Measurement of Jet Production with the ZEUS Detector

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für
Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Denys Lontkovskyi

Hamburg
2015

Gutachter/in der Dissertation:	Prof. Dr. Brian Foster PD Dr. Olaf Behnke
Gutachter/in der Disputation:	Prof. Dr. Erika Garutti Prof. Dr. Gudrid Moortgart-Pick
Tag der Disputation:	23 November 2015
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Ludwig Mathey
Vorsitzender des Promotionsausschusses:	Prof. Dr. Jan Louis
Dekan des Fachbereichs Physik:	Prof. Dr. Heinrich Graener

Abstract

In this thesis the single-differential and double-differential inclusive-jet cross sections have been measured in neutral current deep inelastic ep scattering for exchanged-boson virtualities $Q^2 > 125\,\text{GeV}$ and inelasticity $0.2 < y < 0.6$. The measurements were performed using the ZEUS detector at HERA and are based on the data sample accumulated during the 2004–2007 HERA-II running period with a total integrated luminosity of 295 pb^{-1}. Jets were identified in the Breit frame using the k_T-clustering algorithm in the longitudinally invariant inclusive mode. The cross sections refer to each jet of hadrons with transverse energy in the Breit frame $E_{T,B}^{jet} > 8$ GeV and pseudorapidity in the laboratory frame $-1 < \eta_{lab}^{jet} < 2.5$. The cross sections were measured inclusively as functions of several kinematic variables and differentially in Q^2. Next-to-leading-order QCD calculations obtained using the NLOJET++ program describe the measurements well. The predictions based on different proton PDF sets were also compared to the measurements and the potential of the new data to constraint the proton PDFs is demonstrated. The value of the strong coupling at the scale of the mass of the Z^0 boson, $\alpha_s(M_Z)$, was determined from the measured jet cross sections and the energy-scale dependence of α_s was demonstrated. The obtained value is $\alpha_s(M_Z) = 0.1216 \pm 0.0026(\text{exp.})^{+0.0093}_{-0.0074}(\text{th.})$

Kurzfassung

In dieser Doktorarbeit wurde die inklusive Jetproduktion in tiefinelastischer ep Streuung mit Austausch neutraler Eichbosonen (Photonen und Z^0) bei HERA untersucht. Es wurden sowohl einfach- als auch doppeltdifferenzielle Produktionswirkungsquerschnitte der Jetproduktion gemessen in einem kinematischen Bereich von hohen Virtualitäten $Q^2 > 125$ GeV^2 des ausgetauschten Eichboson und mit der sogenannten Inelastizität y im Bereich von $0.2 < y < 0.6$. Die Messungen wurden mit den ZEUS Detektor erzielt und basieren auf Daten die in der HERA-II Runperiode von 2004–2007 aufgenommen wurden und eine integrierte Luminosität von 295 pb^{-1}haben. Jets wurden im sogenannten "Breit-frame" mit den "k_T-Clustering" Algorithmus im longitudinalen invarianten inklusiven Modus identifiziert. Die Wirkungsquerschnitte zählen jeden hadronischen Jet mit einer transversalen Energie $E_{T,B}^{jet}$ im Breit-frame $E_{T,B}^{jet} > 8$GeV und mit einer Pseudorapidität η_{lab}^{jet} im Bereich $-1 < \eta_{lab}^{jet} < 2.5$ Die Wirkungsquerschnitte wurden einfach-differenziell als Funktion verschiedener kinematischer Observablen des gesamten Ereignis und der Jets gemessen und doppeltdifferenziell als Funktion von Q^2 und anderen kinematischen Observablen. QCD Vorhersagen in nächst-führender Ordnung der Störungsreihe der starken Kraft wurden mit dem NLOJET++ Programm gewonnen und mit den gemessenen Resultaten verglichen, wobei eine gute Übereinstimmung beobachtet wurde. Die QCD Vorhersagen wurden auch mit verschiedenen Parameterisierungen der "Parton Distribution Functions" im Proton bestimmt um das Potenzial der Daten aufzuzeigen diese Parametrisierungen festzulegen. Durch Vergleich der Daten und Vorhersage konnte die Skalenabhängigkeit des Wertes der starken Kopplung, α_s, demonstriert werden sowie ein präziser Wert bei der Skala der Masse des Z Bosons bestimmt werden von $\alpha_s(M_Z) = 0.1216 \pm 0.0026(\text{exp.})^{+0.0093}_{-0.0074}(\text{th.})$.

To my family

Contents

CHAPTER 1

Introduction

Since ancient times humanity has tried to identify the basic building blocks of Nature and find guiding principles that govern all observed phenomena. Beginning from the early 20th century, scattering experiments have played an increasingly important role in revealing the microscopic structure of matter. Thus, for example, pioneering studies of the scattering of α-particles on Gold led Rutherford [1] to the discovery of the atomic nucleus. Within the following decades it was realised that the nucleus is composed of protons and neutrons [2], which were generically named “nucleons”. Soon after that, numerous experiments, dedicated to the measurement of the anomalous magnetic moment of the proton and neutron, gave evidence for nucleon substructure [3–5]. In the 1950s Hofstadter demonstrated non-trivial[1] spatial distribution of electric charge of nucleons [6] and approximately ten years later, the investigation of high energy inelastic electron-nucleon scattering in a series of MIT-SLAC experiments [7–11] provided key evidence for nucleon substructure in terms of pointlike particles. Studies of the internal structure of nucleon culminated in the high-precision determination of the proton content performed at the ep collider HERA, where where the proton was probed by electroweak gauge bosons $\gamma, Z^0, W^{\pm}$.

The proton constituents: *quarks* and *gluons*, which are generally called *partons*, do not appear as free particles in experiment but are tightly bound inside hadrons. However, they manifest themselves in high-energy scattering experiments as ’sprays’ of hadrons, called *jets*. At HERA the production of jets can

---

[1] It was deduced from these experiments that the rms charge radius of the proton is about $\langle r_p \rangle^2 \approx 0.7\,\mathrm{fm}$

be investigated in a wide kinematic phase space, offering a unique opportunity to constrain the proton parton density functions (PDFs) which describe the effective probability density to find a parton with a fraction of longitudinal momentum, x, of the proton.

In recent decades, all experimental and theoretical findings about the interaction of the elementary constituents of matter were unified in the modern concept of the *Standard Model* (SM) [12] of particle physics. Within the SM picture quarks interact strongly via gluon exchange and the strong interaction binds the quarks and gluons in the proton. The investigation of processes involving jets can be regarded as a lab for testing the theory of the strong sector of the Standard Model — Quantum chromodynamics (QCD).

At HERA, jet production has been measured in wide variety of reactions including neutral current [13–24], charged current [25, 26] and photoproduction [27–36] processes. The data have been used for the extraction of the strong coupling as well as for constraining the proton PDFs [37].

The content of this thesis is organised as follows. An outline of the theoretical framework for the jet production and the kinematics of the deep inelastic scattering are described in Chapter 2. The HERA machine as well as the ZEUS detector are introduced in Chapter 3. The details of the final-state reconstruction including the event and jet selection criteria are presented in Chapters 4 and 5. The corrections and reweightings applied to the data and Monte Carlo are described in Chapter 6. The comparison of the MC simulations with the data is also given there. In Chapter 7, the cross section determination procedure is explained. Finally, in Chapters 8, 9 the results of the measurements of inclusive-jet cross sections and QCD analysis of the data are discussed. The thesis conduces with the summary and possible directions for future studies.

CHAPTER 2

Theoretical Framework

All microscopic phenomena observed to date can be described in the framework of the Standard Model (SM). The SM is a renormalisable quantum field theory of the strong, electromagnetic and weak interactions. According to the SM, matter particles are represented by spin-$\frac{1}{2}$ Dirac fermions coming in three generations while interaction between matter fields is mediated by spin-1 gauge bosons. The SM is based on the combined local $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)$ symmetry representing strong, weak and electromagnetic sectors, respectively. The Dirac fermions are distinguished by the quantum numbers corresponding to the gauge groups. Quark fields are endowed with electroweak and colour charges, while leptons (electron e, muon μ, tau τ and three corresponding neutrinos ν_e, ν_μ, ν_τ) carry only electroweak charge. As a consequence of exact local gauge symmetry, the mediators of the strong force, gluons, are massless quanta. In contrast to the strong interaction, three electroweak gauge bosons $W^\pm, Z^0$ acquire mass as a result of spontaneous $\mathrm{SU}_L(2) \otimes \mathrm{U}(1)$ symmetry breaking, while the photon γ stays massless. According to the Higgs mechanism, elementary particles acquire mass due to coupling to the scalar field, quanta of which have been recently discovered at the LHC [38, 39].

In this work, interpretation of the data is based on the Standard Model picture of particle physics. Subsequent sections briefly overview the theoretical framework used for the description of the strong sector of the Standard Model and hard interactions at HERA in particular.

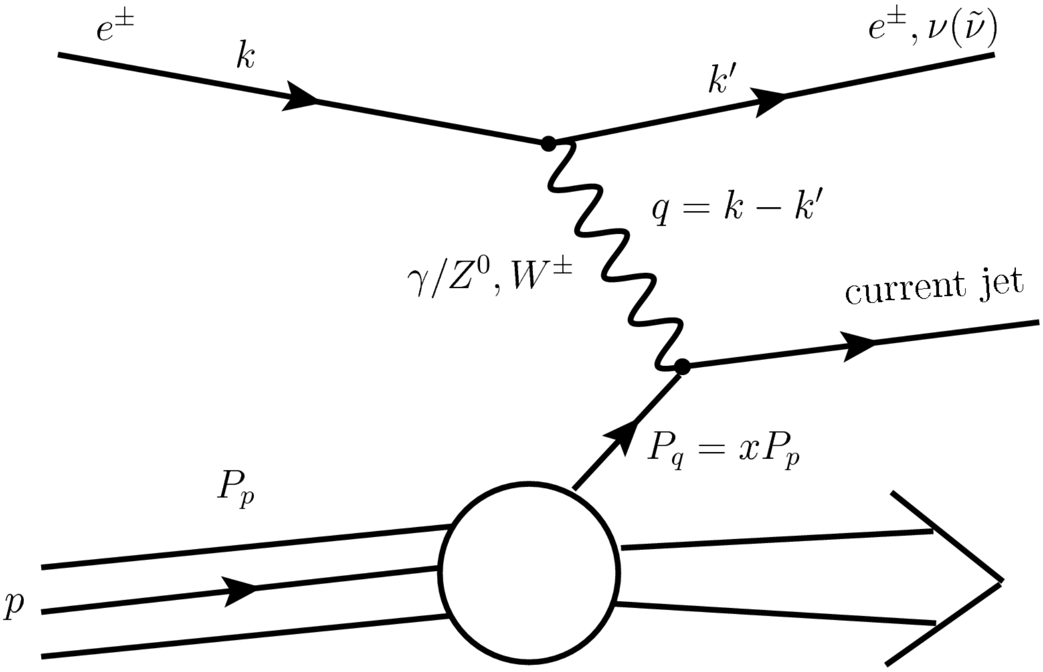


Figure 2.1.: The leading-order Feynman diagram for the deep inelastic scattering process.

2.1. Deep Inelastic Scattering and Jet Production

Deep inelastic scattering is a process in which a high-energy lepton (l) scatters on a nucleon[1] or a nucleus (h) with large momentum transfer. The formal equation for such reactions reads:

$$l(k) + h(P) \to l'\left(k'\right) + X\left(P'\right), \tag{2.1.0.1}$$

where the symbol in brackets indicates the momentum of the particle and X denotes the hadronic final state. At leading order, the interaction between lepton and hadron is mediated by electroweak bosons. For virtual γ or Z^0 exchange, the process is called Neutral Current (NC) DIS, while for $W^{\pm}$ exchange the process is called Charged Current (CC) DIS. At HERA the CC DIS process is characterised by the transformation of the initial-state electron (positron) into a final-state (anti-) neutrino. The leading-order Feynman diagram for the NC and CC process is illustrated in Figure 2.1.

Because of the fixed centre-of-mass energy,

$$\sqrt{s} = \sqrt{\left(k + p\right)^2}, \tag{2.1.0.2}$$

---

[1] The proton in case of H1 and ZEUS experiments.

two independent Lorentz-invariant scalar variables are sufficient to describe the basic scattering process at HERA. The following quantities[2] are typically used:

$$Q^2 = -q^2 = -\left(k - k'\right)^2, \tag{2.1.0.3}$$

$$x = \frac{Q^2}{2p \cdot q}, \tag{2.1.0.4}$$

$$y = \frac{p \cdot q}{p \cdot k}, \tag{2.1.0.5}$$

where Q^2 is the negative square of the four-momentum transfer or the virtuality of the exchange boson. Two kinematic regions are formally distinguished at HERA: $Q^2 < 1$ GeV^2, typically $Q^2 \approx 0$ GeV^2, called the photoproduction region; $Q^2 > 1$GeV^2, called the deep inelastic scattering regime[3]. The scaling variable, x, introduced by Bjorken [40], in the Quark-Parton Model (QPM) [41,42] can be interpreted as the longitudinal momentum fraction of the parton inside the proton that takes part in the hard scattering[4]. The variable y represents the fraction of the lepton energy carried by the gauge boson in the hadron rest frame. When electron and proton masses are ignored, the following equation relating the introduced variables holds:

$$Q^2 = sxy. \tag{2.1.0.6}$$

Choosing Q^2 and x as independent variables, the deep inelastic scattering cross section can be written in terms of the proton structure functions $F_i\left(x, Q^2\right)$:

$$\frac{d^2\sigma\left(e^{\pm}p\right)}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[Y_+ F_2\left(x, Q^2\right) - y^2 F_L\left(x, Q^2\right) \mp Y_- x F_3\left(x, Q^2\right)\right], \tag{2.1.0.7}$$

where α is the fine-structure constant and $Y_{\pm} = 1 \pm \left(1 - y\right)^2$. The dominant contribution to the scattering cross section is given by $F_2\left(x, Q^2\right)$, which in the QPM is directly related to the quark content of the proton:

$$F_2\left(x, Q^2\right) = F_2\left(x\right) = \sum_i e_i^2 x f_i\left(x\right). \tag{2.1.0.8}$$

---

[2] In the equations the masses of initial-state lepton and hadron are ignored

[3] Formally, the deep inelastic scattering regime is achieved when $\Lambda^2_{\mathrm{QCD}}/Q^2 \to 0$. The exact boundary between the photoproduction and deep inelastic scattering defined by $Q^2 = 1$ GeV^2 is conventional. This convention was adopted within the ZEUS collaboration.

[4] The estimator of the longitudinal momentum fraction of the initial-state parton valid beyond the QPM approximation is denoted by ξ and has to be deduced from the momentum of the final-state hadronic system, for example jets.

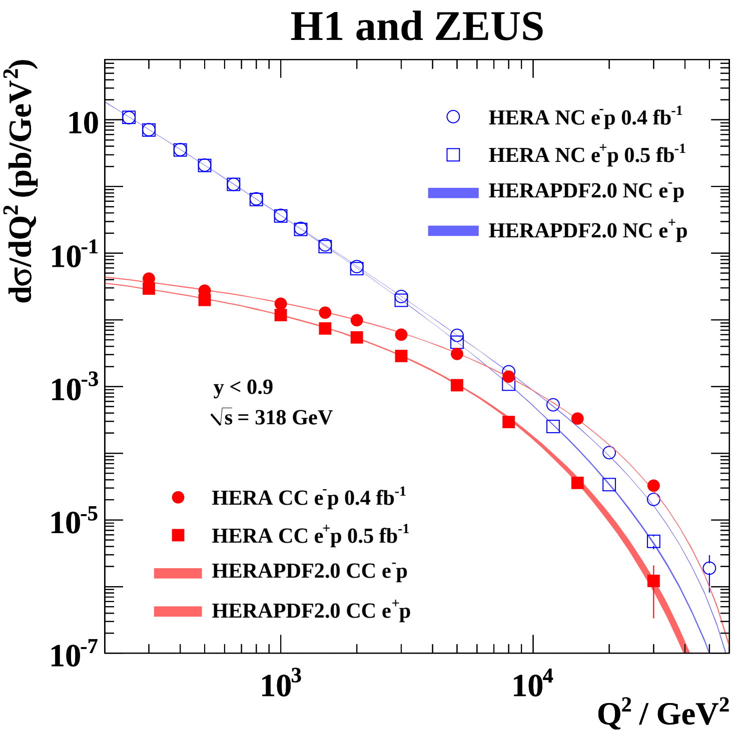


Figure 2.2.: The combined HERA NC and CC e^-p and e^+p cross sections, $d\sigma/dQ^2$ together with theoretical predictions. The plot is taken from [43].

In this equation e_i is the fractional charge of the quark and $f_i(x)$ is the proton parton density function (PDF) describing the density of quarks of different flavours in the nucleon. The Figure 2.2 shows the $d\sigma/dQ^2$ cross sections for NC and CC e^-p and e^+p reactions at HERA together with theoretical predictions. In case of NC DIS cross sections the predictions are obtained by integrating the Eq. (2.1.0.7) over x. In the calculations, a modern PDF parametrisation [43] was used.

In the Eq. (2.1.0.7), the longitudinal structure function, $F_L\left(x,Q^2\right)$, has significant contribution to the cross section only at high values of y and can be related to the cross section, σ_L, for the absorption of longitudinally polarised virtual photons:

$$F_L = \frac{Q^2}{4\pi^2\alpha}\cdot\sigma_L. \tag{2.1.0.9}$$

The structure function $F_3\left(x,Q^2\right)$ arises from Z^0-exchange and γZ^0-interference and has significant size only for $Q^2 \gtrsim \frac{1}{2}M_Z^2$ because processes induced by Z^0 interaction are suppressed by the mass of the Z^0-boson. The difference between the e^-p and e^+p NC DIS cross section caused by electroweak effects is clearly visible in Figure 2.2.

According to the factorisation theorem (see Section 2.2.3), the proton PDFs

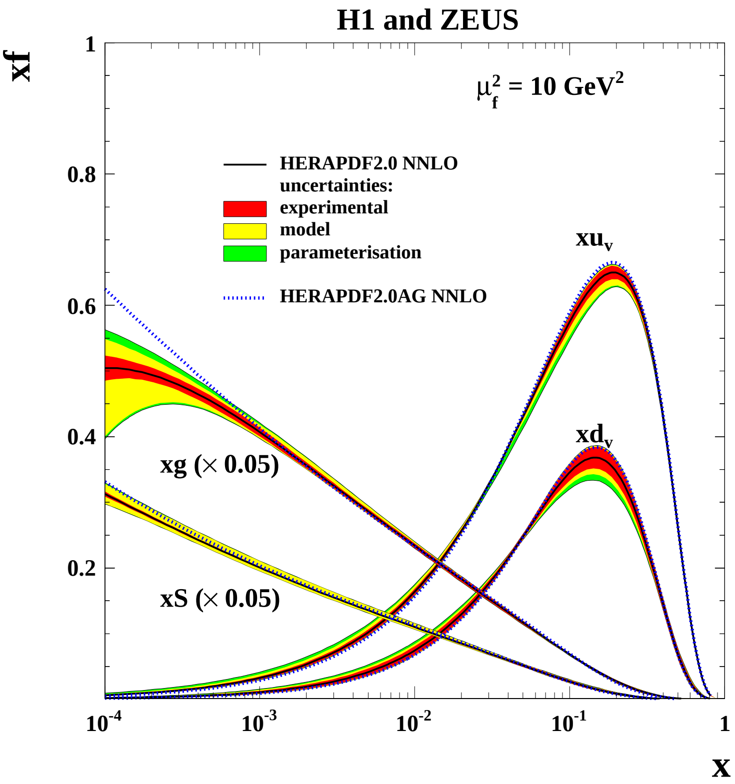


Figure 2.3.: The parton distribution functions xu_v, xd_v, $xS = 2x\left(\bar{U}+\bar{D}\right)$ and xg of HERAPDF2.0 NNLO. The gluon and sea distributions are scaled for better visibility. The plot is taken from [43].

are universal and independent of the process under consideration. Currently, PDFs cannot be predicted reliably from first principles and have to be determined by experiment. The state-of-the-art extraction of the proton PDFs from the combined deep inelastic scattering data from HERA [43] is show in Figure 2.3. As can be seen, the valence quark distributions xu_v and xd_v have a peak at $x \approx 0.1$ which approximately corresponds to the QPM expectations. However, when the proton is examined by the high-energy probe, its dominating gluon and sea quark[5] content is revealed.

The measurement of the inclusive DIS cross section at HERA provides direct access to the proton PDFs. Investigation of various sub-processes contributing to the inclusive DIS cross section can help to understand the details of the hard scattering process. One example is an important class of processes containing hadronic jets. The precise definition of a jet and significance of the studies of such processes is emphasized in the following.

---

[5] Sea quarks arise from the vacuum fluctuations of QCD fields (see Section 2.2)

2.1.1. Jet Definition

As was mentioned, bare partons do not appear as free particles because of the nature of the strong interaction. However high-energy quarks and gluons manifest themselves as collections of hadrons with approximately collinear momenta. Such hadronic final states localised in the kinematic phase space are called jets. Investigation of jet production provides access to the details of the underlying hard interaction as well as to the parton dynamics and the mechanism of parton showering and hadronisation. Provided the kinematics of the final-state jets, important quantities describing the kinematics of the hard scattering can be estimated. For example the longitudinal momentum fraction of the struck parton, ξ, can be calculated using:

$$\xi = x\left(1 + \frac{M^2}{Q^2}\right), \tag{2.1.1.1}$$

where x is the Bjorken scaling variable defined in the Eq. (2.1.0.4) and M is the invariant mass of two or more identified jets. Jets are important objects with which the test of perturbative QCD predictions is possible (see Section 2.2.5).

In the leading-order picture, jets correspond to individual partons emerging in high-energy collisions. An example of the basic diagrams contributing to the jet production in DIS is demonstrated in Figure 2.4. Since the flavour of the struck parton cannot be distinguished in NC DIS reactions, formally two types of processes contributing at leading order in the strong coupling can be distinguished, namely, the boson-gluon fusion (BGF) Figure 2.4(a) and QCD Compton (QCDC) scattering Figure 2.4(b), with gluons and quarks in the initial state, respectively.

The interplay of these two processes allows the effects attributed to the strong coupling and various PDF components to be disentangled, a value of $\alpha_s(M_Z)$ to be extracted and the proton PDFs to be constrained.

In order to give a rigorous definition of the jet, an algorithm for assignment of the particles to a jet must be provided. The proper combination of the particles has to fulfil the following general conditions:

- infrared and collinear safety (see Section 2.2.5);
- conservation of factorisation properties of the hard and soft processes;
- little sensitivity to the hadronisation effects;
- relative insensitivity to the soft interactions of the hadron remnant;
- invariance under longitudinal Lorentz boosts;

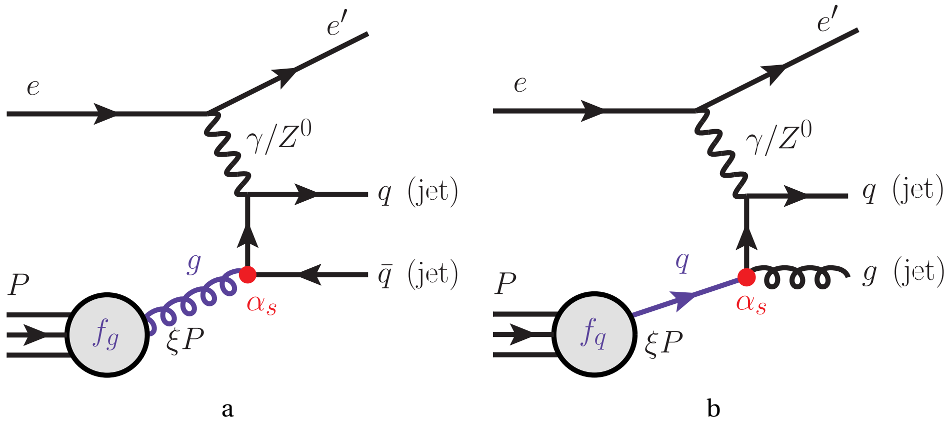


Figure 2.4.: Leading-order Feynman diagrams contributing to the jet production cross section in NC DIS. (a) Boson-gluon fusion; (b) QCD-Compton scattering processes.

- easy implementation at the particle level in experimental analyses as well as at the parton/hadron level in perturbative theoretical calculations.

Among others the recombination-type generalised k_T-algorithm satisfies all mentioned requirements and is defined by the following iterative procedure[6] (see Figure 2.5).

1. A distance measure, d_{ij}, quantifying the phase-space separation of two objects i and j, is defined for each pair of particles:

$$d_{ij} = \min\left(E_{\mathrm{T},i}^{2n}, E_{\mathrm{T},j}^{2n}\right)\frac{\Delta R_{ij}^2}{R_0^2}, \tag{2.1.1.2}$$

   where $\Delta R_{ij}^2 = \left(\eta_i - \eta_j\right)^2 + \left(\phi_i - \phi_j\right)^2$ is the angular separation between objects. The dimensionless parameter R_0 determines the jet radius.

2. A quantity, d_i, defining the distance to the beam-axis is calculated for each object i:

$$d_i = E_{\mathrm{T},i}^{2n}. \tag{2.1.1.3}$$

3. Two objects i and j are merged according to the Snowmass [44] conven-

---

[6] The input objects may refer to the energy deposits in the calorimeter cells; the set of partons in MC or fixed-order predictions or the set of stable hadrons appearing at the hadron level of MC simulations.

tion[7], whenever some d_{ij} is minimal among all d_{ij} and d_i:

$$E_\mathrm{T} = E_{\mathrm{T},i} + E_{\mathrm{T},j} \qquad \eta = \frac{\eta_i E_{\mathrm{T},i} + \eta_j E_{\mathrm{T},j}}{E_\mathrm{T}} \qquad \phi = \frac{\phi_i E_{\mathrm{T},i} + \phi_j E_{\mathrm{T},j}}{E_\mathrm{T}}. \tag{2.1.1.4}$$

When d_i is the smallest, the object is called jet and removed from the list.

4. The algorithm is repeated until no objects remain in the list.

The parameter n in Eq. (2.1.1.2) defines three types of algorithm:

- *n = -1*: the inclusive anti-k_T algorithm [45], which is now extensively used at the LHC. This algorithm results in jets of circular shape. The recombination process is characterised by first assigning particles with largest E_T to the jets;
- *n = 0*: the Cambridge-Aachen [46] algorithm, which takes into account only angular separations between objects, was mostly used in e^+e^- collider experiments;
- *n = 1*: the inclusive k_T algorithm [47], which produces jets of irregular shape and, in contrast to anti-k_T, recombines particles with small E_T first.

It has been shown that the k_T and anti-k_T have similar performance in photoproduction [48] and DIS [49]. The study [17] has demonstrated that $R = 1$ is the optimal choice of the radius parameter at HERA. Taking this into account, the choice of the k_T algorithm with $R = 1$ was adopted in this thesis. Taking advantage of the longitudinal invariance of the algorithm, the jet search was performed in the Breit frame, which is described below.

2.1.2. Breit Frame

In this analysis, the jet search was performed in the so-called Breit[8] frame [50, 51]. It is defined such that the exchanged boson collides with a proton without transverse momentum transfer. In this frame the momenta of the proton, P, and exchange boson, q, satisfy the equation:

$$2x\vec{P} + \vec{q} = 0. \tag{2.1.2.1}$$

In this frame the boson momentum is aligned along the positive Z direction and has only one space-like component *i.e.* $q = (0, 0, 0, -Q)$. The Breit frame is

---

[7] Other conventions exist. The Snowmass prescription results in massless jets.

[8] Also called the brick-wall reference frame.

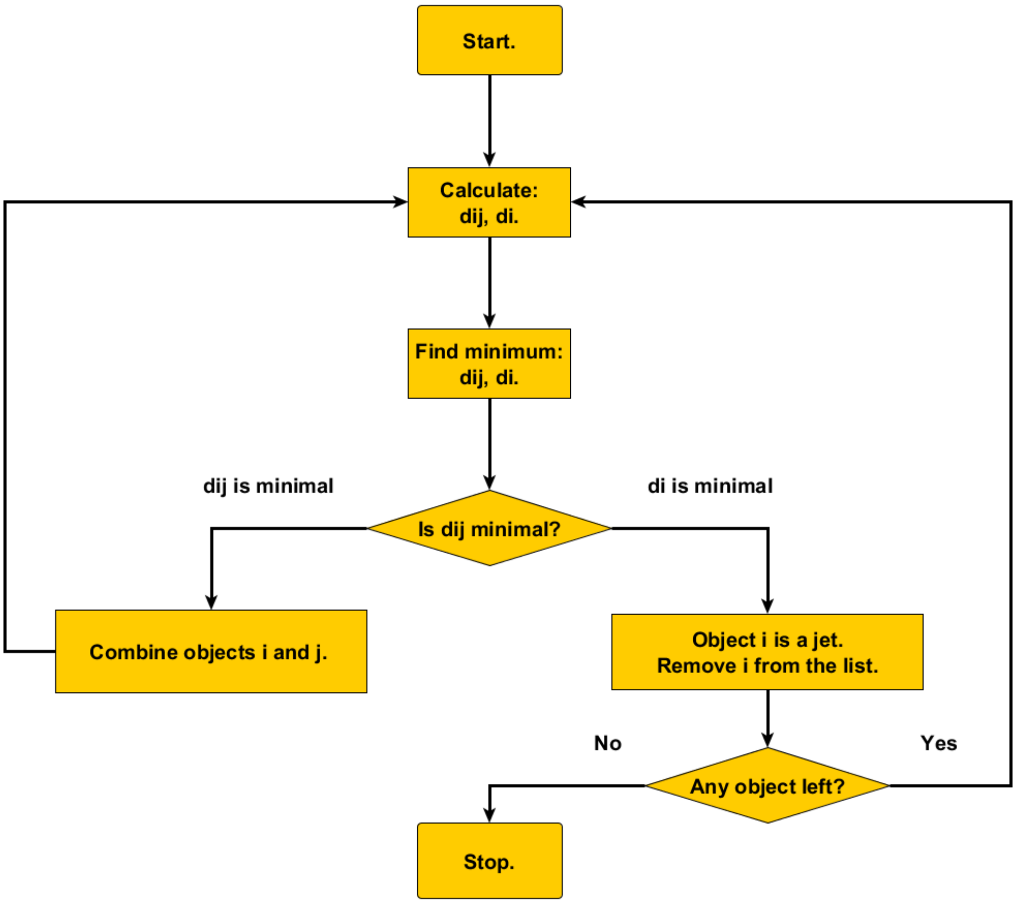


Figure 2.5.: Recombination-type jet algorithm flow char.

constructed in such a way that it corresponds in the QPM process to the back-scattering of the struck quark, maintaining the absolute value of its momentum $|xP|$. The schematic illustration of the QPM and QCDC and BGF processes in the Breit frame is demonstrated in Figure 2.6. The presence of non-zero transverse momentum in the Breit frame is a distinct feature of a QCD process that can be easily identified experimentally. As a result, the requirement of a jet in the Breit frame with sufficiently high transverse energy is related to the generation of a parton in the lowest-order QCD hard process at order $\mathcal{O}(\alpha\alpha_s)$.

2.2. Quantum Chromodynamics

Quantum Chromodynamcs emerged as a model to describe hadron spectra and the absence of observations of free hadron constituents. The QCD Lagrangian density can be derived from local $\mathrm{SU}(3)_C$ invariance and reads as follows

$$\begin{aligned}\mathscr{L}(x)_{\mathrm{QCD}} = &-\frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a + i\sum_{j=1}^{n}\overline{\psi}_j^{\alpha}\gamma_\mu\left(D^\mu\right)_{\alpha\beta}\psi_j^{\beta} - \sum_{j=1}^{n} m_j\overline{\psi}_j^{\alpha}\psi_{j,\alpha} \\ &-\frac{1}{2\alpha_G}\partial^\mu\mathscr{A}_\mu^a\partial_\mu\mathscr{A}_a^\mu - \partial_\mu\overline{\varphi}_a D^\mu\varphi^a,\end{aligned} \tag{2.2.0.2}$$

where

$$G_{\mu\nu}^a \equiv \partial_\mu\mathscr{A}_\nu^a - \partial_\nu\mathscr{A}_\mu^a + g_s f_{abc}\mathscr{A}_\mu^b\mathscr{A}_\nu^c, \qquad a = 1\ldots 8 \tag{2.2.0.3}$$

are the Yang-Mills field-strength tensor [52] constructed from gluon fields $\mathscr{A}_\mu^a$ in the adjoint representation of $\mathrm{SU}(3)_C$; g_s denotes the QCD coupling parameter[9] and the value of α_G defines the gauge. Quarks of different flavours are described by ψ_j fields in the fundamental representation of $\mathrm{SU}(3)_C$ while φ^a are eight anti-commuting scalar Faddeev-Popov ghost fields required in the quantisation procedure [53, 54]. The covariant derivative, $(D^\mu)_{\alpha\beta} = \delta_{\alpha\beta}\partial_\mu - ig_s\sum_a\frac{1}{2}\lambda_{\alpha\beta}^a\mathscr{A}_\mu^a$, is a generator of infinitesimal transformations in colour space acting on quark fields. Gell-Mann 3 × 3 matrices, $\lambda_{\alpha\beta}^a$, are the generators of the $\mathrm{SU}(3)_C$ algebra and f_{abc} are its real structure constants, defined by:

$$[T_a, T_b] = if_{abc}T_c, \qquad T_a = \frac{1}{2}\lambda_a. \tag{2.2.0.4}$$

An important feature that can be immediately observed is a non-linear term, $g_s f_{abc}\mathscr{A}_\mu^b\mathscr{A}_\nu^c$, in the definition of the field-strength tensor (Eq. (2.2.0.3)). This

---

[9] Similarly to electric charge, e, in Quantum Electrodynamics, the parameter g_s corresponds to the colour charge.

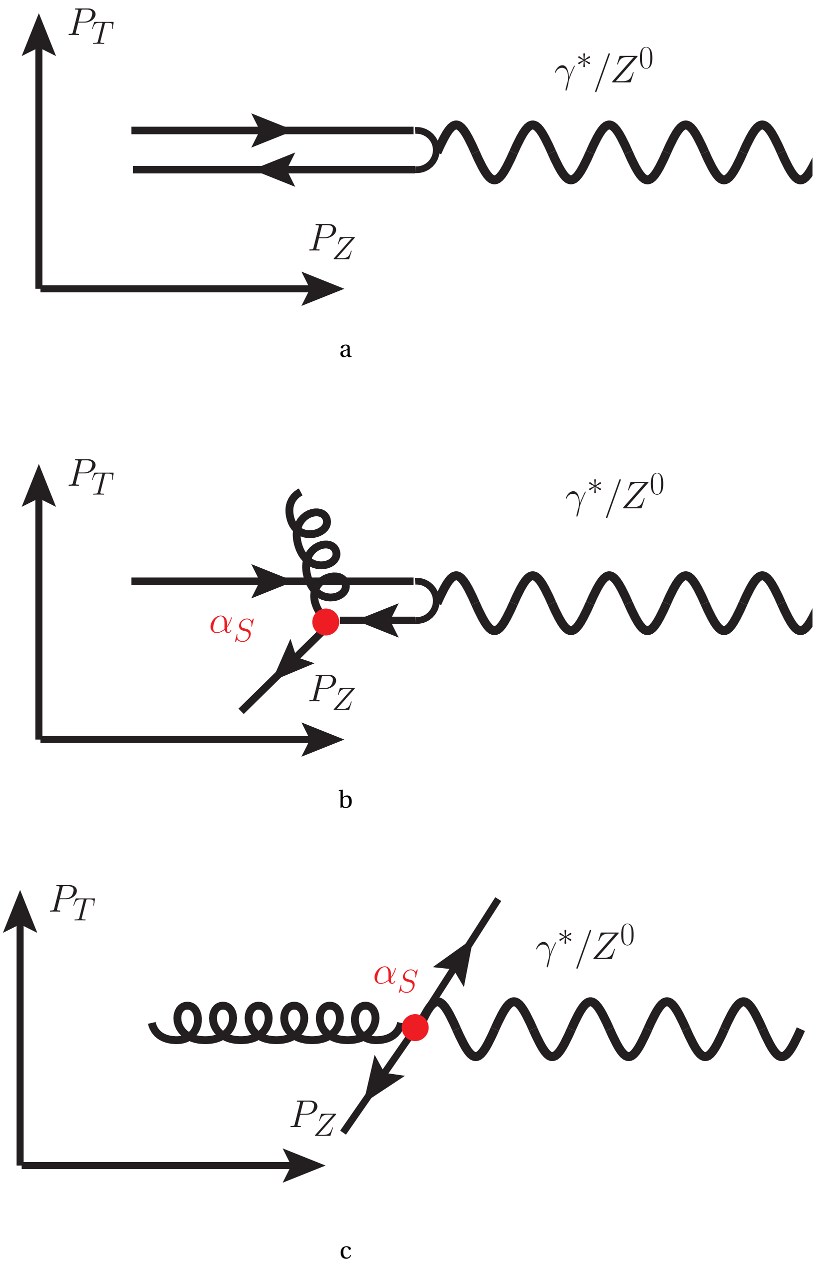


Figure 2.6.: Schematic illustration of the Born (a); QCD Compton (b); boson-gluon fusion (c) processes in the Breit frame in the (p_T, p_Z)-plane. In the quark-parton-model process, the incoming exchanged boson and parton have collinear momenta. The contribution from QCD processes results in non-zero outgoing parton transverse momentum.

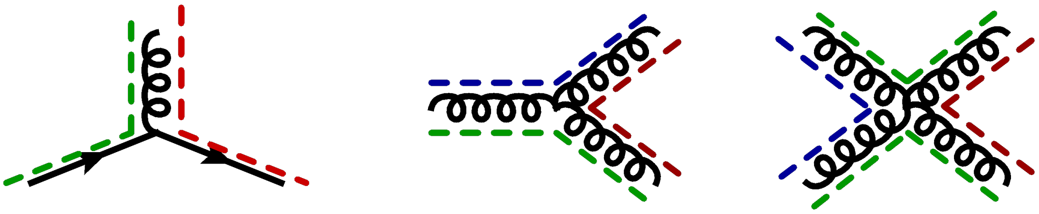


Figure 2.7.: The interaction vertices of the Feynman rules of QCD and schematic colour-flow interpretation for quark-gluon, three-gluon and four-gluon vertices.

term is the result of the non-Abelian structure of the symmetry group and determines the self-interaction of the force carriers. In contrast to the electromagnetic interaction, gluons carry two colour charges. Given the Lagrangian density, the Feynman rules and diagrams for QCD can be derived. Feynman graphs representing interaction vertices of the fundamental fields of QCD are depicted in Figure 2.7.

A theory with such an economical structure is extremely successful describing the vast variety of experimental data collected up to now. In particular, the strong interaction has two distinct features: at large energy scales,[10] hadron constituents behave as free particles and the strength of the coupling decreases ("asymptotic freedom"); conversely, at low energy scales, the strength of the coupling grows, binding quarks and gluons inside hadrons ("confinement").

Direct solution of the Yang-Mills equations is an impossible task. Almost all quantitative QCD predictions are based on three first-principle approaches: perturbative QCD (pQCD), lattice QCD and effective theories. The perturbative approach exploits the smallness of the strong coupling constant in the high-energy regime and develops successive approximations to the solution. The next section summarises basic information about the pQCD approach.

2.2.1. Structure of Perturbative Calculations

In perturbative QCD, the predictions for a physical observable, σ, are calculated order-by-order as a power series in small coupling $\alpha_s \equiv \frac{g_s^2}{4\pi} \ll 1$

$$\sigma = \sigma_0 \alpha_s^k + \sigma_1 \alpha_s^{k+1} + \sigma_2 \alpha_s^{k+2} + \ldots \tag{2.2.1.1}$$

The perturbation series starts at some power of the expansion parameter and only a few terms of the series are usually calculated. The expansion coefficients σ_i in the above series are usually calculated by summing up Feynman diagrams

---

[10] In the limit $\Lambda_{\text{QCD}}/E \to 0$, where $\Lambda \approx 225$ MeV is QCD characteristic energy scale.

or similar techniques [55]. The number of diagrams grows as $\sim i!$ with increasing perturbative order, therefore such series are, in general, divergent and have to be treated as an asymptotic expansion [56]. However, it is commonly assumed that the few first terms in the series provide a reasonable approximation to the exact solution. Leading-order graphs always have a tree-like structure. Calculation of higher-order corrections requires consideration of field configurations which emerge due to quantum fluctuations. Such configurations can be formally divided into two classes according to the topology of the corresponding Feyman diagrams. An example of next-to-leading-order diagrams contributing to jet production in deep inelastic scattering are illustrated in Figure 2.8. The real corrections (Figure 2.8a) are characterised by an increased number of particles leaving the interaction with respect to lower-order graphs, while in virtual contributions (Figure 2.8b) the fields form closed loops.

