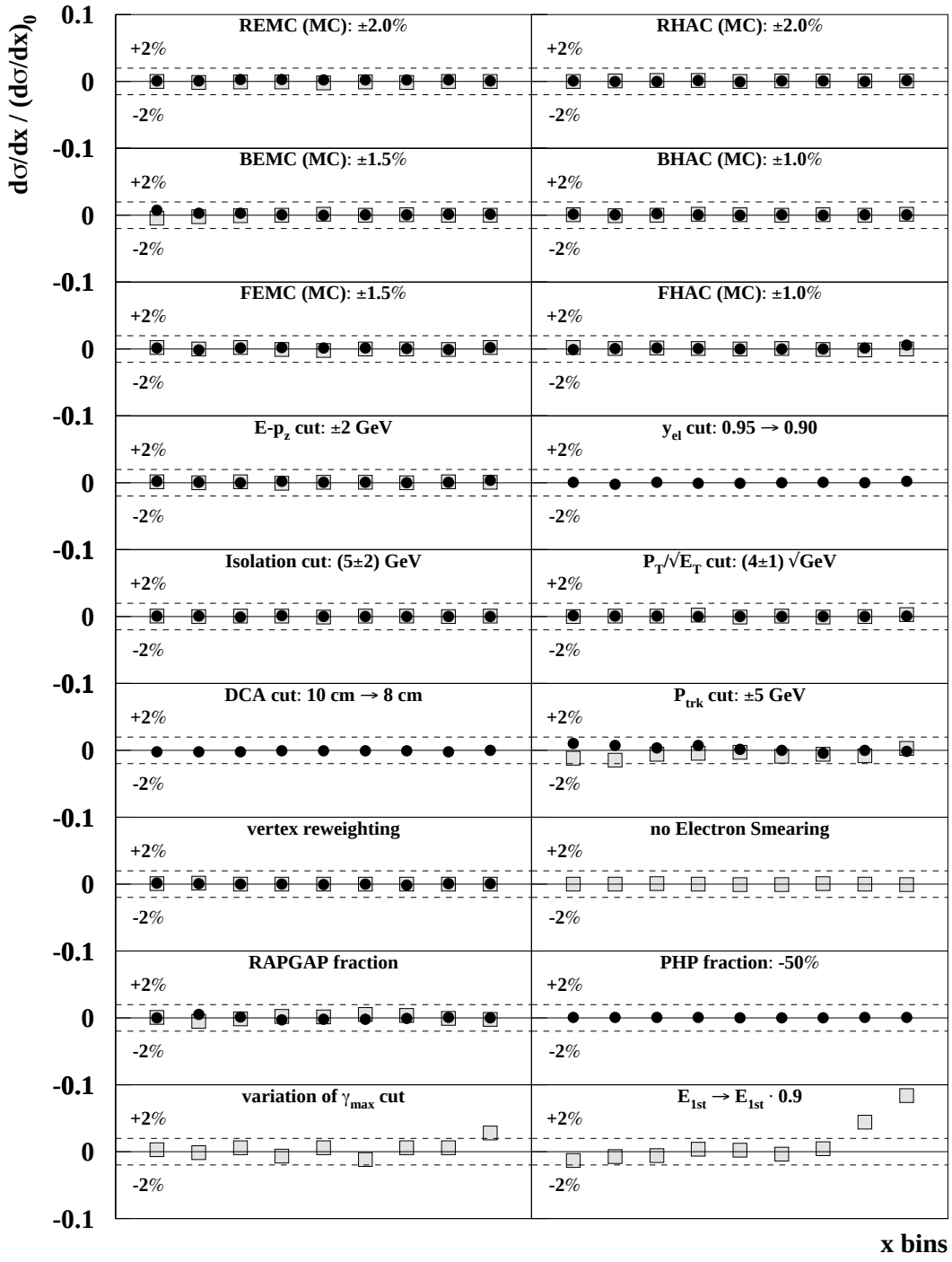


Figure C.2: *Ratio between the nominal $d\sigma/dx$ cross section, $(d\sigma/dx)_0$, and the same cross section after applying the systematic check, as a function of the x bin for $Q^2 > 200\,GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. The dashed lines indicate a variation of $\pm 2\%$.*

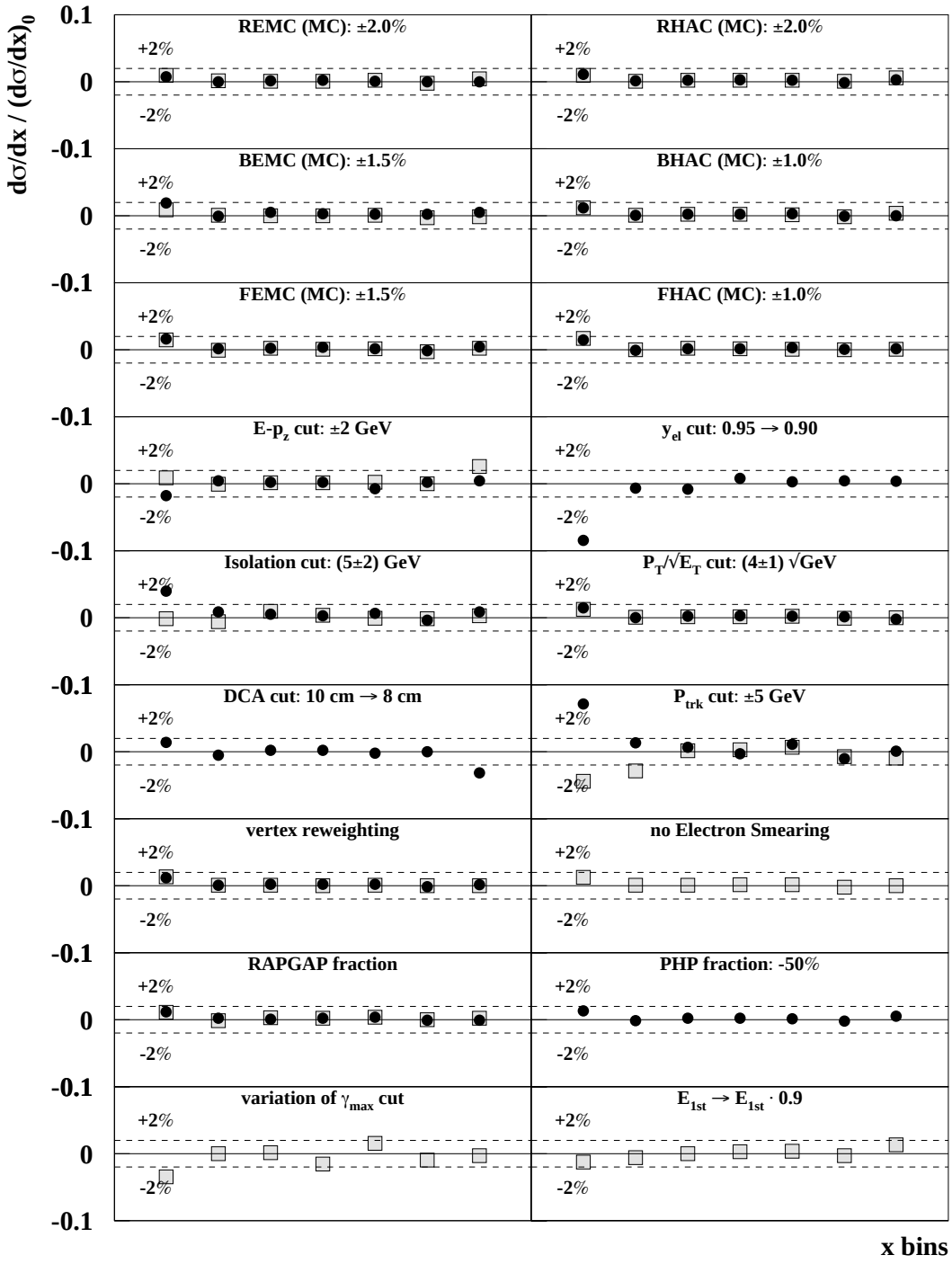


Figure C.3: *Ratio between the nominal $d\sigma/dx$ cross section, $(d\sigma/dx)_0$, and the same cross section after applying the systematic check, as a function of the x bin for $Q^2 > 2\,500\,GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. The dashed lines indicate a variation of $\pm 2\%$.*

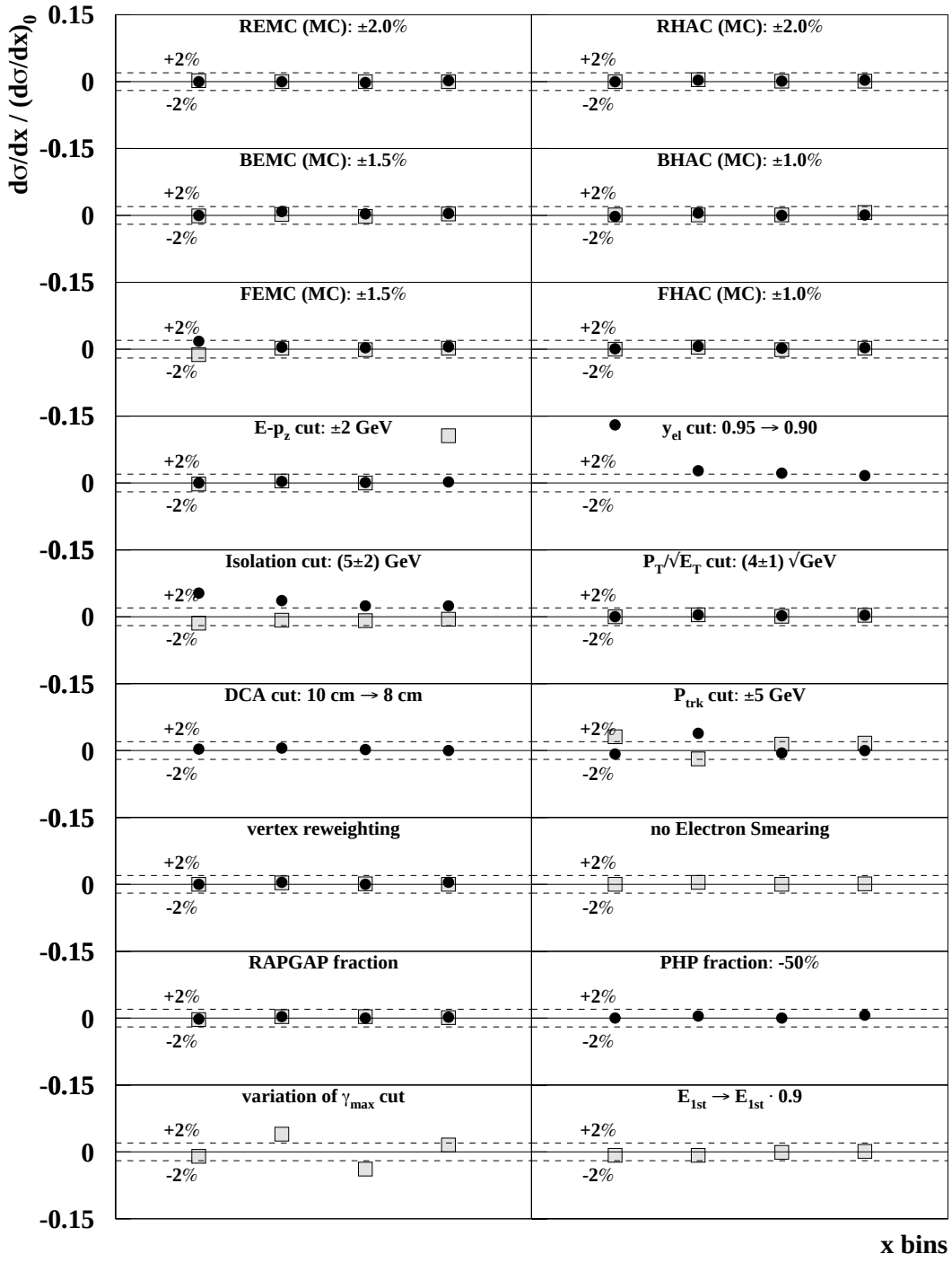


Figure C.4: *Ratio between the nominal $d\sigma/dx$ cross section, $(d\sigma/dx)_0$, and the same cross section after applying the systematic check, as a function of the x bin for $Q^2 > 10\,000\ GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. The dashed lines indicate a variation of $\pm 2\%$.*