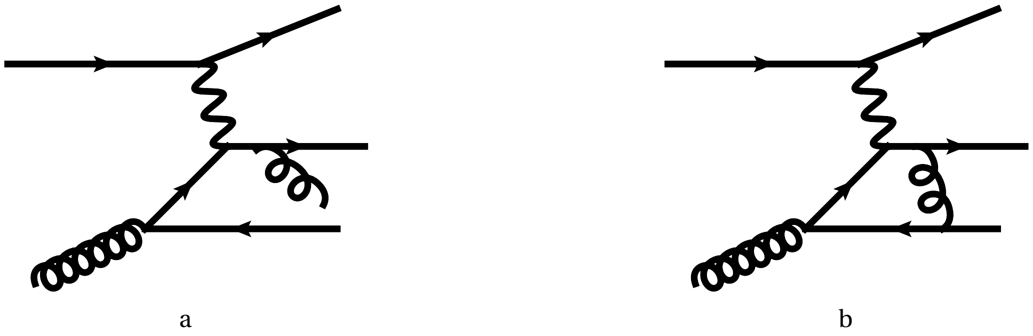


Figure 2.8.: Next-to-leading-order corrections to the jet production include (a) real (b) virtual contributions.

Intuitively, as the spatial scale at which the process is considered decreases, more fluctuations of the quantum fields must be taken into account. Such fluctuations lead to (anti-)screening of the colour charge or self interaction of particles; as a result, the field couplings or masses have to be interpreted as effective parameters of the theory, which take into account these effect. Nevertheless, it is unnatural that fluctuations occurring at scales much smaller than that corresponding to the typical energy scale of the process in question should have a significant influence. Remarkably, QCD admits redefinition of the couplings, fields and masses that incorporate contributions from fields fluctuations occurring in the limit of infinite energy. Practically, such a procedure involves singular transformations consisting of expressing physical observables in terms of a finite number of measured quantities and absorbing singularities emerging in the intermediate calculations into parameters of the theory. Such a process is called renormalisation.

In the calculations used in this thesis the so-called $\overline{\text{MS}}$ [57] renormalisation scheme was utilised. It consists of analytic continuation of the results of the calculations to un-physical $D = 4+2\varepsilon$ dimensions[11] and representing the results as a Laurent series in ε. The renormalisation proceeds with subtraction of the $1/\varepsilon^n$ poles, when a physical limit $\varepsilon \to 0$ is taken, to obtain finite quantities. The procedure introduces an additional parameter μ which has the dimension of energy. The dependence of the results on this parameter is discussed below.

2.2.2. Renormalisation and Renormalisation Group Equation

When the problem is treated perturbatively, the class of quantum fluctuations contributing to the process is naturally restricted. As mentioned above, beyond the tree level this results in the dependence of the calculations on the parameter μ, which approximately represents the spatial scale beyond which the effect of quantum fluctuations are absorbed into the dependence of the theory parameters on the scale μ. The value of the parameter can be chosen arbitrarily, but the physical quantity calculated in perturbation theory, e.g. the jet production cross section, cannot depend on an arbitrary parameter. This requirement can be formulated as follows[12]:

$$\frac{\mathrm{d}}{\mathrm{d}\log\mu^2}\tilde{\sigma}\left(Q^2/\mu^2, \alpha_s\right) = \frac{\partial\tilde{\sigma}}{\partial\log\mu^2} + \frac{\partial\alpha_s}{\partial\log\mu^2}\frac{\partial\tilde{\sigma}}{\partial\alpha_s} = \mathcal{O}\left({\alpha_s}^{k+1}\right), \qquad (2.2.2.1)$$

where for simplicity $\tilde{\sigma}$ is chosen to be a dimensionless observable[13]. An explicit μ dependence of $\tilde{\sigma}\left(Q^2/\mu^2, \alpha_s\right)$ has to be compensated by that of the coupling. An equation for the scale dependence of the strong coupling can be derived (see [58] and references therein):

$$\frac{\mathrm{d}\alpha_s(\mu)}{\mathrm{d}\ln\mu^2} = \beta\left(\alpha_s(\mu)\right), \qquad \beta\left(\alpha_s(\mu)\right) = -{\alpha_s}^2\left(\beta_0 + \beta_1\alpha_s + \beta_2{\alpha_s}^2 + \ldots\right). \qquad (2.2.2.2)$$

---

[11] This procedure is also called dimensional regularisation.

[12] The condition must hold up to terms proportional to ${\alpha_s}^{k+1}$ if the expansion of an observable $\tilde{\sigma}$ is known to $\mathcal{O}\left({\alpha_s}^k\right)$.

[13] An observable $\tilde{\sigma}$ can depend only on dimensionless ratio such as $\frac{Q^2}{\mu^2}$, where Q^2 represents the characteristic energy scale of the problem e.g. virtuality of exchanged boson in DIS.

This equation is called the Renormalisation Group Equation (RGE). The few first terms in the β-function were calculated in perturbation theory to be

$$\beta_0 = \frac{11C_A - 2n_f}{12\pi} = \frac{33 - 2n_f}{12\pi^2}, \tag{2.2.2.3}$$

$$\beta_1 = \frac{17C_A^2 - n_f T_R (C_A + 6C_F)}{24\pi^2} = \frac{153 - 19n_f}{24\pi^2}, \tag{2.2.2.4}$$

$$\beta_2 = \frac{2857 - \frac{5033}{9} n_f + \frac{325}{27} n_f^2}{128\pi^3}, \tag{2.2.2.5}$$

where C_A, C_F are $SU(3)$ structure coefficients, while $T_R = \frac{1}{2}$ and n_f is the number of active flavours[14, 15].

The equation (2.2.2.2) can be solved analytically. Taking into account only the first term involving β_0, the solution is:

$$\alpha_s\left(\mu^2\right) = \frac{\alpha_s\left(\mu_0^2\right)}{1 + \beta_0 \alpha_s\left(\mu_0^2\right) \ln\left(\mu^2/\mu_0^2\right)} = \frac{1}{\beta_0 \ln\left(\mu/\Lambda^2\right)} \tag{2.2.2.6}$$

The initial condition for the solution is specified by the value of the coupling at the starting scale $\alpha_s\left(\mu_0\right)$ or alternatively the integration constant Λ. The positivity of β_0 in the SM results in the coupling constant vanishing when the energy scale μ increases or correspondingly shorter time scales are considered. Quarks and gluons behave as non-interacting free particles in the high-energy limit. On the other hand, in processes characterised by long time intervals or equivalently, small momenta, the coupling grows. Eventually, the coupling becomes undefined near the pole $\mu = \Lambda$. In this region, the theory becomes essentially nonperturbative and the series expansion is no longer valid.

2.2.3. Factorisation

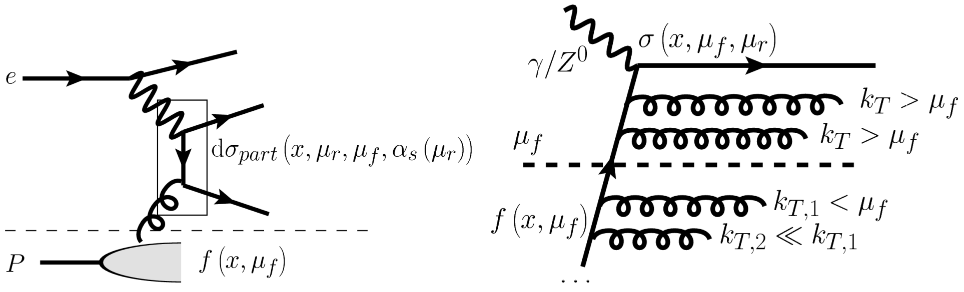
Another remarkable feature of QCD is factorisation of short- and long-time-scale processes. For example, the cross section for jet production in DIS can be represented in factorised form [59] as:

$$\mathrm{d}\sigma_{\mathrm{jet}} = \int f\left(x, \mu_f\right) \mathrm{d}\sigma_{\mathrm{part}}\left(x, \mu_r, \mu_f, \alpha_s\left(\mu_r\right)\right). \tag{2.2.3.1}$$

---

[14] In general, the β-function coefficients, β_i depend on the employed renormalisation scheme. Only β_0 and β_1 are scheme independent. The β_2 term specified here refers to the widely used $\overline{\mathrm{MS}}$ renormalisation scheme.

[15] It is assumed that heavy quark flavours decouple from the theory below energy scales much smaller than the heavy quark mass $\mu \ll m_h$.



a b

Figure 2.9.: Schematic illustration of the factorisation of the hard ep-process into non-perturbative proton PDFs and hard-scattering partonic cross sections (a). An example diagram with strongly ordered parton emission contributing to the NC DIS process (b).

In this expression $f(x, \mu_f)$ represents the nonperturbative proton parton distribution function and $\mathrm{d}\sigma_{\text{part}}$ is the hard-scattering partonic cross section that is calculated in perturbation theory. A schematic illustration representing this equation is depicted in Figure 2.9a. Technically, in the calculation an additional factorisation scale, μ_F is introduced. The parameter μ_F defines approximately the virtuality of the intermediate (virtual) states that contribute to the hard scattering while the long-distance physics is absorbed in universal nonperturbative parameters[16]. Factorisation leads to the calculations being usually performed in two steps. The perturbative part can be evaluated as a series expansion in the strong coupling constant, described above, while parton distributions have to be determined experimentally. In this procedure, singularities attributed to the long-distance processes e.g. soft or collinear radiation of partons, are absorbed into nonperturbative terms. The factorisation scale, μ_f, serves as a reference point at which the subtraction of the singularities is performed. The subtraction scheme defines the prescription for reshuffling of finite terms between partonic cross section and PDFs. The employed factorisation scheme must be consistent with that used for renormalisation. In this analysis the modified variant of the minimal subtraction, $\overline{\text{MS}}$, scheme [57] was used.

2.2.3.1. DGLAP Equations

In analogy to RGE, an evolution equation for the parton distribution functions can be derived. The factorisation-scale dependence of the PDFs is governed by

---

[16] For DIS, the processes with energy scale $Q^2 \gg \mu_F^2$ are attributed to the perturbative part, while those with $Q^2 \ll \mu_F^2$ are absorbed into PDFs.

the DGLAP equations [60–63]:

$$\frac{\mathrm{d}}{\mathrm{d}\log\mu_F^2}\begin{pmatrix}\mathbf{f}_{\mathbf{q_i}}\left(x,\mu_F\right)\\ f_g\left(x,\mu_F\right)\end{pmatrix}=\sum_j^{2n_f}\int_x^1\frac{\mathrm{d}z}{z}\begin{pmatrix}\mathbf{P}_{\mathbf{q_i}\leftarrow\mathbf{q_j}}(z) & \mathbf{P}_{\mathbf{q_i}\leftarrow\mathbf{g}}(z)\\ P_{g\leftarrow q_j}(z) & P_{g\leftarrow g}(z)\end{pmatrix}\begin{pmatrix}\mathbf{f}_{\mathbf{q_j}}\left(x/z,\mu_F\right)\\ f_g\left(x/z,\mu_F\right)\end{pmatrix}, \quad (2.2.3.2)$$

where the summation runs over the number of active quark and antiquark flavours. The kernels of these equations are splitting functions $\mathbf{P}_{\mathbf{a}\leftarrow\mathbf{b}}(z)$ representing the probability of the spitting of a single parton into two particles carrying fractional momentum z and $(1-z)$, respectively. The splitting function $\mathbf{P}_{\mathbf{a}\leftarrow\mathbf{b}}(z)$ can be calculated from the collinear singularity of any hard-scattering process as a power series in α_s:

$$P\left(z,\alpha_s\left(\mu_f\right)\right)=\alpha_s\left(\mu_f\right)P_0(z)+{\alpha_s}^2\left(\mu_f\right)P_1(z)+{\alpha_s}^3\left(\mu_f\right)P_2(z)+\ldots \quad (2.2.3.3)$$

At the moment the splitting functions are known to next-to-next-to-leading order [64, 65]. Intuitively, this system of equations states how sensitive the probe becomes to the low momentum partons as the resolution scale μ_f increases.

Evolution of the scale-dependent parton distributions according to the DGLAP equations effectively resums the Feynman diagrams with parton emission strongly ordered in transverse momentum $\mu_{F,0} \ll \ldots \ll k_{T,i} \ll k_{T,i+1} \ll \ldots \ll \mu_F$ (see Figure 2.9b), where $\mu_{F,0}$ denotes the nonperturbative cut-off scale. Each parton emission in such an approximation is accompanied by a term $\alpha_s \cdot \ln\left(\mu_F^2/\mu_{F,0}^2\right)$ in the matrix element, therefore such resummation is also called the "leading log approximation".

Alternative approaches for the PDF evolution exist. The so-called BFKL approach [63, 66] is focused on ressumation of large $\ln(1/x)$ terms and is valid in the low-x region of phase-space. It does not require strong transverse-momentum ordering of emitted partons and leads to rather uniform sampling in k_T. The CCFR [67–70] approach was developed to be valid both at low x and high Q^2 and is based on angular ordering of the radiation pattern.

2.2.4. Scale Choice

The size of the unknown higher-order terms in the perturbative series Eq. (2.2.1.1) is usually one of the dominant sources of uncertainty in theoretical predictions. These contributions can be estimated from the dependence of the perturbative expansion on the renormalisation and factorisation scales. Using the renormalisation group equation it can be demonstrated that the scale dependence of higher-order coefficients accompanied by logarithmic terms are fully determined by lower-order coefficients [71]. For example, in the case of a dimensionless observable with perturbative expansion of the form Eq. (2.2.1.1), the following expressions can be obtained:

$$f_1\left(\frac{\mu}{Q}\right) = f_1(1) - k\beta_0 f_0 \log\frac{\mu}{Q}, \tag{2.2.4.1}$$

$$f_2\left(\frac{\mu}{Q}\right) = f_2(1) - \left[(k+1)\beta_0 f_1(1) + k\beta_1 f_0\right]\log\frac{\mu}{Q}, \tag{2.2.4.2}$$

$$+\frac{k(k+1)}{2}\beta_0^2 f_0 \log^2\frac{\mu}{Q}, \tag{2.2.4.3}$$

$$\ldots \tag{2.2.4.4}$$

Thus, the μ variation in the $\mathcal{O}\left({\alpha_s}^n\right)$ expression corresponds to the higher-order terms of the form:

$${\alpha_s}^{n+1}(\mu)\sum_{i=1}^{n+1-k}(\text{know part})\cdot\log^i\frac{\mu}{Q} + \mathcal{O}\left({\alpha_s}^{n+1}\right). \tag{2.2.4.5}$$

However, terms that are not accompanied by the logarithms e.g. $f_2(1)$, require explicit calculation. Therefore the reliability of an estimate of the size of the truncated terms depends on whether $f_i(1), i \geq 1$ are of similar order as f_0. Notably, the leading-order coefficient f_0 is independent of the renormalisation scale, therefore the scale dependence of the LO approximation is completely governed by the scale dependence of the strong coupling. Therefore a realistic estimate of the size of unknown terms is possible starting at least at NLO.

Besides that, the sensitivity of perturbative predictions to the μ_f scale variation has to be taken into account. Although formally the DGLAP equations perform all-orders resummation of the ladder diagrams, the residual dependence on the factorisation scale of order $\mathcal{O}\left({\alpha_s}^{k+1}\right)$ persists in the pQCD calculations. Similarly to the renormalisation-scale dependence, the dependence of the perturbative coefficients on μ_f can be recovered:

$$f_1\left(\frac{\mu_f}{Q}, \frac{\mu_r}{Q}\right) = f_1\left(1, \frac{\mu_r}{Q}\right) + P_0 \otimes f_0 \log\frac{Q^2}{\mu_f^2}, \tag{2.2.4.6}$$

where P_0 is the LO splitting function and the convolution symbol denotes

$$P_0 \otimes f_0 = \int \frac{\mathrm{d}z}{z} P_0(x/z)\, f_0(z). \tag{2.2.4.7}$$

For the higher-order coefficients, a pattern similar to the renormalisation case is obtained.

The DGLAP equations involve the strong coupling evaluated at the factorisation scale μ_f. Whenever this scale differs from μ_r, the RGE evolution from factorisation to renormalisation point has to be performed. Since the RGE is determined by incomplete series for the β-function, the strong-coupling evolution can be unreliable if calculation over a wide interval of scales is required. Mathematically this can be formulated as follows. Considering a Taylor expansion of $\alpha_s\left(\mu_f\right)$ around $\alpha_s\left(\mu_r\right)$ and substituting the RGE expression for the derivatives of the strong coupling $\mathrm{d}\alpha_s/\mathrm{d}\log\frac{\mu_f^2}{\mu_r^2}$ gives [72]:

$$\alpha_s\left(\mu_f\right)=\alpha_s\left(\mu_r\right)-\beta_0\log\frac{\mu_f^2}{\mu_r^2}\alpha_s{}^2\left(\mu_r\right)-\left(\beta_1\log\frac{\mu_f^2}{\mu_r^2}-\beta_2^2\log^2\frac{\mu_f^2}{\mu_r^2}\right)\alpha_s{}^3\left(\mu_r\right)+\mathcal{O}\left(\alpha_s{}^4\right) \tag{2.2.4.8}$$

$$\begin{aligned}\alpha_s{}^2\left(\mu_f\right)&=\alpha_s{}^2\left(\mu_r\right)-2\beta_0\log\frac{\mu_f^2}{\mu_r^2}\alpha_s{}^3\left(\mu_r\right)+\mathcal{O}\left(\alpha_s{}^4\right)\\ \alpha_s{}^3\left(\mu_f\right)&=\alpha_s{}^3\left(\mu_r\right)+\mathcal{O}\left(\alpha_s{}^4\right).\end{aligned} \tag{2.2.4.9}$$

Thus, to ensure convergence of such an expansion, the factorisation scale must be closely related to the renormalisation scale. In addition, as in the case of renormalisation, the μ_f scale must be much larger than Λ_{QCD} to justify the applicability of perturbation theory results for the PDF evolution.

There is no general method to estimate the size of the contribution from missing terms in perturbative series. However, it is widely assumed that the corresponding uncertainty can be estimated from the variation of the renormalisation and factorisation scales up and down by a factor of two. The resulting variation of the observable depends on the central values μ_f^0, μ_r^0 around which the variation is performed. It is desirable to choose central values such that the difference between the nominal result and the one with scaled values of μ_f and μ_r is minimised i.e. $\partial\sigma/\partial\mu=0$. This method is called the "principle of minimum sensitivity" (PMS) [73]. However, straightforward application of this method can result in a very large inclusive-jet cross section at the lowest Q^2 and E_T^{jet} values [74]. Alternative prescriptions for the scale choice can be found in [56]. The proposed methods emphasise different aspects of the perturbative expansion. However, it should be noted that all are related to the behaviour of logarithmically enhanced terms.

In this analysis, the traditional prescription of choosing the scale corresponding to the typical energy scale of the process was adopted. More details are provided in Chapter 8.

2.2.5. Calculation of Next-to-Leading-Order Jet Cross Section

As described in Section 2.2.3, the predictions for the jet-production cross sections in ep collisions have a factorised form (see Eq. (2.2.3.1)). The partonic cross section is calculated perturbatively, as a power series in the strong coupling, α_s. The predictions for the jet cross sections are finite at each order according to the KNL theorem [75, 76], provided an infrared- and collinear-safe jet-algorithm is used. However, the parton configurations with soft or collinear radiation have divergent matrix elements; after dimensional regularisation soft and collinear (overlapping) divergences appear as $1/\epsilon$ $(1/\epsilon^2)$ poles in the expressions. These divergences cancel exactly with those arising from the virtual contributions.

The differential jet cross section is calculated according to the expression [12]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}X} = \frac{1}{\text{flux}} \sum_n \frac{1}{n!} \int \mathrm{d}\Phi^n \overline{\sum} \left|\mathcal{M}^{(n)}\left(p_i\right)\right|^2 \delta\left(X - \mathcal{X}_n\left(p_i\right)\right), \tag{2.2.5.1}$$

where $\mathrm{d}\Phi^n = \prod_{i=1}^{n} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i}$ is an element of n-body phase space and $\mathcal{M}$ denotes the Lorentz-invariant matrix element. The first summation is performed over all n-parton final states, assuming that quarks, antiquarks and gluons are indistinguishable ($1/n!$ is a symmetrisation factor). The inner sum represents the averaging over possible colour and spin configurations. The jet-function $\mathcal{X}_n\left(p_i\right)$ of the momenta of n partons represents the measurement observable e.g. $E_{T,B}^{jet}, \eta_B^{jet}$ etc. In order to ensure cancellation of real and virtual divergences, the jet algorithm must be independent of the number of soft and collinear partons in the final state. The cancellation of divergences holds only if the observable satisfies the following conditions:

$$\left.\begin{array}{l} \mathcal{X}_{n+1}\left(p_1, \ldots, \lambda p_n, (1-\lambda)\, p_{n+1}\right) \\ \mathcal{X}_{n+1}\left(p_1, \ldots, \lambda p_n, 0\right) \end{array}\right\} = \mathcal{X}_n\left(p_1, \ldots, p_n\right), \tag{2.2.5.2}$$

where $\lambda \in [0;1]$ is a parameter used to implement smooth transition from from $n+1$ to n-parton configuration. The jet-functions $\mathcal{X}_{n+1}$ and $\mathcal{X}_n$ must be equal in collinear and soft limits. The algorithm must produce identical results if a single particle is replaced by a pair of collinear particles carrying the same total momentum, or if the energy of one of the particles vanishes.

In this analysis the infrared- and collinear-safe k_T jet algorithm was used for the reconstruction of jets from the final-state partons. Since fixed-order QCD predictions refer to the jets of partons while the measurements refer to hadronic jets, the calculations were corrected to the hadron level using Monte Carlo predictions (see Section 8.1.1).

Practical calculations, suitable for the comparison with experimental results involving cuts (e.g. phase-space restrictions or detector-acceptance limitations) utilise numerical techniques for the calculation of the phase space integrals. General schemes for the calculation of the jet production cross section at next-to-leading order, suitable for numerical calculations and independent of experimental requirements, exist. One such scheme [77] is briefly described in the following.

2.2.6. Subtraction Scheme

The NLO n-jet partonic cross section is a sum:

$$\sigma = \sigma_{\mathrm{LO}} + \sigma_{\mathrm{NLO}} = \int_n \mathrm{d}\sigma^B + \left[\int_{n+1} \mathrm{d}\sigma^R + \int_n \mathrm{d}\sigma^V \right], \tag{2.2.6.1}$$

where σ^B is the Born-level cross section, σ^R is the real-radiation correction and σ^V is the virtual correction. In order to remove explicit divergences from the the real and virtual parts, specially constructed counter-terms are added and subtracted from Eq. (2.2.6.1). The counter-term is an approximation to the real-radiation contribution in the region of the phase-space containing a singularity and has the same point-wise singular behaviour. Each singular parton configuration requires a corresponding counter term[17]. The real-radiation contribution with subtracted counter-term, σ^A, becomes a regular function that can be integrated in $D = 4$ dimensions:

$$\int_{n+1} \mathrm{d}\sigma^R \rightarrow \left[\int_{n+1} \mathrm{d}\sigma^R - \int_{n+1} \mathrm{d}\sigma^A \right]. \tag{2.2.6.2}$$

The virtual contribution term is modified as follows:

$$\begin{aligned} \int_n \mathrm{d}\sigma^V \rightarrow \left[\int_n \mathrm{d}\sigma^V + \int_{n+1} \mathrm{d}\sigma^A \right] = \\ \left[\int_n \mathrm{d}\sigma^V + \int_1 \mathrm{d}\sigma^A \right] \end{aligned} \tag{2.2.6.3}$$

The divergence in the virtual contribution appears as a pole in ϵ but this pole is exactly cancelled by that resulting from one-parton phase-space analytic integration of the counter term. After the cancellation, the integration of the virtual part can be carried out numerically in physical $D = 4$ dimensions. Since the net effect of adding and subtracting counter-terms is zero, this scheme results only in reshuffling of the divergences.

2.3. Monte Carlo Models

A precise theoretical description of the final state of ep scattering from first principles is currently an intractable problem. It requires calculations in regions of phase space where perturbative techniques are not applicable or have

---

[17] The infrared structure persists also at higher orders in perturbative expansion. Partial $\mathcal{O}\left(\alpha_s{}^3\right)$ corrections were worked out in [78]

to be performed to high orders. Phenomenological models were developed in order to describe such processes. Typically, such models are implemented in the form of event generators and utilise Monte Carlo calculations. Some models used for the description of DIS hadronic final state are described in the following.

2.3.1. QCD Parton Showers

The parton-shower approach is used to simulate higher-order perturbative QCD contributions when a complete calculation is infeasible or unknown. For example, the DGLAP approach can be utilised to describe initial-state and final-state radiation. The probability for a branching, $\mathscr{P}_{a\to bc}$, during the evolution is governed by the equation:

$$\frac{\mathrm{d}\mathscr{P}_{a\to bc}}{\mathrm{d}Q^2} = \int_0^1 \mathrm{d}z \frac{\alpha_s\left(Q^2\right)}{2\pi} P_{a\to bc}\left(z\right), \tag{2.3.1.1}$$

where $P_{a\to bc}\left(z\right)$ are the Altarelli-Parisi splitting kernels (see Section 2.2.3.1).

Such an approximation is usually used in general-purpose event generators where the successive radiation is simulated until the evolution parameter, e.g. virtuality of the daughter partons, reaches some low energy scale $\mathcal{O}$ (1 GeV). At this point the showering process is stopped and partons are recombined into colourless hadrons.

In order to improve the leading-logarithmic accuracy of the parton-shower approach, hard emissions are described using complete matrix elements. In this case an additional intermediate scale is introduced. At this energy scale regions dominated by parton shower or hard-scattering dynamics are matched. Nowadays most of the event generators are based on LO matrix elements. However, NLO calculations with matched parton showers are starting to appear [79, 80].

Another approximation for QCD radiation that is widely used to describe DIS-related processes is the *colour dipole model* (CDM) [81–86]. It is assumed in this model that the quark–antiquark pairs form colour dipoles with a corresponding dipole radiation pattern. The gluons themselves are interpreted as pairs of colour charges that also build colour dipoles. The schematic illustration corresponding to the CDM picture is shown in Figure 2.11. The radiation from each dipole is assumed to be independent. It proceeds iteratively until some stopping criterion is reached, for example the invariant mass of a dipole falls below some cut-off value. The CDM is based on leading-order matrix elements in the soft gluon approximation. The cross section for the parton emission with transverse momentum p_T and rapidity y (see Eq. (3.2.0.2)) in CDM

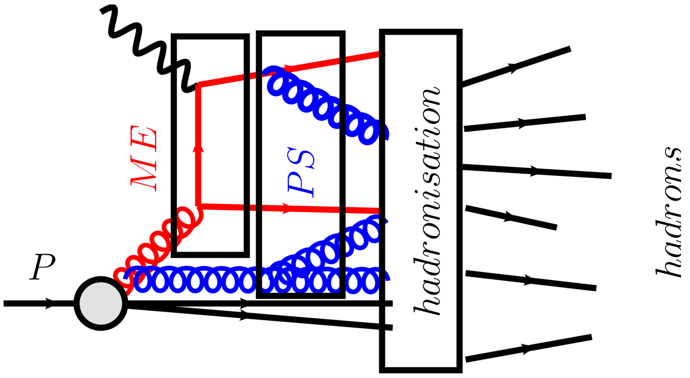


Figure 2.10.: Schematic demonstration of the matrix element + parton shower approach.

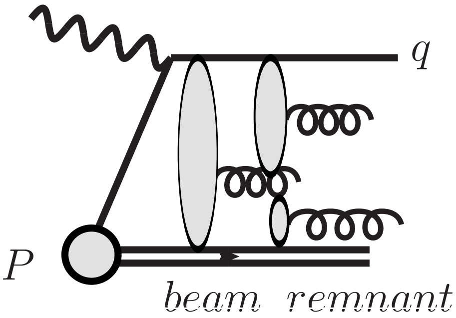


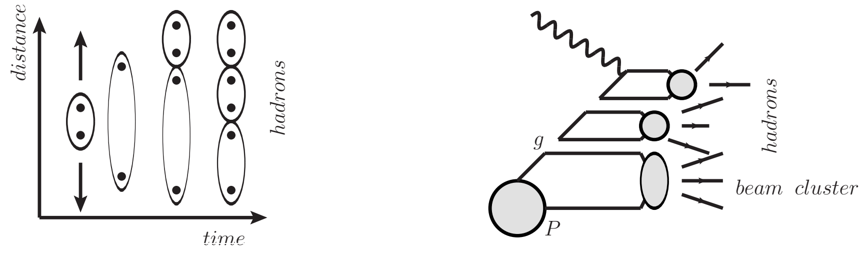
Figure 2.11.: The radiation pattern from the colour-dipole model.

reads:

$$d\sigma = \frac{n_c \alpha_s}{2\pi} \frac{dp_T^2}{p_T^2} dy. \tag{2.3.1.2}$$

In contrast to the leading-logarithm DGLAP-based parton-shower algorithm there is no k_T-ordering for the gluon radiation. Emitted partons are rather uniformly distributed in k_T, thus the CDM approach is somewhat similar to the BFKL evolution.

Another important issue in the simulation of the parton showers is quantum-mechanical interference of the initial-state and final-state radiation or the interference between the partons emitted either in the initial or final state. These effects are naturally taken into account in the complete perturbative calculations, however special care must be taken in the resummed calculations like



a b

Figure 2.12.: Schematic illustration of the string fragmentation (a) and cluster fragmentation (b) model

those based on DGLAP evolution, because they are based on a probabilistic description of the whole process in contrast to quantum-mechanical probability amplitudes.

2.3.2. Fragmentation

In order to be able to compare pQCD predictions to experimental results the calculations have to be defined in terms of experimentally observable quantities, which usually are functions of the momenta of the final-state hadrons. The formation of hadrons, called *hadronisation*, is essentially a non-perturbative process and first-principle calculations are impossible. Therefore phenomenological hadronisation models are used to correct partonic predictions in order to obtain a consistent observable definitions. In practice, the transition from partonic quantities to those defined in terms of hadrons is usually modelled by means of general-purpose event generators. Two widely used hadronisation models are described below.

2.3.2.1. String Fragmentation Model

It is assumed in the Lund string model [87] that the flux of the colour field between two quarks is confined within a tube of finite transverse size. This string-like object has a constant energy-density per unit length of $\mathcal{O}(1\ \text{GeV/fm})$ and the potential energy of the string increases with increasing separation between the quarks. When the tension exceeds the quark–anti-quark production threshold, the $q\bar{q}$-pair is picked up from the vacuum and the string breaks up. Loose ends of the string are terminated by newly created q and $\bar{q}$ and the process is iterated until the potential energy of the daughter strings fall below a cut-off $\mathcal{O}(1\ \text{GeV})$. The gluons have two colours and are represented as a joint between two strings or a kink in the colour flux of the $q\bar{q}$-system in this model.

A schematic illustration of the Lund picture of the hadronisation process is shown in Figure 2.12a.

2.3.2.2. Cluster Fragmentation Model

In the cluster model [88, 89] all partons after the parton-shower step are combined into colourless objects. If the invariant mass of the cluster is large enough it can decay into lighter clusters, which subsequently decay into hadrons. The gluons in this model are converted into $q\bar{q}$-pairs and do not appear in the hadron formation process. This model was inspired by the “preconfinement” [90] idea according to which the colour-connected partons group in the phase space towards the end of perturbative evolution. The cluster model process is schematically depicted in Figure 2.12b.

2.3.3. General-Purpose Event Generators

General-purpose event generators are indispensable tool in high-energy physics because they provide full access to the details of the event final state. Using event generators and detector simulations, the detector performance can be investigated (see Chapter 7) or effects related to the background contributions can be estimated (see Section 5.6).

The generation of events proceeds through Monte Carlo sampling of the processes according to the probability of their occurrence. An ensemble of MC events must resemble the characteristic features of the data. These programs usually have several levels naturally corresponding to the processes separated by different time-scales. The simulation of the hard interaction, occurring over the shortest time intervals, is usually based on the leading-order contribution that can be relatively easy calculated in perturbation theory (see Section 2.2.1). The higher perturbative orders in MC generators are approximated by parton-shower models, as was briefly described in Section 2.3.1. The last step, corresponding to the formation of colour-neutral hadrons is implemented in hadronisation models (see Section 2.3.2), which use the result of the parton-shower stage as an input. The output of the event generators is usually provided in the form of a table containing list of particles and their four-momentum vector components. The output available after the parton-shower and hadronisation steps are called the parton and hadron levels, respectively.

In this work, the NC DIS events were generated using the HERACLES program [91] with the DJANGOH [92] interface to the LEPTO [93] and ARIADNE [94, 95] parton-shower simulation programs. The DJANGOH code implements higher-order QED corrections i.e. real- and virtual-photon radiation as well as two-photon exchange. As an input in the MC, the CTEQ5D [96] proton PDF sets were utilised. Basic information about LEPTO and ARIADNE generators is summarised below.

2.3.3.1. LEPTO

The LEPTO event generator combines the leading-order QCD matrix elements (ME) for the hard-scattering process together with the DGLAP parton shower (PS) for the soft-gluon emission. In order to ensure colour coherence during the showering process, angular ordering is imposed. The Lund string model as implemented in JETSET [97] is used to simulate the hadronisation process. This generator also includes the LO electroweak processes necessary for the description of high-Q^2 DIS. The higher-order QED effects are obtained through the interface to the HERACLES program. The LEPTO generator is also often called MEPS and is used as a reference MC generator in this analysis.

2.3.3.2. ARIADNE

The colour-dipole pattern for QCD radiation is implemented in the ARIADNE event generator. Since this model naturally includes only the QCD Compton scattering diagram, the BGF graph contribution was introduced in addition. The hadronisation is performed using the same JETSET interface as used for LEPTO. This event generator was used in the analysis mainly to estimate systematic effects attributed to the choice of the parton-shower model.

CHAPTER 3

Experimental Setup

This chapter introduces the most important features of the experimental setup that was utilised in this thesis. At the beginning the relevant properties of the accelerator complex is briefly introduced. Later the relevant components of the ZEUS detector are discussed.

3.1. HERA Machine

The Hadron-Electron Ring Accelerator (HERA), the only ep collider in the world, was built in Hamburg, Germany at the national accelerator research centre DESY. At HERA, electrons[1] or positrons of energy 27.5 GeV collided with protons of energy up to 920 GeV, resulting in a centre-of-mass energy up to $\sqrt{s} = 319$ GeV. Four experiments took data at different interaction points along the HERA ring. The ZEUS and H1 experiments, devoted to the study of the internal structure of the proton and searches for phenomena beyond the Standard Model, were operating with colliding beams. The HERMES experiment, dedicated to the investigation of the spin structure of nucleons, was a fixed-target experiment utilising the electron beam only, whereas HERA-B used only the proton beam, aiming at the measurement of CP-violation in the $B\bar{B}$-system. The HERA machine operated during the period 1996–2007 with a shut-down in 2000–2002. This shut-down marks the separation between the so called HERA I and HERA II data-taking periods.

---

[1] In what follows, the term “electron” is used for both electrons and positrons, unless otherwise mentioned.

The acceleration of electrons to their nominal energies was achieved in several stages. A schematic view of the acceleration chains is presented in Figure 3.1. Electrons were initially accelerated in LINAC I/II to 200 MeV. After injection into the DESY II synchrotron the electron energy was increased up to 7.5 GeV. Then, after reaching 12 GeV in PETRA, electrons were finally transported to HERA. The positron beam was obtained by pair production from bremsstrahlung emission of electrons.

The proton beam was obtained in several steps from a H^- ion source. At the first stage 50 MeV ions from a LINAC were transported to DESY III, where they underwent acceleration to 7.5 GeV and stripping off the electrons. Later, after achieving an energy of 40 GeV in the PETRA ring, the protons were finally injected into HERA.

3.1.1. Beam Structure

The usage of Radio-Frequency acceleration cavities at HERA lead to a disctinct time structure of the beams. Protons and electrons were grouped into bunches separated by ~ 28.8 m, which corresponds to 96 ns time intervals. Not all bunches were filled. The so called pilot bunches, for which either the electron or proton "bucket" was not filled, were used for the study of the interaction of the beam with residual gas in the beam vacuum pipe. Bunches in which both proton and electron "buckets" were empty were used for the study of the cosmic event rate and other non-*ep* background.

3.1.2. Luminosity

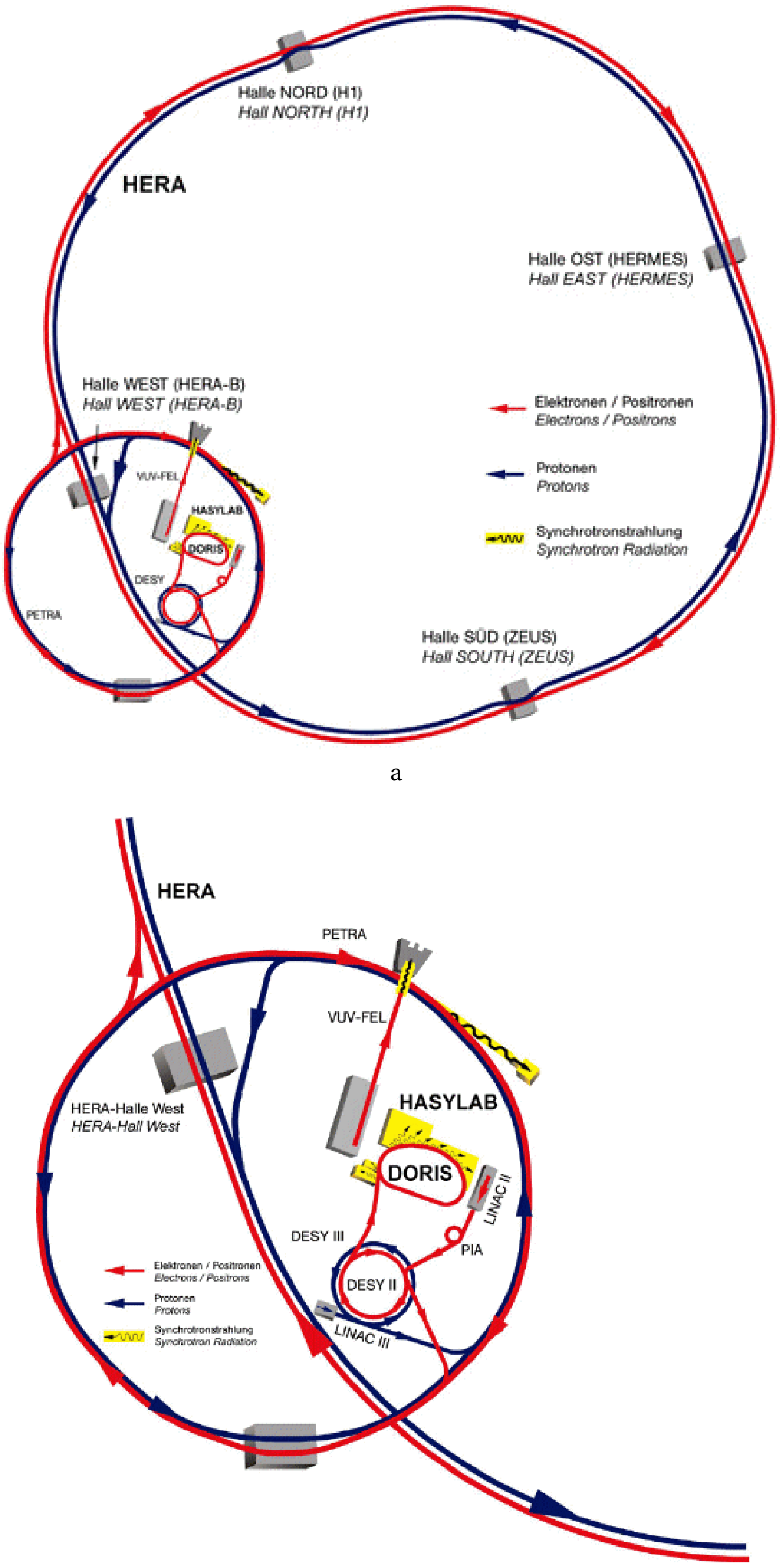
The crucial parameter of the collider that determines the rate of the collisions is the luminosity. It is related to the rate, R, of a process via the following expression:

$$R = \mathscr{L}\sigma, \tag{3.1.2.1}$$

where $\mathscr{L}$ is the instantaneous luminosity and σ is the cross section. The luminosity is related to the parameters of the colliding beams:

$$\mathscr{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}, \tag{3.1.2.2}$$

where f is the bunch-crossing rate, n_1, n_2 are the numbers of particles in the bunches and σ_x, σ_y the width parameters for beams with Gaussian profiles. An increase of the luminosity [98] at HERA II was achieved mainly by reducing the transverse size of the beams by installing additional focusing magnets close to the interaction points.



a

b

Figure 3.1.: Schematic view of electron and proton acceleration chains.

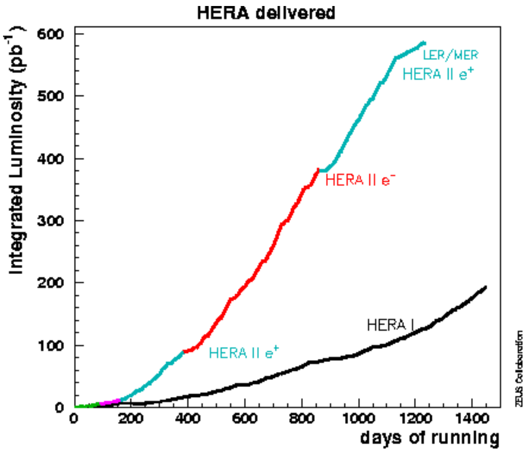


Figure 3.2.: HERA delivered luminosity

Period	Lepton Type	Delivered Luminosity	Gated Luminosity
2004 – 2005	e^-	204.8 pb^{-1}	152.26 pb^{-1}
2006	e^-	86.1 pb^{-1}	61.23 pb^{-1}
2006 – 2007	e^+	180.54 pb^{-1}	145.9 pb^{-1}

Table 3.1.: Information about HERA running periods used in the analysis.

To relate the number of events, N, to the reaction rate, the instantaneous luminosity has to be integrated:

$$N = \int R\,\mathrm{dt} = \sigma \int \mathscr{L}\,\mathrm{dt} = \sigma L, \tag{3.1.2.3}$$

where L is called the integrated luminosity and is often used to denote the amount of collected data. A comparison of the increase of the delivered integrated luminosity during the HERA I and HERA II running periods is presented in Figure 3.2. A typical fraction of 60% of the delivered integrated luminosity was available for the data analysis. The other 40% were lost due to various reasons: not fully operational detector components, inefficiencies of the data-acquisition system, specific trigger problems etc. A summary of the values of delivered and gated (recorded physical events) luminosity for the data-taking periods relevant for this analysis is given in Table 3.1.

3.1.3. Polarisation

The spin of the electron is naturally transversely polarised at storage rings like HERA due to the Sokolov-Ternov effect [99, 100]. Between the HERA I and

HERA II data-taking periods the accelerator was upgraded to provide longitudinally polarised lepton beams at the ZEUS and H1 experiments. The characteristic polarisation build-up time for HERA was about 40 min. To obtain longitudinally polarised beams at the interaction points, a chain of horizontal and vertical dipole magnets (spin rotators [101]) were installed on either side of ZEUS and H1. A typical longitudinal polarisation value of 30% – 40% was achieved.

The most important parameters of the upgraded HERA storage ring are summarised in Table 3.2.

Parameter	Electron beam	Proton Beam
Energy	27.5 GeV	920 GeV
Beam Current	60 mA	160 mA
Particle per bunch	3.5×10^{10}	10^{11}
Maximum number of bunches	210	210
Bunch length	7.8 mm	110 – 150 mm
Beam size	112×30 mm^2	112×30 mm^2
Polarisation time	30 min	–
Maximum instantaneous luminosity	5×10^{31} cm^{-2}s^{-1}	–

Table 3.2.: The HERA storage-ring parameters.

3.2. The ZEUS Detector

The data for this analysis were collected by the ZEUS detector, which was a general-purpose 4π-detector, designed for the measurement of the dynamics of ep interactions. The ZEUS detector consisted of the tracking system, the calorimeter system and muon chambers; the decision to store the data for further off-line analysis was taken by the three-level trigger system. In this section the main characteristics of the detector components relevant for this analysis are briefly described. A detailed description of the detector can be found in [102]. The schematic layout of the ZEUS detector is presented in Figure 3.3. A right-handed Cartesian coordinate system (CS) with the origin at the nominal interaction point was adopted in ZEUS (see Figure 3.4). The incoming proton momentum vector defines the positive Z direction. It was also called the "forward" direction. The positive direction of the X-axis pointed towards the centre of the HERA ring, while the positive Y-axis pointed upwards. The azimuthal angle, ϕ, was defined in the transverse X–Y plane, while the polar angle was measured with respect to the +Z axis.

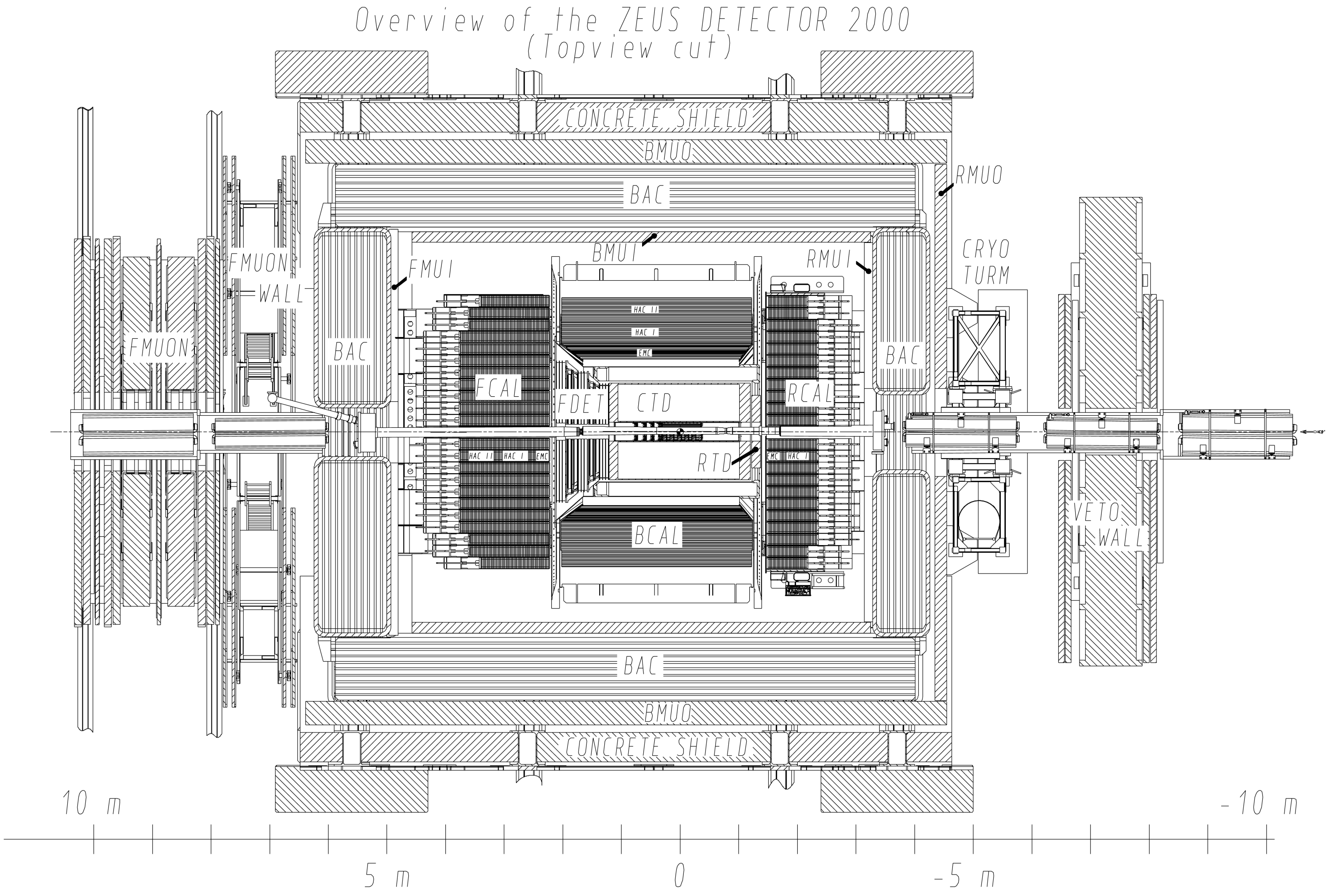


Figure 3.3.: Top view of the ZEUS detector

The pseudorapidity was an important quantity used in this analysis and is defined by the following expression:

$$\eta = -\ln\tan\frac{\theta}{2}. \tag{3.2.0.1}$$

For massless particles the numerical value of the pseudorapidity coincides with the normal rapidity defined as

$$y = \lim_{m\to 0}\frac{1}{2}\ln\left(\frac{E+p_Z}{E-p_Z}\right) \approx \lim_{m\to 0}\frac{1}{2}\ln\frac{\cos^2\theta/2+m^2/4p^2}{\sin^2\theta/2+m^2/4p^2} = -\ln\tan\frac{\theta}{2} = \eta. \tag{3.2.0.2}$$

The rapidity has simple transformation properties under Lorentz boosts. Transformation rules for the pseudorapidity can be also derived. In particular, the difference $\Delta\eta$ is a Lorentz invariant. This property was used in the definition of longitudinally invariant jet algorithm (see Section 2.1.1).

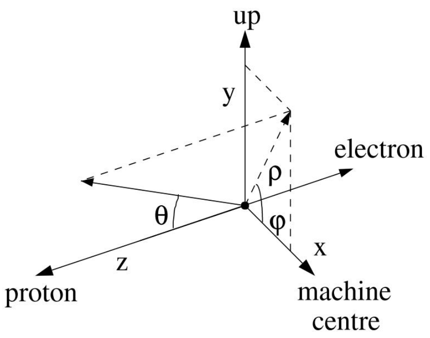


Figure 3.4.: The ZEUS coordinate system.

3.2.1. Tracking Detectors

The main parts of the tracking system of the ZEUS detector were the silicon microvertex detector (MVD) and the central tracking detector (CTD). These tracking detectors were used for the measurements of momenta and positions of charged particles as well as for the identification of the interaction and secondary decay vertices.

3.2.1.1. Microvertex Detector

The MVD [103] was a silicon-strip detector located in the vicinity of the beampipe for achieving an excellent resolution for secondary vertex tagging. It was installed during the 2000–2001 shut-down before the HERA II running period. The microvertex detector was divided into a forward (FMVD) and barrel (BMVD) parts (see Figure 3.5). In the barrel part, the silicon sensors were arranged in ladder structures grouped in two to three cylindrical layers surrounding the beam-pipe. In the forward direction, the sensors were assembled into four circular discs (wheels) oriented perpendicularly to the beam direction. Large number of read-out channels and high resolution allowed reliable separation of tracks emerging from hadronic jets. The main MVD characteristics are summarised in Table 3.3.

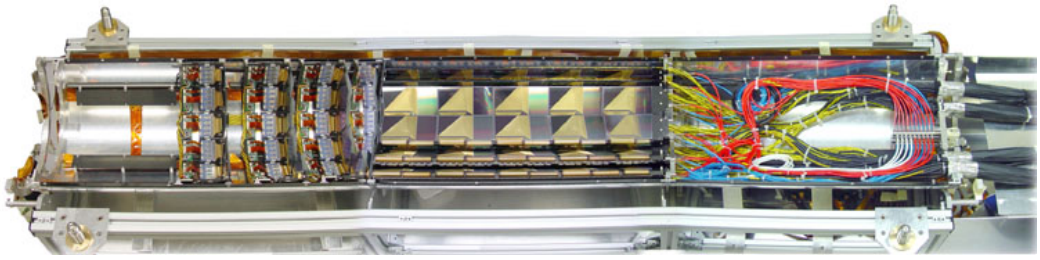


Figure 3.5.: One half of the silicon Microvertex Detector.

Parameter	Value
Polar-angle coverage	7°–160°
Read-out pitch	$120\,\mu m$
Single-hit resolution	$24\,\mu m$
Two-track separation	$200\,\mu m$

Table 3.3.: Silicon Microvertex Detector parameters

Other details of the sensor characteristics and performance can be found in [103–105].

3.2.1.2. Central Tracking Detector

The CTD [106–108] was a multi-wire cylindrical drift chamber used for the determination of charged-particle positions and momenta. The operation principle of CTD was based on detection of the ionisation of a gas mixture by the charged particles traversing the volume of CTD. The transverse momentum of the particle was determined from the curvature of the track in the solenoid magnetic field (see below). The dependence of energy losses within the volume of CTD on particle velocity was used to identify the particle type. Figure 3.6 shows one octant of the CTD. The wires were organised in nine superlayers (SL). Using different orientation of wires with respect to the CTD axis in odd and even superlayers made it possible to also accurately determine the Z position of the hit. The inclination angle in even superlayers was about ±5%, while it

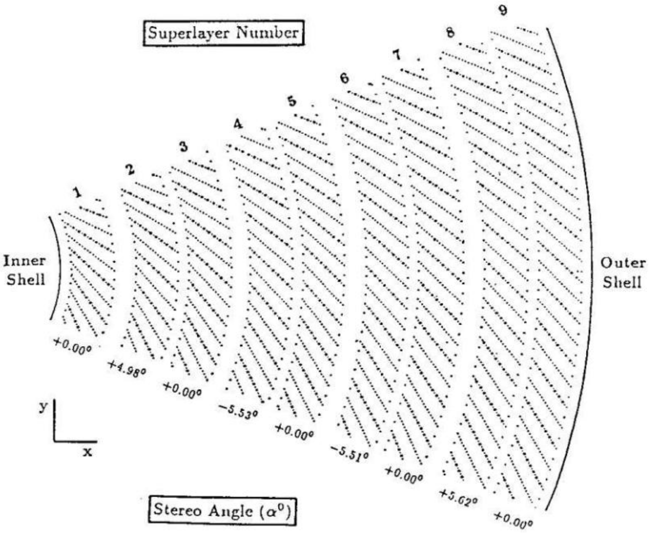


Figure 3.6.: Layout of a CTD octant.

was 0 in odd superlayers. For fast determination of the Z coordinate at the trigger level, the Z-by-timing technique was used. Specialised electronics to detect the distance along the Z-axis from the difference in the arrival time of pulses at each end of the wire was installed in SL3 and 5. The precision achieved in the Z position was about 3 cm.

Both CTD and MVD operated in a 1.43 Tesla magnetic field parallel to the Z-axis and produced by a thin superconducting solenoid surrounding the drift chamber. For CTD-MVD tracks that pass through all nine CTD superlayers, the momentum resolution was $\sigma(p_T)/p_T = 0.0029 p_T \oplus 0.0081 \oplus 0.0012/p_T$, with p_T in GeV. Other parameters characterising the CTD are collected in Table 3.4.

Parameter	Value
Inner radius	16.2 cm
Outer radius	85.0 cm
Length	241 cm
Polar-angle coverage	11.3°–168.2°
Position resolution	270 μm
Z resolution	1.4 mm (stereo) / 30 mm (timing)
Two track resolution	< 2.5 mm

Table 3.4.: Central Tracking Detector parameters

3.2.2. The Uranium Calorimeter

The uranium-scintillator compensating calorimeter (CAL) [109–113] was used for the measurement of the energy of the scattered electrons and positrons and of hadronic jets. The CAL covered 99.7% of the solid angle and consisted of the forward (FCAL), barrel (BCAL) and rear (RCAL) parts. The boost of the hadronic system determined the depth of the calorimeter necessary for the absorption of particles of different maximum energy in various parts of the detector.

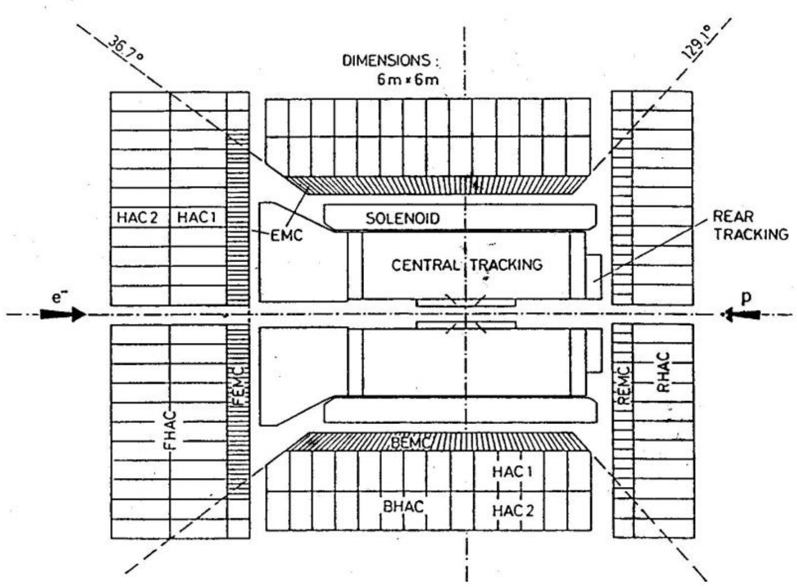


Figure 3.7.: Schematic view of the CAL along the beam axis.

The requirement to absorb a maximum energy of about 800 GeV resulted in $\sim 7\lambda$ depth of the FCAL section, where λ is the hadronic interaction length. The FCAL was longitudinally segmented into towers, each consisting of a single electromagnetic (EMC) and two hadronic (HAC) sections with a front-surface of 20×5 cm^2 and 20×20 cm^2, respectively. The FCAL towers were grouped into 23 modules. The EMC and HAC sections, also called cells, consisted of interleaved layers of depleted uranium (98.1% U^{238}, 1.7% Nb and 0.2% U^{235}) and scintillator (SCSN28) of 2.6 mm and 3.3 mm thickness, respectively. The thickness of the absorber and active plates was chosen in order to achieve equal response of the calorimeter to electromagnetic and hadronic showers of the same energy. The FCAL covered the polar angle range $2.2° < \theta < 39.9°$.

The BCAL had a very similar internal structure as the FCAL, but was composed of 32 wedge-shaped modules forming a cylindrical barrel surrounding the tracking detectors and superconducting coil. Due to lower hadronic activity in the BCAL, its thickness was $\sim 5\lambda$. The BCAL covered the angular range $36.7^\circ < \theta < 129.1^\circ$.

The rear part of the calorimeter (RCAL) consisted of 23 modules of $\sim 4\lambda$ depth. Because of much lower energy in the electron direction, the RCAL had only one HAC section.

All three parts of the calorimeter had a symmetric layout with respect to the beam axis and covered the complete azimuthal range.

In total the CAL had 5918 cells. Each cell was read out by two photomultipliers. The light from the scintillator plates was guided to the photomultipliers through wavelength shifters attached to the opposite sides of a cell. Usage of two photomultipliers helped to avoid events with spontaneous discharge of one of the photomultipliers and minimised the amount of dead cells if one of the PMTs was not functioning.

The natural radioactivity of the uranium made it possible to perform a calibration of individual channels of the calorimeter on a daily basis by providing a stable reference signal. The CAL energy response was calibrated to ±1% using this technique. The calibration of the electronic readout was performed using test pulses simulating photomultiplier signals. A timing resolution of 1 ns was achieved for energy deposits > 4.5 GeV. The energy resolution of the CAL measured under test beam conditions was

$$\frac{\sigma(E)}{E} = \frac{18\%}{\sqrt{E}} \oplus 1\% \quad \text{for electrons} \tag{3.2.2.1}$$

and

$$\frac{\sigma(E)}{E} = \frac{35\%}{\sqrt{E}} \oplus 1\% \quad \text{for hadrons,} \tag{3.2.2.2}$$

where E was the incident particle energy. The 1 ns time resolution of the calorimeter was utilised for rejection of non-*ep* background by providing fast signals to the trigger system. The particle incident angles, determined by the direction of the shower in the CAL with respect to the primary vertex position, were measured with about 10 mrad precision.

Other characteristics of the ZEUS calorimeter are summarised in Table 3.5.

	FCAL	BCAL	RCAL
Polar angle coverage	$2.2^\circ < \theta < 36.7^\circ$	$36.7^\circ < \theta < 129.1^\circ$	$129.1^\circ < \theta < 176.2^\circ$
Pseudorapidity coverage	$4.0 > \eta > 1.1$	$1.1 > \eta > -0.74$	$-0.74 > \eta > -3.4$
EMC section depth	$25.9X_0$	$22.7X_0$	$25.9X_0$
Total module depth	7.14λ	5.1λ	3.99λ

Table 3.5.: Geometric dimensions of the calorimeter modules, where X_0 and λ are the radiation and interaction lengths, respectively.

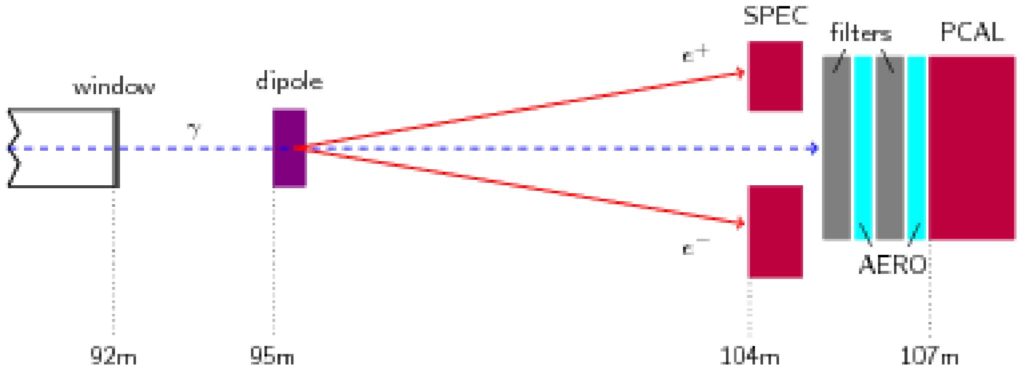


Figure 3.8.: Layout of the ZEUS luminosity monitor

3.2.3. Backing Calorimeter

The high-energy hadronic showers, that cannot be fully contained within the volume of the UCAL, deposited the remaining energy in the backing calorimeter (BAC) [114] consisting of proportional gaseous tube detectors located within the volume of the magnet iron yoke.

3.3. Luminosity Measurement System

The ep-brems strahlung process [115]

$$e + p \to e' + \gamma + p \tag{3.3.0.1}$$

was used to determine the instantaneous luminosity at ZEUS. The large cross section of this process allows rapid accumulation of large event samples in a relatively short time. Furthermore, the theoretical predictions for the rate of this reaction have a precision of better than 0.5%. The detection of the photon or the electron emerging from the interaction was used as an experimental signature of this process. The schematic layout of the luminosity monitor is shown in Figure 3.8. The photons from the reaction are emitted at small angles $\theta \lesssim 0.5$ mrad and leave the beam-pipe through a thin window located at 92.5 m from the interaction point. Approximately 9% of the photons were converted into e^+e^- pairs in the window. Electrons and positrons were deflected by a dipole magnet into the luminosity spectrometer (SPEC) [116], which was installed at 104 m distance downstream. The remaining photons were detected in the photon calorimeter (PCAL) [117–119] located at 107 m.

The instantaneous luminosity was determined from the formula:

$$\mathscr{L} = \frac{R_{ep\to ep\gamma}}{\sigma_{ep\to ep\gamma}}. \tag{3.3.0.2}$$

The background from the interaction of electrons with residual beam gas was estimated using electron pilot bunches and subtracted in the calculations.

The information from SPEC was not available for all runs, while the PCAL functioned continuously, so the luminosity value determined using the PCAL was used on a run-by-run basis. The systematic uncertainty on the measurement obtained with SPEC was lower than that for PCAL. It amounts to 1.8% and was used as the resulting precision of the luminosity value.

3.4. Polarisation Measurement System

As described in Section 3.1.3, at the interaction point the lepton beam is longitudinally polarised. In order to determine the lepton beam polarisation, two independent detectors LPOL [120] and TPOL [121] were used. The dependence of the Compton cross section on the orientation of the electron spin was utilised. In both cases a laser system was used as a source of incident photons.

Circularly polarised continuous green argon-ion laser light was used for the measurement of the transverse polarisation. The photons from the laser beam collided with the transversely polarised electrons. The scattered photons were detected in a dedicated calorimeter. The asymmetry of the photon-scattering-angle distribution was used to determine the polarisation value.

A similar measurement was performed with longitudinally polarised electrons. A Nd:YAG laser pulse was transported to the collision region, where circularly polarised photons backscatter from the electron beam. Switching the circular polarisation of the photon beam from left-handed to right-handed, the asymmetry in the total energy of the scattered photon was determined and the electron beam polarisation value was derived.

3.5. Data Quality and Trigger System

Due to limited data processing speed and storage capabilities, not every ep collision could be recorded. Furthermore, the total event sample is dominated by non-ep background from beam-gas interactions or cosmic-ray showers. In order to reduce the event rate to an acceptable level and efficiently reject the background, a sophisticated three-level trigger system [122–124] was used. The architecture of the ZEUS trigger system is shown in Figure 3.9.

3.5.1. First-Level Trigger

Each detector component was coupled to its own hardware first-level trigger operating with general information such as regional energy sums (CAL) [125], track multiplicity (CTD) [126] or muon tracks. The information from 26 consecutive bunch crossings (2.5 μs) was stored in a 46-event-deep pipe-line and was analysed in parallel streams. The combined information from each detector component was sent to the programmable Global First Level Trigger (GFLT), which selected the events that should be kept for the consideration at the second level. The decision was taken within 1.9 μs, which corresponds to 20 bunch crossings. If an event was accepted, the analogue information from different detector components was digitised and transferred from the pipelines to the data buffers. The GFLT had 64 bits, so-called "slots", corresponding to different event categories (see Section 5.4.1 for the description of the FLT slots used in the analysis). By using the FLT the event rate was reduced to approximately 1 kHz. The time interval during which data taking was disabled while the FLT was processing detector information and therefore was not operational amounted to approximately 1% and was automatically accounted for by disabling the luminosity monitors when the trigger was busy.

3.5.2. Second-Level Trigger

The second-level trigger (SLT) [127] had more time to process information because events were arriving at a reduced rate. The information from each detector component was combined at the GSLT [128–130], which was based on a reconfigurable network of transputers [108]. The additional time available for the SLT allowed a better estimation of the position of the primary vertex, identification of calorimeter clusters and reconstruction of tracks. The timing information from the calorimeter system was used to reject non-ep background events efficiently at the SLT. The output rate of the second level was in range 50 – 100 Hz. The full information for accepted events was sent on to the event builder.

3.5.3. Third-Level Trigger

The third-level trigger (TLT) [131, 132] was a cluster of computer servers running complex algorithms for the vertex reconstruction, electron identification and reconstruction of event kinematic variables. The highly configurable architecture of the TLT made it possible to utilise of up-to-date calibration information as well as fine tuning of the selection algorithms. The output rate of the TLT of about 5 Hz was compatible with the storage capabilities, thus the information from the event builder was converted in ADAMO format [133] and written on magnetic tape.

3.6. Detector Simulation

The detector response to the particles traversing the detector volume was simulated using the MOZART program (based on GEANT 3.21 [134]). In MOZART the propagation of particles in the detector volume including motion in the magnetic field, multiple scattering, energy losses, particle decays and particle showers was implemented. The four-momentum components of the initial- and final-state particles used as an input were generated by one of the general-purpose event generators e.g. LEPTO or ARIADNE (see Section 2.3). These and other event generators had a common front-end interface program called AMADEUS. The behaviour of the trigger system was simulated with the ZGANA program. The ZGANA program kept the event even if it was not accepted by the trigger chain, thus allowing estimation of the trigger efficiency using the simulated events. The output of ZGANA was compatible with the format for the raw data from the detector and can subsequently be analysed by the same programs. The reconstruction of the real and simulated events was provided by the ZEPHYR program. The output of ZEPHYR was stored in ADAMO system format GAF (Generic ADAMO File). The information specific to each program was supplied through a set of steering and GAF files containing, for example, the magnetic field map or the shape of the distribution of the longitudinal component of the interaction-vertex position. All programs keep track of the generated particles at each step of the simulation process, thus providing access to the generator level.

The high-level generic routines for the event reconstruction relevant in most of the analysis were collected in the so-called ORANGE/PHANTOM library. The analysis ntuples in PAW or ROOT format containing the variables necessary for the analysis can be produced using the ORANGE or so-called EAZE jobs. These ntuples can be subsequently analysed by means of user-specific C++ or FORTRAN codes.

In this thesis the v08b version of the Common Ntuples was used for both data and MC. The diagram of the data flow in the simulation process is demonstrated in Figure 3.10.

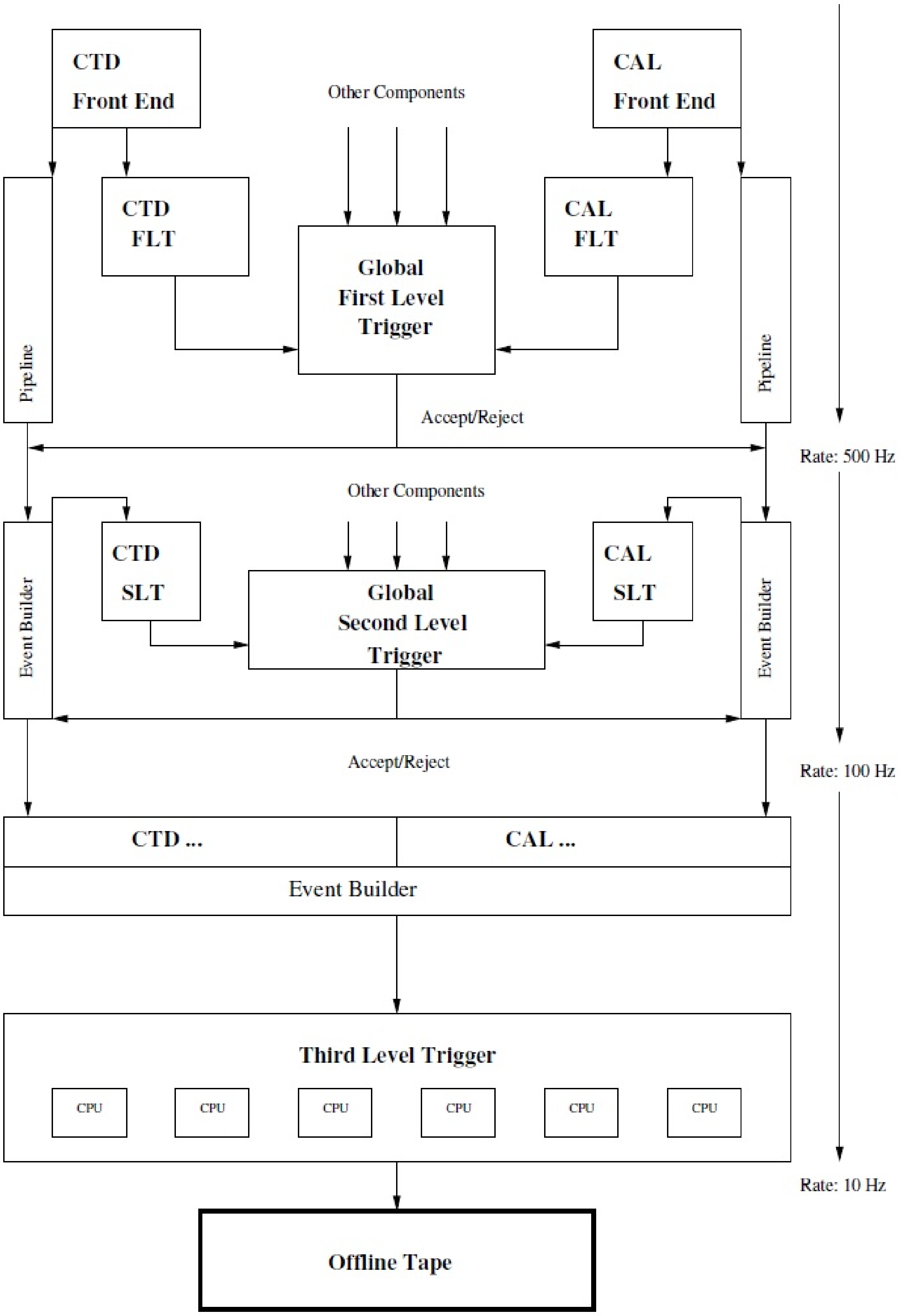


Figure 3.9.: The ZEUS trigger-system architecture.

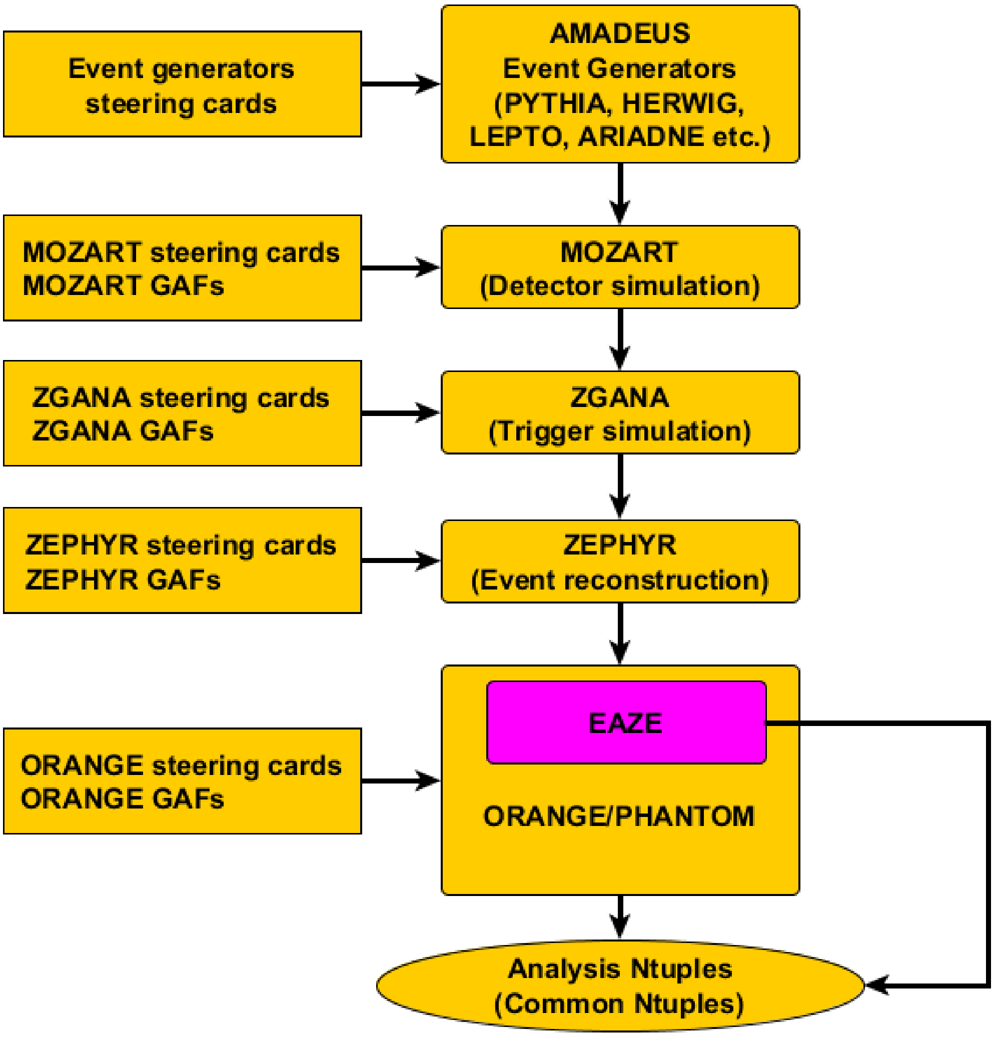


Figure 3.10.: The ZEUS detector-simulation data flow.

CHAPTER 4

Event Reconstruction

A precise reconstruction of the event kinematics and the hadronic final state is crucial for the jet analysis. This chapter provides a brief overview of the details of the track and vertex reconstruction, electron candidate identification, describes various techniques for determination of kinematic variables and introduces the objects necessary for the measurement of the hadronic final state.

4.1. Track and Vertex Reconstruction

As mentioned in Chapter 3, in the ZEUS detector tracks of charged particles are identified by means of the tracking detectors i.e. CTD and MVD. Track identification is initialised by the VCTRACK algorithm [135, 136]. In the pattern recognition stage, VCTRACK identifies 3 hits in the outermost layers of the CTD that may belong to a single track candidate. This allows estimation of the curvature and charge of the candidate. The algorithm proceeds inwards in the direction of the production point, identifying potential hit candidates within the search window defined by the uncertainty of the parameters of the track candidate. In order to improve the precision of the track-parameter estimate and remove outliers, the VCTRACK output hit candidates are supplied to the Kalman-filter based [137] algorithm RTFIT [138]. The output of the RTFIT contains the information necessary for determination of the track momentum and the point of origin[1].

---

[1] The track production point is defined for a given run as the position of the primary (secondary) vertex for vertex-fitted tracks or otherwise the point of closest approach to the centre of the beam distribution.

The location of the production vertex is determined by means of the same VCTRACK code. The set of tracks originating from the vicinity of the proton beam is identified and the weighted average of the coordinates of the individual track production points is determined. The calculation is performed by minimising a corresponding χ^2 function. In this process, tracks providing a large contribution to the χ^2 are typically discarded as outliers. The precision on the position of primary vertex is further improved by applying the iterative "deterministic annealing filter" (DAF) [139] algorithm and taking into account the beam-spot[2] constraint. Secondary decay vertices were also identified but not used in this analysis.

4.2. Hadronic Final-State Reconstruction

The hadronic final-state of the DIS process contains charged and neutral particles that are detected in the different components of the ZEUS detector. The neutral particles can only be detected in the calorimeter while charged particles produce tracks in MVD and CTD and also give a signal in the CAL. In this analysis the properties of the hadronic final state were determined using calorimeter cells and islands. Both approaches are briefly described below.

Calorimeter cells. The energy deposits in the elementary calorimeter units i.e. cells (see Section 3.2.2), were used as an input to the jet finder. In order to suppress noise signals, the minimum energy threshold $E_{\mathrm{EM}} > 50$ MeV and $E_{\mathrm{HAD}} > 100$ MeV was imposed on energy deposits in the electromagnetic and hadronic parts, respectively. Besides that, to avoid signals from spontaneous high-voltage discharge of the photomultipliers, the cells with a difference of the two photomultiplier signals exceeding 90% were also excluded. Since the scattered DIS electron does not contribute to the hadronic final state the cells attributed to the electron were excluded from consideration.

Calorimeter islands. Since single particles can give rise to signals in more than one cell, the energy deposits from adjacent cells satisfying the afore-mentioned requirements were combined, using a dedicated algorithm [140], into so-called calorimeter "islands". The energy of the island was calculated from the energy sum of the corresponding cells E_k and the position of the island $\vec{r}_{isl}$ was defined as the energy-weighted average of the coordinates $\vec{r}_i$ of the centres

---

[2] The beam-spot is a volume defined by the size of the intersection of the electron and proton bunches. The position of the beam-spot is defined as a centre of gravity of the coordinates of many primary vertices in a given run. The cross section of the beam-spot in the XY plane amounts to approximately $80 \times 20\ \mu m^2 \times 12$ cm.

of individual cells belonging to an island:

$$\vec{r}_{isl} = \frac{\sum_i w_i \vec{r}_i}{\sum_i w_i}, \tag{4.2.0.1}$$

$$w_i = \max\left(0, W_0 + \ln \frac{E_i}{\sum_k E_k}\right), \tag{4.2.0.2}$$

where w_i is a weight assigned to the energy deposit and W_0 is a tunable parameter defining the minimum fraction of the energy deposit contributing to the cell-island position. Given the energy of the calorimeter island or cell the four-momentum components of the vector corresponding to the energy deposit were determined using following equations:

$$E = E_{\mathrm{cell(isl)}}, \tag{4.2.0.3}$$

$$P_X = E \sin\theta \cos\phi, \tag{4.2.0.4}$$

$$P_Y = E \sin\theta \sin\phi, \tag{4.2.0.5}$$

$$P_Z = E \cos\theta, \tag{4.2.0.6}$$

where $E_{cell(isl)}$ is the cell (island) energy; the angles θ, ϕ are determined from the direction from the primary vertex to the centre of the calorimeter cell or cluster. The energy of the cell (island) was corrected for losses in the dead material in front of the calorimeter and/or in the inter-module gaps as well as for the "backsplash" effect. The backsplash process is characterised by particle energy deposits in locations different from their expected direction e.g. backscattering from the CAL surface or showering of the particles in front of the calorimeter. Non-uniformities of the response in different parts of the CAL were also taken into account. In this thesis the island information was mainly used for the reconstruction of kinematic quantities (see below).

4.3. Electron Identification

Unlike in photoproduction and CC DIS events, the presence of the scattered electron in the final state is a distinct feature of high-Q^2 NC DIS, therefore the identification of an electron candidate in an event is used to discriminate between different event classes. The kinematic variables characterising an event, like Q^2, y and x can be determined from the measured electron quantities E'_e and θ_e (see Section 4.4), therefore an unambiguous identification of the scattered electron as well as a precise and unbiased measurement of the electron variables is crucial. Two methods of electron identification based on different combinations of the information from various detector components were employed in this thesis.

The SINISTRA algorithm [141] was used as a nominal electron-identification algorithm in this analysis while the EM algorithm [142, 143] was used for the estimation of the systematic uncertainty attributed to the electron identification procedure (see Section 8.3). The former algorithm is based on a neural-network pattern-recognition technique [144]. The cell islands, described above are used as an input to the electron finders. The information about the longitudinal and lateral distribution of the energy deposits in the CAL is used by SINISTRA as an input in order to discriminate between electromagnetic- and hadron-induced showers. All calorimeter clusters are ordered according to their probability of being of electromagnetic type ($P = 1$ - electromagnetic; $P = 0$ - hadronic).

The EM algorithm, on the other hand, combines the information about the energy distribution in the calorimeter with the information from the tracking detectors as well as kinematic features of the NC DIS events in order to estimate the probability, that a particular calorimeter cluster is the true scattered electron. A detailed comparison of two algorithms can be found in [145].

The energy of the electron candidate is determined from the sum of the energy deposits in the calorimeter and corrected for non-uniformities and energy losses using a dedicated electron-energy correction procedure. When a matched track pointing to the electromagnetic cluster was found, the position of the candidate was replaced by the more precise information from the tracking system.