C.3 Single-differential cross sections: $d\sigma/dy$

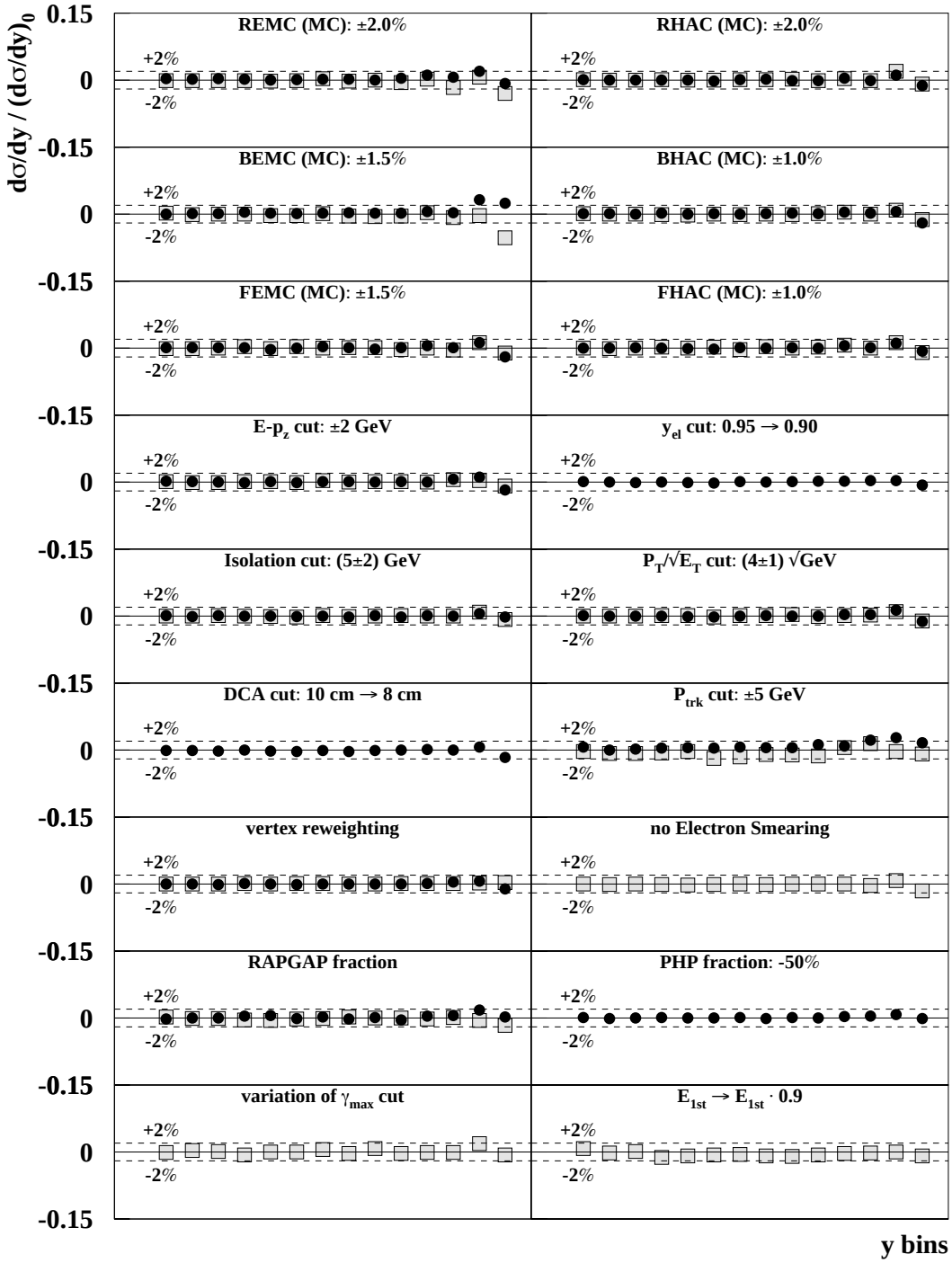


Figure C.5: *Ratio between the nominal $d\sigma/dy$ cross section, $(d\sigma/dy)_0$, and the same cross section after applying the systematic check, as a function of the y bin for $Q^2 > 200\ GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. The dashed lines indicate a variation of ±2%.*

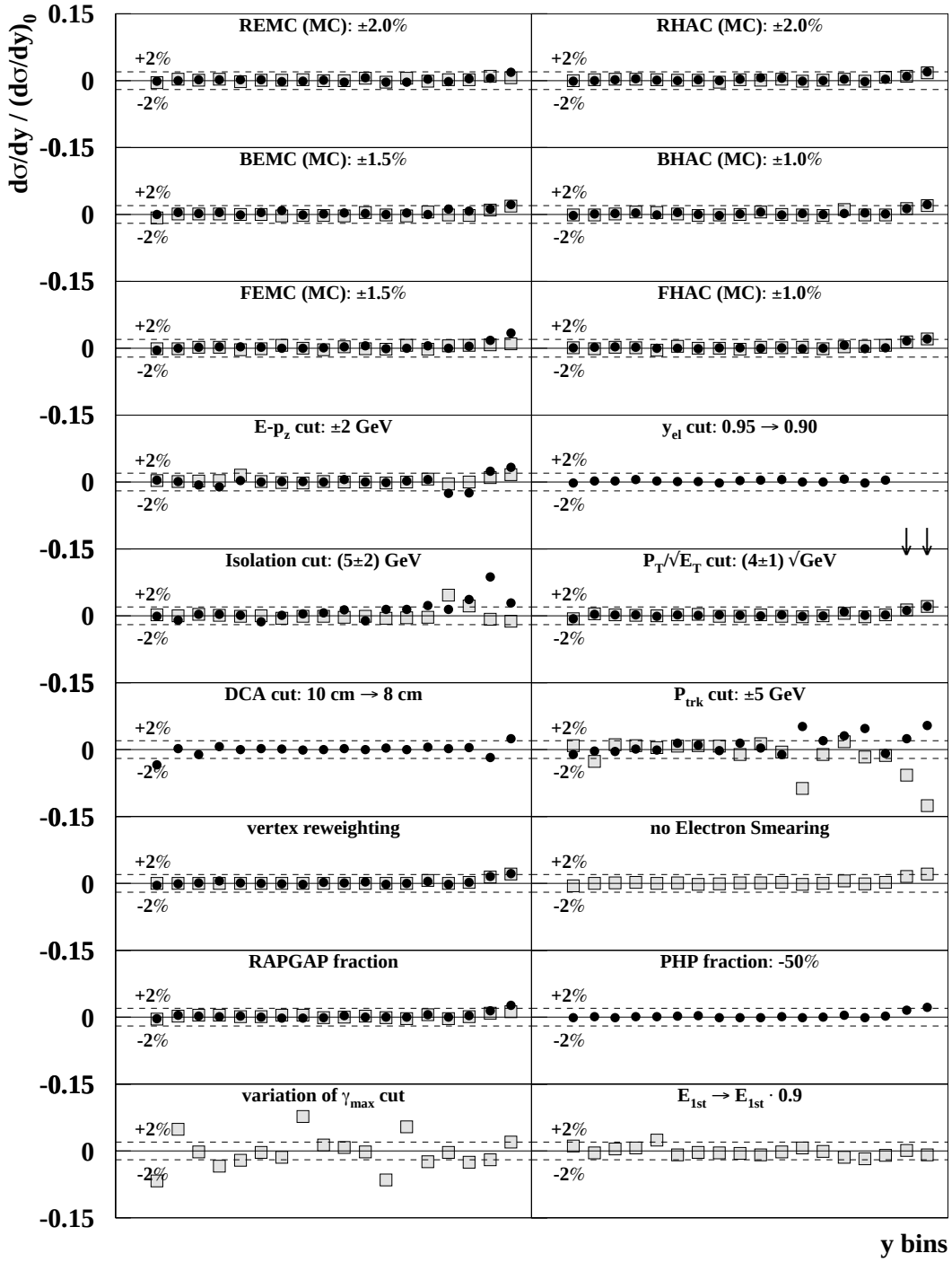


Figure C.6: *Ratio between the nominal $d\sigma/dy$ cross section, $(d\sigma/dy)_0$, and the same cross section after applying the systematic check, as a function of the y bin for $Q^2 > 2\,500\,GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. Variations beyond the ordinate range are marked with an arrow. The dashed lines indicate a variation of ±2%.*

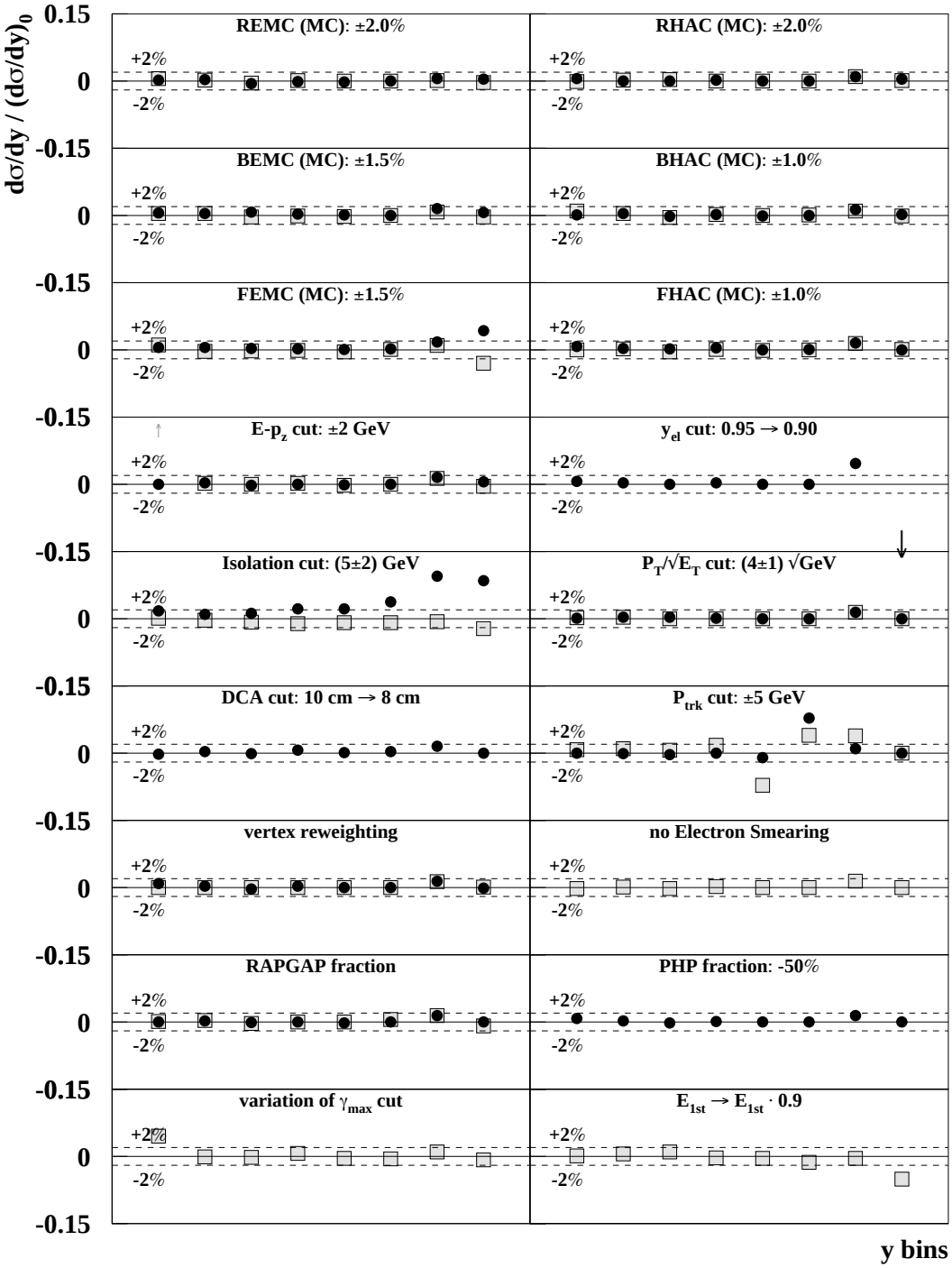


Figure C.7: *Ratio between the nominal $d\sigma/dy$ cross section, $(d\sigma/dy)_0$, and the same cross section after applying the systematic check, as a function of the y bin for $Q^2 > 10\,000\ GeV^2$. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. Variations beyond the ordinate range are marked with an arrow. The dashed lines indicate a variation of $\pm 2\%$.*

C.4 Double-differential cross sections: $d^2\sigma/dxdQ^2$

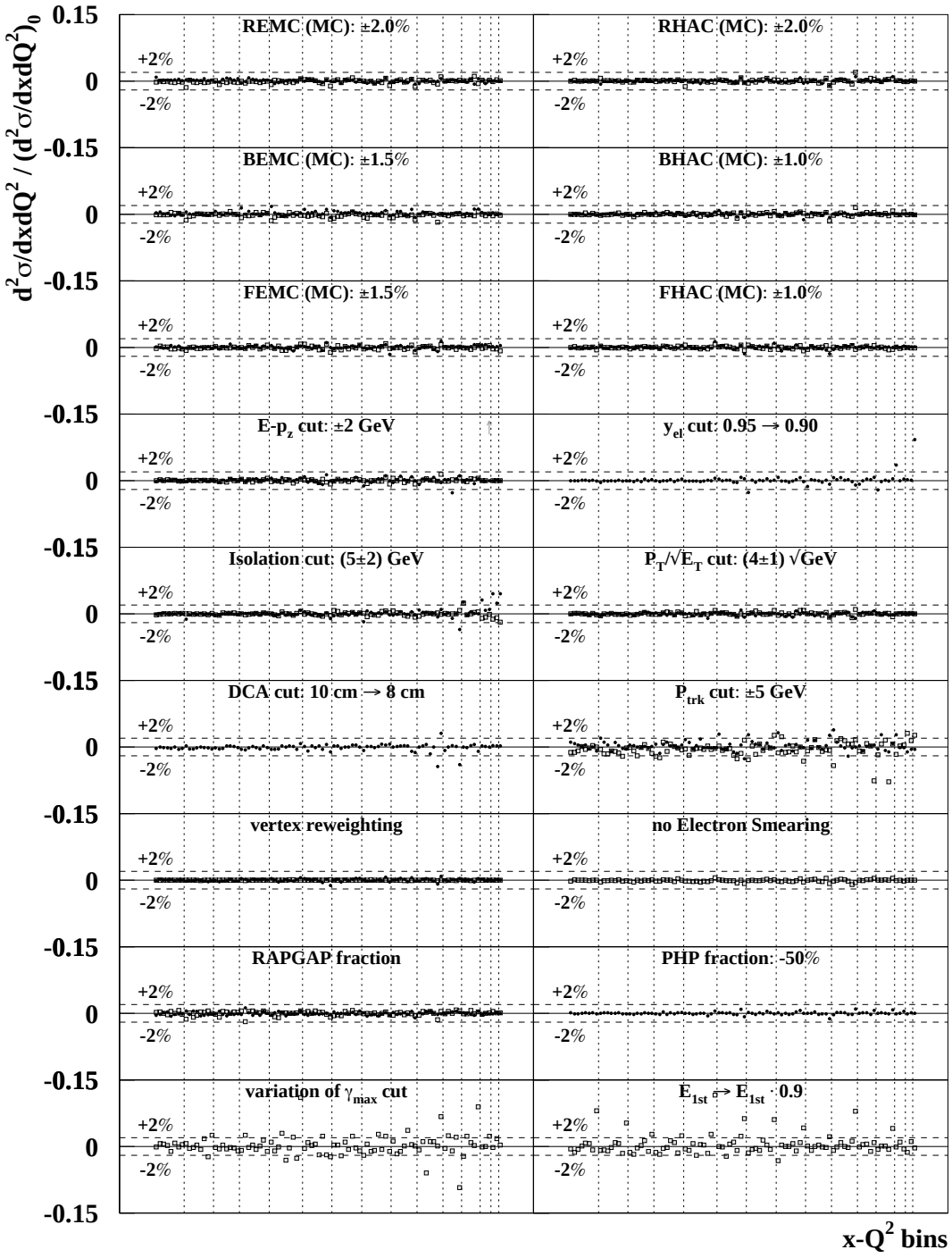


Figure C.8: *Ratio between the nominal $d\sigma/dxdQ^2$ cross section, $(d\sigma/dxdQ^2)_0$, and the same cross section after applying a systematic check, as a function of the x-Q^2 bin. The bins are ordered by increasing x within a Q^2 bin, marked with the vertical dashed lines. In plots where a quantity is varied in two directions the shaded squares represent the variation up and the full circles the variation down. Variations beyond the ordinate range are marked with an arrow. The dashed horizontal lines indicate a variation of* $\pm 2\%$.

Appendix D

Measured cross sections and xF_3 values

D.1 Single-differential cross sections: $d\sigma/dQ^2$

Q^2 range [GeV^2]	Q_c^2 [GeV^2]	$d\sigma/dQ^2$ [pb / GeV^2] measured	$d\sigma/dQ^2$ [pb / GeV^2] SM	N_{obs}	N_{bg}	$\mathcal{A}$
200.0 – 300.0	250	$11.250 \pm 0.103\,^{+0.101}_{-0.094}$	11.220	16080	2.3	0.83
300.0 – 400.0	350	$5.054 \pm 0.071\,^{+0.020}_{-0.060}$	5.022	6827	2.0	0.80
400.0 – 475.7	440	$2.855 \pm 0.059\,^{+0.026}_{-0.015}$	2.890	2851	0.7	0.77
475.7 – 565.7	520	$1.887 \pm 0.047\,^{+0.014}_{-0.014}$	1.924	1946	0.9	0.68
565.7 – 672.7	620	$1.194 \pm 0.038\,^{+0.014}_{-0.017}$	1.251	1214	0.8	0.56
672.7 – 800.0	730	$(8.85 \pm 0.28\,^{+0.14}_{-0.05}) \cdot 10^{-1}$	$8.36 \cdot 10^{-1}$	1234	1.3	0.65
800.0 – 951.4	870	$(5.27 \pm 0.17\,^{+0.06}_{-0.04}) \cdot 10^{-1}$	$5.41 \cdot 10^{-1}$	1197	1.0	0.88
951.4 – 1131.0	1040	$(3.39 \pm 0.12\,^{+0.03}_{-0.02}) \cdot 10^{-1}$	$3.47 \cdot 10^{-1}$	992	0.9	0.95
1131.0 – 1345.0	1230	$(2.16 \pm 0.09\,^{+0.03}_{-0.02}) \cdot 10^{-1}$	$2.28 \cdot 10^{-1}$	764	0.8	0.99
1345.0 – 1600.0	1470	$(1.54 \pm 0.07\,^{+0.01}_{-0.02}) \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$	648	0.5	0.99
1600.0 – 1903.0	1740	$(9.89 \pm 0.48\,^{+0.14}_{-0.06}) \cdot 10^{-2}$	$9.46 \cdot 10^{-2}$	498	1.0	0.97
1903.0 – 2263.0	2100	$(5.35\,^{+0.32}_{-0.30}\,^{+0.08}_{-0.02}) \cdot 10^{-2}$	$5.86 \cdot 10^{-2}$	335	0.7	0.98
2263.0 – 2691.0	2500	$(3.83\,^{+0.25}_{-0.23}\,^{+0.08}_{-0.02}) \cdot 10^{-2}$	$3.75 \cdot 10^{-2}$	284	0.6	0.97
2691.0 – 3200.0	2900	$(2.48\,^{+0.19}_{-0.18}\,^{+0.04}_{-0.03}) \cdot 10^{-2}$	$2.56 \cdot 10^{-2}$	198	0.3	0.97
3200.0 – 4525.0	3800	$(1.38\,^{+0.08}_{-0.08}\,^{+0.01}_{-0.03}) \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$	308	0.1	0.98
4525.0 – 6400.0	5400	$(5.14\,^{+0.44}_{-0.41}\,^{+0.10}_{-0.07}) \cdot 10^{-3}$	$5.08 \cdot 10^{-3}$	160	0.1	0.97
6400.0 – 9051.0	7600	$(2.19\,^{+0.25}_{-0.23}\,^{+0.04}_{-0.03}) \cdot 10^{-3}$	$2.03 \cdot 10^{-3}$	93	0.2	0.94
9051.0 – 12800.0	10800	$(5.9\,^{+1.2}_{-1.0}\,^{+0.3}_{-0.1}) \cdot 10^{-4}$	$7.6 \cdot 10^{-4}$	35	0.0	0.93
12800.0 – 18100.0	15200	$(3.2\,^{+0.8}_{-0.6}\,^{+0.2}_{-0.1}) \cdot 10^{-4}$	$2.7 \cdot 10^{-4}$	25	0.0	0.90
18100.0 – 25600.0	21500	$(7.0\,^{+3.5}_{-2.5}\,^{+0.6}_{-0.4}) \cdot 10^{-5}$	$8.3 \cdot 10^{-5}$	7	0.1	0.93
25600.0 – 36200.0	30400	$(2.0\,^{+2.0}_{-1.1}\,^{+0.3}_{-0.0}) \cdot 10^{-5}$	$2.0 \cdot 10^{-5}$	3	0.0	0.92
36200.0 – 51200.0	43100	$(5.4\,^{+12.3}_{-4.3}\,^{+0.8}_{-0.1}) \cdot 10^{-6}$	$3.3 \cdot 10^{-6}$	1	0.0	0.91

Table D.1: *Table of cross sections $d\sigma/dQ^2$. The first column lists the bin boundaries, whereas in column two Q_c^2 denotes the quoted Q^2 value for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the MC predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data (N_{obs}), the number of background events (N_{bg}) and the acceptance ($\mathcal{A}$) in each bin.*