4.4. Kinematics Reconstruction

As described in Section 2.1, deep inelastic scattering at HERA can be characterised by two independent variables, for example, Q^2 and y. The values of the kinematic variables can be determined from the components of the four-momenta of the scattered electron or the hadronic system or from the combination of the two. Below the different approaches for the measurement of the DIS kinematics are described.

4.4.1. Electron Method

The electron method utilises the energy, E'_e, and the polar angle, θ_e, of the scattered electron only[3]. The kinematic quantities characterising an event are given

---

[3] It is assumed that the electron does not undergo initial- and/or final-state radiation that effectively reduces the electron energy.

by the following expressions (see Section 2.1):

$$Q^2_{\mathrm{el}} = 2E_eE'_e\,(1+\cos\theta_e)\,, \tag{4.4.1.1}$$

$$y_{\mathrm{el}} = 1 - \frac{E'_e}{2E_e}\,(1-\cos\theta_e)\,, \tag{4.4.1.2}$$

$$x_{\mathrm{el}} = \frac{Q^2_{\mathrm{el}}}{s y_{\mathrm{el}}}, \tag{4.4.1.3}$$

where, E_e, is the energy of the initial-state electron. It has been shown [146] that this method accurately reconstructs the event kinematic variables at high values of y, while at low values of inelasticity it has poor resolution.

4.4.2. Jacquet-Blondel Method

Transverse-momentum conservation and the almost complete hermiticity of the ZEUS calorimeter enables the reconstruction of the event kinematics based exclusively on the energy deposits attributed to the hadronic system. Jacquet and Blondel proposed a method [147] for the kinematic reconstruction for CC DIS events in which the final-state neutrino escapes the detector and cannot be measured. The kinematic variables are obtained from the following expressions:

$$y_{\mathrm{JB}} = \frac{\sum (E-P_Z)}{2E_e}, \tag{4.4.2.1}$$

$$Q^2_{\mathrm{JB}} = \frac{(\sum P_X)^2 + (\sum P_Y)^2}{1-y_{JB}}, \tag{4.4.2.2}$$

$$x_{\mathrm{JB}} = \frac{Q^2_{\mathrm{JB}}}{s y_{\mathrm{JB}}}. \tag{4.4.2.3}$$

The sum in these expressions runs over the reconstructed final-state objects (see above) excluding those belonging to the scattered electron. It is assumed that the target remnants escaping the detector volume have small transverse momentum and are therefore suppressed in the expressions. The accuracy of this method is limited by the hadron-calorimeter energy resolution, the presence of dead material and backsplash/backscattering in the calorimeter. Since this method does not require the detection of the scattered electron, it is the only choice for photoproduction events[4].

---

[4] In photoproduction events the electron is scattered at small angles and escapes detection in the beam pipe.

4.4.3. Double-Angle Method

The Double-angle method [148] benefits from exploiting the combined information on the scattered electron and the hadronic final state. The kinematic variables can be expressed in terms of the electron scattering angle, θ_e, and γ_{had}, which in the quark-parton model corresponds to the scattering angle of the struck quark and can be obtained from the following equation

$$\cos\gamma_{\mathrm{had}} = \frac{(\sum P_X)^2 + (\sum P_Y)^2 - (\sum E - P_Z)^2}{(\sum P_X)^2 + (\sum P_Y)^2 + (\sum E - P_Z)^2}, \tag{4.4.3.1}$$

where the sum runs over the energy deposits attributed to the hadronic final state.

The kinematic variables are obtained as follows:

$$y_{\mathrm{DA}} = \frac{\sin\theta_e \left(1 - \cos\gamma_{\mathrm{had}}\right)}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin\left(\gamma_{\mathrm{had}} + \theta_e\right)}, \tag{4.4.3.2}$$

$$Q^2_{\mathrm{DA}} = 4E_e^2 \frac{\sin\gamma_{\mathrm{had}} \left(1 + \cos\theta_e\right)}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin\left(\gamma_{\mathrm{had}} + \theta_e\right)}, \tag{4.4.3.3}$$

$$x_{\mathrm{DA}} = \frac{E_e}{E_p} \frac{\sin\gamma_{\mathrm{had}} + \sin\theta_e + \sin\left(\gamma_{\mathrm{had}} + \theta_e\right)}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin\left(\gamma_{\mathrm{had}} + \theta_e\right)}, \tag{4.4.3.4}$$

where E_p, is the proton initial energy.

Using these relations an expression for the energy of the scattered electron can be derived as

$$E_{\mathrm{el}}^{\mathrm{DA}} = \frac{2E_e \sin\gamma_{\mathrm{had}}}{\sin\gamma_{\mathrm{had}} - \sin\theta_e - \sin\left(\gamma_{\mathrm{had}} + \theta_e\right)}. \tag{4.4.3.5}$$

The measurements of the scattering angles usually have better resolution and are approximately independent of the absolute energy-scale of the calorimeter. It has been shown in [149] that this method is optimal in the phase space of this measurement and therefore this was used as a default in this analysis.

CHAPTER 5

Event Selection

In order to define the phase space of the measurement, ensure low fraction of background contamination and reasonable description of detector acceptance by MC simulations, selection cuts were applied to the data and MC events. In this chapter data quality as well as (on-)offline event-selection criteria are described. The chapter starts with a discussion of characteristic features of signal events. Then background sources are elucidated. The chapter proceeds with the discussion of the run selection, data-quality requirements and the list of online and offline selection cuts. It finishes with the description of the final event sample and comparison of data and MC distributions for selected events.

5.1. Signal Characteristics

As can be seen from the Eq. (4.4.1.1), large values of Q^2 in NC DIS events correspond to large scattering angles of the electron. This produces a well isolated electromagnetic shower in the calorimeter with a charged track pointing from the interaction vertex to the electromagnetic cluster (see Figure 5.1). At low and medium values of $Q^2 < 500\,\text{GeV}^2$, the electron typically scatters into the RCAL. For larger values of Q^2, the electron can be found in BCAL and at very high-Q^2 the electron scatters to the FCAL. In NC DIS at HERA the initial-state transverse momentum is zero, therefore final-state transverse momentum calculated from the sum over all energy deposits in the CAL must vanish:

$$P^2_{T,tot} = P^2_{X,tot} + P^2_{Y,tot} = \left(\sum_i E_i \sin\theta_i \cos\phi_i\right)^2 + \left(\sum_i E_i \sin\theta_i \sin\phi_i\right)^2 \approx 0\,\mathrm{GeV}^2. \tag{5.1.0.1}$$

Momentum conservation implies that the electron recoils from the hadronic system, which balances the scattered electron in the transverse-momentum plane.

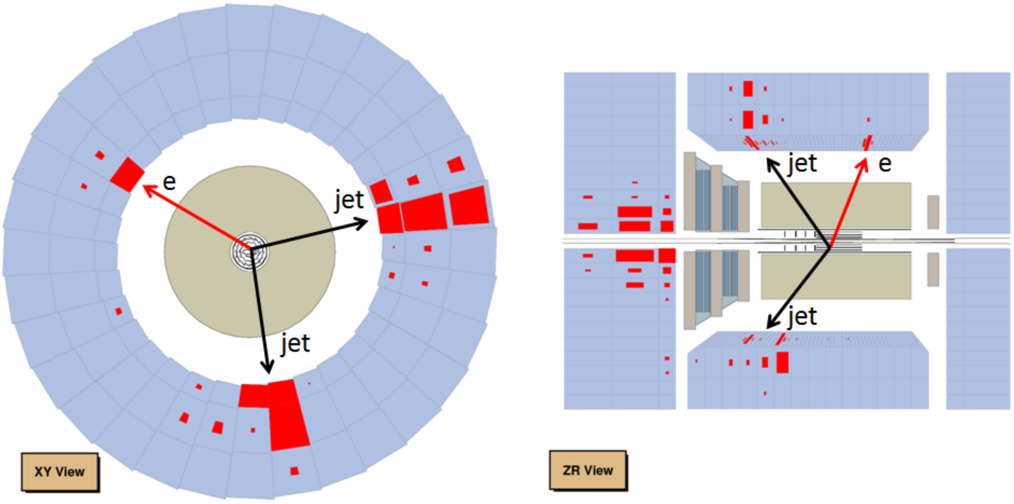


Figure 5.1.: A typical NC DIS event with $E - P_Z$ = 51 GeV, P_T = 2 GeV and Q^2 = 965 GeV^2 and two jets identified in the Breit frame with $E^{jet}_{T,B}$ > 8 GeV. Red blocks represent energy deposits in the calorimeter. The final-state hadronic jets (wide energy clusters in the BCAL) and the scattered electron are indicated with arrows.

Furthermore, in NC DIS processes the quantity δ, defined as

$$\delta = \delta_h + \delta_e = \left(E_p - P_{z,p}\right) + \left(E_e - P_{z,e}\right) = 2E_e = 55\,\mathrm{GeV}, \tag{5.1.0.2}$$

is also conserved[1], i.e. the same variable calculated from the final-state energy deposits has to fulfill

$$\delta = \sum_i \left(E_i - P_{Z,i}\right) = \sum_i \left(E_i - E_i \cos\theta_i\right) \approx 55\,\mathrm{GeV}. \tag{5.1.0.3}$$

In this equation the photon remnants provide vanishing contribution to δ since they move approximately parallel to z-axis in the negative direction and

$$E^{\gamma\text{-remnant}} - P_Z^{\gamma\text{-remnant}} \approx 0\,\mathrm{GeV}. \tag{5.1.0.4}$$

---

[1] In the following this is referred as longitudinal momentum balance.

Deviations of $P_{T,tot}$ and δ from the nominal quantity can be caused by e.g. undetected particles escaping the detector volume and/or due to the finite energy and spatial resolution of the calorimeter.

As was explained in Section 2.1.2, at leading order ($\mathcal{O}(\alpha_s)$) in the Breit frame a process with hard QCD interaction always has at least two hadronic jets. In experiment, however, one of the jets can fail to pass the selection criteria due to limited detector acceptance and/or resolution, therefore the signal events in this analysis were required to have at least one jet appearing in the Breit frame within the fiducial volume of the detector and having transverse energy exceeding some minimum energy threshold.

5.2. Characteristics of Background Processes

5.2.1. Photoproduction

The cross section of inelastic electron-proton scattering with $Q^2 \approx 0$ GeV^2 is large and serves as a potential source of background to the NC DIS analysis. At low Q^2, the scattered electron is deflected by a small angle and escapes in the beam pipe. Isolated electromagnetic clusters from e.g. π^0 decays can be misidentified as the final-state electron thus mimicking the signature of NC DIS processes. In photoproduction, as in NC DIS, the total transverse momentum vanishes but the quantity δ calculated from the final-state energy deposits will be smaller than 55 GeV by approximately twice the energy of the escaped electron. Besides restrictions on δ, the photoproduction events can be removed by applying additional cuts on isolation of the electron candidate because electromagnetic clusters from hadron decays are often accompanied by other hadrons from the hadronic final state.

5.2.2. Beam-Gas Interactions, Cosmics and Halo Muons

The electron and proton beams can interact with the residual gas within the volume enclosed by the beam pipe. The proton-gas collisions are characterised by large hadronic activity and multiple tracks emerging at low polar angles upstream from the nominal interaction point (see Figure 5.2). In such events, signals in the FCAL follow those from the RCAL and are separated by a time interval. Using the information about particles arrival time in forward and rear parts of the CAL as well as signals from the specific iron-scintillator detector (Veto-Wall [102]) located in the rear part of the ZEUS, these events can be efficiently rejected. In addition, the suppression of beam-gas events can be achieved by imposing restrictions on the longitudinal momentum balance, because such processes are characterised by a lower value of δ than that for NC DIS events.

In contrast to proton-gas collisions, the rate of electron beam interactions was found in a previous study [150] to be negligible.

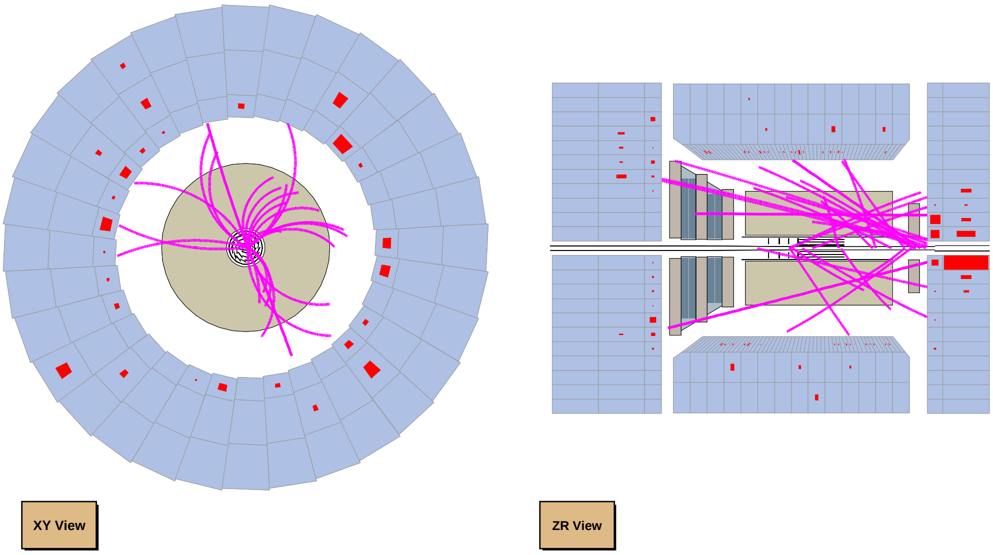


Figure 5.2.: An example of proton beam – residual gas collision identified using pilot-bunch crossing number. Many tracks emerge from the interaction point in a Z-region corresponding to the rear part of the ZEUS detector outside the acceptance of the tracking system.

Additional sources of background come from cosmic and beam-halo muons. Cosmic muons are produced in high-energy interactions of cosmic rays with the earth's atmosphere. Passage of cosmic muons though the volume of the detector results in a time difference between the signals in the upper and lower parts of the CAL (see Figure 5.3). Such a timing signature and a typically low track multiplicity were used to detect cosmic events.

In contrast to cosmic rays, beam-halo muons (see Figure 5.4), emerge from decays of pions produced in beam-gas interactions upstream of the detector and travel parallel to the beam axis, therefore the time difference between the signals in the RCAL and FCAL can be used to eliminate such processes, similarly to beam-gas interactions described above.

Moreover, cosmic and beam-halo events have non-zero total transverse momentum due to asymmetry of the energy deposits in the transverse plane. This information can also be used for the suppression of these background contributions.

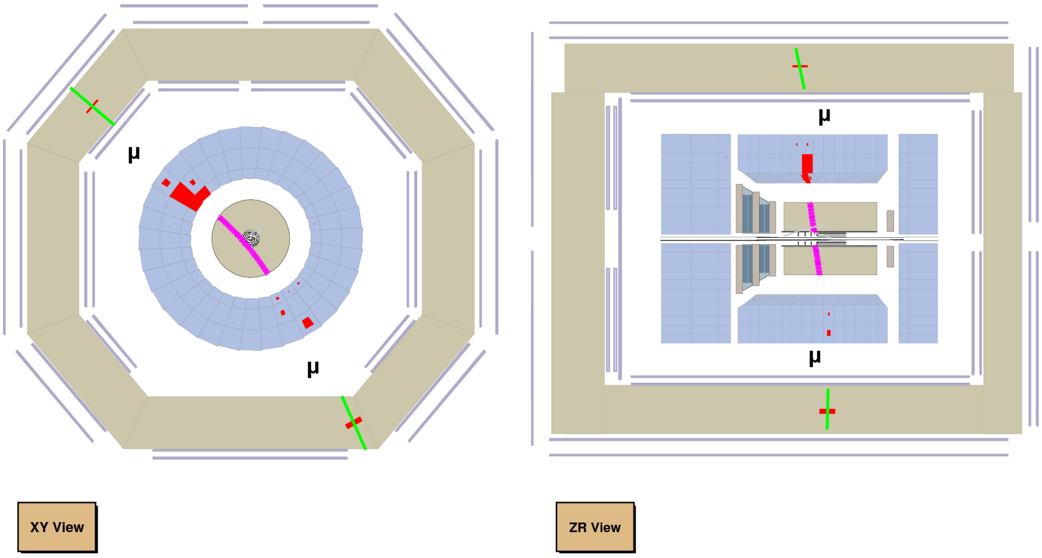


Figure 5.3.: A high-energy cosmic muon traversing the the volume of the ZEUS detector. Muon track segments are present in the upper and lower halves of the backing calorimeter and muon chambers. Energy clusters from the interaction of the muon with the material of the CAL can mimic signatures of hadronic jets.

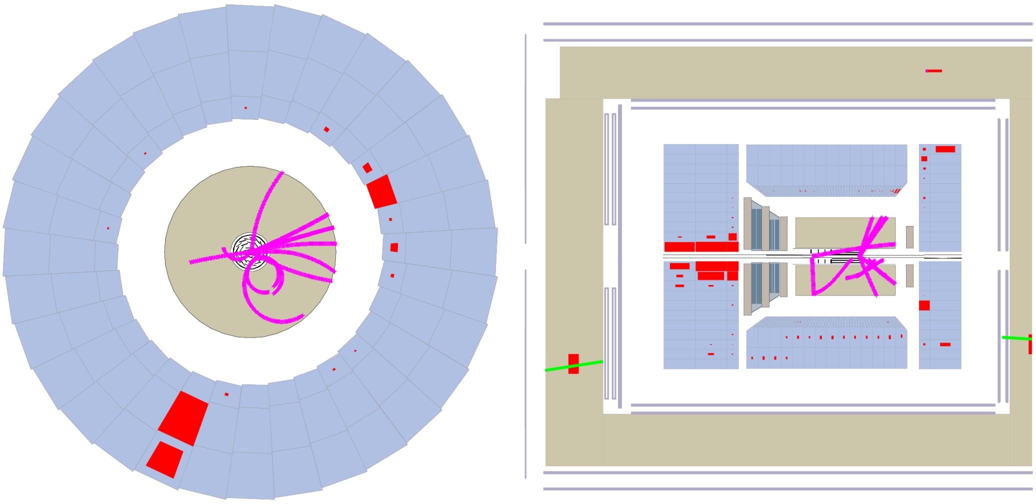


Figure 5.4.: A typical DIS event overlaid with a high-energy muon from the beam-gas interaction downstream from the interaction point. A beam halo can be seen muon traversing lower part of the detector.

5.2.3. QED-Compton scattering

The reactions of the type $ep \rightarrow e'\gamma p'$ or $ep \rightarrow e'\gamma X$ in which the initial or the final-state electron radiates a high-energy photon are called elastic and inelastic QED-Compton (QEDC) scattering, respectively. In case of inelastic reaction the proton breaks up resulting in hadronic activity in the forward direction while in elastic processes the proton escapes down the beam pipe staying intact.

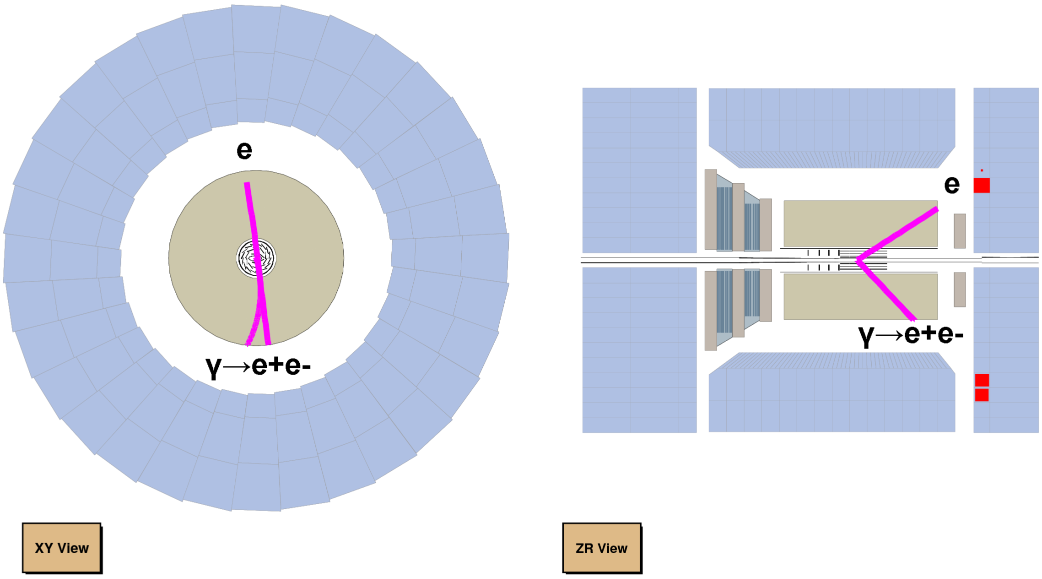


Figure 5.5.: An elastic QED-Compton scattering event as recorded by the ZEUS detector. The electromagnetic clusters from the electron and photon (two oppositely charged tracks from $\gamma \rightarrow e^+e^-$ conversion) are detected in the RCAL. The particles balance each other in P_T and there is no energy deposit in the forward direction.

QEDC events can be misidentified as high-Q^2 NC DIS processes when the electromagnetic cluster from the photon is wrongly reconstructed as the final-state electron and leads to incorrect determination of event kinematics. The characteristic features of such events are the presence of two isolated electromagnetic energy deposits in the calorimeter with approximately equal transverse momentum and low or vanishing hadronic activity in the FCAL (see Figure 5.5). As was demonstrated in the study [150] inelastic QEDC processes are well described by the DJANGOH MC [92] while elastic scattering is not well reproduced. Therefore such events have to be removed from the data sample by imposing dedicated cuts (see below).

5.3. Event Selection

The pre-selection of the NC DIS events relevant for the analysis starts naturally during the data-taking phase when the the ZEUS trigger system is used. Only those events for which a positive trigger decision was taken are stored on tape and can be further analysed offline. In the following sections general information about analysed data and MC samples as well as selection criteria applied to event samples are described.

5.3.1. Data and MC Sets.

This analysis was performed using the data recorded during the 2004 – 2007 HERA running period. The ep centre-of-mass collision energy during this period was $\sqrt{s} = 318$ GeV. The total integrated luminosity of the processed data sample amounts to 295 pb^{-1}. The luminosity values for the analysed data-taking periods are summarised in Table 5.1.

Period	Integrated Luminosity
2004 – 2005 e^-	133.6 pb^{-1}
2006 e^-	53.0 pb^{-1}
2006 – 2007 e^+	108.5 pb^{-1}

Table 5.1.: Luminosity values for the final data samples used for the jet cross section measurements.

Monte Carlo simulated events were used for the estimation of detector effects and the size of various contributions not accounted for in the perturbative QCD calculations (see Chapters 7, 8). The MC events for the mentioned data-taking periods were generated separately taking varying experimental conditions into account. The program packages LEPTO [93] and ARIADNE [94, 95], described in Section 2.3 were utilised for this purpose. In total, about 32 million NC DIS events with $Q^2 > 100$ GeV^2 were generated, leading to statistical uncertainty in the MC distributions smaller than 1% across the measured phase space.

Additional samples of photoproduction events with $Q^2 \approx 0$ GeV^2 generated using HERWIG [151, 152] and PYTHIA [153] programs were also used. These samples were utilised in the jet-photoproduction analysis [48] for the estimation of detector effects. Another sample of simulated events that almost fills the gap in Q^2 between the photoproduction and high-Q^2 regions was taken from the analysis of [154] and covers the interval 4 GeV $< Q^2 <$ 100 GeV. The photoproduction and low-Q^2 MC samples were used for the estimation of background contribution (see below).

5.4. On-line Selection

As described in Section 3.5 the ZEUS trigger has a three-level architecture. At each level, every event can be classified into different categories[2] according to event characteristics such as total and/or regional energy sums, presence of electromagnetic clusters, etc. Technically, when an event is recorded on storage-tape it is marked with specific flags, the so-called trigger slots or bits, which correspond to different event classes. The requirements imposed by specific trigger bits are listed below. Any of the specified trigger bits was required to be fired[3] in order to keep the event for further analysis.

5.4.1. FLT Trigger

The FLT selection is centred around the idea of having an event with large transverse momentum and significant electromagnetic energy deposit corresponding to the DIS electron. Besides that, every FLT bit imposes additional restrictions (not listed here) on the CAL timing and information from other detector components like Veto-Wall in order to suppress background processes.

- **FLT bit 40:** The total energy in the electromagnetic section of the calorimeter, $E_{\mathrm{EMC}}^{\mathrm{CAL}}$, exceeds 20 GeV and a condition on the track multiplicity, the so called *track veto* (see Section 6.3).
- **FLT bit 43:** The total transverse energy in the calorimeter, $E_{\mathrm{T}}^{\mathrm{CAL}}$, exceeds 20 GeV and at least one *good track*[4] is required.
- **FLT bit 50:** The total energy, E^{CAL}, greater than 15 GeV or $E_{\mathrm{EMC}}^{\mathrm{CAL}} > 10$ GeV or $E_{\mathrm{EMC}}^{\mathrm{BCAL}} > 3.5$ GeV or $E_{\mathrm{EMC}}^{\mathrm{RCAL}} > 2$ GeV and $E_{\mathrm{T}}^{\mathrm{CAL}} > 1$ GeV and a good track is required.

5.4.2. SLT Trigger

The second-level trigger was used to suppress further beam-related background. The information from the ZEUS global tracking trigger was utilised at the SLT for the reconstruction of the interaction vertex position and to reject events originating from background processes. Moreover, the calorimeter

---

[2] A single event can be attributed to **several** categories.

[3] Which corresponds to a logical **OR** between individual bits.

[4] FLT Track class > 1 (see Section 6.3).

timing information was used extensively to suppress beam-gas and cosmic-ray background. For example, hadrons emerging in the vicinity of the nominal interaction point are characterised by approximately the same arrival time in different parts of the CAL, in contrast to cosmic-shower events for which the signals from the upper part of the BCAL will arrive earlier than those from the bottom part. Thus, given the high timing resolution of the CAL, the background processes can be efficiently discriminated at SLT.

The SLT bits DIS01, 04, 07 were used in the analysis. Besides the required FLT bits, the following restrictions were imposed:

- **SLT bit DIS01:** The total energy in the electromagnetic sections of the calorimeter, $E_{\mathrm{EMC}}^{\mathrm{RCAL}} > 2.5$ GeV or $E_{\mathrm{EMC}}^{\mathrm{BCAL}} > 2.5$ GeV or $E_{\mathrm{EMC}}^{\mathrm{FCAL}} > 10$ GeV or hadronic energy in the forward region $E_{\mathrm{HAC}}^{\mathrm{FCAL}} > 10$ GeV and $E - P_Z > 29$ GeV;

- **SLT bit DIS04:** The total energy in the electromagnetic sections of the calorimeter, $E_{\mathrm{EMC}}^{\mathrm{RCAL}} > 2.5$ GeV or $E_{\mathrm{EMC}}^{\mathrm{BCAL}} > 2.5$ GeV or $E_{\mathrm{EMC}}^{\mathrm{FCAL}} > 10$ GeV or hadronic energy in the forward region $E_{\mathrm{HAC}}^{\mathrm{FCAL}} > 10$ GeV and $E - P_Z > 19$ GeV;

- **SLT bit DIS07:** The energy of the SLT electron candidate $E_e^{SLT} > 5$ GeV.

5.4.3. TLT Trigger

The final stage of the trigger selection was based on the identification of the scattered DIS electron. The reduced read-out rate at the TLT allowed application of the complex reconstruction algorithms much more closely related to those used in the offline analysis. The TLT bit DIS03 together with DST bit 12 imposing loose cuts on the electron candidate were used in the selection chain and are summarised in the Table 5.2, where E'_{el} denotes the energy of the scattered electron candidate and R_{el}^{CAL} denotes the distance of the centre-of-gravity of the electron cluster to the beam-line in the XY plane.

scattered electron energy	$E'_{el} > 4\,\mathrm{GeV}$
distance from the origin of the coordinate system	$R_{el}^{CAL} > 35\,\mathrm{cm}$
longitudinal momentum balance	$30 < E - p_Z < 100\,\mathrm{GeV}$

Table 5.2.: The requirements imposed on the events at the TLT.

5.5. Offline Selection

The trigger-level event selection cannot fulfil the signal-to-background discrimination requirements necessary for precision analysis because of limited processing time and significant complexity of the reconstruction algorithms for the objects such as secondary vertices, particle decays or jets. In addition, information about detector operating conditions such as the number of dead channels, changes in high voltage, CTD gas pressure, temperature, etc. are difficult or impossible to take into account at the trigger level. Thus, after the trigger-based pre-selection, a set of cuts is applied offline to ensure a low level of background and a high signal purity of the selected sample. Additional requirements are typically imposed on a data sample to restrict the phase space of the measurement to a region of well understood detector acceptance and efficiency.

5.5.1. Data-Quality Requirements.

The offline selection starts with removing runs with inappropriate detector operation. For this purpose during the data taking the functioning of the detector components was continuously monitored. The status of each of component was stored for every run in the so-called, “EVTAKE” flag. In this analysis EVTAKE=1 was required, indicating that all main detector components e.g. CTD, MVD, CAL were fully operational.

In addition to the EVTAKE flag, the LPOLTAKE and TPOLTAKE status records were used in order to verify the availability of the information from the corresponding lepton-beam polarisation detectors. At least one of the detectors was required to have good status in order to determine the polarisation value for particular runs.

5.5.2. Electron Selection

The next key step in the selection procedure is the electron-based identification of the high-Q^2 DIS event. To ensure high purity and reliability of the reconstruction of the scattered electron, the following requirements were imposed on electron candidates:

- **Probability:** The electron identification probability, as given by the SINISTRA algorithm, was required to be greater than 90%. If several electron candidates satisfied this criterion, the candidate with the highest probability was used for the reconstruction of event kinematics.

- **Energy:** The electron energy, E'_e was required to be greater than 10 GeV, to ensure the best electron-candidate reconstruction efficiency and high acceptance. Moreover, this cut helps to reduce the amount of wrongly identified low-energy electron candidates arising due to electromagnetic showers from $\pi^0 \to \gamma\gamma$, $\eta \to \gamma\gamma$ decays.

- **Isolation:** In order to remove events in which the electromagnetic shower of the electron candidate is contaminated by the energy deposits from hadrons, the fraction of the energy within a cone of radius of 0.7 units in the pseudorapidity-azimuth plane, not associated to the electron, was required to be less than 10%. The cone axis was defined by the electron momentum direction.

- **Track Matching:** The tracking system can be used to validate the electron identification, because electromagnetic clusters within the acceptance of the tracking system[5] that have no matching track are most likely photons. Moreover, the tracking information can be used to determine much more precisely the polar angle of the scattered electron than the determination from the electron cluster and primary vertex position only. Thus, if the electron candidate was within the tracking system acceptance region, the presence of a matched track was imposed. This track was required to have a distance of closest approach between the track extrapolation point at the front surface of the CAL and the cluster centre-of-gravity-position of less than 10 cm^2. The track energy as measured by the tracking system had to be greater than 3 GeV, taking into account energy losses by bremsstrahlung. In case the electron track was outside the acceptance region of the tracking detectors, the information from the calorimeter system was used to determine the position of the electron candidate.

- **Position:** To ensure full containment of the electromagnetic shower inside the fiducial volume of the calorimeter system and to avoid regions poorly described by the MC simulations, additional requirements on the position of the electromagnetic shower were imposed. The events in which the electron was found in the following regions were rejected:
  - $-104\,\text{cm} < Z_e < -98.5\,\text{cm}$ and $164\,\text{cm} < Z_e < 174\,\text{cm}$, where Z_e is the longitudinal coordinate of the centre of gravity of the of the electromagnetic cluster attributed to the electron — the so called “super-crack” regions between the BCAL and the RCAL or between the BCAL and the FCAL;
  - $|X_e| < 10$ cm and $Y_e > 80$ cm, where X_e and Y_e are the coordinates of the position of the electron energy deposit in the CAL. In this region

---

[5] The tracking system covers the region of polar angles restricted to $0.3 < \theta_e < 2.85$.

some of the calorimeter cells were removed to make room for the pipes transporting liquid helium to the superconducting solenoid;

- $36\,\mathrm{cm} < R_\mathrm{e}^\mathrm{RCAL} < 170\,\mathrm{cm}$, where $R_\mathrm{e}^\mathrm{RCAL}$ is the distance in the $X - Y$ plane from the origin of the ZEUS CS to the electron. The leakage of the electromagnetic showers from the electrons hitting the RCAL close to the inner or outer edges is not well simulated, especially at the trigger level [155]. These regions were therefore excluded.

5.5.3. Primary-Vertex Selection

Proper identification of the interaction point is important for the reconstruction of the kinematic variables. The longitudinal extent of the interaction region is determined by the length of the interacting bunches and the time structure of the beam. Beam-gas or cosmic-ray events are approximately evenly distributed along the longitudinal coordinate, while ep events appear with higher rate in the vicinity of the nominal beam-beam interaction point. The distribution of the longitudinal component of the position of the primary vertex has a Gaussian-like shape (see Section 6.2). Selecting the events with the primary interaction vertex fit satisfying $\chi^2 < 10$ and Z_vtx being within $\sim \pm 3\sigma$ of the width of the distributions suppresses the non-physics background contribution and ensures good understanding of the dependence of the acceptance of the calorimeter and tracking systems on the position of the interaction vertex. The mean value and the width of the Z_vtx changes between different data-taking periods. The final restrictions imposed on Z_vtx are detailed in Table 5.3.

Data-taking perod	imposed cut
2004-2005 e^-	$-32\,\mathrm{cm} < Z_\mathrm{vtx} < 30.1\,\mathrm{cm}$
2006-2007 e^+	$-28.5\,\mathrm{cm} < Z_\mathrm{vtx} < 26.7\,\mathrm{cm}$

Table 5.3.: Requirements imposed on the longitudinal cordinate of the position of the interaction vertex.

5.5.4. Longitudinal Momentum Balance

As discussed previously in Section 5.2.2, longitudinal momentum balance can be used to discriminate against photoproduction and beam-gas background. As mentioned, the actual value of the quantity $\delta = \sum_i \left(E_i - p_{z,i}\right)$ for a particular event may deviate from the nominal $\delta = 55$ GeV due to finite energy resolution and/or ISR effects[6]. Therefore the following requirement was implied on δ:

$$38 < \delta < 65\,\text{GeV}. \tag{5.5.4.1}$$

The lower value was chosen in order to reject photoproduction and beam-gas events, while the upper cut is required for the suppression of events with significant energy deposits in the rear direction, e.g. due to backspash processes that are poorly simulated in MC.

5.5.5. Transverse Momentum Balance

Due to finite resolution, the total transverse momentum of an NC DIS event may be greater that zero. The energy resolution of the calorimeter scales approximately as $1/\sqrt{E}$. In order to suppress beam-gas-related, cosmic-ray and charged current processes (see Section 2.1) with a misidentified electron, the following ratio was required to be small:

$$\frac{p_T}{\sqrt{E_T}} < 2.5\,\sqrt{\text{GeV}}. \tag{5.5.5.1}$$

5.5.6. Event Inelasticity

The DST bit 12 has a requirement on the inelasticity of the event. As was mentioned, the online information at the trigger level can be less precise than that available offline. In order to have the offline selection consistent with the trigger requirements, an additional cut

$$y_{el} < 0.75 \tag{5.5.6.1}$$

was imposed. Besides other restrictions listed in Table 5.2, this particular requirement was included into the definition of the DST bit 12 in order to rejects photoproduction background more efficiently, but since the measurement is performed in a much more restricted phase space (see Section 5.5.11), the effect of this offline cut was found to be very small.

5.5.7. Elastic QED-Compton

As described above, elastic Compton scattering processes ($ep \to ep\gamma$) are not well described by Monte Carlo simulations. Therefore, events containing two electron candidates satisfying the requirements listed in Section 5.5.2 and having transverse-momentum vectors opposite to each other, $\left|\phi^1 - \phi^2\right| > 3$, the ratio of transverse momenta, $0.8 < P_T^1/P_T^2 < 1.2$, and the total event-energy sum, excluding the contribution from the two electron clusters, $E_T^{CAL} < 3\text{GeV}$, were rejected.

---

[6] ISR leads to smaller values of δ.

5.5.8. Higher-Order QED Predictions

In order to ensure the validity of higher-order QED corrections (see Section 7.4) the region of phase space given by the requirement

$$y_{JB} \cdot (1 - x_{DA})^2 > 0.004 \tag{5.5.8.1}$$

was excluded because Monte Carlo predictions were found to be unreliable in this region [92].

5.5.9. Hadronic Scattering Angle

In order to suppress events with large hadronic activity in the forward region, where the simulations were found to be inaccurate [156], a cut on the projection of the hadronic scattering angle in the FCAL $R^{\mathrm{FCAL}}_{\gamma^{had}} > 18$ cm was implied. This constraint had only a marginal effect because inelasticity requirements suppress events with large hadronic activity in the FCAL.

5.5.10. Track Multiplicity

The residual beam-gas contamination was minimised by requiring the presence of at least one track, which crossed a minimum of three CTD superlayers and had transverse momentum $p_T > 0.2$ GeV. It was found that this cut had only a minor effect on the signal acceptance, as expected, since events with jets composed solely of neutral particles are very unlikely.

5.5.11. Phase Space

Phase-space cuts were performed to select the kinematic region of the reaction of interest. As mentioned earlier, two variables completely determine the kinematics of deep inelastic scattering. In this analysis, the exchanged boson virtuality Q^2 and inelasticity y were used for the NC DIS phase-space definition.

- **Photon virtuality**, $125 < {Q^2}_{DA} < 20000$ GeV^2; deep-inelastic-scattering processes with large four-momentum transfer were selected.

- **Inelasticity**, $0.2 < y_{DA} < 0.6$; the lower cut was imposed to reject events with large hadronic activity in the forward direction for which hadronisation corrections were found to be inaccurate. The upper cut was applied to ensure reliability of the detector acceptance corrections, since the region $y_{DA} > 0.6$ was poorly described in the MC [149].

5.5.12. Jet Selection

In order to identify jets, the k_T clustering-algorithm in the longitudinally invariant inclusive mode was used. The jet search was performed in the Breit frame, therefore the momentum vectors corresponding to the energy deposits in the CAL were transformed accordingly.

The phase space for jet production was limited by the following requirements:

- **Transverse energy** of the jets in the Breit frame was required to be greater than 8 GeV. Such a relatively high energy-threshold was chosen to ensure applicability of perturbative QCD, statistical significance of the event sample and high purity of the jet signal.

- **Pseudorapidity** of the jets had to be in the region $-1 < \eta_{lab}^{jet} < 2.5$. The lower cut corresponds approximately to the transition region between the barrel and rear parts of the CAL; since the RCAL has only one HAC section (see Section 3.2.2), the high-energy hadronic showers cannot be fully absorbed, leading to the increase of the absolute energy-scale uncertainty of the jets in that region. The upper η_{lab}^{jet} cut was motivated by the trigger limitations since events with large hadronic activity in the forward direction occur at a rate that exceeds the capabilities of the trigger.

The purity of the jet sample was enhanced by applying the following jet-cleaning cuts:

- good isolation of the electron candidate from hadronic jets was ensured by requiring the distance between electron and jet in the pseudorapidity-azimuth plane to be $\Delta R > 1$;

- initial-state electron radiation hitting the RCAL was often reconstructed as a jet. Such events were removed from the sample if a jet with $E_{T,B}^{jet} >$ 5 GeV and $\eta_{lab}^{jet} < -1$ was identified;

- jets with low transverse energy in the laboratory frame have large relative energy-scale uncertainty, therefore jets with $E_{T,lab}^{jet} <$ 3 GeV were excluded.

5.6. Final Event Sample

Summarising, the imposed requirements restrict the kinematic phase space of the measurements to

$$125 < Q^2 < 20000\,\mathrm{GeV}^2 \qquad 0.2 < y < 0.6, \tag{5.6.0.1}$$

$$E_{T,B}^{jet} > 8\,\mathrm{GeV} \qquad -1 < \eta_{lab}^{jet} < 2.5. \tag{5.6.0.2}$$

A comparison of the data and ARIADNE (LEPTO) MC distributions for the most important event observables after full selection is demonstrated in Figures 5.7 and 5.8, respectively. The presented plots include the distribution of the reconstructed longitudinal coordinate of the primary vertex, Z_{vtx}; kinematic observables: y_{DA}, ${Q^2}_{\mathrm{DA}}$; energy, azimuthal and polar angle distributions for the NC DIS electron candidate as well as the polar angle of the hadronic final state, $\cos\gamma_{\mathrm{had}}$; observables characterising momentum balance: $p_T/\sqrt{E_T}$, $(E - P_Z)$. In all figures the bin content of the MC distributions was rescaled to the number of events in the data sample by applying a constant factor:

$$\mathscr{R} = \frac{N_{\mathrm{data}}}{N_{\mathrm{MC}}}, \tag{5.6.0.3}$$

where N_{data} and N_{MC} are the number of events in the corresponding samples.