D.2 Single-differential cross sections: $d\sigma/dx$

Q^2 cut [GeV^2]	x range	x_c	$d\sigma/dx$ [pb]		N_{obs}	N_{bg}	$\mathcal{A}$
			measured	SM			
200	$(0.63 - 1.00) \cdot 10^{-2}$	$0.794 \cdot 10^{-2}$	$(8.00 \pm 0.15\,^{+0.11}_{-0.15}) \cdot 10^4$	$8.02 \cdot 10^4$	3859	3.0	0.80
	$(0.10 - 0.16) \cdot 10^{-1}$	$0.126 \cdot 10^{-1}$	$(5.59 \pm 0.09\,^{+0.05}_{-0.10}) \cdot 10^4$	$5.40 \cdot 10^4$	4565	4.1	0.83
	$(0.16 - 0.25) \cdot 10^{-1}$	$0.200 \cdot 10^{-1}$	$(3.40 \pm 0.06\,^{+0.03}_{-0.03}) \cdot 10^4$	$3.38 \cdot 10^4$	4396	1.9	0.86
	$(0.25 - 0.40) \cdot 10^{-1}$	$0.316 \cdot 10^{-1}$	$(2.03 \pm 0.03\,^{+0.02}_{-0.02}) \cdot 10^4$	$2.02 \cdot 10^4$	4519	1.3	0.89
	$(0.40 - 0.63) \cdot 10^{-1}$	$0.501 \cdot 10^{-1}$	$(1.16 \pm 0.02\,^{+0.01}_{-0.01}) \cdot 10^4$	$1.17 \cdot 10^4$	4034	0.2	0.89
	$(0.63 - 1.00) \cdot 10^{-1}$	$0.794 \cdot 10^{-1}$	$(6.43 \pm 0.12\,^{+0.04}_{-0.10}) \cdot 10^3$	$6.65 \cdot 10^3$	3750	0.2	0.94
	0.10 – 0.16	0.126	$(3.56 \pm 0.07\,^{+0.03}_{-0.02}) \cdot 10^3$	$3.66 \cdot 10^3$	3500	0.1	0.98
	0.16 – 0.25	0.200	$(1.97 \pm 0.04\,^{+0.08}_{-0.02}) \cdot 10^3$	$1.89 \cdot 10^3$	2940	0.1	0.98
	0.25 – 0.40	0.316	$(8.37 \pm 0.30\,^{+0.73}_{-0.02}) \cdot 10^2$	$8.15 \cdot 10^2$	999	0.0	0.49
2 500	$(0.25 - 0.40) \cdot 10^{-1}$	$0.316 \cdot 10^{-1}$	$(2.82\,^{+0.45}_{-0.39}\,^{+0.25}_{-0.27}) \cdot 10^2$	$2.59 \cdot 10^2$	54	0.8	0.98
	$(0.40 - 0.63) \cdot 10^{-1}$	$0.501 \cdot 10^{-1}$	$(4.23\,^{+0.38}_{-0.35}\,^{+0.06}_{-0.11}) \cdot 10^2$	$3.67 \cdot 10^2$	149	0.1	0.96
	$(0.63 - 1.00) \cdot 10^{-1}$	$0.794 \cdot 10^{-1}$	$(2.80\,^{+0.24}_{-0.22}\,^{+0.04}_{-0.02}) \cdot 10^2$	$3.03 \cdot 10^2$	166	0.1	0.96
	0.10 – 0.16	0.126	$(2.02\,^{+0.16}_{-0.15}\,^{+0.03}_{-0.03}) \cdot 10^2$	$2.06 \cdot 10^2$	198	0.1	0.96
	0.16 – 0.25	0.200	$(1.22\,^{+0.10}_{-0.09}\,^{+0.03}_{-0.01}) \cdot 10^2$	$1.18 \cdot 10^2$	186	0.1	0.97
	0.25 – 0.40	0.316	$(5.74\,^{+0.54}_{-0.49}\,^{+0.03}_{-0.08}) \cdot 10^1$	$5.25 \cdot 10^1$	139	0.0	0.97
	0.40 – 0.63	0.501	$(1.31\,^{+0.21}_{-0.18}\,^{+0.04}_{-0.04}) \cdot 10^1$	$1.25 \cdot 10^1$	51	0.0	0.96
10 000	0.10 – 0.16	0.126	$4.49\,^{+4.35}_{-2.40}\,^{+1.92}_{-0.32}$	13.20	3	0.0	0.87
	0.16 – 0.25	0.200	$(2.06\,^{+0.49}_{-0.40}\,^{+0.12}_{-0.04}) \cdot 10^1$	$1.66 \cdot 10^1$	25	0.1	0.90
	0.25 – 0.40	0.316	$(1.10\,^{+0.27}_{-0.22}\,^{+0.04}_{-0.04}) \cdot 10^1$	$0.96 \cdot 10^1$	25	0.0	0.92
	0.40 – 0.63	0.501	$2.73\,^{+1.14}_{-0.82}\,^{+0.28}_{-0.01}$	2.53	10	0.0	0.93

Table D.2: *Table of cross sections $d\sigma/dx$ for lower Q^2 cuts of 200, 2 500 and 10 000 GeV^2. The second column lists the bin boundaries, whereas in column two x_c denotes the quoted Q^2 value for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data (N_{obs}), the number of background events (N_{bg}) and the acceptance ($\mathcal{A}$) in each bin.*

D.3 Single-differential cross sections: $d\sigma/dy$

Q^2 cut [GeV^2]	y range	y_c	$d\sigma/dy$ [pb]		N_{obs}	N_{bg}	$\mathcal{A}$
			measured	SM			
200	0.05 – 0.10	0.075	$(7.46 \pm 0.11\,^{+0.08}_{-0.03}) \cdot 10^3$	$7.50 \cdot 10^3$	5764	0.1	0.89
	0.10 – 0.15	0.125	$(5.27 \pm 0.10\,^{+0.02}_{-0.05}) \cdot 10^3$	$5.21 \cdot 10^3$	3899	0.0	0.87
	0.15 – 0.20	0.175	$(4.17 \pm 0.09\,^{+0.02}_{-0.04}) \cdot 10^3$	$4.02 \cdot 10^3$	2956	0.1	0.84
	0.20 – 0.25	0.225	$(3.31 \pm 0.08\,^{+0.03}_{-0.05}) \cdot 10^3$	$3.28 \cdot 10^3$	2325	0.2	0.83
	0.25 – 0.30	0.275	$(2.67 \pm 0.07\,^{+0.02}_{-0.04}) \cdot 10^3$	$2.75 \cdot 10^3$	1874	0.2	0.83
	0.30 – 0.35	0.325	$(2.43 \pm 0.07\,^{+0.02}_{-0.05}) \cdot 10^3$	$2.36 \cdot 10^3$	1662	0.0	0.82
	0.35 – 0.40	0.375	$(2.11 \pm 0.06\,^{+0.02}_{-0.03}) \cdot 10^3$	$2.06 \cdot 10^3$	1467	0.9	0.84
	0.40 – 0.45	0.425	$(1.77 \pm 0.06\,^{+0.01}_{-0.03}) \cdot 10^3$	$1.82 \cdot 10^3$	1213	0.4	0.81
	0.45 – 0.50	0.475	$(1.67 \pm 0.06\,^{+0.02}_{-0.03}) \cdot 10^3$	$1.62 \cdot 10^3$	1140	0.8	0.81
	0.50 – 0.55	0.525	$(1.38 \pm 0.05\,^{+0.02}_{-0.03}) \cdot 10^3$	$1.45 \cdot 10^3$	920	0.6	0.77
	0.55 – 0.60	0.575	$(1.30 \pm 0.05\,^{+0.03}_{-0.01}) \cdot 10^3$	$1.31 \cdot 10^3$	852	1.2	0.77
	0.60 – 0.65	0.625	$(1.19 \pm 0.05\,^{+0.03}_{-0.02}) \cdot 10^3$	$1.20 \cdot 10^3$	700	1.0	0.69
	0.65 – 0.70	0.675	$(1.17 \pm 0.05\,^{+0.07}_{-0.01}) \cdot 10^3$	$1.10 \cdot 10^3$	647	1.2	0.63
	0.70 – 0.75	0.725	$(9.08 \pm 0.48\,^{+0.30}_{-0.82}) \cdot 10^2$	$10.08 \cdot 10^2$	458	1.6	0.59

Table D.3: *Table of cross sections $d\sigma/dy$ for a lower Q^2 cut of $200\,GeV^2$. The second column lists the bin boundaries, whereas in column three y_c denotes the quoted Q^2 value for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data (N_{obs}), the number of background events (N_{bg}) and the acceptance ($\mathcal{A}$) in each bin.*

Q^2 cut [GeV^2]	y range	y_c	$d\sigma/dy$ [pb]		N_{obs}	N_{bg}	$\mathcal{A}$
			measured	SM			
2 500	0.05 – 0.10	0.075	$(7.57\,^{+1.16}_{-1.02}\,^{+0.11}_{-0.51}) \cdot 10^1$	$6.59 \cdot 10^1$	57	0.0	0.92
	0.10 – 0.15	0.125	$(1.04\,^{+0.13}_{-0.11}\,^{+0.05}_{-0.03}) \cdot 10^2$	$0.98 \cdot 10^2$	88	0.0	0.99
	0.15 – 0.20	0.175	$(1.04\,^{+0.12}_{-0.11}\,^{+0.02}_{-0.01}) \cdot 10^2$	$0.99 \cdot 10^2$	88	0.0	0.98
	0.20 – 0.25	0.225	$(8.41\,^{+1.10}_{-0.98}\,^{+0.17}_{-0.33}) \cdot 10^1$	$9.22 \cdot 10^1$	75	0.0	0.97
	0.25 – 0.30	0.275	$(9.12\,^{+1.16}_{-1.04}\,^{+0.26}_{-0.18}) \cdot 10^1$	$8.45 \cdot 10^1$	79	0.0	1.01
	0.30 – 0.35	0.325	$(7.46\,^{+1.05}_{-0.93}\,^{+0.13}_{-0.12}) \cdot 10^1$	$7.72 \cdot 10^1$	65	0.0	0.96
	0.35 – 0.40	0.375	$(8.76\,^{+1.12}_{-1.01}\,^{+0.11}_{-0.11}) \cdot 10^1$	$7.06 \cdot 10^1$	78	0.0	0.99
	0.40 – 0.45	0.425	$(4.05\,^{+0.79}_{-0.67}\,^{+0.51}_{-0.06}) \cdot 10^1$	$6.47 \cdot 10^1$	36	0.0	0.98
	0.45 – 0.50	0.475	$(5.97\,^{+0.95}_{-0.83}\,^{+0.13}_{-0.08}) \cdot 10^1$	$5.96 \cdot 10^1$	53	0.0	0.96
	0.50 – 0.55	0.525	$(7.12\,^{+1.07}_{-0.94}\,^{+0.13}_{-0.06}) \cdot 10^1$	$5.51 \cdot 10^1$	59	0.0	0.95
	0.55 – 0.60	0.575	$(6.24\,^{+0.98}_{-0.86}\,^{+0.07}_{-0.08}) \cdot 10^1$	$5.11 \cdot 10^1$	54	0.0	1.02
	0.60 – 0.65	0.625	$(3.61\,^{+0.77}_{-0.65}\,^{+0.26}_{-0.52}) \cdot 10^1$	$4.76 \cdot 10^1$	31	0.0	1.00
	0.65 – 0.70	0.675	$(4.58\,^{+0.88}_{-0.75}\,^{+0.27}_{-0.06}) \cdot 10^1$	$4.45 \cdot 10^1$	38	0.0	0.92
	0.70 – 0.75	0.725	$(3.90\,^{+0.83}_{-0.70}\,^{+0.19}_{-0.12}) \cdot 10^1$	$4.19 \cdot 10^1$	31	0.2	0.96
	0.75 – 0.80	0.775	$(4.27\,^{+0.87}_{-0.73}\,^{+0.27}_{-0.14}) \cdot 10^1$	$3.95 \cdot 10^1$	34	0.0	0.89
	0.80 – 0.85	0.825	$(4.01\,^{+0.83}_{-0.70}\,^{+0.15}_{-0.15}) \cdot 10^1$	$3.75 \cdot 10^1$	32	0.1	0.98
	0.85 – 0.90	0.875	$(3.82\,^{+0.86}_{-0.72}\,^{+0.37}_{-0.65}) \cdot 10^1$	$3.57 \cdot 10^1$	28	0.4	0.95
	0.90 – 0.95	0.925	$(3.67\,^{+1.04}_{-0.83}\,^{+0.37}_{-3.44}) \cdot 10^1$	$3.41 \cdot 10^1$	19	0.4	0.70
10 000	0.20 – 0.30	0.250	$2.34\,^{+1.87}_{-1.12}\,^{+0.73}_{-0.01}$	2.78	4	0.0	1.03
	0.30 – 0.40	0.350	$6.24\,^{+2.61}_{-1.88}\,^{+0.09}_{-0.02}$	4.77	10	0.0	0.93
	0.40 – 0.50	0.450	$6.96\,^{+2.74}_{-2.01}\,^{+0.11}_{-0.08}$	5.68	11	0.0	0.92
	0.50 – 0.60	0.550	$6.43\,^{+2.68}_{-1.94}\,^{+0.18}_{-0.07}$	5.97	10	0.0	0.88
	0.60 – 0.70	0.650	$6.78\,^{+2.67}_{-1.96}\,^{+0.13}_{-0.43}$	5.95	11	0.0	0.92
	0.70 – 0.80	0.750	$6.89\,^{+2.71}_{-1.99}\,^{+0.50}_{-0.09}$	5.80	11	0.0	0.92
	0.80 – 0.90	0.850	$3.82\,^{+2.33}_{-1.55}\,^{+0.69}_{-0.05}$	5.61	5	0.1	0.93
	0.90 – 1.00	0.950	$1.99\,^{+4.59}_{-1.60}\,^{+0.51}_{-5.42}$	5.41	1	0.0	0.68

Table D.4: *Table of cross sections $d\sigma/dy$ for lower Q^2 cuts of 2 500 and 10 000 GeV^2. The second column lists the bin boundaries, whereas in column three y_c denotes the quoted Q^2 value for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data (N_{obs}), the number of background events (N_{bg}) and the acceptance ($\mathcal{A}$) in each bin.*

D.4 Double-differential cross sections: $d^2\sigma/dxdQ^2$

Q^2 range [GeV^2]	x range	Q_c^2 [GeV^2]	x_c	$d^2\sigma/\left(dx\,dQ^2\right)$ [pb / GeV^2] measured	SM	$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{A}$
185.0 – 240.0	$(0.37 - 0.60)\cdot 10^{-2}$	200	$0.50\cdot 10^{-2}$	$1.155 \pm 0.032\ ^{+0.019}_{-0.015}$	1.125	1746	1.6	0.90
	$(0.60 - 1.00)\cdot 10^{-2}$	200	$0.80\cdot 10^{-2}$	$(9.65 \pm 0.26\ ^{+0.13}_{-0.17})\cdot 10^{-1}$	$9.49\cdot 10^{-1}$	1829	0.2	0.91
	$(0.10 - 0.17)\cdot 10^{-1}$	200	$0.13\cdot 10^{-1}$	$(8.08 \pm 0.22\ ^{+0.07}_{-0.09})\cdot 10^{-1}$	$8.01\cdot 10^{-1}$	1779	0.1	0.90
	$(0.17 - 0.25)\cdot 10^{-1}$	200	$0.21\cdot 10^{-1}$	$(7.07 \pm 0.24\ ^{+0.05}_{-0.04})\cdot 10^{-1}$	$6.79\cdot 10^{-1}$	1239	0.0	0.92
	$(0.25 - 0.37)\cdot 10^{-1}$	200	$0.32\cdot 10^{-1}$	$(5.89 \pm 0.21\ ^{+0.08}_{-0.09})\cdot 10^{-1}$	$5.90\cdot 10^{-1}$	1155	0.0	0.97
	$(0.37 - 0.60)\cdot 10^{-1}$	200	$0.50\cdot 10^{-1}$	$(5.07 \pm 0.17\ ^{+0.06}_{-0.04})\cdot 10^{-1}$	$5.13\cdot 10^{-1}$	1296	0.0	0.99
	$(0.60 - 1.20)\cdot 10^{-1}$	200	$0.80\cdot 10^{-1}$	$(4.24 \pm 0.12\ ^{+0.04}_{-0.04})\cdot 10^{-1}$	$4.43\cdot 10^{-1}$	1663	0.0	1.09
	0.12 – 0.25	200	0.18	$(3.33 \pm 0.11\ ^{+0.27}_{-0.04})\cdot 10^{-1}$	$3.28\cdot 10^{-1}$	1386	0.0	1.06
240.0 – 310.0	$(0.37 - 0.60)\cdot 10^{-2}$	250	$0.50\cdot 10^{-2}$	$1.11 \pm 0.05\ ^{+0.03}_{-0.03}$	1.16	757	0.7	0.59
	$(0.60 - 1.00)\cdot 10^{-2}$	250	$0.80\cdot 10^{-2}$	$(9.90 \pm 0.32\ ^{+0.12}_{-0.19})\cdot 10^{-1}$	$9.73\cdot 10^{-1}$	1330	0.2	0.90
	$(0.10 - 0.17)\cdot 10^{-1}$	250	$0.13\cdot 10^{-1}$	$(8.68 \pm 0.28\ ^{+0.07}_{-0.19})\cdot 10^{-1}$	$8.18\cdot 10^{-1}$	1361	0.0	0.89
	$(0.17 - 0.25)\cdot 10^{-1}$	250	$0.21\cdot 10^{-1}$	$(6.96 \pm 0.27\ ^{+0.05}_{-0.09})\cdot 10^{-1}$	$6.90\cdot 10^{-1}$	910	0.0	0.92
	$(0.25 - 0.37)\cdot 10^{-1}$	250	$0.32\cdot 10^{-1}$	$(6.11 \pm 0.24\ ^{+0.08}_{-0.08})\cdot 10^{-1}$	$5.98\cdot 10^{-1}$	885	0.0	0.95
	$(0.37 - 0.60)\cdot 10^{-1}$	250	$0.50\cdot 10^{-1}$	$(5.22 \pm 0.20\ ^{+0.11}_{-0.08})\cdot 10^{-1}$	$5.17\cdot 10^{-1}$	964	0.0	0.95
	$(0.60 - 1.20)\cdot 10^{-1}$	250	$0.80\cdot 10^{-1}$	$(4.37 \pm 0.15\ ^{+0.03}_{-0.16})\cdot 10^{-1}$	$4.44\cdot 10^{-1}$	1218	0.0	1.01
	0.12 – 0.25	250	0.18	$(3.20 \pm 0.11\ ^{+0.19}_{-0.02})\cdot 10^{-1}$	$3.26\cdot 10^{-1}$	1134	0.0	1.16
310.0 – 410.0	$(0.60 - 1.00)\cdot 10^{-2}$	350	$0.80\cdot 10^{-2}$	$1.01 \pm 0.05\ ^{+0.01}_{-0.01}$	1.01	677	0.6	0.59
	$(0.10 - 0.17)\cdot 10^{-1}$	350	$0.13\cdot 10^{-1}$	$(8.56 \pm 0.32\ ^{+0.05}_{-0.23})\cdot 10^{-1}$	$8.43\cdot 10^{-1}$	978	0.1	0.84
	$(0.17 - 0.25)\cdot 10^{-1}$	350	$0.21\cdot 10^{-1}$	$(7.30 \pm 0.32\ ^{+0.06}_{-0.15})\cdot 10^{-1}$	$7.07\cdot 10^{-1}$	750	0.1	0.91
	$(0.25 - 0.37)\cdot 10^{-1}$	350	$0.32\cdot 10^{-1}$	$(5.91 \pm 0.27\ ^{+0.05}_{-0.06})\cdot 10^{-1}$	$6.09\cdot 10^{-1}$	683	0.0	0.93
	$(0.37 - 0.60)\cdot 10^{-1}$	350	$0.50\cdot 10^{-1}$	$(5.28 \pm 0.22\ ^{+0.11}_{-0.02})\cdot 10^{-1}$	$5.23\cdot 10^{-1}$	786	0.0	0.94
	$(0.60 - 1.20)\cdot 10^{-1}$	350	$0.80\cdot 10^{-1}$	$(4.40 \pm 0.17\ ^{+0.03}_{-0.05})\cdot 10^{-1}$	$4.46\cdot 10^{-1}$	997	0.1	1.01
	0.12 – 0.25	350	0.18	$(3.22 \pm 0.13\ ^{+0.09}_{-0.06})\cdot 10^{-1}$	$3.23\cdot 10^{-1}$	905	0.0	1.09
410.0 – 530.0	$(0.60 - 1.00)\cdot 10^{-2}$	450	$0.80\cdot 10^{-2}$	$1.02 \pm 0.05\ ^{+0.02}_{-0.03}$	1.04	593	1.0	0.81
	$(0.10 - 0.17)\cdot 10^{-1}$	450	$0.13\cdot 10^{-1}$	$(8.63 \pm 0.53\ ^{+0.27}_{-0.29})\cdot 10^{-1}$	$8.63\cdot 10^{-1}$	339	0.5	0.45
	$(0.17 - 0.25)\cdot 10^{-1}$	450	$0.21\cdot 10^{-1}$	$(6.85 \pm 0.42\ ^{+0.12}_{-0.04})\cdot 10^{-1}$	$7.20\cdot 10^{-1}$	347	0.0	0.67
	$(0.25 - 0.37)\cdot 10^{-1}$	450	$0.32\cdot 10^{-1}$	$(6.32 \pm 0.35\ ^{+0.11}_{-0.08})\cdot 10^{-1}$	$6.18\cdot 10^{-1}$	424	0.0	0.81
	$(0.37 - 0.60)\cdot 10^{-1}$	450	$0.50\cdot 10^{-1}$	$(5.14 \pm 0.25\ ^{+0.09}_{-0.04})\cdot 10^{-1}$	$5.28\cdot 10^{-1}$	531	0.0	0.93
	$(0.60 - 1.00)\cdot 10^{-1}$	450	$0.80\cdot 10^{-1}$	$(4.45 \pm 0.22\ ^{+0.04}_{-0.05})\cdot 10^{-1}$	$4.48\cdot 10^{-1}$	508	0.0	0.96
	0.10 – 0.17	450	0.13	$(3.90 \pm 0.21\ ^{+0.07}_{-0.06})\cdot 10^{-1}$	$3.74\cdot 10^{-1}$	459	0.0	0.95
	0.17 – 0.30	450	0.25	$(2.77 \pm 0.15\ ^{+0.07}_{-0.03})\cdot 10^{-1}$	$2.59\cdot 10^{-1}$	442	0.0	1.08
530.0 – 710.0	$(0.60 - 1.00)\cdot 10^{-2}$	650	$0.80\cdot 10^{-2}$	$(9.76 \pm 0.60\ ^{+0.33}_{-0.17})\cdot 10^{-1}$	$10.79\cdot 10^{-1}$	332	0.7	0.73
	$(0.10 - 0.17)\cdot 10^{-1}$	650	$0.13\cdot 10^{-1}$	$(9.35 \pm 0.43\ ^{+0.13}_{-0.21})\cdot 10^{-1}$	$8.93\cdot 10^{-1}$	590	0.9	0.87
	$(0.17 - 0.25)\cdot 10^{-1}$	650	$0.21\cdot 10^{-1}$	$(7.99\ ^{+0.53}_{-0.51} \pm 0.08\ ^{+0.08}_{-0.13})\cdot 10^{-1}$	$7.41\cdot 10^{-1}$	291	0.0	0.58
	$(0.25 - 0.37)\cdot 10^{-1}$	650	$0.32\cdot 10^{-1}$	$(6.00\ ^{+0.52}_{-0.49} \pm 0.19\ ^{+0.19}_{-0.13})\cdot 10^{-1}$	$6.31\cdot 10^{-1}$	176	0.0	0.42
	$(0.37 - 0.60)\cdot 10^{-1}$	650	$0.50\cdot 10^{-1}$	$(5.15\ ^{+0.43}_{-0.40} \pm 0.10\ ^{+0.10}_{-0.18})\cdot 10^{-1}$	$5.36\cdot 10^{-1}$	190	0.0	0.41
	$(0.60 - 1.00)\cdot 10^{-1}$	650	$0.80\cdot 10^{-1}$	$(4.05\ ^{+0.35}_{-0.32} \pm 0.03\ ^{+0.03}_{-0.03})\cdot 10^{-1}$	$4.51\cdot 10^{-1}$	178	0.0	0.45
	0.10 – 0.17	650	0.13	$(3.42\ ^{+0.28}_{-0.26} \pm 0.07\ ^{+0.07}_{-0.08})\cdot 10^{-1}$	$3.73\cdot 10^{-1}$	196	0.0	0.52
	0.17 – 0.30	650	0.25	$(2.60\ ^{+0.21}_{-0.20} \pm 0.07\ ^{+0.07}_{-0.07})\cdot 10^{-1}$	$2.55\cdot 10^{-1}$	198	0.0	0.60
	0.30 – 0.53	650	0.40	$(1.27\ ^{+0.19}_{-0.17} \pm 0.01\ ^{+0.21}_{-0.01})\cdot 10^{-1}$	$1.36\cdot 10^{-1}$	63	0.0	0.42