The primary vertex distribution (Figures 5.7(a), 5.8(a)) has a proper Gaussian-like shape reflecting the charge distribution and the timing structure of colliding bunches, explained in Section 3.1.1. Distributions of $y_{\mathrm{DA}}$and ${Q^2}_{\mathrm{DA}}$ (Figures 5.7(b),(a), 5.8(b),(c)) approximatively demonstrate the behaviour predicted by Eq. (2.1.0.7). In the phase space of the measurement, the total NC DIS cross section is dominated by the F_2 term, thus neglecting $F_{L,3}$ and, in addition, small y^2F_2 terms, an approximately linear dependence on y is obtained; the steeply falling nature of the Q^2 distribution is related to the $1/Q^2$ and $\left(Q^2 + M_Z^2\right)^{-1}$ nature of the photon and Z^0 propagators, respectively. Distributions of electron variables (Figures 5.7(b)–(d), 5.8(d)–(f)) on one hand reflect event kinematics related to Q^2 and y distributions and on the other hand the performance of electron identification and reconstruction. In particular, the flatness of the ϕ_{ele} distribution is a manifestation of the azimuthal symmetry of NC DIS scattering and the drop around $\phi_{ele} = \pi/2$ is attributed to low efficiency of the uninstrumented region of the RCAL, described in Section 5.5.2. The peak of the $(E - P_Z)$ distribution (Figures 5.7(g), 5.8(g)) is located at the value $\delta \approx 55\mathrm{GeV}$ predicted by longitudinal momentum conservation (see Section 5.1). The origin of the observed discrepancy in the description of the data by MC simulations is unclear and was taken into account in the systematic uncertainty (see Section 8.3) The distributions of missing transverse momentum and $\cos\gamma_{\text{had}}$ (Figures 5.7(h),(i) and 5.8(h),(i)) are a convolution of event kinematics and detector effects.

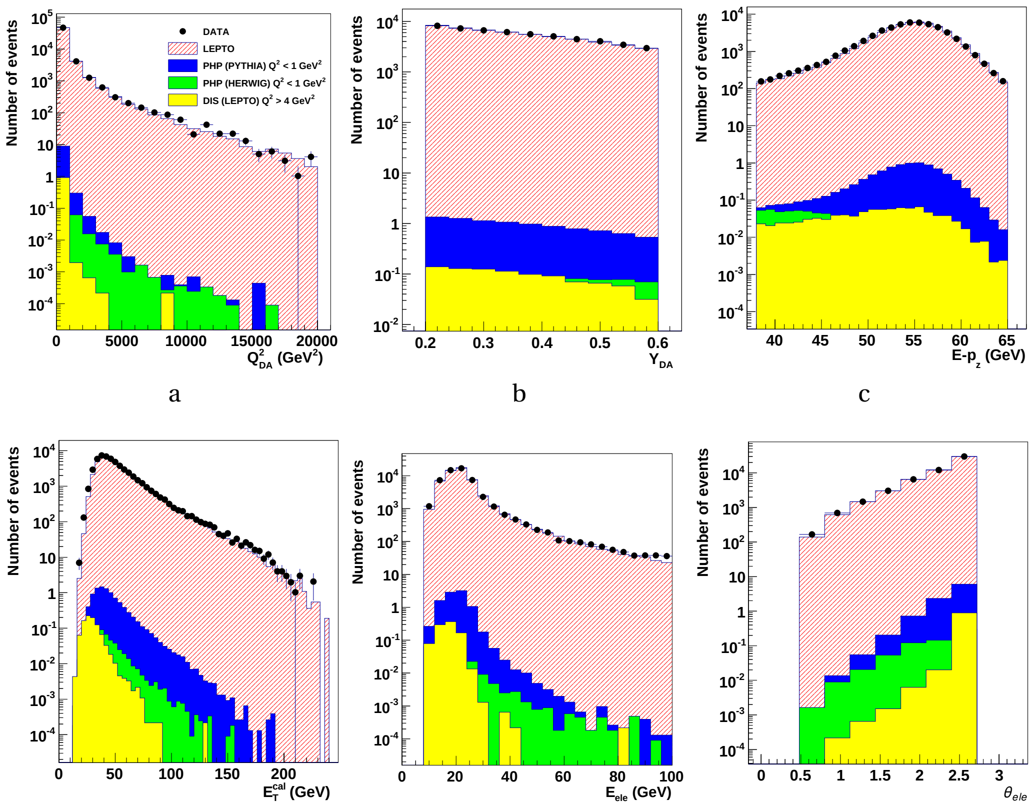
In general, the observed distributions are very well described by the Monte Carlo simulations when all corrections and reweightings (see Chapter 6) are applied to the data and MC, thus verifying the accuracy of the determination of the detector acceptance (Section 7.1) and systematic effects (Section 8.3) using the MC samples.

The residual background contribution from beam-gas interactions was estimated by applying the described selection procedure to events with a proton bunch crossing only. In this pilot-bunch sample, only a small number of events persisted after the selection. The visual inspection of such collisions revealed that selected events were typical NC DIS reactions with hard hadronic jets. It was concluded that for some runs the bunch crossing number was set incorrectly during the data taking[7]. Similarly, the full off-line selection chain was applied to events identified at the trigger level as cosmic. No events from this sample survived final selection, thus it was concluded that the background due to cosmic showers can be neglected.

The photoproduction and low-Q^2 NC DIS contamination was investigated utilising MC samples prepared for dedicated studies of jet production at low and very low exchanged-boson virtualities. The estimated contribution of these events to the final jet sample is illustrated in Figure 5.6. The photoproduction component estimate varies by a factor of two depending on the generator used. The PYTHIA MC predicts a cross section approximately twice as large as that given by HERWIG for the kinematic region $Q^2 < 1$ GeV^2. Due to the steeply falling nature of the Q^2 distribution, the NC DIS cross section obtained using LEPTO for a much wider region $4 < Q^2 < 125$ GeV^2 appeared to be approximately of the same magnitude as for the $Q^2 < 1$ GeV^2 region. As demonstrated in Figures 5.6(a, d) the background populates the low-Q^2 and low-E_T region of phase space and arises mostly due to incorrect identification of the DIS electron candidate. In total, the contamination from events with $Q^2 < 125$ GeV^2 to the measurement sample amounted to less than 0.2% and was neglected.

---

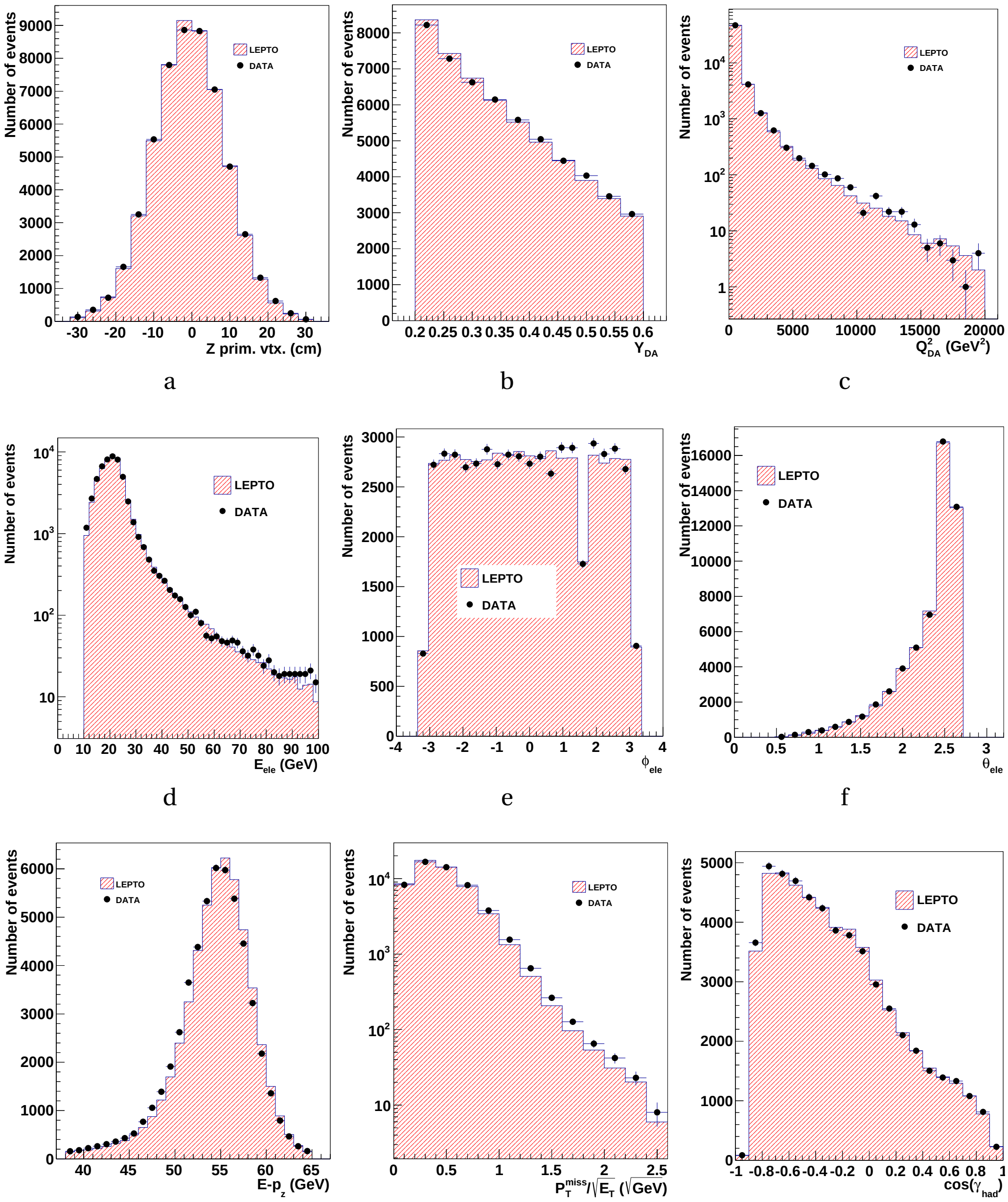
[7] No requirements on bunch-crossing number were applied for selections of the final sample.



a b c

d e f

Figure 5.6.: Comparison of corrected MC (LEPTO) and data distributions for event variables after full inclusive-jet selection. Various background components estimated using MC simulations are also shown.

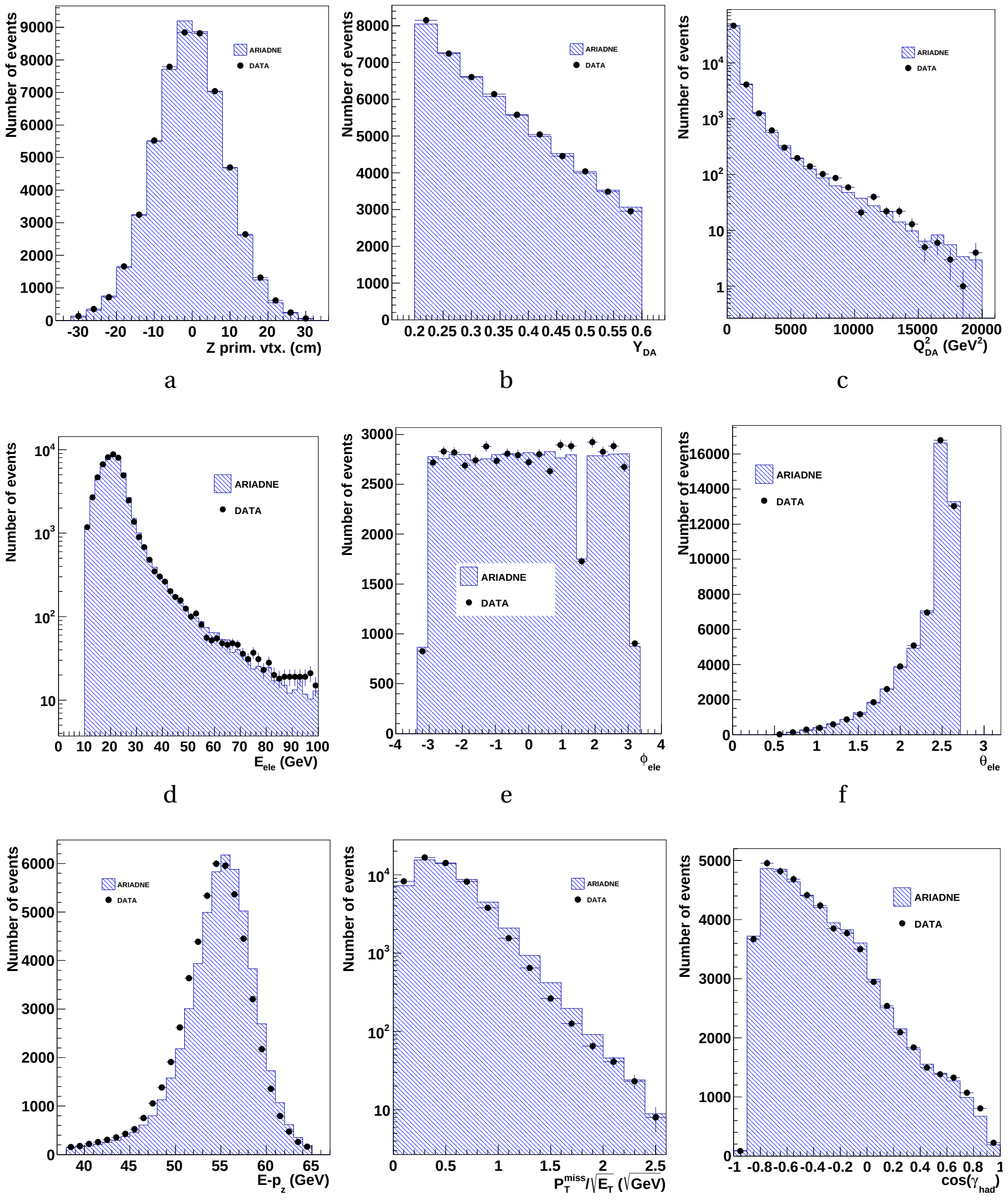


a b c

d e f

g h i

Figure 5.7.: Comparison of corrected MC (LEPTO) and data distributions for event variables after full inclusive-jet selection.



a b c

d e f

g h i

Figure 5.8.: Comparison of corrected MC (ARIADNE) and data distributions for event variables after full inclusive-jet selection.

CHAPTER 6

Calibrations and Corrections

In many high energy physics analyses, the estimation of detector effects is often based on Monte Carlo simulations. However, as will be described in Chapter 7, a reliable estimate of the influence of the detector response is only possible if the MC simulations accurately describe all relevant distributions. A discrepancy between data and MC may originate from two basic sources: inadequacy of the modelling of the ep physics process up to the generation of the final-state hadrons or of the response of the detector to the generated particles.

Improvement in the simulations is achieved by using more accurate predictions for the physical process and employing better detector simulation algorithms. However, obtaining a detailed description of the complete physical process can be a formidable task, therefore, often, a more feasible approach is used.

This chapter describes the corrections applied to the samples of MC generated events in order to obtain an accurate description of observed spectra by MC simulations. At the beginning, a brief account of the reweighting method is introduced. Then, the details of the longitudinal-vertex-position reweighting and the correction procedure for MC trigger efficiency are presented. The chapter proceeds with the details of the Q^2 and $E_{T,B}^{jet}$ spectra reweighting and the comparison of electromagnetic and hadronic energy scales in data and simulations. Afterwards, jet-energy-scale and jet-transverse-energy corrections are elucidated. The chapter concludes with the effect of the aforementioned corrections.

6.1. Reweighting Method

An improvement of the description of the data distribution is obtained by assigning weights to the MC events (*reweighting method*). The weights are usually functions of the kinematic variables and are adjusted in such a way that the simulations reasonably describe the data. The size of the weights is usually determined from an empirical fit to the observed spectra. The reweighting procedure is required to be independent of the reconstruction, therefore it must be based on MC *true* level information. When several quantities are reweighted, the final weight factor applied to the MC event is a product of individual weights, $w = \prod_i w_i$.

6.2. Reweighting of the Longitudinal Vertex Position

The detection and reconstruction of the scattered electron depend on the longitudinal position of the primary vertex, Z_{vtx}. In particular the detector and trigger acceptances vary with Z_{vtx}. The shape of the distribution of the primary vertex position is determined by the length[1] of the interacting bunches and thus depends on machine conditions. In order to suppress non-ep background, restrictions were applied on the primary vertex position (see Chapter 5). Therefore, since the luminosity measurements refer to the whole ep interaction region, any cuts on Z_{vtx} have an effect on the overall normalisation.

An accurate simulation of the Z_{vtx} distribution in the MC samples is therefore very important. The Z_{vtx}-reweighting procedure adopted here was developed in [149] and consisted of the following steps:

- in order to avoid a possible bias from the jet selection as well as from tracking restrictions at the trigger level, the FLT30 bit was required instead of the FLT40, 43, 50 bits that were used as standard in the analysis;

- selection cuts on the longitudinal position of the interaction vertex were removed;

- the Z_{vtx} distributions in the data and MC were fitted to the sum of four Gaussian functions,

$$f(\vec{a}) = \sum_{i=1}^{4} a_N^{(i)} \exp\left[-\left(Z_{\mathrm{vtx}} - a_\mu^{(i)}\right)^2 / \left(a_\sigma^{(i)}\right)^2\right]. \qquad (6.2.0.1)$$

---

[1] The space-charge distribution within a bunch typically has a Gaussian shape.

The reweighting factors, w_Z, were determined from the fit of the ratio of the normalised data and MC distributions to the function

$$w = f(\vec{a}_1)/f(\vec{a}_2) \tag{6.2.0.2}$$

using the parameters of individual fits $\vec{a}_{\mathrm{Data}}, \vec{a}_{\mathrm{MC}}$ as seed values;

- the weights were determined from the detector-level distributions as function of the reconstructed position, Z_{vtx}. They were then applied to the MC events as a function of the true position, $Z_{\mathrm{vtx}}^{\mathrm{true}}$. This substitution $Z_{\mathrm{vtx}} \mapsto Z_{\mathrm{vtx}}^{\mathrm{true}}$ can be made because the migration effects for the Z_{vtx} distribution were found to be very small and can be neglected.

The existing MC samples for the 2004–2005 e^- and 2006 e^- data taking periods describe the data very well in the whole interaction region. Only for the 2006/2007 e^+ running period did the MC not reproduce the data distribution reasonably[2]. In particular, disagreement between the measured and simulated distributions in the satellite-bunch and transition regions ($|Z_{\mathrm{vtx}}| > 30\,\mathrm{cm}$) was observed. Therefore, a reweighting of the longitudinal position of the primary vertex was implemented only for this MC sample. The comparison between data and MC distributions for the 2006/2007 e^+ data taking period before the reweighting is shown in Figure 6.1. Although, individual fits have relatively large χ^2/N_{df} values, which, in principle, may introduce a bias, the final fit to the ratio of the normalised distributions has $\chi^2/N_{\mathrm{df}} \approx 1$, and justifies using the fit results in the reweighting. After correcting the MC distribution (see Figure 6.2) good agreement between data and simulations was observed.

6.3. Track Veto Efficiency Correction

An accurate description of the trigger efficiency is an important ingredient in this analysis. As described in Chapter 3, the ZEUS trigger was used to select hard *ep* collisions with high acceptance and to reject non-*ep* background. At the FLT, most of the trigger bits utilised CTD information to veto events characterised by specific combinations of all and vertex-fitted tracks. For example, events with large track-multiplicity and few tracks fitted to the primary vertex (corresponding event classes are 1, 2, 8) originate typically from beam-gas collisions and thus have to be discarded. Cosmic-ray events with low track multiplicity (corresponding to event class 3) are also typically excluded. The definitions of all event classes is illustrated in Figure 6.3. This kind of trigger track-multiplicity requirements were generally called trigger track-veto.

---

[2] The primary vertex distribution for the HERA II running period was measured in a dedicated unbiased study of low-Q^2 NC DIS events [157] and was implemented in the MC production software

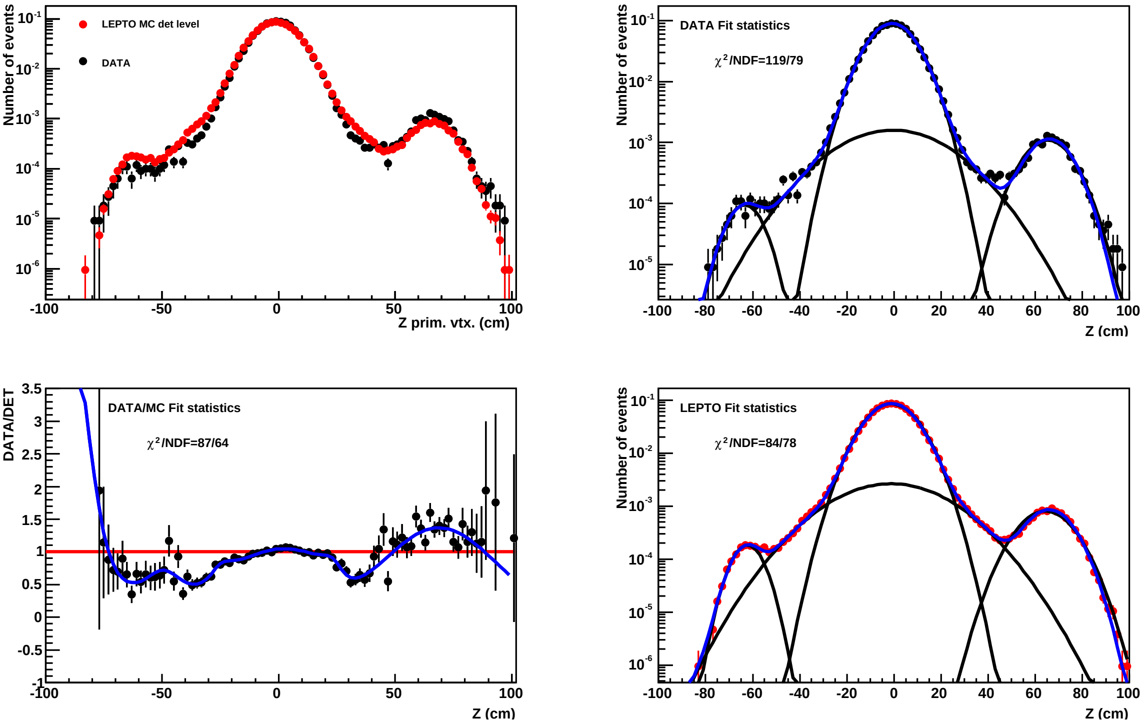


Figure 6.1.: The Z_{vtx} distributions in the data and LEPTO MC before the reweighting (top left panel) together with the individual fits of the data (top right panel) and MC (bottom right panel) to the function Eq. (6.2.0.1). The ratio DATA/MC and the fit to Eq. (6.2.0.2) is also shown (bottom left panel).

Two track-veto types were relevant for this analysis: "semi-loose" defined by the conditions: track-class ≤ 2 or track class = 8 and track multiplicity ≥ 26, and "tight" track-veto requiring track-class ≤ 2.

In order to check the description of the track-veto in MC simulations, a monitor trigger was used. The FLT30 required an isolated electromagnetic cluster in the RCAL and therefore was independent of the CTD information. The track-veto efficiency, expressed as the ratio

$$\epsilon_{\text{trk}} = \frac{N(\text{track veto} \wedge \text{FLT30})}{N(\text{FLT30})}, \tag{6.3.0.3}$$

where $N(\text{FLT30})$ is the number of events triggered by the trigger bit FLT30 and $N(\text{track veto} \wedge \text{FLT30})$ is the number of events in a subset satisfying additional track-veto requirements, was studied separately in data and MC for different data-taking periods. To determine $N(\text{track veto} \wedge \text{FLT30})$, the track veto was simulated offline by imposing additional restrictions on track quantities available at the FLT. The efficiency was investigated as a function of y_{DA} because

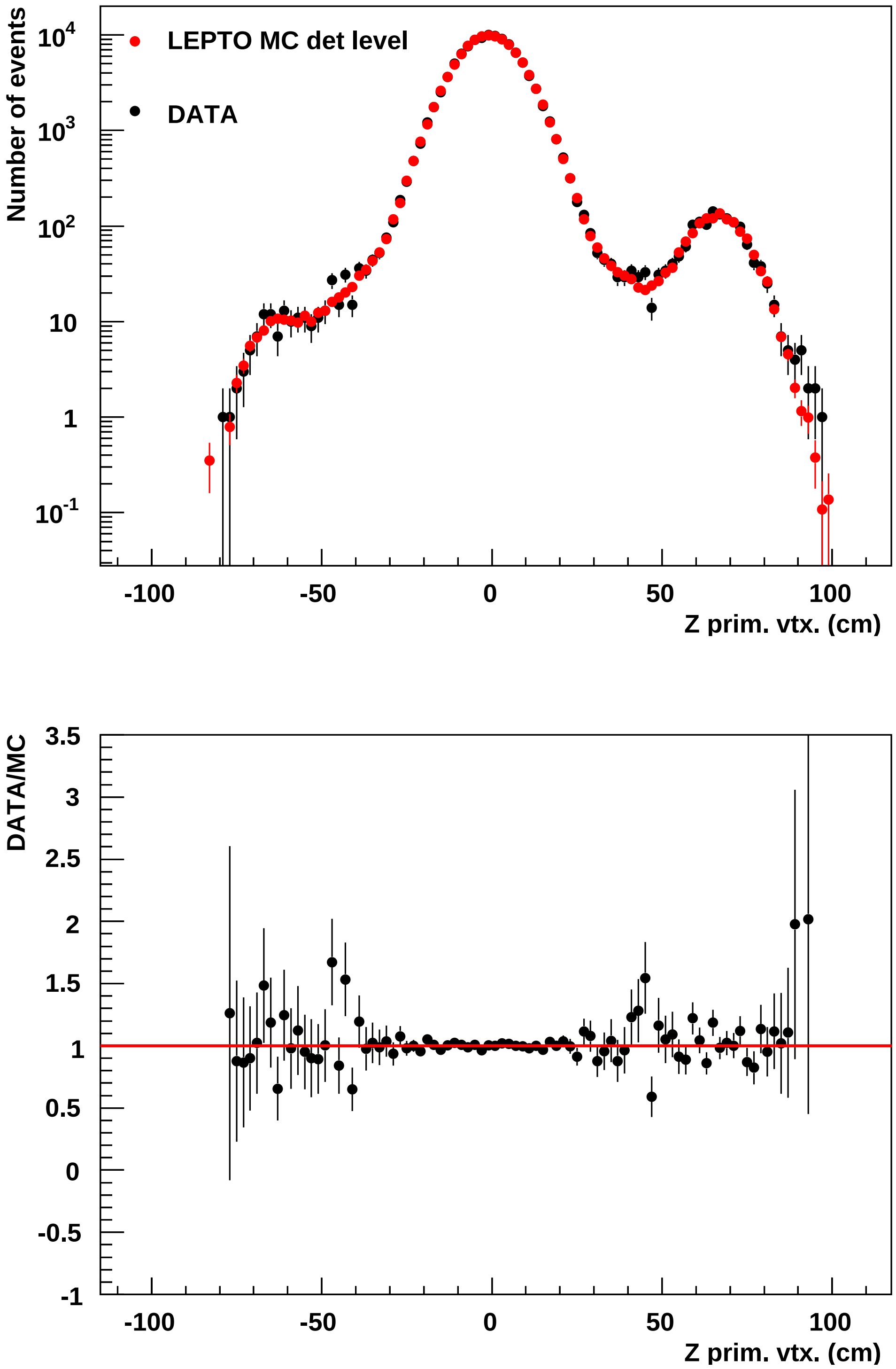


Figure 6.2.: Comparison of the Z_{vtx} in data and LEPTO MC distribution after reweighting (top panel) and the ratio of the two distributions (bottom panel).

this variable was strongly correlated with the amount of hadronic activity and thus with the track multiplicity. Only for the 2006/2007 e^+ period was a significant discrepancy observed (see Figures 6.4(a)–(b)). The corresponding ratios in the data and MC are shown in Figures 6.4(c)–(d). The disagreement between data and MC simulations was attributed to a bad description of the track-class distribution in the MC. In order to compensate for higher efficiency in the simulations, an additional correction was implemented. As the efficiency observed in MC was higher than that in the data it can be corrected by rejecting excess MC events. Therefore, for each MC event a uniformly distributed random number, r, was generated and the event was rejected if $r > f\left(y_{DA}\right)$. The quantity $f\left(y_{DA}\right)$ was obtained from a fit of the ratio of efficiencies in the data and MC to the function

$$f\left(y_{DA}\right) = a_0 + a_1 \cdot y_{DA}. \tag{6.3.0.4}$$

For both MC generators, reasonable fit quality was obtained (see Figures 6.4(c)–(d)), however the quality of the description of the data after implementing the correction was more important.

The correction was implemented in the LEPTO and ARIADNE samples, separately for different data-taking periods. As shown by the fits, the size of the correction depends approximately linearly on the value of y_{DA} and, on average, was typically less than 0.5% for 2004–2005 e^- and 2006 e^- samples and less than 3% for the 2006-2007 e^+ sample. It was observed that for the "semi-loose" track-veto, the same correction as for "tight" track-veto can be applied. The comparison of the track-veto efficiencies in the data and simulations after applying the correction is illustrated in Figures 6.5 (a)–(d). After the correction the data efficiency was very well described by the MC.

The systematic effects attributed to the MC track-veto correction were examined by investigating the trigger efficiency as a function of the CTD-FLT track multiplicity. The results of these studies are detailed in Section 8.3.

6.4. Virtuality and Jet-Transverse-Energy Reweighting

After the inclusive-jet selection described in Chapter 5 and after applying the Z_{vtx} and track-veto corrections explained above, the distributions for kinematic variables were still not very well described by the MC. In order to obtain a reliable estimation of the detector effects to be corrected for in the cross-section determination procedure (see Chapter 7), a further reweighting of the Monte Carlo distributions was employed.

The kinematic distributions in LEPTO and ARIADNE featured different properties, for example, an excess of events was observed in the high-Q^2 re-

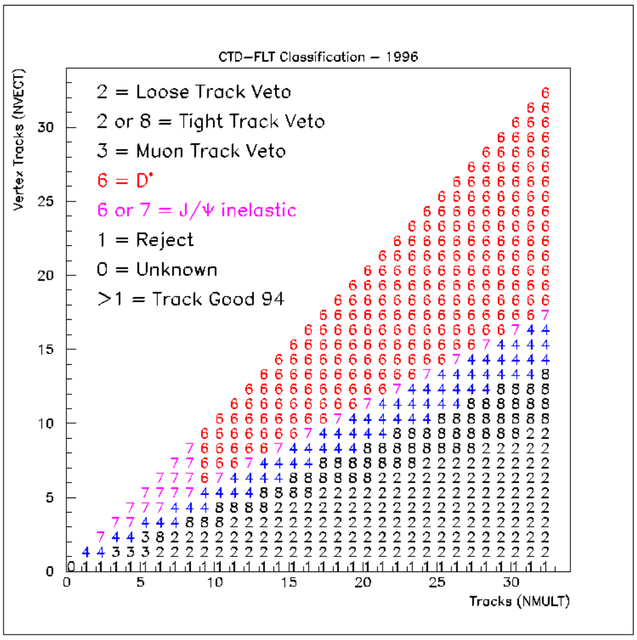


Figure 6.3.: The definition of track veto classes (taken from [158]).

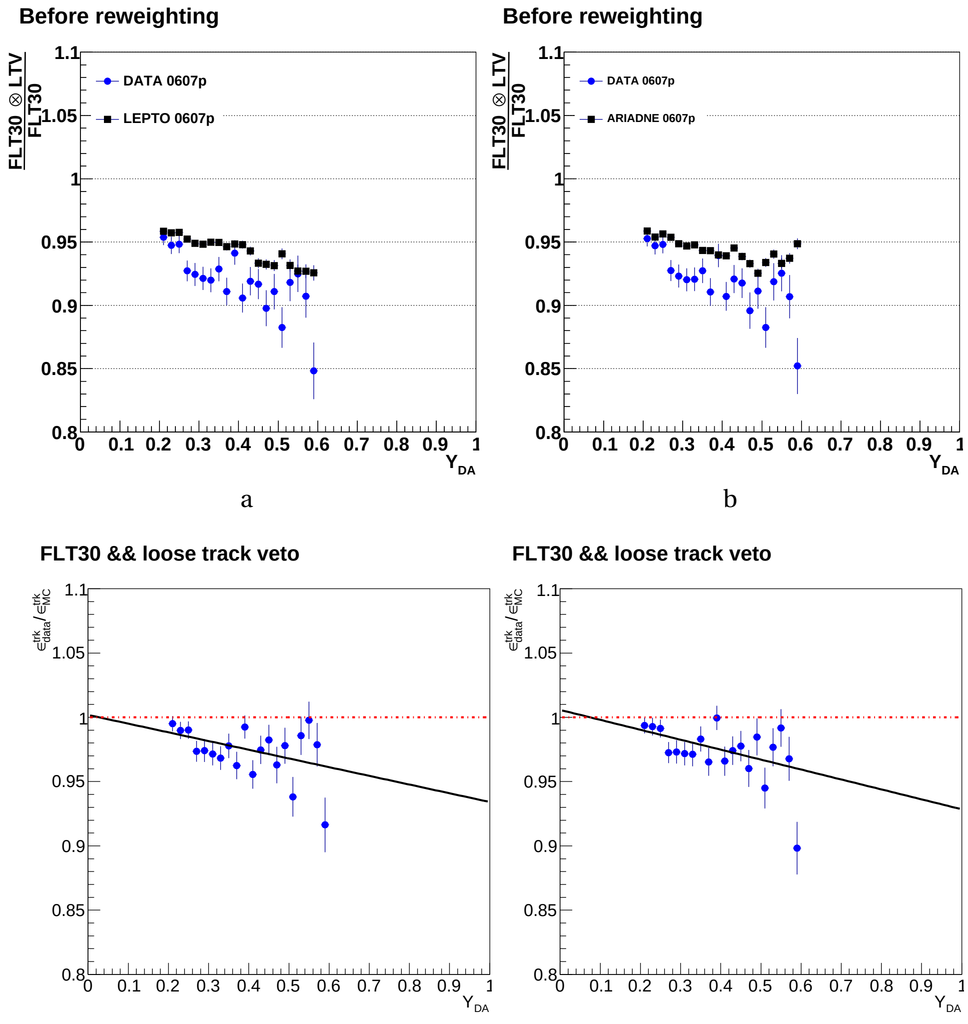
gion in the ARIADNE sample (see Figure 6.6(a)), while LEPTO had a deficiency in this region (see Figure 6.7(a)). In addition, as can be seen from Figure 6.6(b), the ARIADNE MC predicted a slightly harder jet spectrum than in the data. In order to take into account differences between the MC samples, two different reweighting procedures were developed for LEPTO and ARIADNE.

6.4.1. LEPTO Reweighting

A reweighting as a function of Q^2 was imposed on the events simulated using the LEPTO event generator. Two iterations were necessary to achieve an adequate description of the data distribution. In each iteration the ratio of the normalised Q^2 distributions in the data and MC was fitted to empirical functions to determine the weights, $w(Q^2)$. In order to mitigate the influence of detector effects, the data were corrected to the generator level using the acceptance correction factors determined at each reweighting iteration. The ansatz for the reweighting functions was as follows[3]:

---

[3] Other expressions for the reweighting functions were also tested. It was empirically found that these two relatively simple functions provide reasonable results.



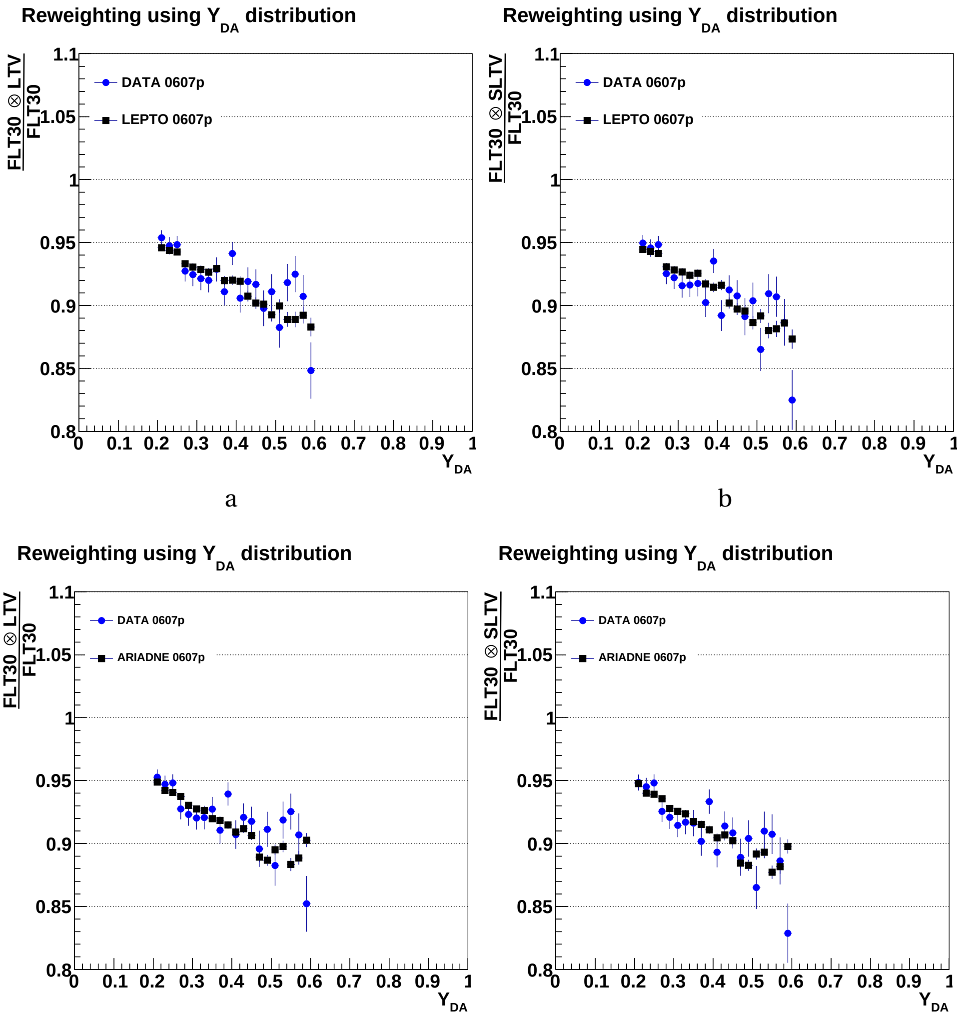
a

b

c

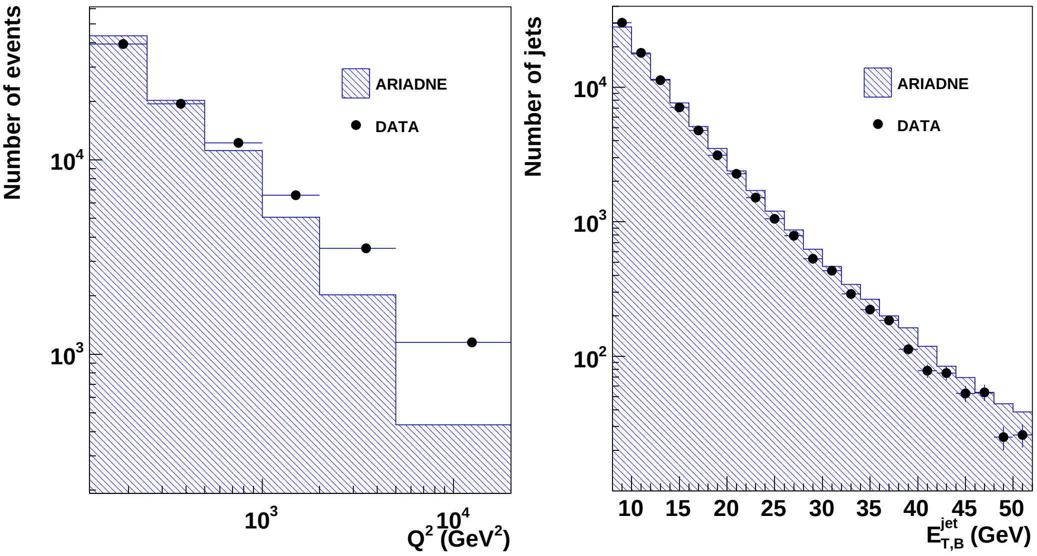
d

Figure 6.4.: Loose track-veto efficiency as a function of y_{DA} in the data and LEPTO MC (a) and ARIADNE MC (b). Distributions of the ratio of the track-veto efficiency in data and LEPTO MC (c), and data and ARIADNE MC (d) and the results of the straight-line fits.



a b c d

Figure 6.5.: Comparison of the loose (a,c) and semi-loose (b,d) track-veto efficiency in data and MC after applying the track-veto correction.



a

b

Figure 6.6.: Control distributions before the reweighting for Q^2 (a) and $E_{T,B}^{jet}$ (b) variables.

$$w\left(Q^2\right) = \begin{cases} a_1 + a_2/\log_{10}\left(Q^2/\mathrm{GeV}^2\right), & \text{(1st iteration)} \quad (6.4.1.1) \\ b_1 + b_2 \cdot \left(Q^2/\mathrm{GeV}^2\right), & \text{(2nd iteration).} \quad (6.4.1.2) \end{cases}$$

In these equations, the Q^2 value determined from the generator-level quantities was used. The product of the weights obtained in each iteration was used for the final reweighting of the LEPTO sample. The effect of each iteration of the Q^2 spectrum reweighting is illustrated in Figure 6.7(a)–(c). The original Q^2 spectrum (Figure 6.7(a)) of the LEPTO generator was characterised by a deficit of events in the lowest and highest Q^2 bins and an excess of events for $250 < Q^2 < 5000\mathrm{GeV}^2$. Due to the irregular shape of the ratio of normalised Q^2 spectra, it was approximated by a smooth function in several iterations. The first iteration (Eq. (6.4.1.1)) improved the Q^2 spectrum in the region $125 < Q^2 < 5000\mathrm{GeV}^2$, while the description of the data in the last Q^2 bin became worse (see Figure 6.7(b)) and had to be further downweighted. Free coefficients $a_{1,2}$ of $w\left(Q^2\right)$ were determined from the fit to the ratio of the data and MC distributions. To minimise the discrepancy between data and simulations in the last Q^2 bin, a second iteration of the reweighting (Eq. (6.4.1.2)) was applied. The coefficients $b_{1,2}$ were also determined from the fit. Overall, after reweighting, a considerable improvement of the description of the Q^2 distribution was observed (see Figure 6.7(c)). The deviation between the data and MC in the last bin of Q^2 after the reweighting was statistically insignificant.