Table D.5: *Table of cross sections $d^2\sigma/dxdQ^2$ for $185\ GeV^2 < Q^2 < 710\ GeV^2$. The first and second column list the bin boundaries in Q^2 and x, whereas in column three and four Q_c^2 and x_c denote the quoted Q^2 and x values for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data ($N_{\rm obs}$), the number of background events ($N_{\rm bg}$) and the acceptance ($\mathcal{A}$) in each bin.*

Q^2 range [GeV^2]	x range	Q_c^2 [GeV^2]	x_c	$d^2\sigma/\left(dx\,dQ^2\right)$ [pb / GeV^2] measured	SM	$N_{\rm obs}$	$N_{\rm bg}$	$\mathcal{A}$
710.0 – 900.0	$(0.90 - 1.70)\cdot 10^{-2}$	800	$1.30\cdot 10^{-2}$	$(8.89 \pm 0.49^{+0.38}_{-0.02})\cdot 10^{-1}$	$9.12\cdot 10^{-1}$	404	1.1	0.94
	$(0.17 - 0.25)\cdot 10^{-1}$	800	$0.21\cdot 10^{-1}$	$(8.19^{+0.56\ +0.11}_{-0.53\ -0.11})\cdot 10^{-1}$	$7.54\cdot 10^{-1}$	276	0.0	0.97
	$(0.25 - 0.37)\cdot 10^{-1}$	800	$0.32\cdot 10^{-1}$	$(6.23^{+0.46\ +0.06}_{-0.43\ -0.15})\cdot 10^{-1}$	$6.40\cdot 10^{-1}$	237	0.0	0.91
	$(0.37 - 0.60)\cdot 10^{-1}$	800	$0.50\cdot 10^{-1}$	$(5.84^{+0.41\ +0.08}_{-0.39\ -0.03})\cdot 10^{-1}$	$5.41\cdot 10^{-1}$	266	0.0	0.80
	$(0.60 - 1.00)\cdot 10^{-1}$	800	$0.80\cdot 10^{-1}$	$(4.89^{+0.39\ +0.06}_{-0.37\ -0.14})\cdot 10^{-1}$	$4.53\cdot 10^{-1}$	207	0.0	0.71
	0.10 – 0.17	800	0.13	$(3.38^{+0.33\ +0.08}_{-0.31\ -0.12})\cdot 10^{-1}$	$3.73\cdot 10^{-1}$	141	0.0	0.62
	0.17 – 0.30	800	0.25	$(2.60^{+0.31\ +0.10}_{-0.28\ -0.06})\cdot 10^{-1}$	$2.53\cdot 10^{-1}$	98	0.0	0.52
	0.30 – 0.53	800	0.40	$(1.33^{+0.24\ +0.09}_{-0.21\ -0.06})\cdot 10^{-1}$	$1.34\cdot 10^{-1}$	45	0.0	0.45
900.0 – 1300.0	$(0.10 - 0.17)\cdot 10^{-1}$	1200	$0.14\cdot 10^{-1}$	$(9.38^{+0.63\ +0.46}_{-0.60\ -0.30})\cdot 10^{-1}$	$9.25\cdot 10^{-1}$	284	1.8	0.94
	$(0.17 - 0.25)\cdot 10^{-1}$	1200	$0.21\cdot 10^{-1}$	$(6.85^{+0.50\ +0.12}_{-0.47\ -0.06})\cdot 10^{-1}$	$7.84\cdot 10^{-1}$	237	0.0	0.97
	$(0.25 - 0.37)\cdot 10^{-1}$	1200	$0.32\cdot 10^{-1}$	$(6.04^{+0.43\ +0.06}_{-0.41\ -0.18})\cdot 10^{-1}$	$6.61\cdot 10^{-1}$	248	0.0	0.97
	$(0.37 - 0.60)\cdot 10^{-1}$	1200	$0.50\cdot 10^{-1}$	$(5.20^{+0.33\ +0.09}_{-0.31\ -0.04})\cdot 10^{-1}$	$5.55\cdot 10^{-1}$	311	0.1	0.97
	$(0.60 - 1.00)\cdot 10^{-1}$	1200	$0.80\cdot 10^{-1}$	$(4.35^{+0.28\ +0.07}_{-0.27\ -0.05})\cdot 10^{-1}$	$4.61\cdot 10^{-1}$	297	0.0	0.99
	0.10 – 0.17	1200	0.13	$(3.61^{+0.25\ +0.11}_{-0.24\ -0.00})\cdot 10^{-1}$	$3.75\cdot 10^{-1}$	264	0.0	0.97
	0.17 – 0.30	1200	0.25	$(2.65^{+0.21\ +0.03}_{-0.19\ -0.03})\cdot 10^{-1}$	$2.51\cdot 10^{-1}$	216	0.0	0.92
	0.30 – 0.53	1200	0.40	$(1.10^{+0.13\ +0.08}_{-0.12\ -0.01})\cdot 10^{-1}$	$1.31\cdot 10^{-1}$	90	0.0	0.90
1300.0 – 1800.0	$(0.17 - 0.25)\cdot 10^{-1}$	1500	$0.21\cdot 10^{-1}$	$(8.59^{+0.77\ +0.43}_{-0.72\ -0.26})\cdot 10^{-1}$	$8.06\cdot 10^{-1}$	156	1.1	0.99
	$(0.25 - 0.37)\cdot 10^{-1}$	1500	$0.32\cdot 10^{-1}$	$(7.77^{+0.65\ +0.19}_{-0.61\ -0.16})\cdot 10^{-1}$	$6.77\cdot 10^{-1}$	180	0.0	0.97
	$(0.37 - 0.60)\cdot 10^{-1}$	1500	$0.50\cdot 10^{-1}$	$(6.01^{+0.46\ +0.05}_{-0.43\ -0.10})\cdot 10^{-1}$	$5.65\cdot 10^{-1}$	208	0.1	1.01
	$(0.60 - 1.00)\cdot 10^{-1}$	1500	$0.80\cdot 10^{-1}$	$(4.49^{+0.37\ +0.03}_{-0.35\ -0.08})\cdot 10^{-1}$	$4.67\cdot 10^{-1}$	178	0.0	0.96
	0.10 – 0.15	1500	0.13	$(3.95^{+0.39\ +0.12}_{-0.36\ -0.06})\cdot 10^{-1}$	$3.78\cdot 10^{-1}$	131	0.0	0.99
	0.15 – 0.23	1500	0.18	$(3.34^{+0.34\ +0.05}_{-0.31\ -0.05})\cdot 10^{-1}$	$3.19\cdot 10^{-1}$	124	0.0	1.01
	0.23 – 0.35	1500	0.25	$(2.70^{+0.34\ +0.10}_{-0.31\ -0.00})\cdot 10^{-1}$	$2.51\cdot 10^{-1}$	85	0.0	0.94
	0.35 – 0.53	1500	0.40	$(1.33^{+0.24\ +0.06}_{-0.21\ -0.06})\cdot 10^{-1}$	$1.30\cdot 10^{-1}$	42	0.0	1.04
1800.0 – 2500.0	$(0.23 - 0.37)\cdot 10^{-1}$	2000	$0.32\cdot 10^{-1}$	$(7.39^{+0.74\ +0.31}_{-0.68\ -0.05})\cdot 10^{-1}$	$7.04\cdot 10^{-1}$	127	0.8	0.97
	$(0.37 - 0.60)\cdot 10^{-1}$	2000	$0.50\cdot 10^{-1}$	$(5.59^{+0.56\ +0.06}_{-0.51\ -0.11})\cdot 10^{-1}$	$5.85\cdot 10^{-1}$	126	0.0	0.96
	$(0.60 - 1.00)\cdot 10^{-1}$	2000	$0.80\cdot 10^{-1}$	$(4.31^{+0.44\ +0.10}_{-0.40\ -0.05})\cdot 10^{-1}$	$4.79\cdot 10^{-1}$	122	0.0	1.00
	0.10 – 0.15	2000	0.13	$(3.85^{+0.46\ +0.03}_{-0.41\ -0.04})\cdot 10^{-1}$	$3.85\cdot 10^{-1}$	92	0.0	0.96
	0.15 – 0.23	2000	0.18	$(2.80^{+0.37\ +0.10}_{-0.33\ -0.04})\cdot 10^{-1}$	$3.22\cdot 10^{-1}$	75	0.0	0.99
	0.23 – 0.35	2000	0.25	$(2.88^{+0.42\ +0.06}_{-0.38\ -0.04})\cdot 10^{-1}$	$2.51\cdot 10^{-1}$	63	0.0	0.95
	0.35 – 0.53	2000	0.40	$(1.07^{+0.25\ +0.04}_{-0.21\ -0.04})\cdot 10^{-1}$	$1.29\cdot 10^{-1}$	27	0.0	1.00
2500.0 – 3500.0	$(0.37 - 0.60)\cdot 10^{-1}$	3000	$0.50\cdot 10^{-1}$	$(6.91^{+0.80\ +0.31}_{-0.73\ -0.29})\cdot 10^{-1}$	$6.29\cdot 10^{-1}$	94	0.2	0.98
	$(0.60 - 1.00)\cdot 10^{-1}$	3000	$0.80\cdot 10^{-1}$	$(4.65^{+0.57\ +0.10}_{-0.51\ -0.01})\cdot 10^{-1}$	$5.09\cdot 10^{-1}$	86	0.0	0.99
	0.10 – 0.15	3000	0.13	$(3.71^{+0.55\ +0.09}_{-0.48\ -0.22})\cdot 10^{-1}$	$4.02\cdot 10^{-1}$	61	0.0	0.96
	0.15 – 0.23	3000	0.18	$(3.07^{+0.48\ +0.06}_{-0.42\ -0.05})\cdot 10^{-1}$	$3.32\cdot 10^{-1}$	55	0.0	0.98
	0.23 – 0.35	3000	0.25	$(2.94^{+0.51\ +0.05}_{-0.44\ -0.06})\cdot 10^{-1}$	$2.56\cdot 10^{-1}$	46	0.0	0.92
	0.35 – 0.53	3000	0.40	$(1.65^{+0.42\ +0.03}_{-0.34\ -0.09})\cdot 10^{-1}$	$1.28\cdot 10^{-1}$	24	0.0	0.92
	0.53 – 1.00	3000	0.65	$(1.85^{+0.94\ +0.22}_{-0.66\ -0.03})\cdot 10^{-2}$	$1.95\cdot 10^{-2}$	8	0.0	0.96
3500.0 – 5600.0	$(0.40 - 1.00)\cdot 10^{-1}$	5000	$0.80\cdot 10^{-1}$	$(6.31^{+0.61\ +0.05}_{-0.56\ -0.08})\cdot 10^{-1}$	$5.81\cdot 10^{-1}$	130	0.1	0.96
	0.10 – 0.15	5000	0.13	$(4.33^{+0.64\ +0.11}_{-0.57\ -0.05})\cdot 10^{-1}$	$4.47\cdot 10^{-1}$	60	0.0	0.99
	0.15 – 0.23	5000	0.18	$(4.33^{+0.60\ +0.06}_{-0.54\ -0.13})\cdot 10^{-1}$	$3.62\cdot 10^{-1}$	67	0.0	0.96
	0.23 – 0.35	5000	0.25	$(2.77^{+0.50\ +0.07}_{-0.43\ -0.03})\cdot 10^{-1}$	$2.72\cdot 10^{-1}$	42	0.0	1.02
	0.35 – 0.53	5000	0.40	$(1.61^{+0.40\ +0.04}_{-0.33\ -0.21})\cdot 10^{-1}$	$1.32\cdot 10^{-1}$	24	0.0	0.99
5600.0 – 9000.0	$(0.70 - 1.50)\cdot 10^{-1}$	8000	$1.30\cdot 10^{-1}$	$(5.54^{+0.81\ +0.18}_{-0.71\ -0.17})\cdot 10^{-1}$	$5.28\cdot 10^{-1}$	60	0.2	0.91
	0.15 – 0.23	8000	0.18	$(4.03^{+0.80\ +0.15}_{-0.68\ -0.03})\cdot 10^{-1}$	$4.17\cdot 10^{-1}$	35	0.0	0.97
	0.23 – 0.35	8000	0.25	$(3.16^{+0.75\ +0.07}_{-0.61\ -0.03})\cdot 10^{-1}$	$3.04\cdot 10^{-1}$	26	0.0	1.02
	0.35 – 0.53	8000	0.40	$(1.54^{+0.55\ +0.03}_{-0.42\ -0.12})\cdot 10^{-1}$	$1.42\cdot 10^{-1}$	13	0.0	1.00
	0.53 – 1.00	8000	0.65	$(1.9^{+1.5\ +0.2}_{-0.9\ -0.0})\cdot 10^{-2}$	$2.0\cdot 10^{-2}$	4	0.0	0.95
9000.0 – 15000.0	0.11 – 0.23	12000	0.18	$(4.5^{+1.0\ +0.2}_{-0.8\ -0.1})\cdot 10^{-1}$	$4.9\cdot 10^{-1}$	28	0.0	0.91
	0.23 – 0.35	12000	0.25	$(2.8^{+1.0\ +0.0}_{-0.8\ -0.0})\cdot 10^{-1}$	$3.5\cdot 10^{-1}$	12	0.0	0.94
	0.35 – 0.53	12000	0.40	$(1.2^{+0.7\ +0.2}_{-0.5\ -0.0})\cdot 10^{-1}$	$1.6\cdot 10^{-1}$	6	0.0	0.99
15000.0 – 25000.0	0.18 – 0.35	20000	0.25	$(4.9^{+1.7\ +0.3}_{-1.3\ -0.1})\cdot 10^{-1}$	$4.4\cdot 10^{-1}$	13	0.1	0.91
	0.35 – 1.00	20000	0.40	$(2.1^{+1.3\ +0.1}_{-0.8\ -0.0})\cdot 10^{-1}$	$1.9\cdot 10^{-1}$	6	0.0	0.92
25000.0 – 50000.0	0.30 – 1.00	30000	0.40	$(2.4^{+1.9\ +0.3}_{-1.1\ -0.1})\cdot 10^{-1}$	$2.2\cdot 10^{-1}$	4	0.0	0.93

Table D.6: *Table of cross sections $d^2\sigma/dxdQ^2$ for $710\,GeV^2 < Q^2 < 50\,000\,GeV^2$. The first and second column list the bin boundaries in Q^2 and x, whereas in column three and four Q_c^2 and x_c denote the quoted Q^2 and x values for the measured cross section. The column "measured" contains the measured cross section values with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for the cross section using CTEQ5D as the PDF parameterization. The last three columns display the number of observed events in data ($N_{\rm obs}$), the number of background events ($N_{\rm bg}$) and the acceptance ($\mathcal{A}$) in each bin.*

D.5 xF_3

Q^2 range [GeV^2]	x range	Q_c^2 [GeV^2]	x_c	xF_3		N_{e^-p}	N_{e^+p}
				measured	SM		
1300 – 2500	0.017 – 0.060	1800	0.037	$(9.7 \pm 2.7 ^{+1.2}_{-0.8}) \cdot 10^{-2}$	$4.0 \cdot 10^{-2}$	799	1112
	0.060 – 0.150	1800	0.100	$(4.6 \pm 6.2 ^{+1.4}_{-1.5}) \cdot 10^{-2}$	$5.0 \cdot 10^{-2}$	523	897
	0.150 – 0.350	1800	0.230	$(0.8 \pm 1.2 ^{+0.3}_{-0.5}) \cdot 10^{-1}$	$0.5 \cdot 10^{-1}$	347	579
	0.350 – 0.530	1800	0.430	$(-3.2 ^{+2.2}_{-2.0} ^{+0.9}_{-1.7}) \cdot 10^{-1}$	$0.2 \cdot 10^{-1}$	69	127
2500 – 5600	0.037 – 0.150	3500	0.100	$(10.1 \pm 3.5 ^{+0.9}_{-0.8}) \cdot 10^{-2}$	$8.4 \cdot 10^{-2}$	432	600
	0.150 – 0.350	3500	0.230	$(21.2 ^{+7.9}_{-7.5} ^{+1.7}_{-1.6}) \cdot 10^{-2}$	$7.4 \cdot 10^{-2}$	210	279
	0.350 – 1.000	3500	0.430	$(2.2 ^{+1.3}_{-1.2} ^{+0.3}_{-0.5}) \cdot 10^{-1}$	$0.3 \cdot 10^{-1}$	56	62
5600 – 15000	0.070 – 0.350	9000	0.230	$(12.9 ^{+3.4}_{-3.2} ^{+0.8}_{-0.5}) \cdot 10^{-2}$	$12.6 \cdot 10^{-2}$	161	162
	0.350 – 1.000	9000	0.430	$(0.8 ^{+7.3}_{-6.1} ^{+1.7}_{-1.7}) \cdot 10^{-2}$	$5.4 \cdot 10^{-2}$	23	36
15000 – 50000	0.180 – 1.000	30000	0.430	$(8.1 ^{+2.9}_{-2.3} ^{+0.7}_{-0.2}) \cdot 10^{-2}$	$7.5 \cdot 10^{-2}$	23	13

Table D.7: *Table of measured xF_3 structure function values. The first and second column list the bin boundaries in Q^2 and x, whereas in column three and four Q_c^2 and x_c denote the quoted Q^2 and x values for the measured cross section. The column "measured" contains the measured structure functions with statistical and systematic errors. If the upper and lower statistical errors differ by less than 5% only one value is given. The column SM lists the SM predictions for xF_3 function using CTEQ5D as the PDF parameterization. The last two columns display the number of observed events in data for e^-p and e^+p, respectively.*

References

[1] ZEUS Coll., M. Derrick et al., Phys. Lett. **B 316**, 412 (1993).

[2] ZEUS Coll., M. Derrick et al., Z. Phys. **C 65**, 379 (1994).

[3] ZEUS Coll., M. Derrick et al., Z. Phys. **C 69**, 607 (1996).

[4] ZEUS Coll., M. Derrick et al., Z. Phys. **C 72**, 399 (1996).

[5] ZEUS Coll., J. Breitweg et al., Phys. Lett. **B 407**, 432 (1997).

[6] ZEUS Coll., J. Breitweg et al., Phys. Lett. **B 487**, 53 (2000).

[7] ZEUS Coll., S. Chekanov et al., Preprint hep-ex-0105090, 2001.

[8] ZEUS Coll., J. Breitweg et al., Z. Phys. **C 74**, 207 (1997).

[9] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 11**, 427 (1999).

[10] ZEUS Coll., *Measurement of high-Q^2 neutral current cross sections in e^+p deep inelastic scattering at HERA*. Abstract 630, International Europhysics Conference on High Energy Physics, Budapest, Hungary, July 12-18, 2001.

[11] ZEUS Coll., *Measurement of High Q^2 Neutral Current Deep Inelastic Scattering Cross Sections in e^-p scattering at HERA*. Contributed paper 549 to EPS '99, Tampere, Finland, 1999.

[12] ZEUS Coll., *Measurement of high-Q^2 neutral current cross sections in e^-p DIS and a first measurement of the structure function xF_3 at HERA*. Abstract 1039, XXXth International Conference on High Energy Physics, Osaka, Japan, 2000.

[13] A. Kappes, *NC/CC deep inelastic ep-scattering at high Q^2 with ZEUS at HERA*, 2000, available on `http://hep.ph.liv.ac.uk/~disproc/PWG3/`. Appeared in *Proc. 8^{th} Int. Workshop on Deep-Inelastic Scattering (DIS2000), Liverpool, April 2000.*

[14] Francis Halzen, Alan D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc, 1984.

[15] P. Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*, Lecture Notes in Physics, Vol. 295. Springer, Berlin, Heidelberg, 1988.

[16] R.P. Feynman, Phys. Rev. **76**, 749 (1949);
R.P. Feynman, Phys. Rev. **76**, 769 (1949).

[17] H.N. Brown et al., Phys. Rev. Lett. **86**, 2227 (2001).

[18] D. Griffiths, *Introduction to elementary particles*. John Wiley & Sons, Inc., 1987.

[19] S.L. Glashow, Nucl. Phys. **22**, 579 (1961).

[20] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[21] A. Salam, in *Elementary Particle Physics*. 1968. Reprinted in Gauge Theory of Weak and Electromagnetic Interactions, C.H. Lai, ed.(Singapore: World Scientific, 1981).

[22] P.W. Higgs, Phys. Lett. **12**, 132 (1964);
P.W. Higgs, Phys. Rev. Lett. **13**, 508 (1964);
P.W. Higgs, Phys. Rev. **145**, 1156 (1966).

[23] P.W. Anderson, Phys. Rev. **130**, 439 (1963).

[24] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[25] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Vol. 8. Cambridge University Press, 1996.

[26] W.A. Bardeen et al., Phys. Rev. **D 18**, 3998 (1978).

[27] G. Altarelli, R.K. Ellis and G. Martinelli, Nucl. Phys. **B 157**, 461 (1979).

[28] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977);
V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972);
L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975);
Yu.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[29] E.D. Bloom et al., Phys. Rev. Lett. **23**, 930 (1969).

[30] M. Breidenbach et al., Phys. Rev. Lett. **23**, 935 (1969).

[31] F. Dydak, in *Proceedings of the 1983 International Symposium on Lepton and Photon Interactions at High Energies (LP83), Cornell, New York, USA, Aug. 1983*, eds. D.G. Cassel and D.L. Kreinick, p. 634. F.R. Newman Laboratory of Nuclear Studies, Ithaca, New York 14853, USA, Dec. 1983.

[32] C.H.L. Smith, Phys. Rep. **3C**, 261 (1972).

[33] F.E. Close, *An Introduction to Quarks and Partons*. Academic Press Inc. (London) Ltd, 24–28 Oval Road, London NW1, 1979.

[34] CTEQ Coll., H.L. Lai et al., Eur. Phys. J. **C 12**, 375 (2000).

[35] J. Pumplin et al., Preprint hep-ph/0201195, 2002.

[36] H.L. Lai et al., Phys. Rev. **D 55**, 1280 (1997).

[37] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. **C 4**, 463 (1998).

[38] A.D. Martin et al., Preprint hep-ph/0201127, 2002.

[39] R.S. Thorne et al., Preprint hep-ph-0106075, 2001.

[40] M. Botje, Eur. Phys. J. **C 14**, 285 (2000).

[41] ZEUS Coll., *The ZEUS NLO QCD fit to determine parton distribution functions and* α_s, 2001. Contribution to EPS01.

[42] H. Spiesberger, J. Phys. **G 19**, 1469 (1993).

[43] G. Kramer and H. Spiesberger, in *Proc. Workshop on Physics at HERA*, Vol. 2, p. 789. W. Buchmüller and G. Ingelman, Hamburg, Germany, DESY, 1991.

[44] A. Courau and P. Kessler, Phys. Rev. **D 46**, 117 (1992).

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

[46] J.J. Sakurai, Phys. Rev. Lett. **22**, 981 (1969).

[47] K. Wille, *Physik der Teilchenbeschleuniger und Synchrotronstrahlungsquellen*. B.G. Teubner Stuttgart, 1992.

[48] ZEUS Coll., U. Holm (ed.), *The ZEUS Detector*. Status Report (unpublished), DESY, 1993, available on `http://www-zeus.desy.de/bluebook/bluebook.html`.