6.4.2. ARIADNE Reweighting

It was found that reweighting as a function of Q^2 only was insufficient to achieve reasonable agreement between jet spectra in the data and the ARIADNE MC; in particular a residual discrepancy in $E_{T,B}^{jet}$ distribution was observed. A direct reweighting as a function of $E_{T,B}^{jet}$ is impossible because, while the jet can be present at the reconstructed level, it can be lost due to, for example, phase-space restrictions at the generator level. Thus, a dedicated simultaneous reweighting as a function of the transverse energy of the hardest jet, $E_{T,B}^{jet1}$, and the process virtuality Q^2 was employed. The reweighting procedure proceeded as follows:

- the data and MC events were classified according to the jet multiplicity into three categories with one, two and three or more jets, respectively;
- for each category the two dimensional Q^2 vs $E_{T,B}^{jet1}$ distribution in the data and MC was measured; for comparison of the data distributions to the generator level MC to be valid, the data were corrected to the hadron level using the acceptance correction factors as described in Section 6.4.1.
- the bin content in the MC was multiplied by

  $$w\left(Q^2, E_{T,B}^{jet1}/\right) = a_1\, e^{\left(-\alpha E_{T,B}^{jet1}/\mathrm{GeV}\right)}\left(1 - e^{\left(-\beta Q^2/\mathrm{GeV}^2\right)} + a_2 E_{T,B}^{jet1}/\mathrm{GeV}\right), \tag{6.4.2.1}$$

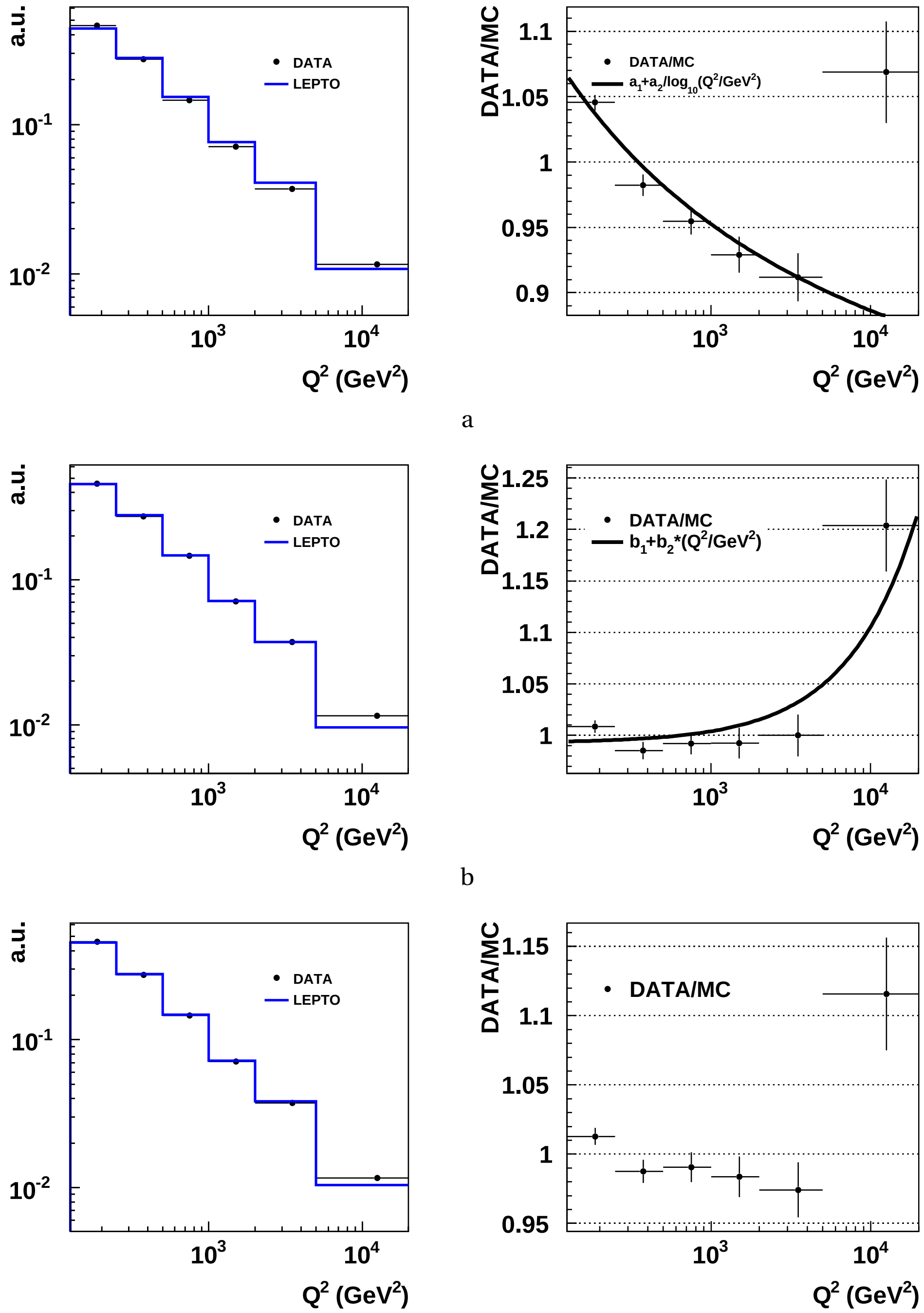
  where a_i, α, β are free parameters and $E_{T,B}^{jet1}$, Q^2 correspond to the generator-level quantities. A 2d-likelihood fit of the shape of the data distribution was performed to determine the free parameters.

The described procedure was iterated until reasonable agreement between data and MC was achieved. Figures 6.8 and 6.9 show the distribution of the exchanged boson virtuality and jet transverse momentum in single-jet events after applying the described reweighting[4]. As can be seen, after the reweighting the Q^2 and $E_{T,B}^{jet}$ spectra are very well described by the simulations. The quality of the fits assessed by the value χ^2/N_{df} is summarised in the Table 6.1.

In addition to the variables used in the reweighting procedure, the description of other jet quantities was verified, e.g. the comparison of the η_B^{jet} and jet multiplicity distributions before and after reweighting is demonstrated in

---

[4] The figures for events with higher jet multiplicity can be found in the Appendix.



a

b

c

Figure 6.7.: Result of two iterations of the Q^2 reweighting in the LEPTO MC sample. Comparison of original Q^2 spectrum in data and MC and the ratio together with the fit (a). Data and MC distributions and their ratio after the first iteration of Q^2 reweighting. The reweighting function determined from the fit is also shown (b). The resulting data and MC distribution and their ratio (c).

jet multiplicity	χ^2/N_{df}
1 jet	0.88
2 jet	0.88
≥ 3 jets	0.86

Table 6.1.: χ^2/N_{df} for likelihood fits for simultaneous Q^2 and $E_{T,B}^{jet}$ reweighting.

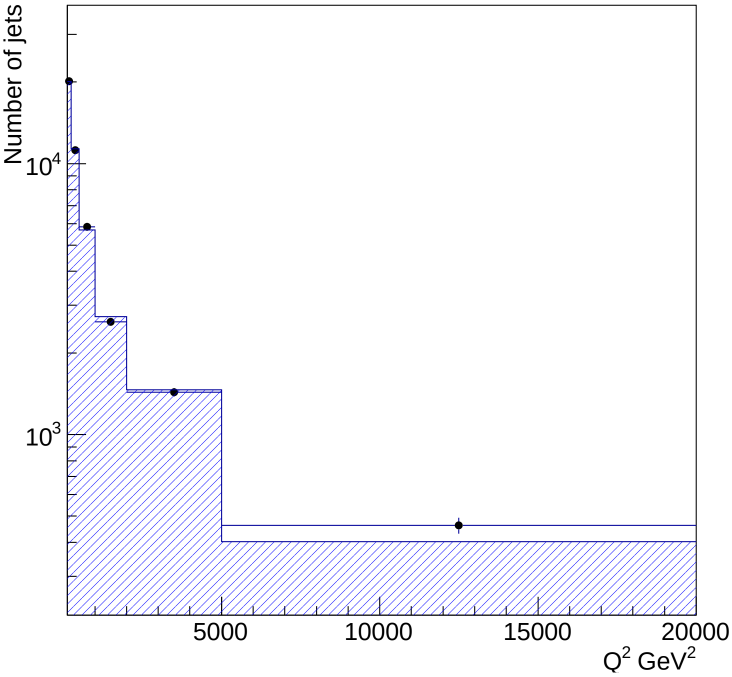


Figure 6.8.: Comparison of Q^2 spectra in data and ARIADNE after MC reweighting for events containing exactly one jet.

Figures 6.10(a)–(d). Significant improvement in the description of the jet multiplicity and η_B^{jet} was achieved. Quantitatively it can be explained as follows. Before the reweighting, an excess of high-energy jets in the simulations resulted in a larger fraction of events with > 1 jet in the event (see Figure 6.10(a)). The reduction of the fraction of high-energy jets after the reweighting led to an increase of the number of events with a single jet and a decrease of those with two. An improvement of the η_B^{jet} distribution was attributed to the changes in the Q^2 spectrum, which affected the jet kinematics in the Breit frame.

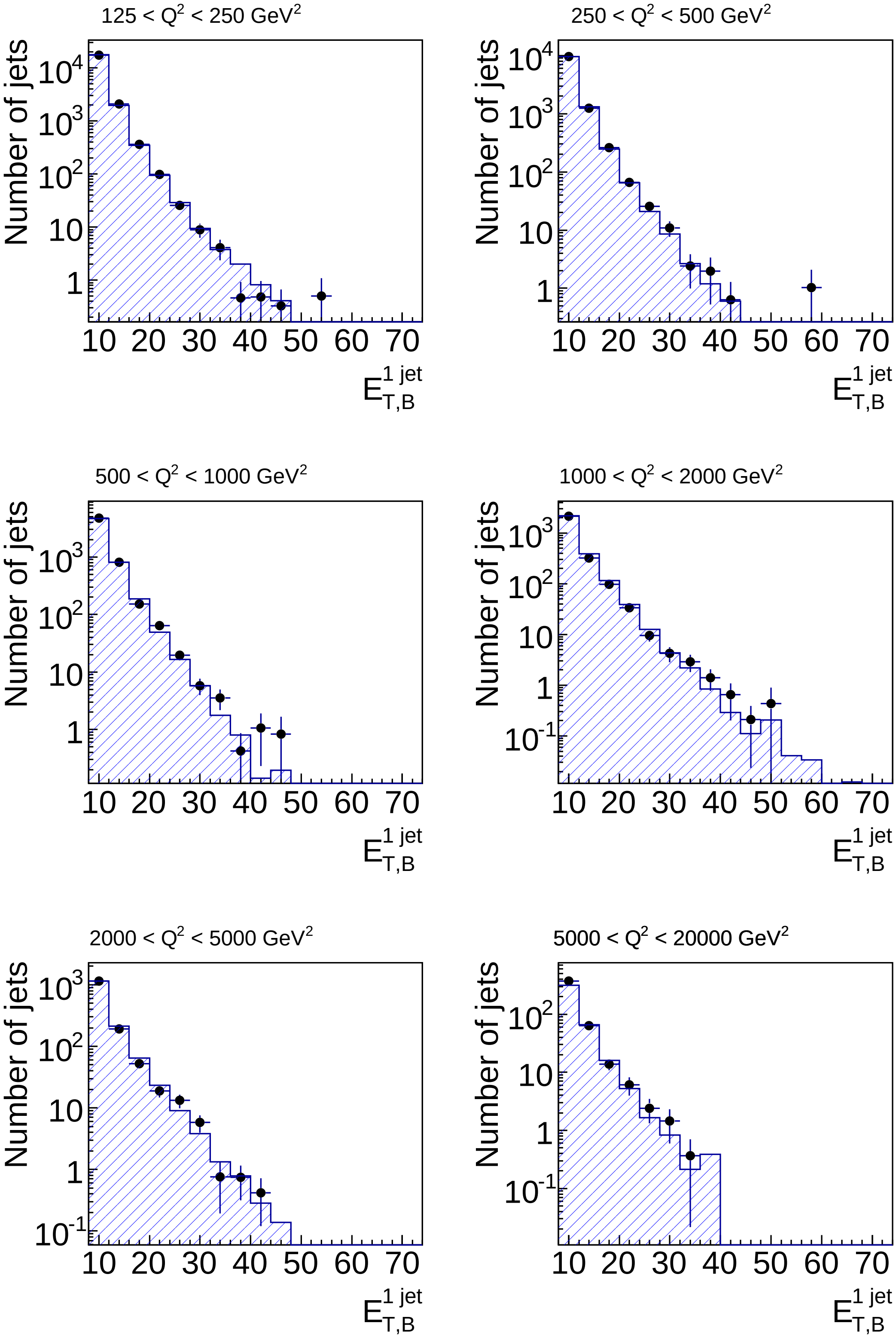
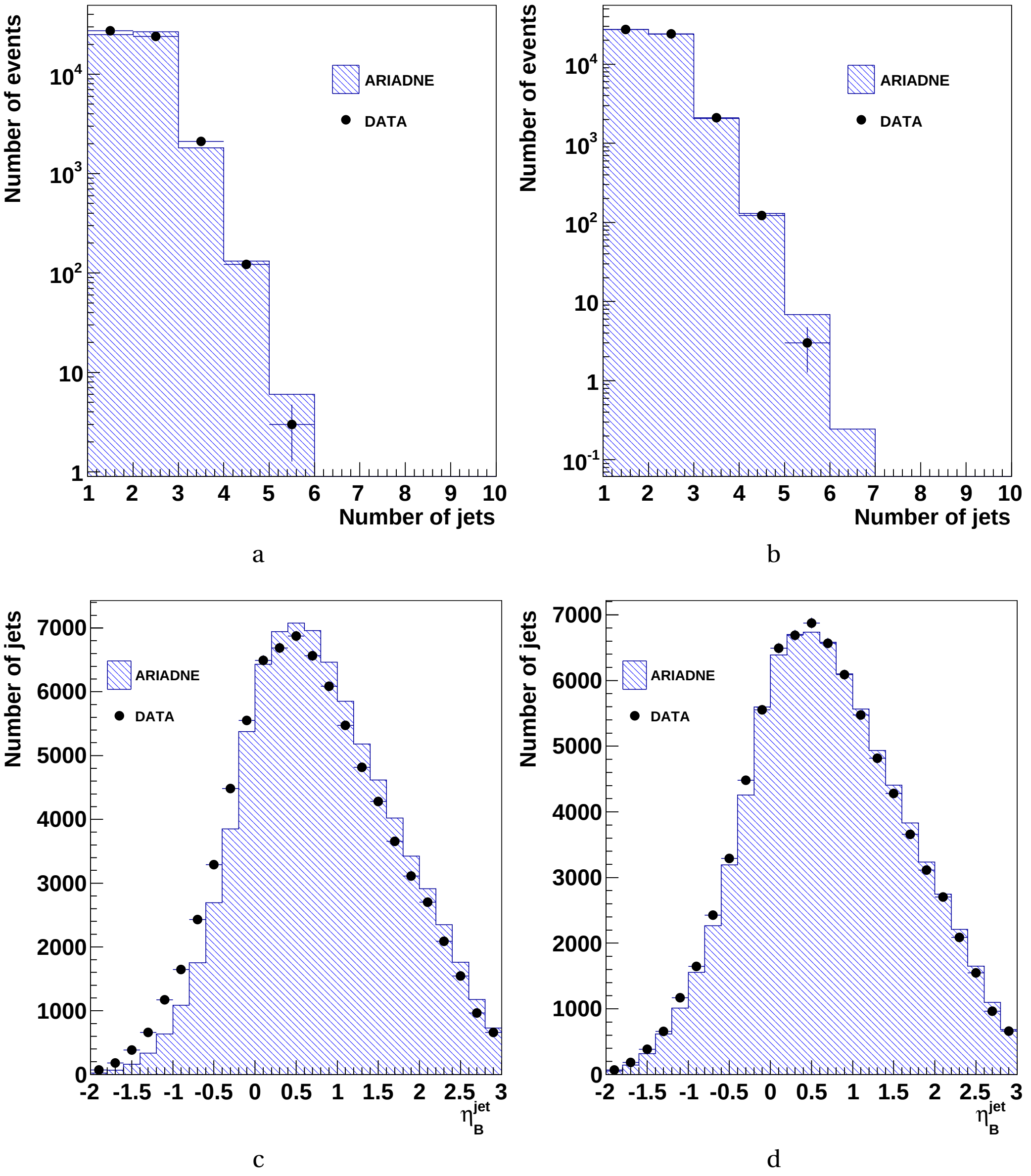


Figure 6.9.: Comparison of transverse energy of the leading-jet spectra in data and ARIADNE in different Q^2 bins after MC reweighting for events containing exactly one jet.



a

b

c

d

Figure 6.10.: Result of two iterations of the Q^2 reweighting in the ARIADNE MC sample. Jet multiplicity and η^{jet} distributions in data and MC before (a,c) and after the reweighting (b,d)

6.5. Electromagnetic Energy Scale

The pre-processing of the data with ORANGE/PHANTOM libraries includes dead-material and non-uniformity corrections in the electron identification algorithms. Nevertheless a residual discrepancy between data and MC simulations was observed. In order to study this discrepancy in detail the data and MC events satisfying the requirements described in Chapter 5 were used.

The resolution and electromagnetic energy scale were investigated in data and MC by taking the ratio $E_{\mathrm{SI}}/E_{\mathrm{DA}}$, where E_{SI} is the measured electron energy including all corrections and E_{DA} the energy measured by the double-angle method (see Section 4.4.3). As was explained, the electron energy determined using the double-angle method is approximately independent of the absolute energy scale of the calorimeter and therefore was used as a reference scale for the comparison. As shown in Figure 6.11 (a) the distribution has a Gaussian-like shape with the full width at half maximum of the data and LEPTO distributions of about 10%. In general, the simulations adequately describe the shape of the data distribution, however a systematic shift of the mean value was observed.

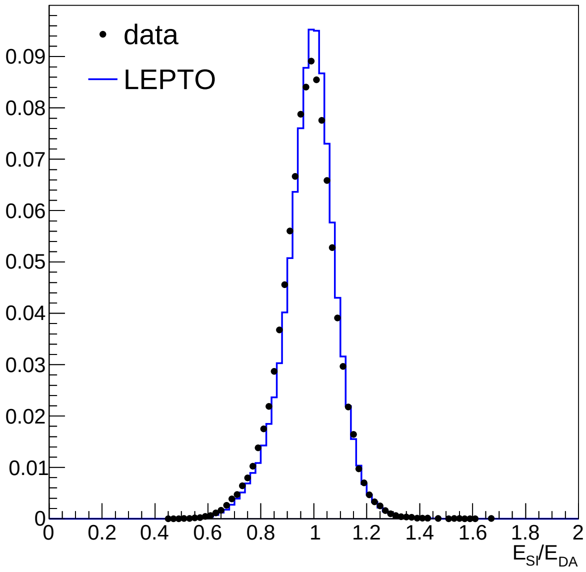
To investigate this shift in more details in each bin of E_{DA}, a Gaussian fit to the $E_{\mathrm{SI}}/E_{\mathrm{DA}}$ distribution was performed. The mean value extracted from the fit was plotted as a function of E_{DA} (see Figure 6.11(b)). It was found that the double ratio between the absolute E_{SI} energy scales in data and Monte Carlo simulations deviated from unity by less than 2%. This discrepancy was taken into account as a systematic uncertainty as described in Section 8.3.

6.6. Jet Corrections

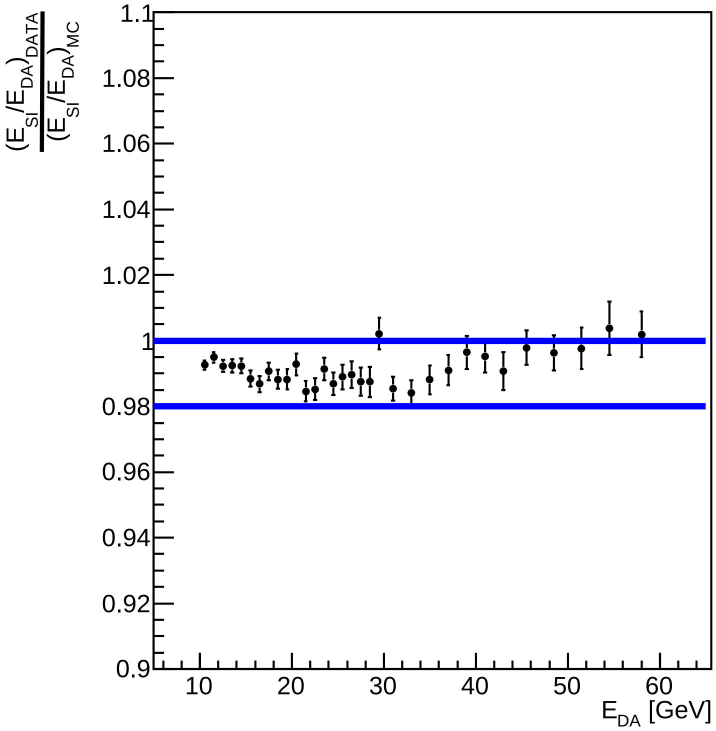
6.6.1. Jet Energy-Scale Calibration

In most jet analyses, the jet energy-scale uncertainty is usually the dominant source of systematic error. Inclusive-jet cross sections are steeply falling functions of jet transverse energy. An uncertainty on the hadronic energy scale affects strongly the precision with which the jet cross sections can be measured. In recent ZEUS publications [48, 159] the jet energy-scale uncertainty was determined to be ±1% for jets with transverse energy $E_T^{jet} > 10$ GeV and ±3% for jets with $3 \text{ GeV} < E_T^{jet} < 10$ GeV, resulting in a systematic uncertainty on the measured jet cross section of about 5 – 10%, depending on the region of phase space. In this section, the study of the hadronic energy scale performed in this analysis is described in detail.

The response of the calorimeter to jets was investigated by comparing the measured jet transverse energy to the transverse energy of the scattered electron. The transverse energy of the jet must be equal to that of the final-state



a



b

Figure 6.11.: Comparison of $E_{\mathrm{SI}}/E_{\mathrm{DA}}$ in data and Monte Carlo (a). Double ratio between the electromagnetic energy scale in data and Monte Carlo simulations as a function of electron energy (b).

electron according to momentum conservation[5] and the following relation must be satisfied

$$r = \frac{E_T^{jet}}{E_T^{DA}}. \tag{6.6.1.1}$$

The deviation of a double ratio calculated in data and Monte Carlo events

$$C_{\text{scale}} = \frac{r^{DATA}}{r^{MC}} \tag{6.6.1.2}$$

indicates a difference in the hadronic energy scale of the calorimeter in data and simulations. This factor can be used to correct the relative difference between data and MC assuming that C_{scale} is independent of jet energy.

The procedure for the extraction of the relative correction factors C_{scale} is based on the single-jet event sample. Therefore the selection requirements described in Section 5 were modified to conform to the assumptions of the method. The required modifications are outlined below:

- in order to avoid problems with imprecisely reconstructed Lorentz boosts to the Breit frame, the jet search was performed in the laboratory frame; moreover, the calorimeter energy scale is related to the distribution of the material within the detector volume, thus the laboratory frame is more natural for this study;
- a single jet with $E_{T,lab}^{jet} > 10$ GeV and no other jets with $E_{T,lab}^{jet} > 5$ GeV were required in order to suppress further hadronic activity not related to the hard scattering;
- the requirement on the pseudorapidity and the transverse energy in the Breit frame was omitted (see Eq. (5.6.0.2));
- to increase the number of events with hadronic activity in the forward direction, the inelasticity cut on y was removed (see Eq. (5.6.0.1)).

Figure 6.12 demonstrates the description of the quantity r by the Monte Carlo simulations in different intervals of η_{lab}^{jet}. In general, MC describes the shape of the data well, however a discrepancy between the mean values of the order of 3% was observed in some bins of η_{lab}^{jet}. The ratio of the mean values in data and MC is illustrated in Figure 6.13(a). These values were used to correct the discrepancy in the hadronic energy scale between data and MC. The transverse energy of the jets in the MC was multiplied such that:

$$E_{T,B}^{jet} \mapsto E_{T,B}^{jet'} = E_{T,B}^{jet} \cdot C_{\text{scale}}. \tag{6.6.1.3}$$

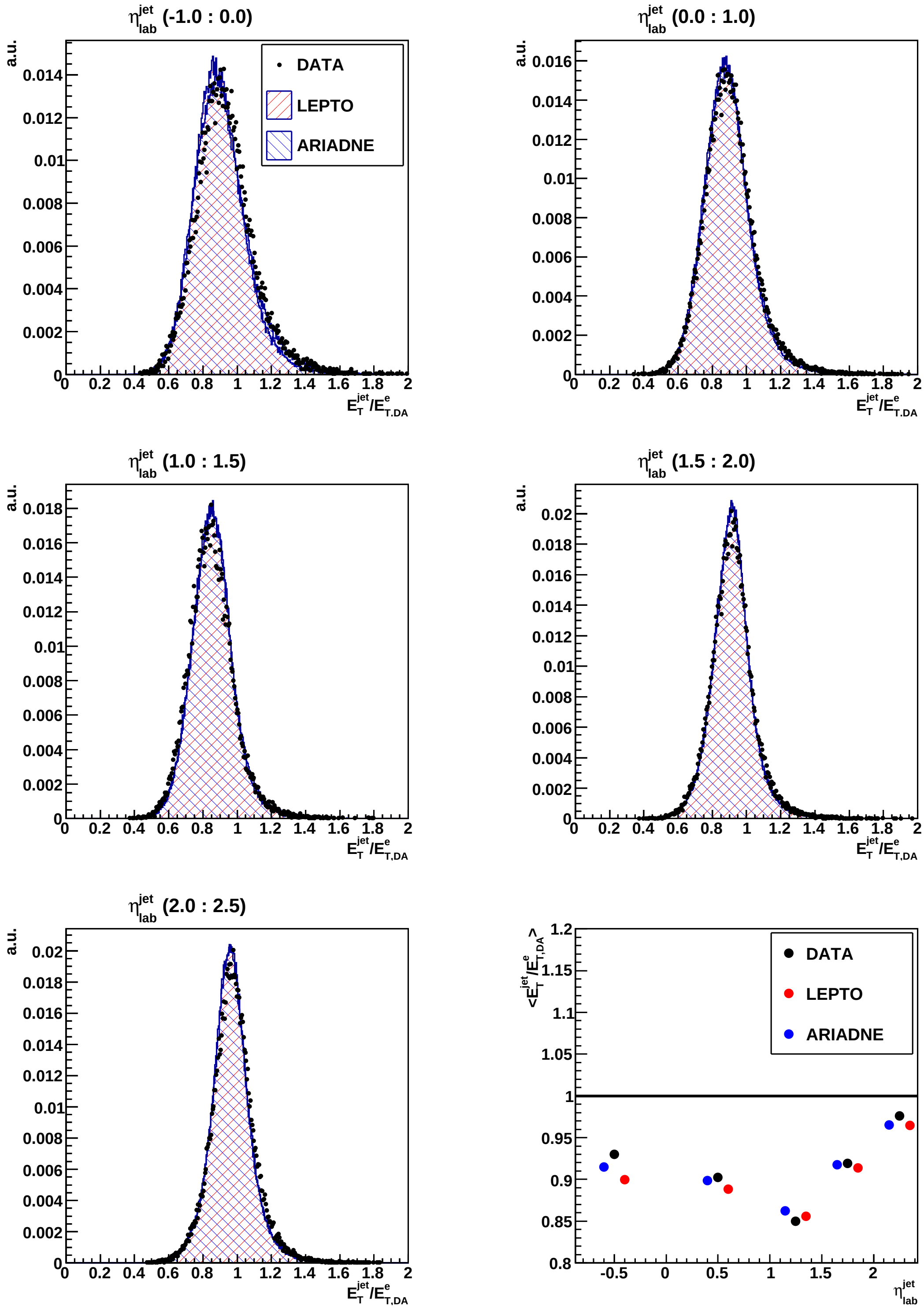
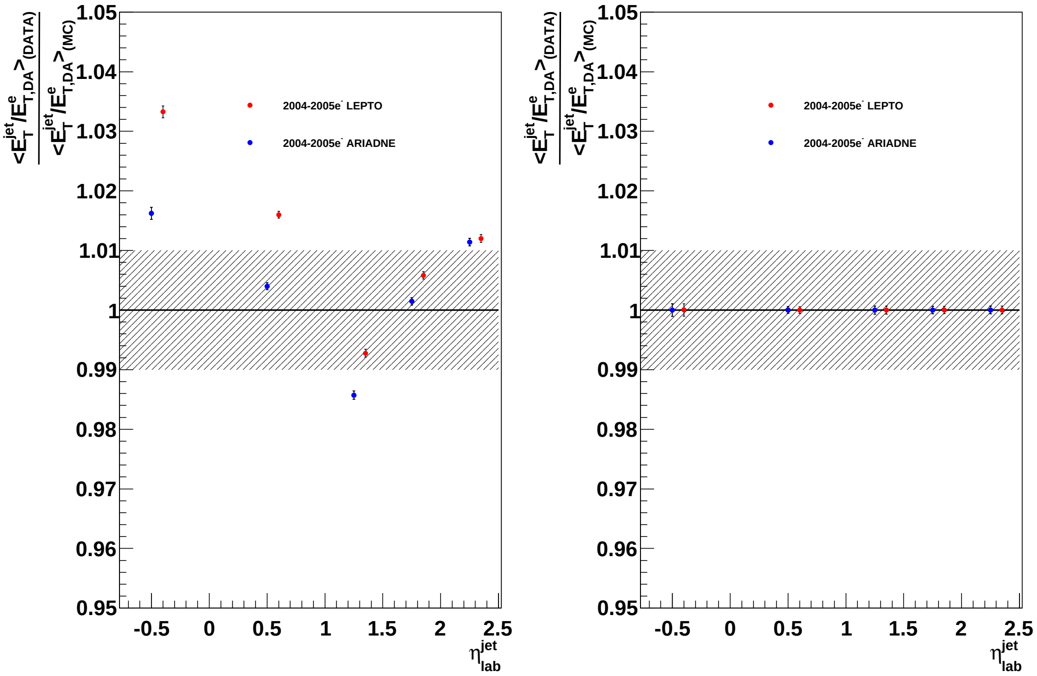


Figure 6.12.: The ratio $r = \frac{E_T^{jet}}{E^e_{T,DA}}$ of the transverse energies of the jet and the electron measured with the double-angle method in different regions of jet pseudorapidity η^{jet}_{lab} and the mean value $\langle r \rangle$ as a function of η^{jet}_{lab} (bottom right panel) in data and MC for the 2004–2005 e^- data-taking period.

The result of such a relative shift was verified and is demonstrated in Figure 6.13(b).



a

b

Figure 6.13.: Double ratio $\langle r^{DATA} \rangle / \langle r^{MC} \rangle$ of the transverse energies of the jet and the electron measured with the double-angle method as a function of the jet pseudorapidity for the 2004–2005 e^- data-taking period before (a) and after (b) jet energy-scale correction.

6.6.2. Jet Energy-Scale Uncertainty

As mentioned previously, the precise determination of the jet energy-scale uncertainty is an important ingredient in many jet analyses, so the accuracy of the jet energy-scale corrections, described above, was investigated in a dedicated study. Assuming a valid description of the fraction of charged and neutral particles constituting a jet in MC simulations, the difference in hadronic energy scale in the data and MC was examined independently using the tracking information, after the jet energy-scale correction, described above, was implemented.

---

[5] It is assumed that the transverse momentum of the beam remnants is negligibly small and the hadronic final state is attributed to a single jet.

For jets with transverse energy in the laboratory frame $E_{T,lab}^{jet} > 10\,\text{GeV}$, the tracks[6] within the tracking-system acceptance attributed to jets were identified according to their proximity to the jet axis

$$R^2 = \left(\eta_{\text{track}} - \eta_{\text{jet}}\right)^2 + \left(\phi_{\text{track}} - \phi_{\text{jet}}\right)^2 < 1. \tag{6.6.2.1}$$

The ratio of transverse energy of the jet, $E_{T,lab}^{jet}$, and the sum of traverse momenta, $p_{T,i}$, of matched tracks

$$r_{\text{tracks}} = \frac{E_{T,lab}^{jet}}{\sum\limits_{\text{tracks}} p_{T,i}}, \tag{6.6.2.2}$$

was compared in the data and MC in different regions of η_{lab}^{jet} and for different running periods (see Figures 6.14). Overall, the MC simulations provide a good description of the data in shape, however the discrepancy between the mean values of the data and MC distributions remains. The double ratio $\langle r_{\text{tracks}}^{\text{DATA}}\rangle / \langle r_{\text{tracks}}^{\text{MC}}\rangle$ measured separately for different data-taking periods is illustrated in Figure 6.15. The relative difference between the hadronic energy scale in data and MC does not exceed 1%. This discrepancy was therefore taken into account in the systematic uncertainty. Exploiting the transverse momentum conservation in ep collisions at HERA (see Section 5.1), it was demonstrated in [149, 154] that jets with pseudorapidity outside the tracking acceptance region also contribute ±1%. The energy-scale uncertainty for jets with transverse energy $3 < E_{T,lab}^{jet} < 10\,\text{GeV}$ was found [149] to be ±3%.

6.6.3. Jet Energy Correction

The energy of hadronic jets reconstructed from energy deposits in the calorimeter is influenced by various effects e.g. particle absorption in uninstrumented material between the production point and the calorimeter, inhomogeneities of the detector, noise etc. Hadron jets in the ZEUS detector typically loose 5-15% of their energy in inactive media (superconducting solenoid, support structures etc.) in front of the calorimeter. This effect may lead to systematic migrations of jets to cross-section bins with lower energy. In principle, such effects must be taken into account in the unfolding procedure (see Chapter 7). Nevertheless, in order to minimise migrations and to avoid a possible bias from the energy loss in inactive detector media, a dedicated jet-energy correction was employed in this analysis.

Two approaches for correcting the jet energy loss exist:

---

[6] A track was required to pass through at least three CTD superlayers and to be fitted to the primary vertex; transverse momentum of the track has to be $p_T > 300\text{MeV}$.

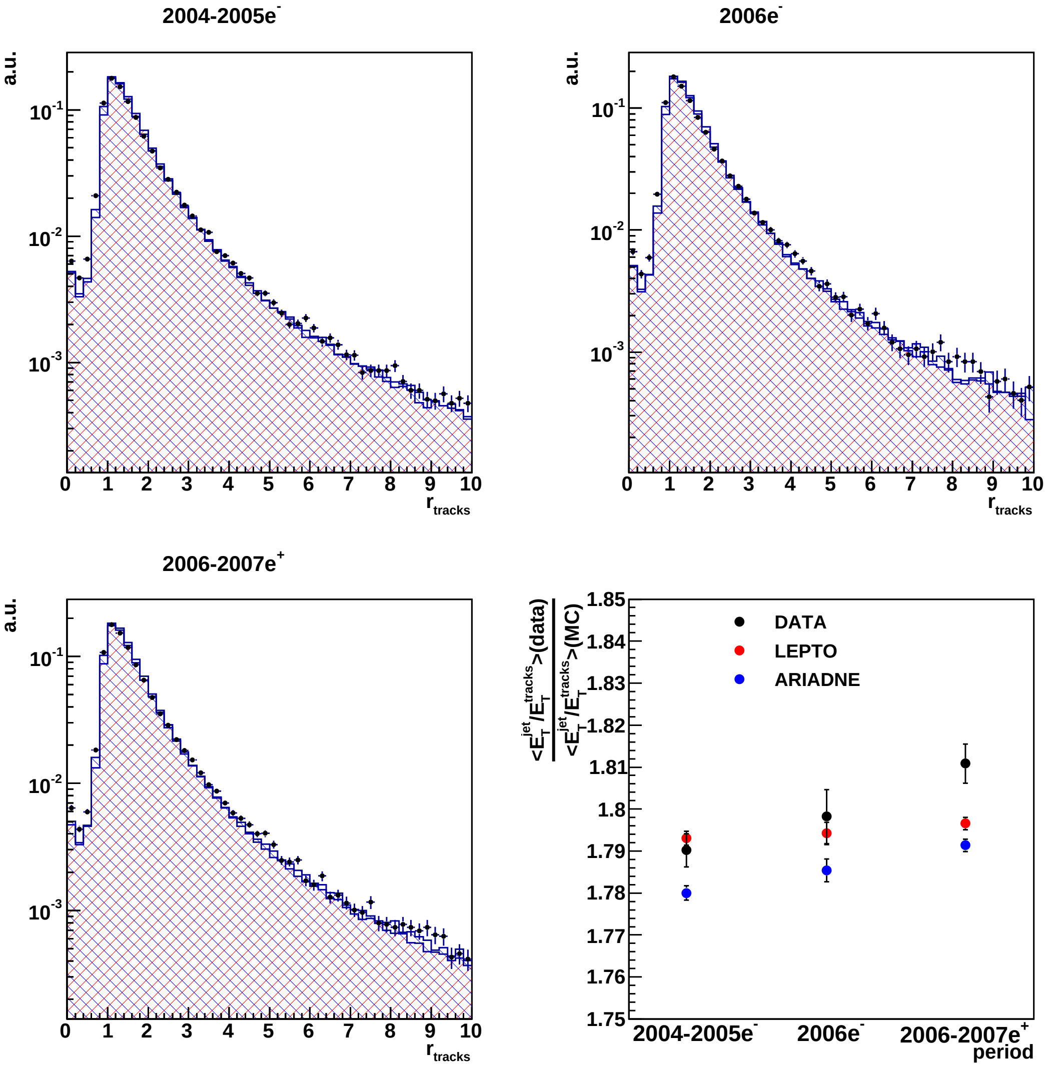


Figure 6.14.: The distribution of the ratio $r_{\text{tracks}} = \frac{E^{jet}_{T,lab}}{\sum\limits_{\text{tracks}} p_{T,i}}$ and of the mean values $\langle r_{\text{tracks}} \rangle$ in data and MC for different data-taking periods.

- the *bottom-up* approach consists of correcting the energy of the input objects (i.e. calorimeter cells in this analysis) to compensate for the energy loss and then using the corrected objects as an input to the jet algorithm;
- in the *top-down* approach the energy of identified jets is corrected directly.

In principle, with the former approach more precise correction can be achieved, because individual jet details can be take into account, while in the later, only

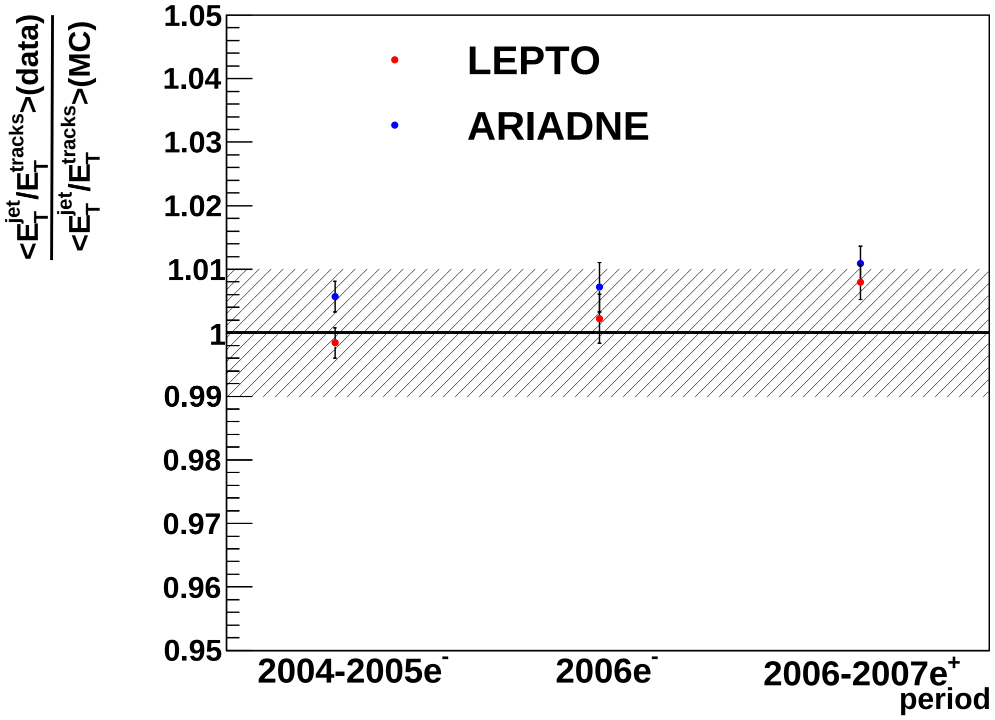


Figure 6.15.: The double ratio $\langle r^{DATA}_{\text{tracks}} \rangle / \langle r^{MC}_{\text{tracks}} \rangle$ for different data-taking periods. The hatched band indicates the attributed jet energy-scale uncertainty.

an average correction is achieved. Nevertheless, the *top-down* approach is much more simple and therefore was used in this thesis.