[49] A. Garfagnini, Nucl. Inst. Meth. **A 435**, 34 (1999).

[50] I.C. Brock, *A Straw-Tube Tracker for ZEUS* (unpublished). ZEUS-98-056, internal ZEUS-note, 1998.

[51] J. Ng and W. Verkerke, *An Overview of SRTD Analysis* (unpublished). ZEUS-95-37.

[52] A. Weber, *Der Silizium-Spurdetektor BPT für kleinste Elektron-Streuwinkel in ZEUS: Alignment mit Teilchenspuren*. Diploma Thesis, Bonn University, 1998.

[53] C. Amelung, *Measurement of the Proton Structure Function F_2 at Very Low Q^2 at HERA*. Doctoral Thesis, Universität Bonn, Bonn, (Germany), Report BONN-IR-99-14, DESY-THESIS-2000-002, 1999.

[54] M. Grothe, *Differentielle Wirkungsquerschnitte diffraktiver tiefinelastischer ep-Streuung bei HERA: Messung mit Nachweis des gestreuten Protons im Leading Proton Spectrometer LPS von ZEUS*. Doctoral Thesis, Bonn University, 1998.

[55] M. Moritz, Private communication, 2001. `http://www-zeus.desy.de/~moritz/ZEUS_ONLY/nc/fpc.html`.

[56] C. Catterall and F. Pelucchi, *Comparison of CTD and Calorimeter Measurement of the DIS Positron Angle with High Q^2 Events* (unpublished). ZEUS-97-56, July 1997.

[57] R. Hall-Wilton et al., *The CTD Tracking Resolution* (unpublished). ZEUS-99-024, internal ZEUS-note, 1999.

[58] A. Bornheim, *Kalibration des Presamplers für die Elektron-Energiemessung im ZEUS-Detektor*. Diploma Thesis, University of Bonn, 1995.

[59] A. Dwurazny et al., Nucl. Inst. Meth. **A 277**, 176 (1989).

[60] K. Piotrzkowski and M. Zachara, *Determination of the ZEUS Luminosity in 1994* (unpublished). ZEUS-95-138, 1995.

[61] K. Olkiewicz and A. Eskreys, *Off-line Luminosity Calculation in the ZEUS Experiment in 1997, 1998 and 1999* (unpublished). ZEUS-99-44, 1999.

[62] Z. Qian et al., Comp. Phys. Comm. **45**, 283 (1987).

[63] H. Spiesberger, HERACLES *and* DJANGOH*: Event Generation for ep Interactions at HERA Including Radiative Processes*, 1998, available on `http://www.desy.de/~hspiesb/djangoh.html`.

[64] H. Spiesberger, *An Event Generator for ep Interactions at HERA Including Radiative Processes (Version 4.6)*, 1996, available on `http://www.desy.de/~hspiesb/heracles.html`.

[65] G. Ingelman, A. Edin and J. Rathsman, Comp. Phys. Comm. **101**, 108 (1997).

[66] L. Lönnblad, Comp. Phys. Comm. **71**, 15 (1992).

[67] T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994).

[68] H. Jung, Comp. Phys. Comm. **86**, 147 (1995).

[69] G. Marchesini et al., Comp. Phys. Comm. **67**, 465 (1992).

[70] R. Brun et al., GEANT3, Technical Report CERN-DD/EE/84-1, CERN, 1987.

[71] M. Martinez, *Calorimeter Noisy Cell Study for 1998-2000 ZEUS Data* (unpublished). ZEUS-00-016, internal ZEUS-note, 2000;
A. Savin, *Study of Calorimeter Noise in the 1996 Data* (unpublished). ZEUS-98-007.

[72] R. Deffner, *Measurement of the Proton Structure Function F_2 at HERA using the 1996 and 1997 ZEUS Data.* Ph.D. Thesis, University of Bonn, 1999.

[73] P. Billoir and S. Qian, Nucl. Inst. Meth. **A 311**, 139 (1992).

[74] A. López-Durán Viani and S. Schlenstedt, *Electron finder efficiencies and impurities. A comparison between SINISTRA95, EM and EMNET* (unpublished). ZEUS-99-077, internal ZEUS note, 1999.

[75] A. Kappes, *Verwendung von Neuronalen Netzen zur Identifikation des gestreuten Elektrons in ep-Ereignissen mit hohem Q^2 bei ZEUS.* Diploma Thesis, Universität Bonn, Bonn, Germany, Report BONN-IB-97-28, 1997.

[76] R. Sinkus and T. Voss, Nucl. Inst. Meth. **A 391**, 360 (1997).

[77] R. Sinkus, *Measurement of the Proton Structure Function F_2 from the 1994 HERA Data using a Neural Network for the Identification of the Scattered Lepton.* Thesis, Hamburg University, 1994.

[78] Ch. Amelung, *Electron Position Reconstruction in ZEUS: Further Update of the ELECPO Package (Based on 1995 Data)* (unpublished). ZEUS-96-093, internal ZEUS note, 1996;
Ch. Amelung, *Electron Position Reconstruction in ZEUS: Another Update of the ELECPO Package (Based on 1995/96 Data)* (unpublished). ZEUS-97-053, internal ZEUS note, 1997.

[79] J. Krüger, *The Uranium Scintillator Calorimeter for the ZEUS Detector at the Electron - Proton collider HERA: The Heart of ZEUS.* Habilitation Thesis, Universität Hamburg, Report DESY-F35-92-02, 1992.

[80] N. Tuning, *Proton Structure Functions at HERA*. Ph.D. Thesis, Amsterdam University, 2001.

[81] W. Liu, *Search for Resonances Decaying to e+jet and Limits on First Generation Leptoquarks in e^+p Interactions at HERA*. Ph.D. Thesis, Columbia University, 2001.

[82] J. Grosse-Knetter, *Corrections for the Hadronic Final State* (unpublished). ZEUS-98-031, internal ZEUS-note, 1998.

[83] A. López-Durán Viani, *Measurement of the Proton Structure Function xF_3 in High Q^2 DIS Events Using ZEUS Data*. Ph.D. Thesis, Humboldt-Universität zu Berlin, 2001.

[84] F. Jacquet and A. Blondel, in *Proceedings of the Study for an ep Facility for Europe*, ed. U. Amaldi, p. 391. Hamburg, Germany, 1979. Also in preprint DESY 79/48.

[85] L. Adamczyk, private communication, Nov. 2001.

[86] M. Moritz, *in preparation*. Doctoral Thesis, University Hamburg, 2001.

[87] H. Spiesberger, DJANGO6 *Version 2.4 – A Monte Carlo Generator for Deep Inelastic Lepton Proton Scattering Including QED and QCD Radiative Effects*, 1996, available on `http://www.desy.de/~hspiesb/django6.html`.

[88] C.M. Cormack, private communication, 2001.

[89] J. Gonçalo and K. Long, *A minimum bias vertex distribution for the 1998/99 e^-p data* (unpublished). ZEUS-00-027, internal ZEUS-note, 2000.

[90] OPAL Coll., G. Abbiendi, Preprint hep-ex/0012018, Nov. 2000.

[91] W. Hollik et al., in *Proc. Workshop on Physics at HERA*, eds. W. Buchmüller and G. Ingelman, Vol. 2, p. 923. Hamburg, Germany, DESY, 1991.

[92] E. Leader and E. Predazzi, *An introduction to gauge theories and modern particle physics*, Vol. 1. Cambridge University Press, 1996.

[93] H. Spiesberger, in *Proc. Workshop on Future Physics at HERA*, eds. G. Ingelman, A. De Roeck and R. Klanner, Vol. 1, p. 227. DESY, Hamburg, Germany, 1996, available on `http://www.desy.de/~hspiesb/eprc.html`.

[94] Particle Data Group, D.E. Groom et al., Eur. Phys. J. **C15**, 1 (2000).

[95] N. Tuning, *Precise determination of the relative normalization using NC DIS data* (unpublished). ZEUS-01-027, internal ZEUS note, 2001.

[96] CCFR/NuTeV Coll., U.K. Yang et al., Preprint hep-ex/0009041, 2000.

[97] L.S. Barabash et al., Preprint hep-ex/9611912v2, 1996.

[98] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 19**, 269 (2001).

[99] R.J. Cashmore et al., in *Proc. the Workshop on Future Physics at HERA*, eds. G. Ingelman, A. De Roeck and R. Klanner, Vol. 1, p. 163. DESY, 1996.

Acknowledgments

Ich möchte an dieser Stelle allen Menschen danken, die zum Gelingen dieser Arbeit beigetragen haben. Zu allererst sei hier mein Doktorvater Erwin Hilger erwähnt. Während der 5 Jahre, die ich in seiner Gruppe mit Diplomarbeit und Promotion zubringen durfte, hatte er immer ein offenes Ohr für Fragen und Probleme und zeigte reges Interesse an meiner Arbeit, ließ mir aber auch immer den nötigen Freiraum, um eigene Ideen zu verwirklichen. Vielen Dank, Herr Hilger.

Als nächstes möchte ich Uli Katz danken. Seine stets kompetenten Ratschläge in physikalischen wie Computer- bzw. TeX-technischen Dingen zu annähernd jeder Tages- und Nachtzeit waren eine unschätzbare Hilfe, ohne die diese Arbeit im wörtlichen Sinne in dieser Form nicht möglich gewesen wäre. Uli, ich wünsche Dir alles Gute für Deine (berufliche) Zukunft.

Für die richtige Arbeitsatmosphäre innerhalb der Gruppe trugen die anderen Gruppenmitglieder bei. Hier sei Ainas Weber erwähnt, mit dem ich für einen großen Teil meiner Promotionszeit das Zimmer teilte. In dieser Zeit konnte unsere Raum um einige interessante 'Features' erweitert werden und die Aktion 'saubere Umwelt' steigerte das Sprachniveau erheblich. Julian Rautenberg möchte ich für sein Engagement bei der gemeinsamen Administration unseres Computerclusters und für die vielen interessanten Diskussionen über Computer- und sonstige Themen danken. Vor allem kurz vor Fertigstellung der Dissertation konnte ich mich dank ihm voll auf meine Arbeit konzentrieren. Robert Kerger trug durch sein gewissenhaftes Korrekturlesen ebenfalls zum Gelingen der Arbeit bei und Detlef Bartsch sorgte dafür, daß unsere allwöchentlichen Meetings nie zu kurz wurden.

Aber auch den anderen ehemaligen und jetzigen Mitgliedern, sowohl FDETlern als auch CALoris, möchte ich für die schöne Zeit in der ZEUS-Gruppe danken. Ich habe mich hier immer sehr wohl gefühlt und dazu trugen auch sicherlich unsere gelegentlich ausgetragenen, äußerst friedfertigen virtuellen Auseinandersetzungen bei (1 Frag left!).

I had the luck to be able to participate in an analysis from the very first data taking up to the final publication. However, this would not have been possible without many people at DESY that contributed either way. Among them, I would like to thank Arnulf Quadt first, who helped me starting with the analysis in autumn 1998 and only with his guidance first results were already available in spring 1999. As today's high-energy physics experiments are very complex it takes many people to perform all the studies that are necessary before results can be published. Directly involved in this work were Chris Cormack, Jason Breitweg, Henning Schnurbusch, Ricardo Gonçalo and Matthias Moritz. Ken Long coordinated the efforts that will hopefully yield fruits soon. Last but not least I would like to thank Amaya López-Durán Viani for staying the course with me during the last months, so that we can bring our analysis to publication soon.

Zuletzt möchte ich aber vor allem meiner Familie meinen tief empfundenen Dank aussprechen. Ohne ihre uneingeschränkte moralische und auch finanzielle Unterstützung während all der Studien- und Promotionsjahre hätte ich mich mit Sicherheit nicht in dem Maße meiner Arbeit widmen können, wie dies der Fall gewesen ist. Vielen, vielen Dank Ingrid, Karl und Sabine!