Monte Carlo simulations were used to estimate the energy loss because they provide the detailed information about hadron propagation in the detector volume. The measured jet energy, $E_T^{jet,det}$, depends approximately linearly on the 'true', E_T^{jet}, value, but the size of the energy loss depends on the thickness of the traversed material and therefore on the jet pseudorapidity in the laboratory frame. For this reason, the complete measurement region $-1 < \eta^{jet}_{lab} < 2.5$ was divided into 14 equal size regions and the correction was determined separately in each η^{jet}_{lab} bin.

The procedure was as follows:

- in the MC events accepted at the generated and reconstructed levels, a pair of jets was identified according to the distance between jets in the $\eta - \phi$ plane. A hadron-level jet was matched to the reconstructed jet if the distance
$$r = \sqrt{\left[\left(\eta_{had} - \eta_{det}\right)^2 + \left(\phi_{had} - \phi_{det}\right)^2\right]} \qquad (6.6.3.1)$$
between the two was small, $r < 0.7$, and no further jets were recon-

structed within the cone. In order to avoid any bias on the correction procedure introduced by the boundaries of the jet phase space, the cuts on the jet transverse energy were relaxed to $E_{T,B}^{jet,had} > 6$ GeV and $E_{T,B}^{jet,det} > 3$ GeV at the hadron and detector levels, respectively;

- in each bin i of η_{lab}^{jet}, a linear fit $\left\langle E_{T,B}^{jet,det} \right\rangle = a_i + b_i \cdot \left\langle E_{T,B}^{jet,had} \right\rangle$ was performed, where $\left\langle E_{T,B}^{jet,det} \right\rangle$ and $\left\langle E_{T,B}^{jet,had} \right\rangle$ were determined using the set of matched jets; the fits for the 2004–2005 e^- running period are illustrated in Figure 6.16;
- the corrections were determined using the ARIADNE and LEPTO sample for each data-taking period separately.

Given the extracted fit parameters, the components of the jet four-momentum were scaled such that the following relation was obtained:

$$E_{T,B}^{jet,det} \mapsto E_{T,B}^{jet,corr} = \frac{E_{T,B}^{jet,det} - a_i}{b_i}. \tag{6.6.3.2}$$

As can be seen in Figure 6.16, the fractional energy loss, represented by the slope of the fitted straight line, varies as a function of jet pseudorapidity. The size of the correction decreases towards the forward region of the detector. Such a behaviour was attributed to the variation of the amount of material in front of the calorimeter, in particular, the presence of superconducting solenoid surrounding the tracking system.

Since the analysis was performed in the Breit frame, the jet pseudorapidity in the laboratory frame was recalculated and the corresponding correction factors were applied. Assuming a valid description of the detector effects in the simulations, the correction was applied to both the data and MC jets. In the simulations the correction was applied on top of that introduced in Section 6.6.1.

6.6.4. Conclusion

Obtaining an accurate description of jet observables is a necessary prerequisite for the measurement of the jet cross sections. In order to improve the description of the jet distributions in MC simulations, three types of corrections were implemented. The described methods include the reweighting of the jet spectrum (Section 6.4), correction of the hadronic energy-scale difference in data and MC (Section 6.6.1) and correction of the jet transverse energy for the looses in inactive detector material (Section 6.6.3). As a result, the description of the

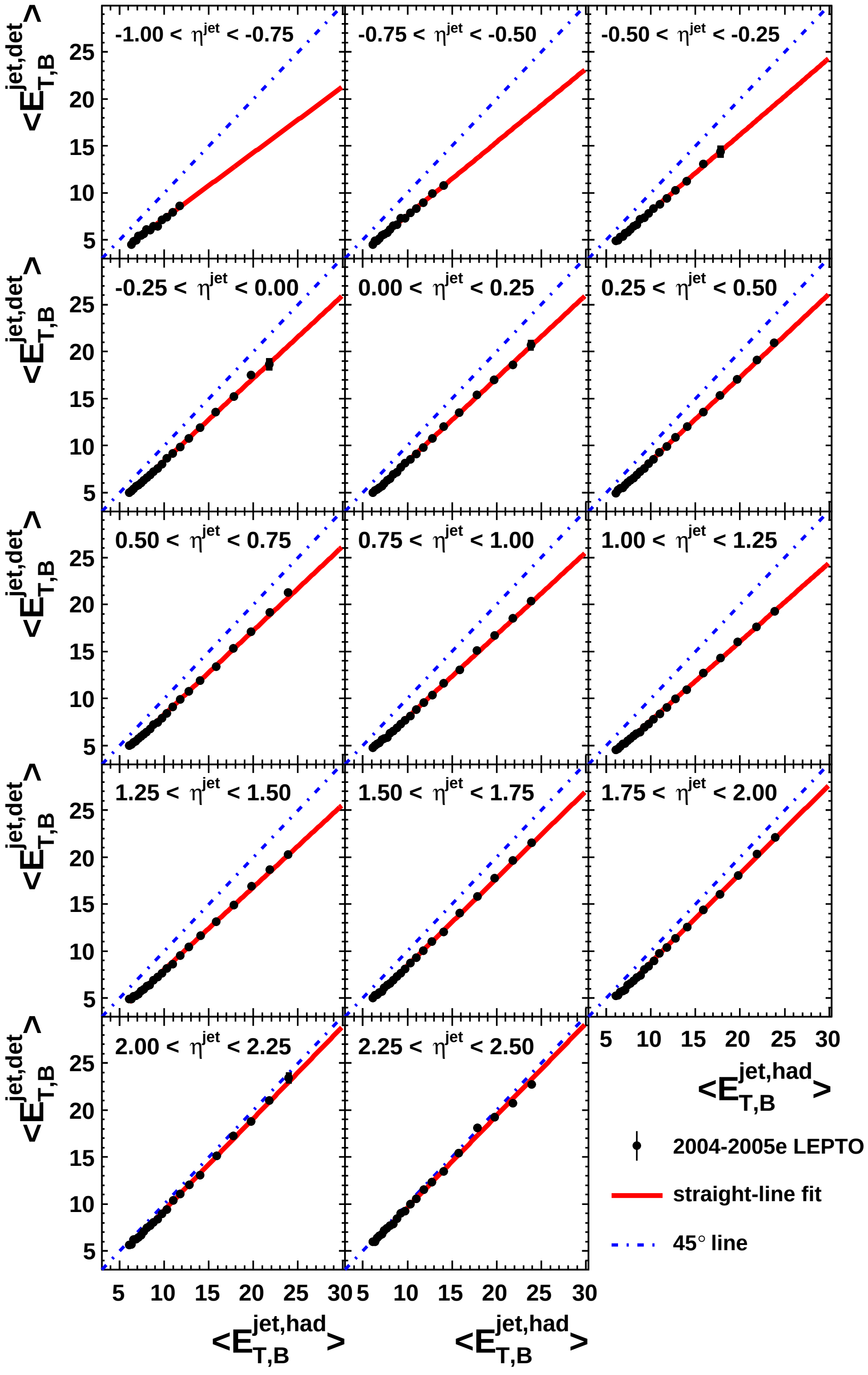
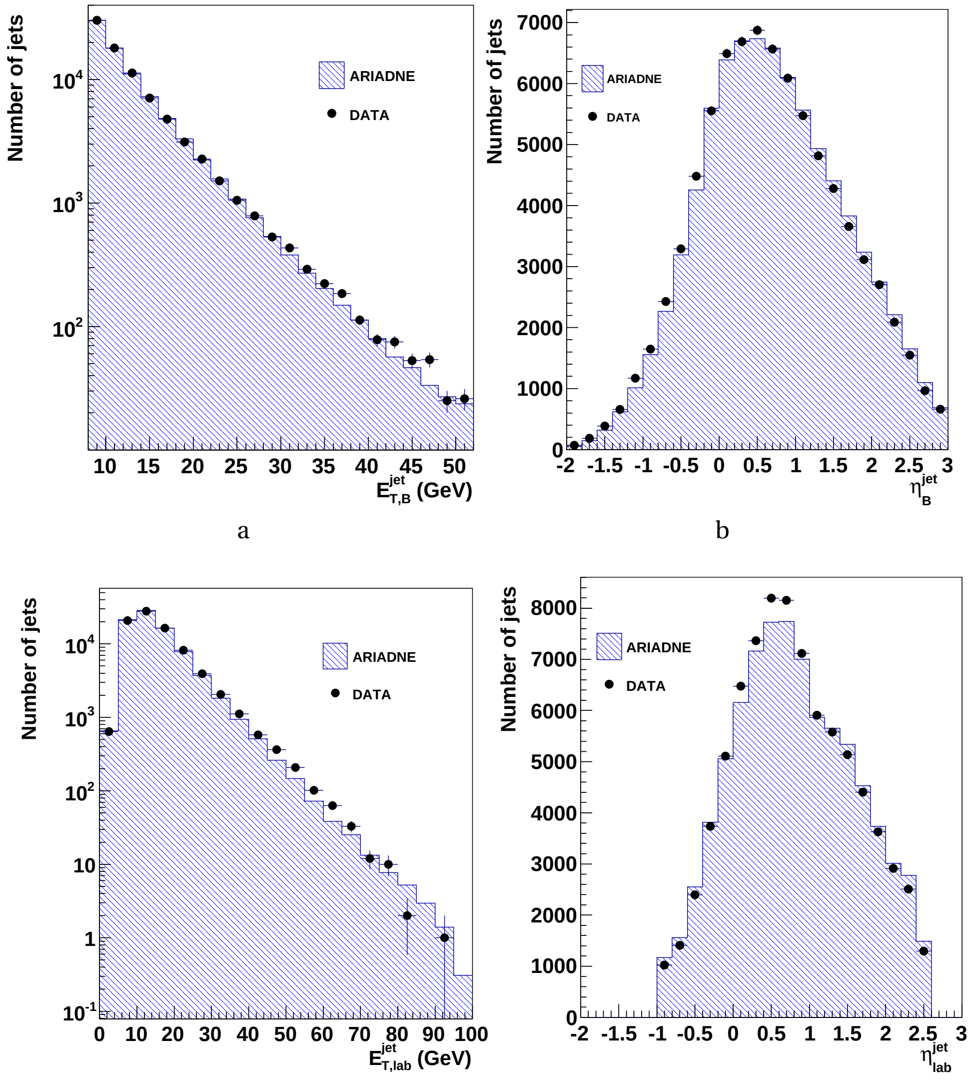


Figure 6.16.: The average measured detector-level jet transverse energy $\left\langle E_{T,B}^{jet,det} \right\rangle$ as a function of $\left\langle E_{T,B}^{jet,had} \right\rangle$ and the corresponding straight-line fits in different regions of η_{lab}^{jet} for the data-taking period 2004–2005 e^- in the LEPTO MC sample.

jet quantities was significantly improved when compared to original distributions. The comparison of resulting distributions in data and MC are show in Figures 6.17, 6.18. In addition, after implementing all jet corrections, the correlation between generated and reconstructed jet quantities was checked. As shown in Figure 6.19, the reconstructed value of $E_{T,B}^{jet}$ provides an unbiased estimator of the truth-level quantity in the complete range of jet transverse energies. The same conclusion can be made about jet angular variables (see Figures 6.20, 6.21), although a slight bias for η_B^{jet} was observed in the region $\eta_B^{jet} < 0.5$. This was attributed to the fact that at the generator level the origin of the ZEUS coordinate system was used instead of the position of the primary vertex for the determination of η_B^{jet}.

After the correcting the difference between the hadronic energy scale in data and simulations, the energy-scale uncertainty for the calibrated jet sample is equal to ±1% for jets with $E_{T,B}^{jet} > 10$ GeV and ±3% for jets with $3 < E_{T,B}^{jet} < 10$ GeV. The effect of this uncertainty on the measured jet cross sections will be discussed in Section 8.3.

Given that an accurate description of the measured distributions in MC has been demonstrated in this chapter, a reliable estimation of detector effects such as trigger and/or reconstruction inefficiencies, migrations etc. can be performed and will be described in Chapters 7 and 8.



a

b

c

d

Figure 6.17.: Comparison of corrected MC (ARIADNE) and data distributions for jet variables. (a) jet transverse energy in the Breit frame; (b) jet transverse energy in the laboratory frame; (c) jet pseudorapidity in the Breit frame; (d) jet pseudorapidity in the laboratory frame η_{lab}^{jet}.

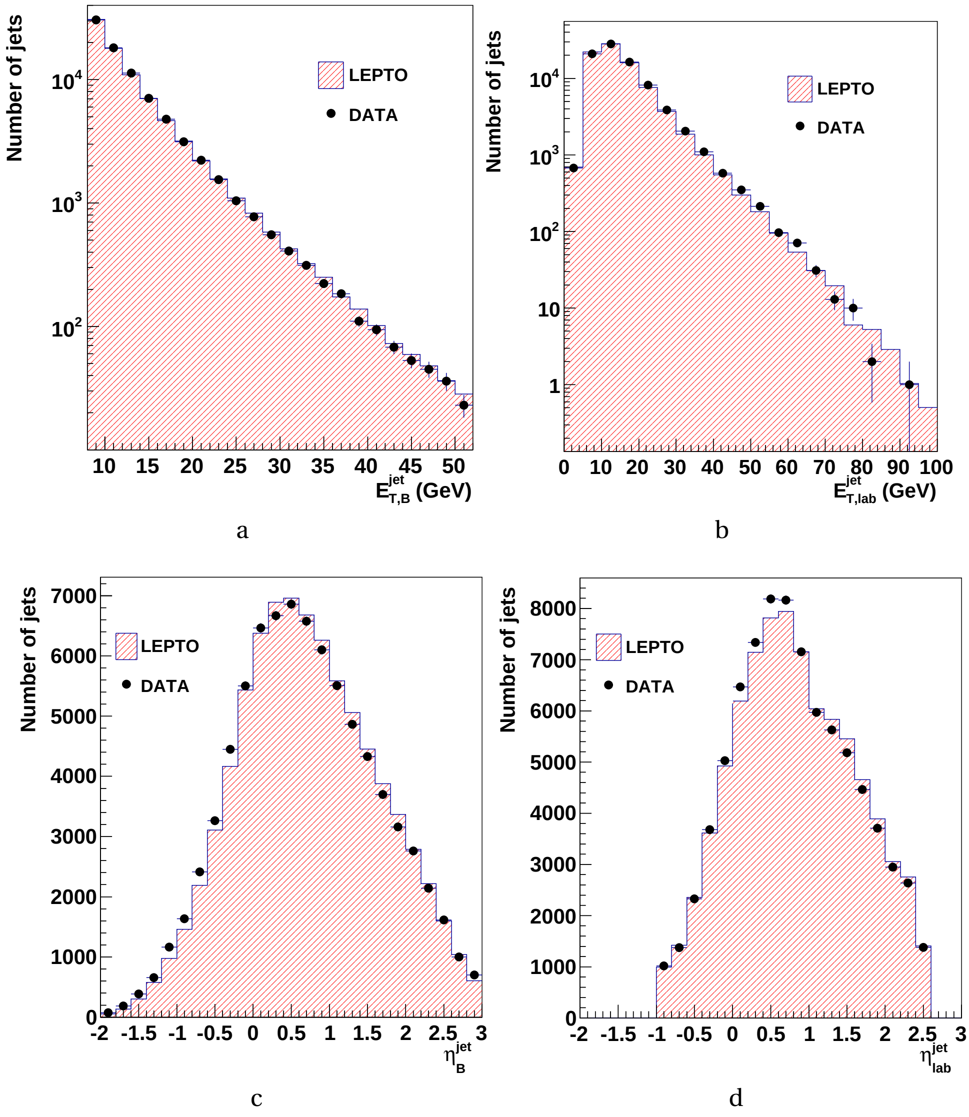


Figure 6.18.: Comparison of corrected MC (LEPTO) and data distributions for jet variables. (a) jet transverse energy in the Breit frame; (b) jet transverse energy in the laboratory frame; (c) jet pseudorapidity in the Breit frame; (d) jet pseudorapidity in the laboratory frame η^{jet}_{lab}.

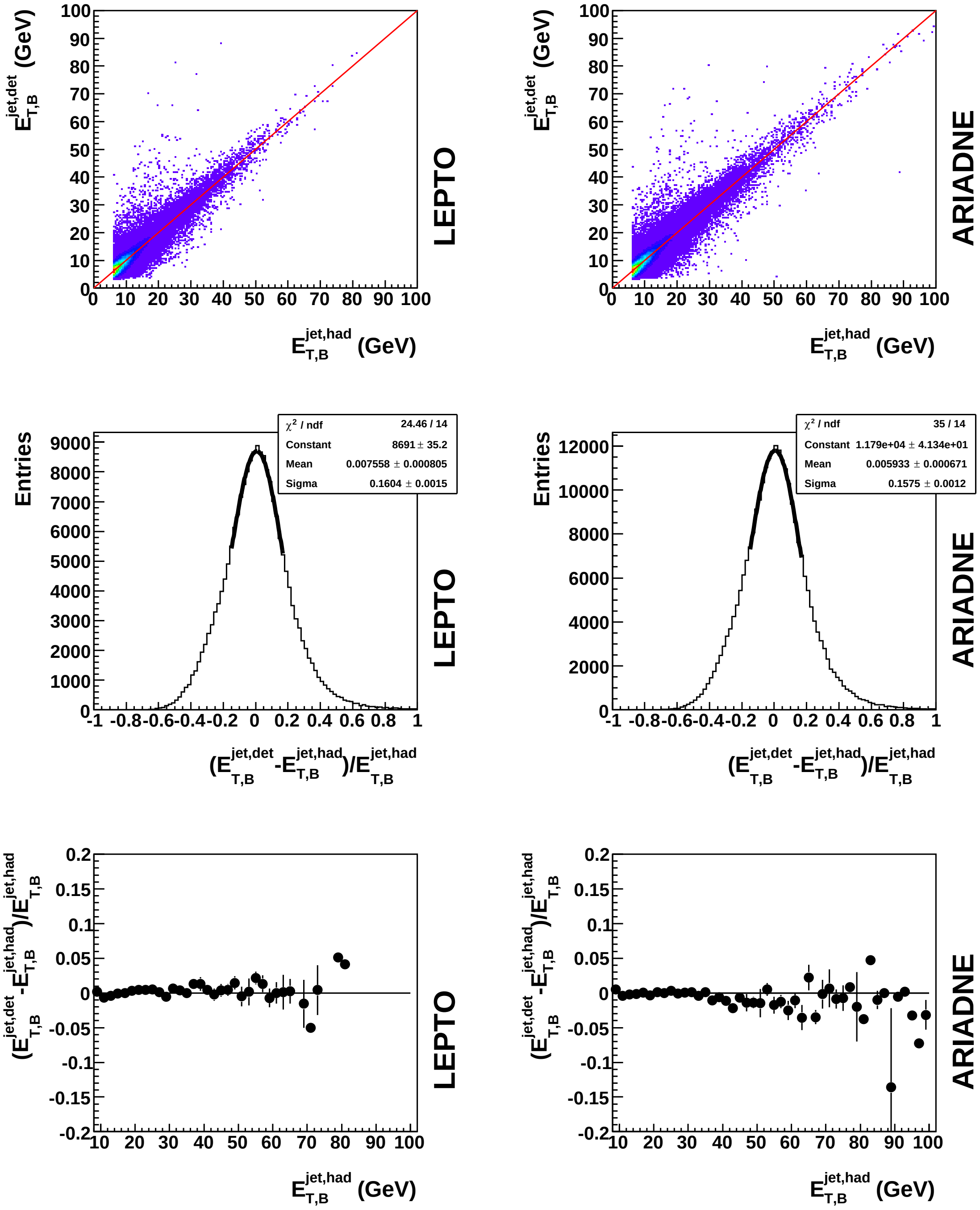


Figure 6.19.: Correlation between generated and reconstructed jet transverse energy $E_{T,B}^{\text{jet,det}}$ vs $E_{T,B}^{\text{jet,had}}$ (top row). Relative difference distribution $\left(E_{T,B}^{\text{jet,det}} - E_{T,B}^{\text{jet,had}}\right)/E_{T,B}^{\text{jet,had}}$ (middle row). Variation of the relative difference with respect to $E_{T,B}^{\text{jet,had}}$ (bottom row).

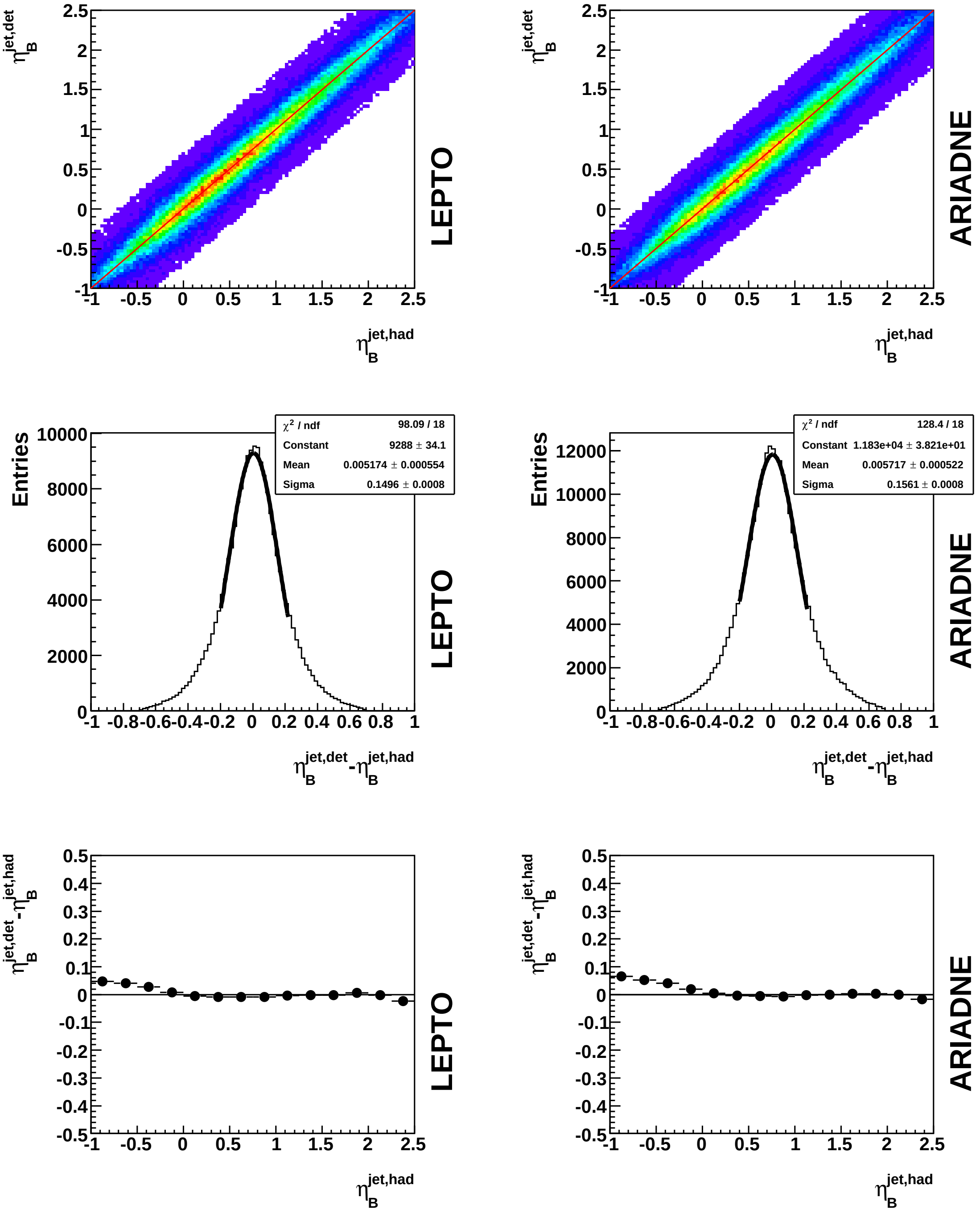


Figure 6.20.: Correlation between generated and reconstructed jet pseudorapidity $\eta_B^{\mathrm{jet,det}}$ vs $\eta_B^{\mathrm{jet,had}}$ (top row). Relative difference distribution $\left(\eta_{T,B}^{\mathrm{jet,det}} - \eta_{T,B}^{\mathrm{jet,had}}\right)/\eta_{T,B}^{\mathrm{jet,had}}$ (middle row). Variation of the relative difference with respect to $\eta_{T,B}^{\mathrm{jet,had}}$ (bottom row).

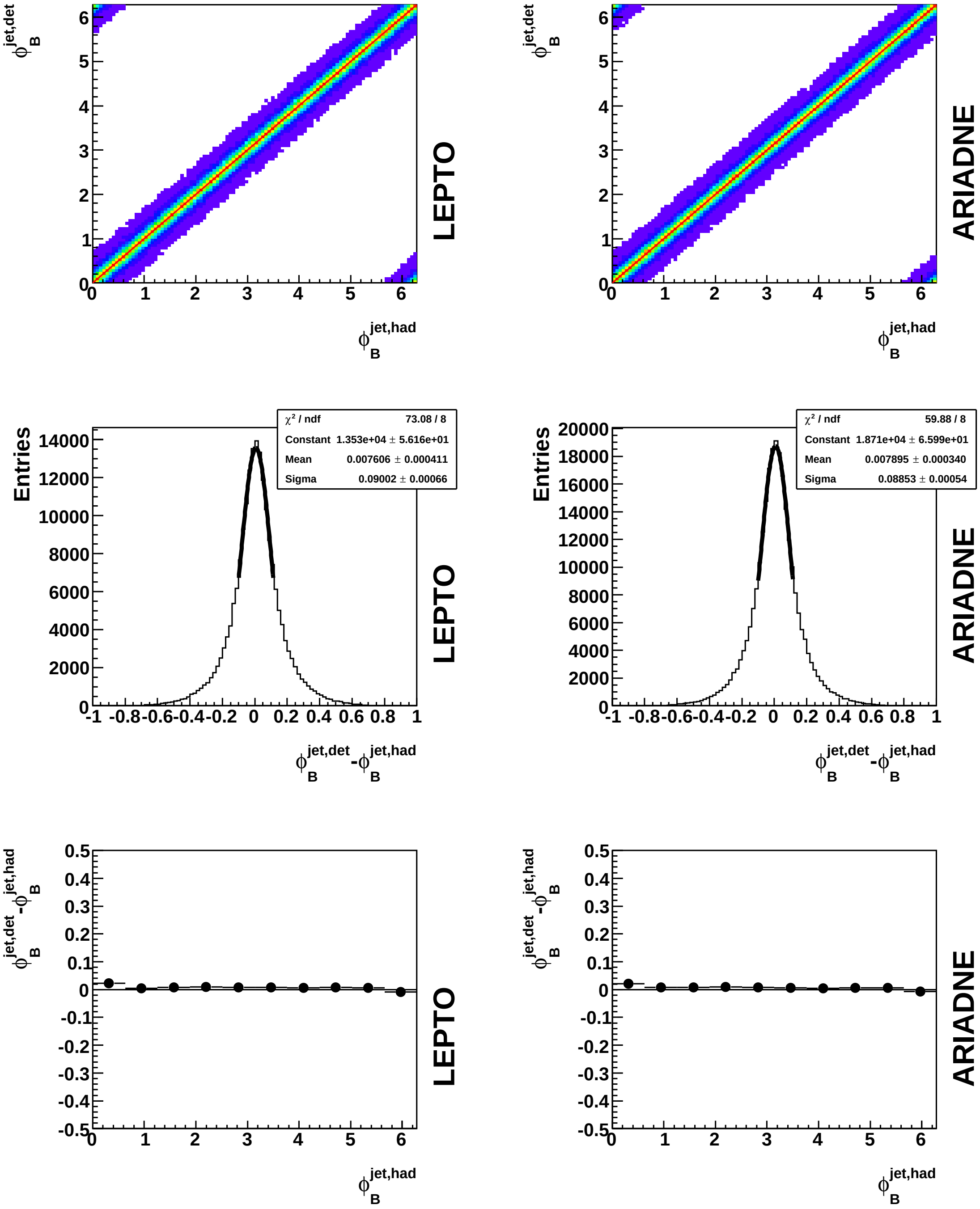


Figure 6.21.: Correlation between generated and reconstructed jet pseudorapidity $\phi_B^{\text{jet,det}}$ vs $\phi_B^{\text{jet,had}}$ (top row). Relative difference distribution $\left(\phi_{T,B}^{\text{jet,det}} - \phi_{T,B}^{\text{jet,had}}\right)/\eta_{T,B}^{\text{jet,had}}$ (middle row). Variation of the relative difference with respect to $\phi_{T,B}^{\text{jet,had}}$ (bottom row).

CHAPTER 7

Cross-Section Determination

The data are subject to various detector effects including non-linear response, finite resolution, limited acceptance and reconstruction inefficiencies. Such distortions in the case of vanishing background contributions are described by the Fredholm integral equation [71]:

$$\int_{\Omega} K(s,t)\, f(t)\, \mathrm{d}t = g(s)\,, \tag{7.0.4.1}$$

where $g(s)$ is the measured distribution, $f(t)$ the underlying true distribution and the kernel $K(s,t)$ describes the response of the detector. Thus the determination of $f(t)$ requires the solution of this equation. The process of extracting the true distribution from the measured one is called *unfolding*. Various techniques were proposed in the literature to solve equation 7.0.4.1. In this thesis the so-called *method of correction factors (or bin-by-bin method)* was employed.

In this chapter the procedure used for the determination of the inclusive-jet cross sections is described. In Section 7.1 the bin-by-bin method is presented. In Section 7.2 the sensitivity of the cross section determination procedure to the choice of MC model is investigated. After discussing the influence of the polarisation and higher-order QED effects on the measurements in Sections 7.3 and 7.4, respectively, the chapter concludes with the binning definition for the jet-production cross sections.

7.1. Acceptance Correction

The differential cross section for inclusive-jet production in some kinematic bin is determined according to:

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right|_{\mathrm{bin},i} = \frac{N_{\mathrm{bin},i}}{\mathscr{L}\cdot\Delta_{\mathrm{bin},i}}\cdot\mathscr{A}^{-1}_{\mathrm{bin},i}\cdot\mathscr{C}_{\mathrm{bin},i}, \tag{7.1.0.2}$$

where $N_{\mathrm{bin},i}$ is the number of jets reconstructed in the data in bin i, $\mathscr{L}$ the integrated luminosity of the data sample, $\Delta_{\mathrm{bin},i}$ the bin width and $\mathscr{A}_{\mathrm{bin},i}$ the acceptance correction factor described below. The effects from higher-order QED processes or those related to the polarisation of the lepton beam need to be included in the definition of the cross section in order to obtain an observable consistent with that provided by existing NLO QCD codes (see Section 8.1). These effects are combined in the additional multiplicative term $\mathscr{C}_{\mathrm{bin},i}$.

The acceptance correction factors $\mathscr{A}_{\mathrm{bin},i}$ are applied to the data in order to correct for detector effects. They are determined using MC samples from the number of jets generated in some kinematic bin, $N^{\mathrm{gen}}_{\mathrm{bin},i}$, and the number of jets reconstructed in the same bin, $N^{\mathrm{rec}}_{\mathrm{bin},i}$:

$$\mathscr{A}_{\mathrm{bin},i} = \frac{N^{\mathrm{rec}}_{\mathrm{bin},i}}{N^{\mathrm{gen}}_{\mathrm{bin},i}}. \tag{7.1.0.3}$$

In order to ensure the validity of this bin-by-bin multiplicative correction, the migrations across the neighbouring bins have to be sufficiently small and MC simulations have to provide a good description of the shape of the measured distributions. Two additional variables can be defined in order to quantify the detector effects, namely the *purity*

$$\mathscr{P} = \frac{N^{\mathrm{rec}\wedge\mathrm{gen}}_{\mathrm{bin},i}}{N^{\mathrm{rec}}_{\mathrm{bin},i}} \tag{7.1.0.4}$$

and the *efficiency*

$$\mathscr{E} = \frac{N^{\mathrm{rec}\wedge\mathrm{gen}}_{\mathrm{bin},i}}{N^{\mathrm{gen}}_{\mathrm{bin},i}}. \tag{7.1.0.5}$$

In these definitions $N^{\mathrm{rec}\wedge\mathrm{gen}}_{\mathrm{bin},i}$ represents the number of jets generated and reconstructed in some particular bin. The purity of a bin is the fraction of jets generated and reconstructed in a given bin divided by the number reconstructed in that bin. It corresponds approximately to the fraction of jets that migrated from the neighbouring bins into the bin under consideration. This can happen, for example, due to the finite resolution of the detector. The efficiency quantifies the fraction of jets that were genereted and reconstructed in the same bin; $1-\mathscr{E}$ is an estimate of jet loss due to migrations outside the measurement bin or due to reconstruction inefficiency or cut requirements. Using purity and efficiency the acceptance can be re-expressed as:

$$\mathscr{A} = \frac{\mathscr{E}}{\mathscr{P}}. \tag{7.1.0.6}$$

The terms appearing in the definitions (7.1.0.3)–(7.1.0.5) are not statistically independent, therefore a correlation between different factors has to be taken into account when the statistical uncertainty attributed to $\mathscr{A}$, $\mathscr{P}$ or $\mathscr{E}$ factors is needed. However, they can be expressed in terms of statistically independent quantities

$$\mathscr{P} = \frac{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}}}{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}} + N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\ulcorner\mathrm{gen}}}, \quad \mathscr{E} = \frac{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}}}{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}} + N_{\mathrm{bin},i}^{\ulcorner\mathrm{rec}\wedge\mathrm{gen}}}, \quad \mathscr{A} = \frac{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}} + N_{\mathrm{bin},i}^{\ulcorner\mathrm{rec}\wedge\mathrm{gen}}}{N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\mathrm{gen}} + N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\ulcorner\mathrm{gen}}}, \tag{7.1.0.7}$$

where $N_{\mathrm{bin},i}^{\ulcorner\mathrm{rec}\wedge\mathrm{gen}}$ and $N_{\mathrm{bin},i}^{\mathrm{rec}\wedge\ulcorner\mathrm{gen}}$ are the number of jets generated but not reconstructed in bin i and reconstructed but not generated in that bin, respectively.

The acceptance correction factors, efficiency and purity for the relevant inclusive-jet cross sections determined using either LEPTO or ARIADNE MC samples are shown in Figures 7.1 (a)–(c). In general, the purity is typically above 40% for all kinematic observables, while the efficiency is typically between 30% and 65%. The decrease of efficiency in the region $250\,\mathrm{GeV}^2 < Q^2 < 500\,\mathrm{GeV}^2$ was observed before [149, 154, 155, 160] and was attributed to the reduced electron identification capabilities in the transition region between the RCAL and BCAL. The acceptance correction factors never exceed 1.6 and are typically below 1.4. A major limitation of the described method is its possibly high sensitivity to the MC true-level distribution [161]. This effect is investigated in the following section.

7.2. MC Studies of the Unfolding Approach

As was demonstrated in Chapter 5, both MC models after the reweightings describe the observed distributions reasonably well and are consistent with each other at the detector level. However, as can be seen in Figure 7.1, the acceptance correction factors estimated using ARIADNE and LEPTO MC models can differ by 10%, implying substantial difference between the generators at the hadron level. In order to investigate the stability of the unfolding procedure with respect to the choice of MC model, an additional study has been performed.

The observed cross-section value in a particular bin can be regarded as a

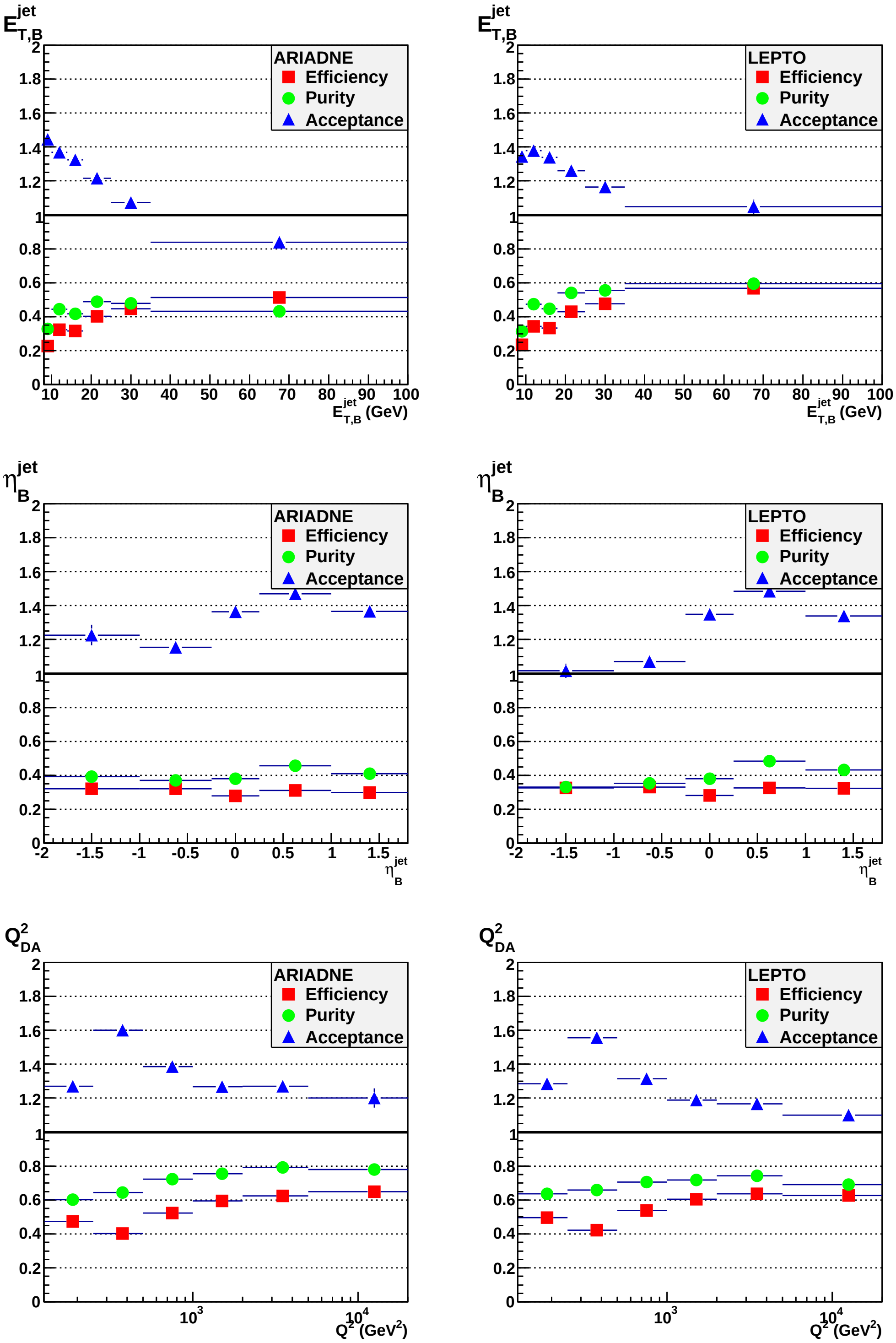


Figure 7.1.: Acceptance correction factors, efficiency and purity for inclusive-jet cross sections as functions of $E_{T,B}^{jet}$, η_B^{jet} and Q^2.

Gaussian-distributed[1] random variable $x = \mathcal{N}\left(\mu, w\right)$ with mean value, μ, provided by the *true* physical cross section and the variance, w, given by statistical uncertainty[2]. In this context, the Eq. (7.1.0.2) has to be considered as a definition of a statistical estimator of the true cross-section value.

The pull distribution provides valuable information when the properties of an estimator such as bias and efficiency are studied. It is defined as

$$p = \frac{x - \mu}{w}. \tag{7.2.0.8}$$

This quantity follows a normal distribution with zero mean and unit variance if an estimator, x, is unbiased and its variance is correctly determined. In practice, however, the true cross-section value, μ, is unknown and has to be determined. Therefore for investigation of the properties of the employed unfolding approach, pseudo-experiments were performed. For this purpose, LEPTO and ARIADNE MC events were split separately in ten equal-size statistical samples. For every sample the true cross-section value, provided by the hadron-level prediction is known, and pulls can be constructed. For every single-differential cross section, the detector-level spectra of both MC samples were unfolded using the acceptance-correction factors determined using the full LEPTO sample[3]. The unfolded cross sections were compared to the corresponding hadron-level values. The statistical uncertainty of the detector-level cross section was used as an estimate of w. The obtained pull distributions for individual cross-section bins were added together and presented in Figure 7.2. Additional figures with individual pull distributions for all single-differential cross-section bins can be found in the Appendix.

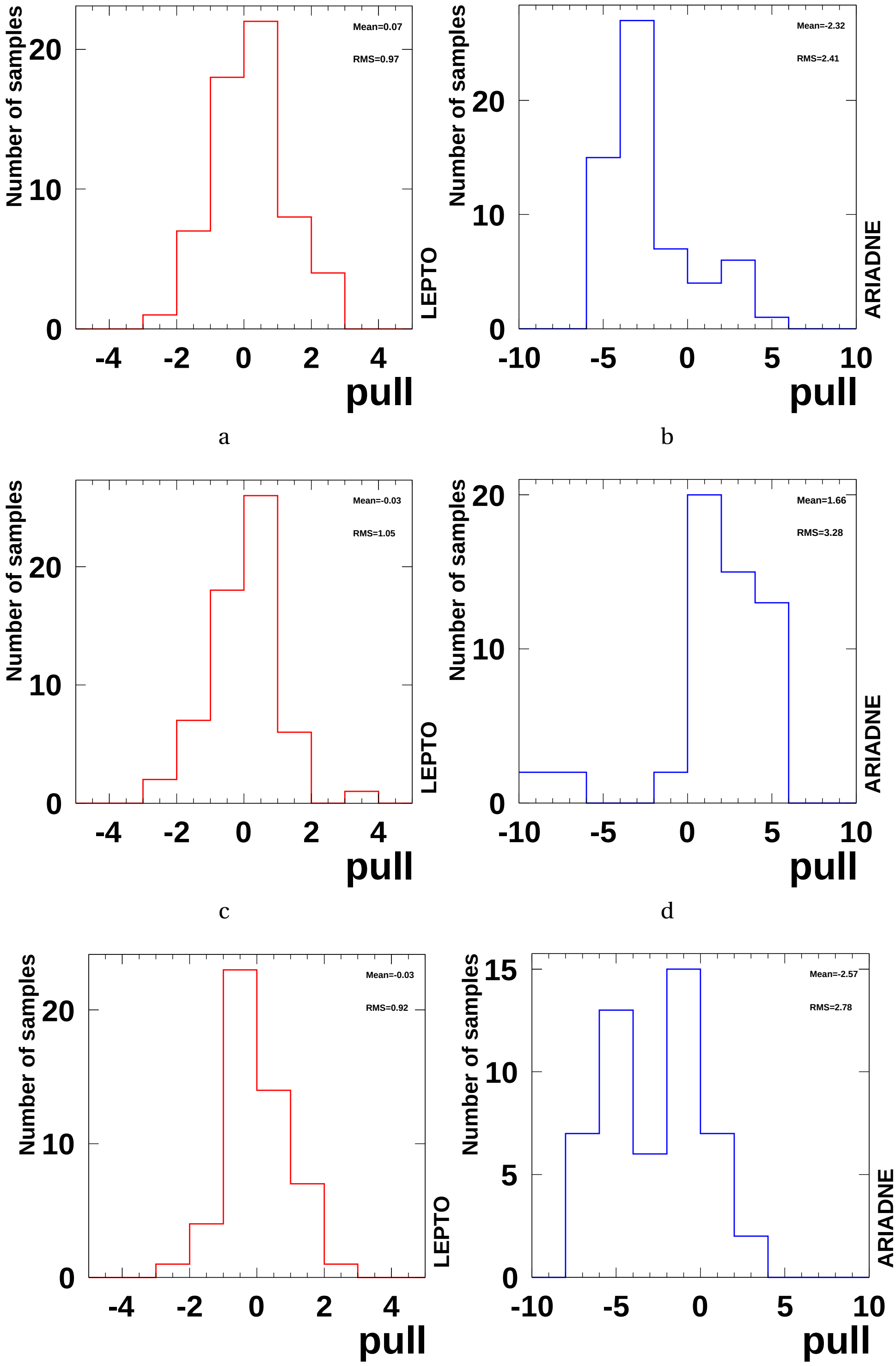
It was observed that the LEPTO acceptance correction applied to LEPTO pseudo-data provides an unbiased estimate of the hadron-level cross section. However, when the same correction factors were applied to ARIADNE, non-Gaussian pull distributions with significant bias and non-unit variance were observed, indicating sensitivity of the cross-section estimate to the choice of MC model. This effect was taken into account in the systematic-uncertainty assessment (see Section 8.3).

---

[1] In general, the number of counts in a particular bin follows a multi-Poisson distribution, as described in Section 8.4, nevertheless the Gaussian approximation can be used in the limit of large number of counts (no bins with less than 55 counts were observed).

[2] Experimental uncertainty can be neglected for simplicity.

[3] The small correlation between the acceptance-correction factors and detector-level number of counts in LEPTO pseudo-experiment samples can be neglected.



a b

c d

e f

Figure 7.2.: Pull distributions for $d\sigma/dQ^2$ (a,b) $d\sigma/dE_{T,B}^{jet}$ (c,d) $d\sigma/\eta_B^{jet}$ (e,f) cross sections in LEPTO and ARIADNE MC.

7.3. Polarisation Correction

The MC samples used in this analysis were generated assuming vanishing polarisation of the lepton beam, $P_e = 0$. In order to take non-zero polarisation of the electrons into account, the MC samples were reweighted using theoretical predictions. For this purpose the HECTOR program [162] interfaced to BASES [163] with the CTEQ5D PDFs [96] was used. The reweighting factors were determined from the ratio of predictions for the unpolarised inclusive DIS cross sections and those for the lepton beam polarisation corresponding to the particular data-taking period. The polarisation correction was implemented as a weight assigned to each MC event according to the Q^2 of the scattering process:

$$w_p\left(Q^2\right) = w_p = \frac{\sigma_{\text{pol}}}{\sigma_{\text{unpol}}}. \tag{7.3.0.9}$$

The average polarisation for different data-taking periods is summarised in Table 7.1. The obtained correction factors for the 2005 e^- sample as a function of Q^2 and using a spline interpolation are illustrated in Figure 7.3. The size of the correction increases with increasing Q^2 but nowhere exceeds 3% and typicaly is below 1%. The sign of the correction depends on the helicity of the lepton beam.

Data-taking period	Average polarisation, P_e
2004-2005 e^-	-0.06184
2006 e^-	0.09386
2006-2007 e^+	-0.06857

Table 7.1.: The average polarisation values for the data samples used in the analysis.

After applying this correction, MC events corresponded to the data sample with correct average polarisation of the lepton beam. Nevertheless, as described in Section 8.1, direct comparison of the measured spectra, unfolded using corrected MC samples, is imposible, because NLO pQCD predictions used in this thesis do not take into account polarisation effects. Therefore, besides the polarisation correction applied to MC events, an inverse of determined factors, $w_p^{-1}\left(Q^2\right)$, were applied to the data in order to obtain jet cross sections corresponding to unpolarised lepton scattering.

7.4. QED Corrections

The theoretical predictions for the measured cross sections obtained using the `NLOJET++` program include only the leading-order QED contribution, while

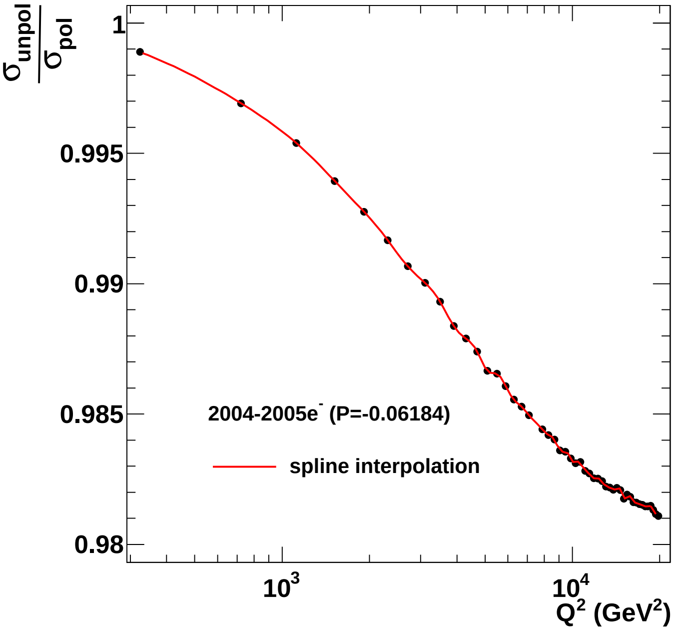
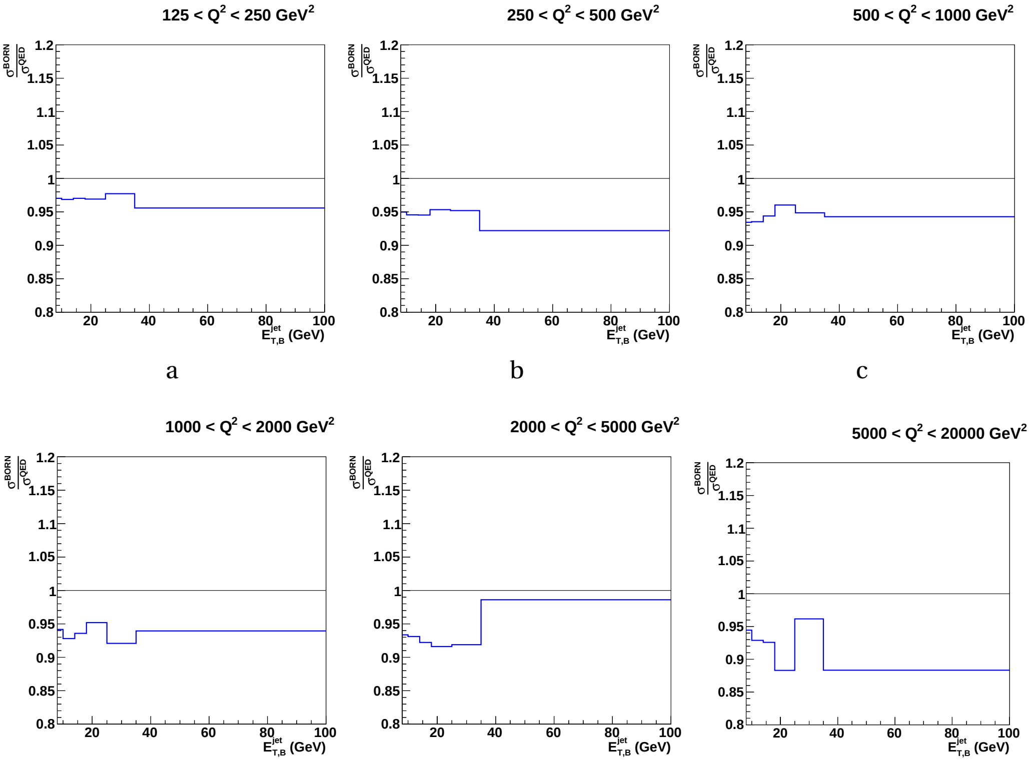


Figure 7.3.: The polarisation reweighting factors for the 2004/2005 e^- data taking period determined using the HECTOR program. The red curve represents the spline interpolation.

the measurements were influenced by higher-order processes like the running of the electromagnetic coupling, initial- and final-state EM radiation etc. To maintain the consistency between the data and theoretical predictions, the measured cross sections were corrected to the Born level using the MC predictions. A multiplicative factor applied to the data was determined using two LEPTO MC samples with higher-order QED processes switched on and off. The corresponding jet cross sections are denoted by σ^{QED} and σ^{BORN}, respectively. The correction factor is equal to the ratio:

$$\mathscr{C}_i^{\text{QED}} = \frac{\sigma_i^{\text{BORN}}}{\sigma_i^{\text{QED}}}. \tag{7.4.0.10}$$

Figure 7.4 illustrates the QED corrections determined in different kinematic bins. In general, the correction is approximately independent of $E_{T,B}^{jet}$ and increases with increasing Q^2. It is about 3% in the lowest Q^2 range and reaches about 10% for $5000 < Q^2 < 20000\,\text{GeV}^2$.



a b c

d e f

Figure 7.4.: QED multiplicative correction factors for inclusive-jet cross sections as functions of $E_{T,B}^{jet}$ in bins of Q^2.

7.5. Bin Definition

The choice of the size of the bin width is limited by two factors. Naturally, in order to obtain maximum information from the measurements, the bin width has to be as small as possible. However, the reduction of the bin size leads to a decrease of the number of entries. Therefore the width must be large enough to obtain a statistically significant number of entries per bin. In addition, because of finite resolution and possibly non-linear response of the detector, the migration effects can be substantial when the bin width is much smaller than the detector resolution. Besides that it is convenient to define the binning for inclusive-jet analysis consistent with that of the dijet and trijet analysis performed at ZEUS [149, 164].

The binning definition for the jet cross sections measured in this analysis is outlined below.

Bin Definition for $\mathrm{d}\sigma/\mathrm{d}Q^2$

For the single-differential cross-section $\mathrm{d}\sigma/\mathrm{d}Q^2$, six bins were defined spanning the measurement phase space from 125 GeV to 20000 GeV. Since the inclusive NC DIS cross section scales approximately as $1/Q^4$, the size of the bins increases with increasing Q^2 in order to obtain a statistically significant jet sample in each bin. Moreover, the bin width was chosen to be exceeding the variable resolution in particular bin, in order to minimise the migration effects.

Bin Definition for $\mathrm{d}\sigma/\mathrm{d}E^{jet}_{T,B}$

For the single-differential inclusive jet cross section as a function of $E^{jet}_{T,B}$, again six bins were chosen with varying width. Although the cross section is large at low $E^{jet}_{T,B}$ the size of the bin width was limited by the purity at low jet transverse energies. At low $E^{jet}_{T,B}$ the measurement is bounded by the phase-space restriction $E^{jet}_{T,B} > 8$ GeV, while the upper limit was chosen as 100 GeV because only very few jets in the sample have an $E^{jet}_{T,B}$ value above 100 GeV.

Bin Definition for $\mathrm{d}\sigma/\mathrm{d}\eta^{jet}_{lab}$

The cross-section $\mathrm{d}\sigma/\mathrm{d}\eta^{jet}_{lab}$ spans the region $-2 < \eta^{jet}_{B} < 1.8$ and has 5 bins. The sizes of the bins were chosen in order to optimise the purity and at the same time get sufficient number of bins for retaining essential features of the shape of the distribution. As can be seen in Figure 7.1 (b), a stable and reasonably high purity was achieved.

Bin Definition for $\mathrm{d}\sigma/\mathrm{d}E^{jet}_{T,B}$ in different regions of Q^2

The single-differential cross-sections $\mathrm{d}\sigma/\mathrm{d}E^{jet}_{T,B}$ were measured in six regions of Q^2. The size of the Q^2 bins corresponds to that of the $\mathrm{d}\sigma/\mathrm{d}Q^2$ cross section and the $E^{jet}_{T,B}$ binning scheme was the same as for the Q^2-integrated $\mathrm{d}\sigma/\mathrm{d}E^{jet}_{T,B}$ cross section.

CHAPTER 8

Results. Inclusive-Jet Cross Sections

8.1. NLO Calculations

The next-to-leading-order QCD $\left(\mathcal{O}\left({\alpha_s}^2\right)\right)$ predictions for the observables measured in this thesis were calculated using the NLOJET++ program [165, 166]. The predictions were performed in the $\overline{\mathrm{MS}}$ renormalisation and factorisation scheme and the dipole-subtraction method was applied to cancel the singularities arising in intermediate calculations from infrared and collinear phase-space regions. The number of active flavours was set to five and renormalisation $\left(\mu_R\right)$ and factorisation $\left(\mu_F\right)$ scales were set to

$$\mu_R = \sqrt{Q^2 + {E_{T,B}^{jet}}^2} \tag{8.1.0.1}$$

$$\mu_F = Q^2, \tag{8.1.0.2}$$

respectively. The strong-coupling evolution was calculated at two loops with $\alpha_s(M_Z) = 0.1176$. The HERAPDF 1.5 proton PDF parametrisation [167] was used in the calculations, however alternative sets were also investigated. The parton-level predictions for the jet cross sections were obtained by applying the k_T algorithm in the Breit frame to the partons generated by the program. In order to obtain predictions for jets of hadrons, the calculations were corrected to the hadron level using the MC models, as described in 8.1.1. The predictions do not include contributions from γZ interference or Z^0 exchange, so they were corrected for these effects; the details of the correction procedure are given in Section 8.1.2.

The uncertainty on the theoretical predictions was investigated and is presented in the Section 8.2.

8.1.1. Hadronisation Correction

In order to directly compare the NLO QCD predictions with the data, the calculations have to be corrected for hadronisation effects, because the measurements refer to jets of hadrons while the predictions relate to partons. To determine hadronisation corrections and to estimate the influence of the parton-shower modelling and hadronisation process, the predictions from ARIADNE and LEPTO event generators were utilised. The hadronisation correction was determined from the ratio

$$\mathscr{C}_i^{\mathrm{hadr}} = \frac{\sigma_i^{\mathrm{hadr}}}{\sigma_i^{\mathrm{part}}} \tag{8.1.1.1}$$

of the jet cross sections at the hadron, σ_i^{hadr}, and parton σ_i^{part} levels, respectively. The parton-level cross section was determined using the partons available as an input to the hadronisation model after the parton-shower simulation step. The hadron level refers to the 'stable'[1] particles available in the MC event record. An average of the correction factors determined from ARIADNE and LEPTO was used to correct for hadronisation effects.

Figure 8.1 illustrates the hadronisation corrections as functions of $E_{T,B}^{jet}$ in different regions of Q^2. Both MC generators predict similar behaviour of the correction factors. In general, the correction does not exceed 15% and its magnitude decreases with increasing $E_{T,B}^{jet}$ and Q^2.

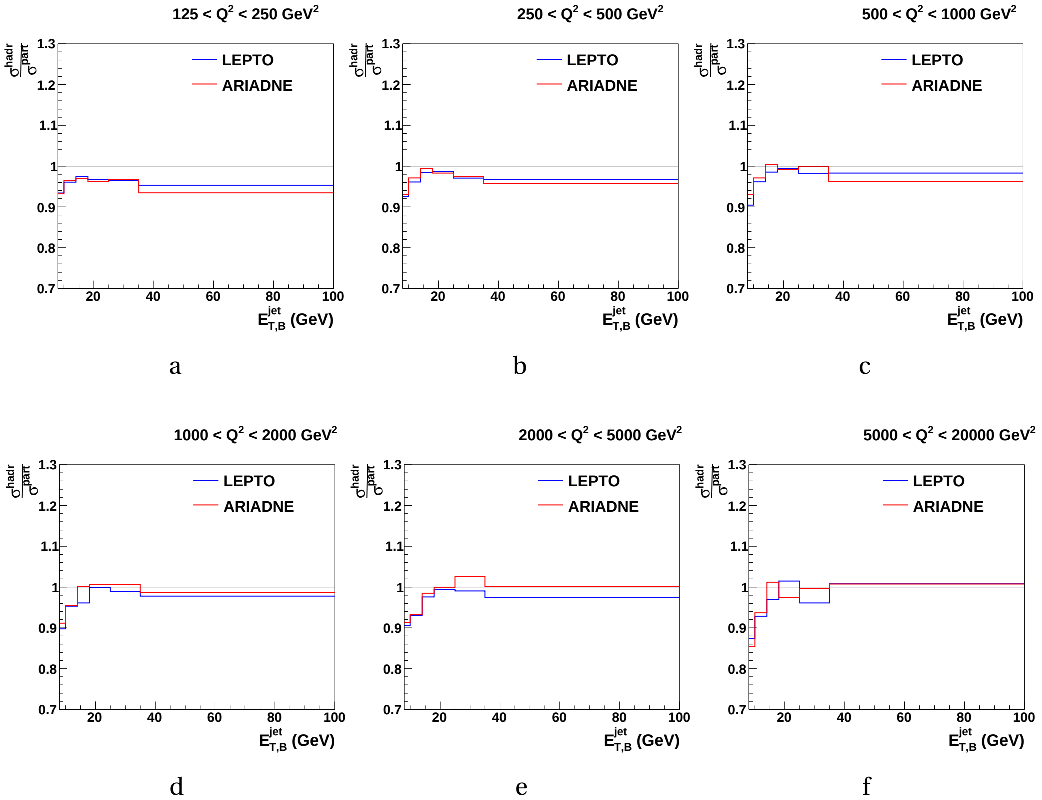
8.1.2. Electroweak Corrections

As mentioned earlier, the fixed-order pQCD predictions from the NLOJET++ include only the single-photon exchange component. However the contribution from γZ-interference and Z^0-exchange become significant in the kinematic region of boson virtualities $Q^2 > 1000\,\mathrm{GeV}^2$. The LEPTO generator was utilised for the estimation of the size of these effects. The ratio of the jet production cross sections calculated including and excluding electroweak effects was used for the correction:

$$\mathscr{C}_i^{Z^0} = \frac{\sigma_i^{Z^0}}{\sigma_i^{\mathrm{no}\ Z^0}}. \tag{8.1.2.1}$$

---

[1] According to the ZEUS convention, all particles with lifetime $\tau > 10$ ns.



a b c

d e f

Figure 8.1.: Hadronisation multiplicative correction factors for inclusive-jet cross sections as function of $E_{T,B}^{jet}$ in bins of Q^2, applied to the NLO pQCD predictions. The correction factors were determined using LEPTO and ARIADNE MC samples.

The γZ^0 interference depends on the charge of the lepton beam, thus the luminosity-weighted average of the correction factors for the $e^+ p$ and $e^- p$ scattering was applied to the pQCD predictions.

The correction factors are shown in Figure 8.2. The size of the correction is approximately independent of jet transverse energy and increases with increasing exchanged boson virtuality, reaching about 20% in the largest Q^2 bin.

8.2. Theoretical Uncertainties

In order to test perturbative QCD predictions it is important to check the stability of the predictions with respect to variations of input parameters, such as the value of α_s, proton PDF, renormalisations and factorisation scales etc. In

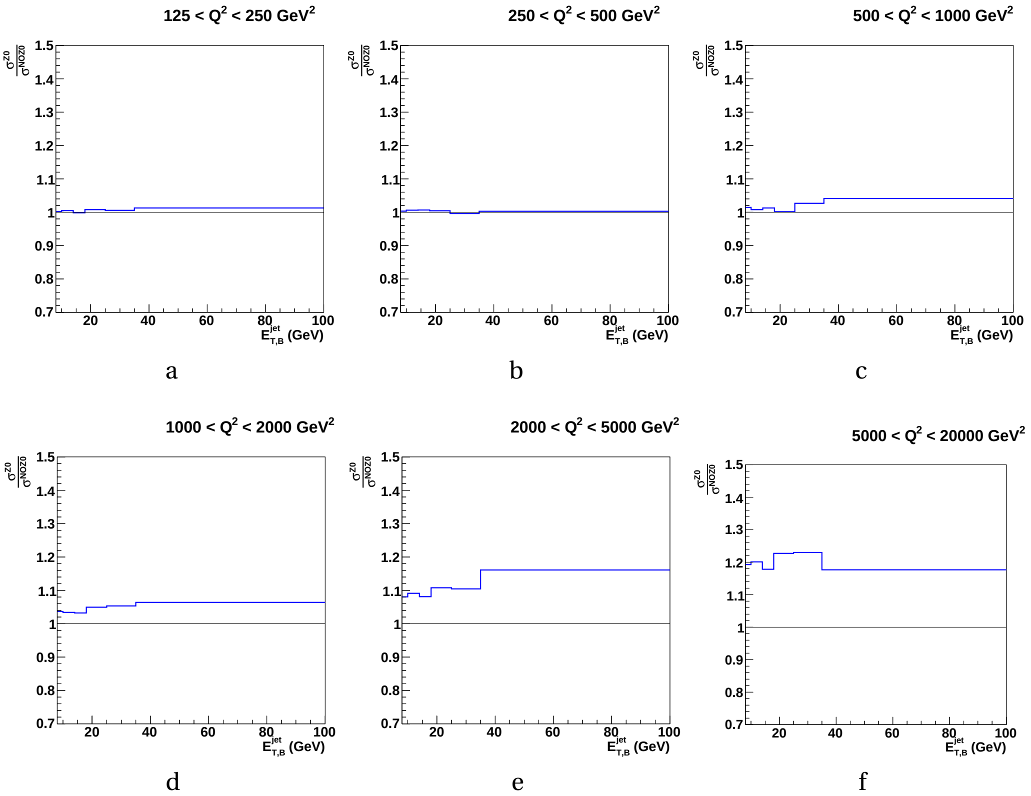


Figure 8.2.: Electroweak multiplicative correction factors for inclusive-jet cross sections as function of $E_{T,B}^{jet}$ in bins of Q^2, applied to NLO pQCD predictions. The correction factors were determined using dedicated LEPTO MC samples.

this analysis, an uncertainty was attributed to each individual variation of theoretical parameters. The following sources of uncertainty were investigated:

- the uncertainty on the predicted inclusive-jet cross section due to the value of $\alpha_s(M_Z)$ was estimated by repeating the calculations assuming different values[2] of $\alpha_s(M_Z)$ =0.1156 and 0.1196 and using the corresponding proton PDF sets from the HERAPDF 1.5 series. The resulting uncertainty was typically about ±2%;
- the uncertainty on the NLO QCD calculations arising from higher-order terms, estimated by varying the renormalisation scale up and down by a factor of 2. It was about ±10% in the low-Q^2, low-$E_{T,B}^{jet}$ region but decreasing to about ±5% at high Q^2 and high $E_{T,B}^{jet}$;

---

[2] These values correspond to the standard deviation of α_s, that was determined in the HERAPDF fit [167]

- the uncertainty of the calculations originating from the dependence on factorisation scale was estimated by repeating the calculations with μ_F scaled up and down by factors 0.5 and 2, resulting in a difference between the predictions $\lesssim \pm 3\%$;

- the uncertainty due to the modelling of the parton shower was estimated as the symmetric relative difference between the predictions for hadronisation correction factors obtained using LEPTO or ARIADNE MC

  $$\delta^{\pm}_{\text{PS},i} = \frac{1}{2}\max\left(\left|\mathscr{C}^{\text{hadr}}_{\text{LEPTO},i} - \mathscr{C}^{\text{hadr}}_{av,i}\right|, \left|\mathscr{C}^{\text{hadr}}_{\text{ARIADNE},i} - \mathscr{C}^{\text{hadr}}_{av,i}\right|\right), \quad (8.2.0.2)$$

  where

  $$\mathscr{C}^{\text{hadr}}_{av,i} = \frac{1}{2}\left(\mathscr{C}^{\text{hadr}}_{\text{LEPTO},i} + \mathscr{C}^{\text{hadr}}_{\text{ARIADNE},i}\right).$$

  It was found to be typically about ±1% in most of the phase space.

- the statistical uncertainty of the theoretical calculations was < 1% and therefore was neglected.

- the theoretical uncertainty due to the proton PDF was estimated according to the HERAPDF 1.5 recommendation [167]. The error was subdivided into two independent components: PDF eigenvectors variations and model and parametrisation variations (see Chapter 9 for detailed discussion). The detailed breakdown of PDF uncertainty for individual double-differential cross-section bins is presented in Figure 8.4. Overall, the uncertainty amounts to about $\lesssim 5\%$ for $Q^2 < 5000$ GeV^2 and $\lesssim 3\%$ in the high-Q^2 region. Thus, the PDF uncertainty is the second-largest contribution to the total theoretical uncertainty. It was observed that the dominant contribution to the PDF error was due to assumptions in the PDF parametrisation[3].

The total theoretical uncertainty was calculated by summing the individual contributions in quadrature. The break-down of the theoretical uncertainties as functions of $E^{jet}_{T,B}$ and Q^2 is summarised in Figure 8.3.

8.3. Systematic Uncertainties

Systematic uncertainties arise due to various effects such as, for example, incomplete understanding of the detector response or deficiencies in the modelling of the ep physics. The systematic uncertainties from identified sources

---

[3] The predictions with the largest deviation from the central value were characterised by specific PDF parametrisation in which parameters E_{u_v} and D_{u_v} were let to be free parameters in the PDF fit. See Section 9.2.1 for more details.

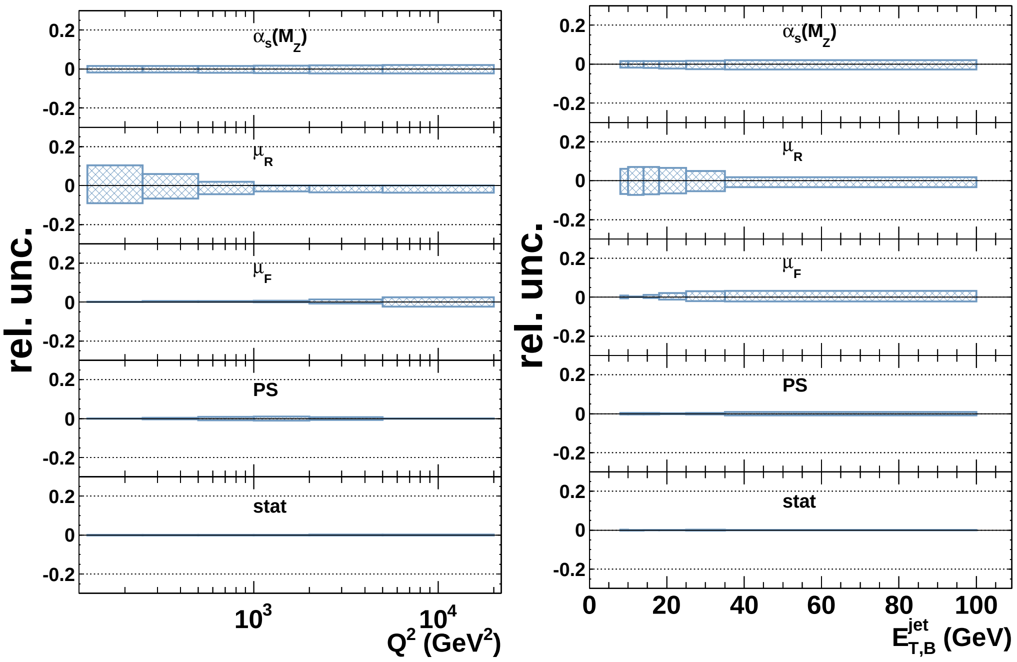


Figure 8.3.: Estimated theoretical uncertainties on NLO pQCD predictions as functions of Q^2 (a) and $E^{jet}_{T,B}$ (b).

were quantified in the the following way. For each source of uncertainty attributed to imperfect knowledge of some continuous parameter, a variation by its estimated uncertainty was performed and the cross sections were reevaluated. In the case when the source of uncertainty is attributed to a specific reconstruction technique, the difference with respect to the cross section calculated using an alternative reconstruction method is taken as an approximation of the systematic uncertainty. The positive, δ_i^+, (negative, δ_i^-) variations of the cross sections were added in quadrature in order to obtain the positive (negative) total systematic uncertainties

$$\delta^2_{(+)} = \sum_i \delta_i^{(+)2},$$

$$\delta^2_{(-)} = \sum_i \delta_i^{(-)2}.$$

In what follows a detailed description of the sources of systematic uncertainties is presented.

- **Hadronic energy scale**
  As demonstrated in Section 6.6.1, the precision to which the absolute jet-energy scale can be calibrated in the data is down to ±1% for jets with

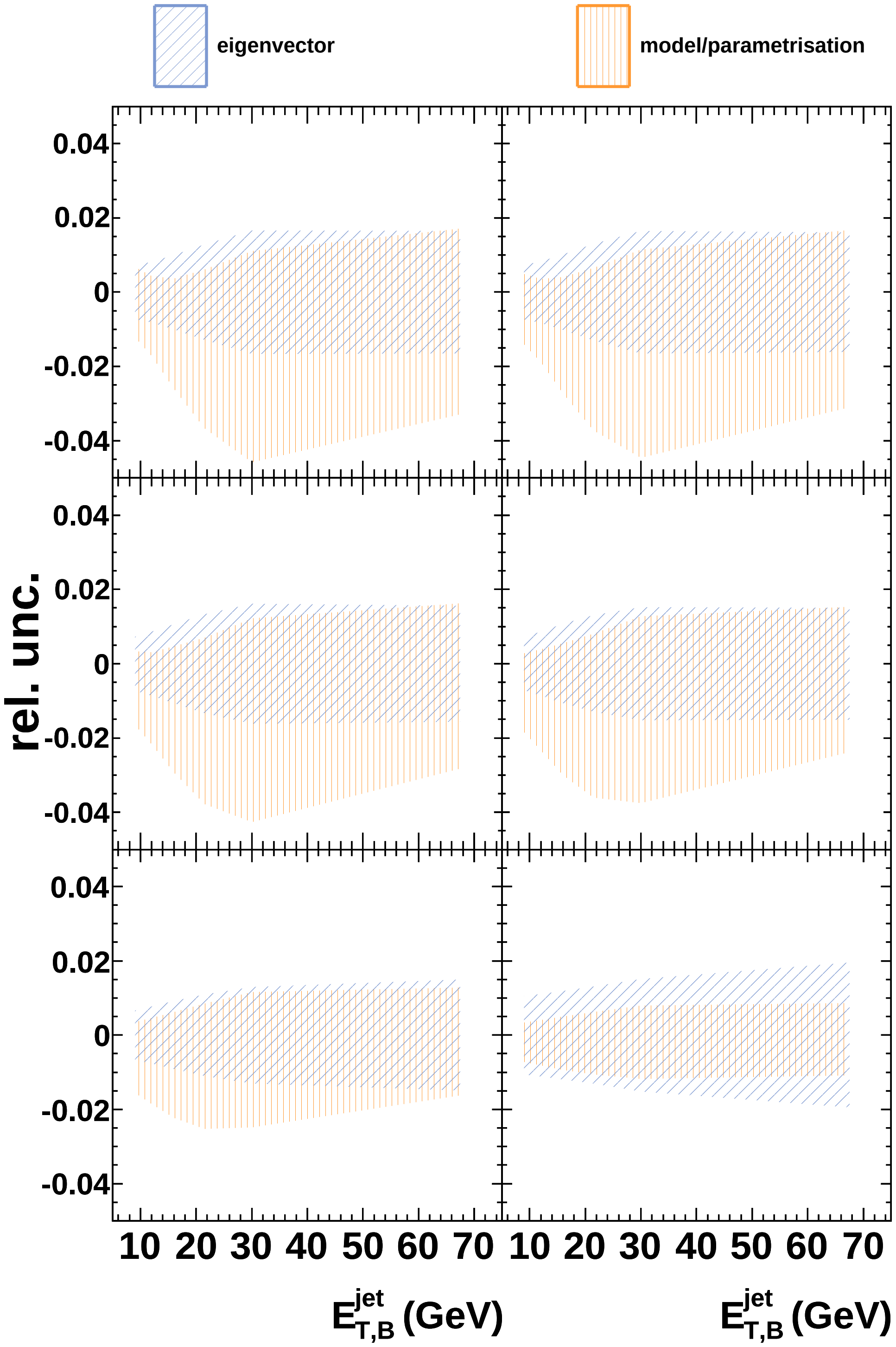


Figure 8.4.: Estimated theoretical uncertainty on NLO pQCD predictions due to PDF as functions of $E_{T,B}^{jet}$ in different bins of Q^2.

$E^{jet}_{T,lab} > 10$ GeV and, as was found in other jet analysis at ZEUS, ±3% for jets with $3 < E^{jet}_{T,lab} < 10\,\text{GeV}$. Therefore, to assess the effect of this uncertainty on the cross sections, the jet transverse energy was varied in the MC simulations up and down within the specified limits. Each variation resulted in a single-sided[4] change of the cross sections in all bins with respect to the nominal values. This uncertainty was monotonically decreasing as a function of Q^2, ranging from 5% at low Q^2 to 3% in the highest Q^2 bin. As a function of $E^{jet}_{T,B}$, it was approximately constant and did not exceed 7%.

- **Jet energy resolution**
  In study [149] the description of the jet energy resolution in data and MC simulations was investigated. For this purpose single-jet events were selected and two quantities:

$$\sigma^{\text{DATA}}_{\text{rel}} = \frac{E^{jet}_{T,lab} - E_{T,DA}}{E_{T,DA}}, \tag{8.3.0.3}$$

$$\sigma^{\text{MC}}_{\text{rel}} = \frac{E^{det,jet}_{T,lab} - E^{had,jet}_{T,lab}}{E^{had,jet}_{T,lab}}. \tag{8.3.0.4}$$

  where $E^{jet}_{T,lab}$ is the jet transverse energy in the laboratory frame and $E_{T,DA}$ is the transverse energy of the scattered electron (see Eq. (4.4.3.5)), while $E^{det,jet}_{T,lab}$ ($E^{had,jet}_{T,lab}$) is the jet transverse energy determined at the reconstructed (generated) level in MC, were examined. Two variables $\sigma^{\text{DATA}}_{\text{rel}}$ and $\sigma^{\text{MC}}_{\text{rel}}$ represent the measure of the jet-energy resolution in data and simulations, respectively. In that study, the agreement between these two quantities was found to be sufficient for not attributing the systematic uncertainty to this source.

- **Acceptance correction**
  The dependence of the acceptance correction factors on the modelling of the parton-shower process was taken into account in the systematic uncertainty. The cross sections were re-evaluated using the acceptance corrections obtained from the ARIADNE sample instead of LEPTO. In order to symmetrize the uncertainty, half of the absolute difference between the obtained cross sections was assigned to positive and negative components of the systematic error. The effect of the change of MC program was below 3% for $Q^2 < 1000$ GeV^2 but increased up to 5% in the largest

---

[4] As the result of the variation, the measured cross sections in *all* bins increase or decrease simultaneously.

Q^2 bin. As a function of $E_{T,B}^{jet}$ it was typically below 5% but reached 10% in the high-$E_{T,B}^{jet}$ regions. In principle, in order to reduce this sensitivity a composition of generated event samples can be used for the determination of the acceptance-correction factors. Nevertheless, in this analysis, the `LEPTO` sample was used as a nominal event generator, as it initially provided a better description of the data.

- **Dependence on the electron finder**
  The SINISTRA electron finder (see Section 4.3) was used for the determination of the nominal cross-section values. In order to estimate the size of the systematic uncertainty attributed to the electron identification, the EM electron finder [142, 143] was used as an alternative[5]. The resulting uncertainty was typically about 0.5% but could reach 2% in regions of high Q^2.

- **Electromagnetic energy scale**
  The uncertainty on the electromagnetic energy scale was determined using a technique similar to that for the calibration of the hadronic energy scale. Exploiting the approximate independence of the energy of the electron reconstructed using the double-angle method, a calibration of the absolute electromagnetic energy scale of the calorimeter can be performed. As was demonstrated in Section 6.5, the residual discrepancy between data and MC simulations was less than 2%. This value was interpreted as an uncertainty on the electron energy scale and was propagated to the cross sections by varying the scattered electron energy up and down by ± 2% in the Monte Carlo simulation. Therefore, this variation also covers the uncertainty due to the determination of the Lorentz boost to the Breit reference frame.

- **Track-veto efficiency correction**
  In order to estimate the size of the effect on the cross sections due to the MC track-veto correction introduced in Section 6.3, the procedure was changed slightly. Instead of the parametrisation of the track-veto efficiency as a function of y_{DA}, the CTD-FLT vertex-track multiplicity was used as an alternative. The effect on the cross section was typically about 0.5%.

- **Photoproduction background**
  The contribution to the high-Q^2 NC DIS jet cross sections from photo-

---

[5] By such a variation the effects due to the characteristics of a particular identification algorithm (SINISTRA is based on neural-network pattern recognition, while EM models directly the probability of a shower to originate from the DIS electron candidate) are taken into account.

production processes can be estimated using the dedicated Monte Carlo event generators e.g. PYTHIA and HERWIG. As was demonstrated in Section 5.6, the photoproduction background can be neglected in the phase space studied here.

- **Selection cuts variation**
  The stability of the results with respect to the choice of particular cut values, introduced in Chapter 5, was investigated. For this purpose the cut thresholds were changed in the data and MC according to the resolution of the considered quantity. As mentioned in Section 5.6, the description of the $(E - p_Z)$-distribution in the data by simulations was inaccurate and requires correction. Hence, for

  - $(E - p_Z)$ cut: the variation by ±6% was performed and resulted in the change of the cross sections by $\lesssim \pm 1\%$

- **Luminosity**
  The luminosity and its uncertainty were determined using dedicated detector components, as described in Section 3.3. The overall normalisation of the measurements depends on the integrated luminosity of the data sample that was determined with ±1.8% for the given data-taking period.

All described systematic studies are summarised in Table 8.1. The total uncorrelated component of the experimental systematic uncertainty was obtained by adding corresponding individual contributions in quadrature. The Figure 8.5 illustrates the individual components of the statistical and systematic uncertainties as functions of various kinematic observables.

Source	**Variation**	**Applied to**	**Correlation**
jet energy scale	±3% for $3 < E^{jet}_{T,lab} < 10$GeV, ±1% for $E^{jet}_{T,lab} > 10$GeV	MC	bin-to-bin correlated
electron energy scale	±2%	MC	bin-to-bin correlated
selection cuts variation			
$E - p_Z$ cut	±6%	data and MC	uncorrelated
Additional			
Acceptance correction	ARIADNE	data	bin-to-bin correlated
Electron identification	EM-algorithm	data and MC	uncorrelated
track veto efficiency correction	as function of CTD-FLT track multiplicity	MC	uncorrelated

Table 8.1.: Summary of variations investigated for the systematic-uncertainty estimation.

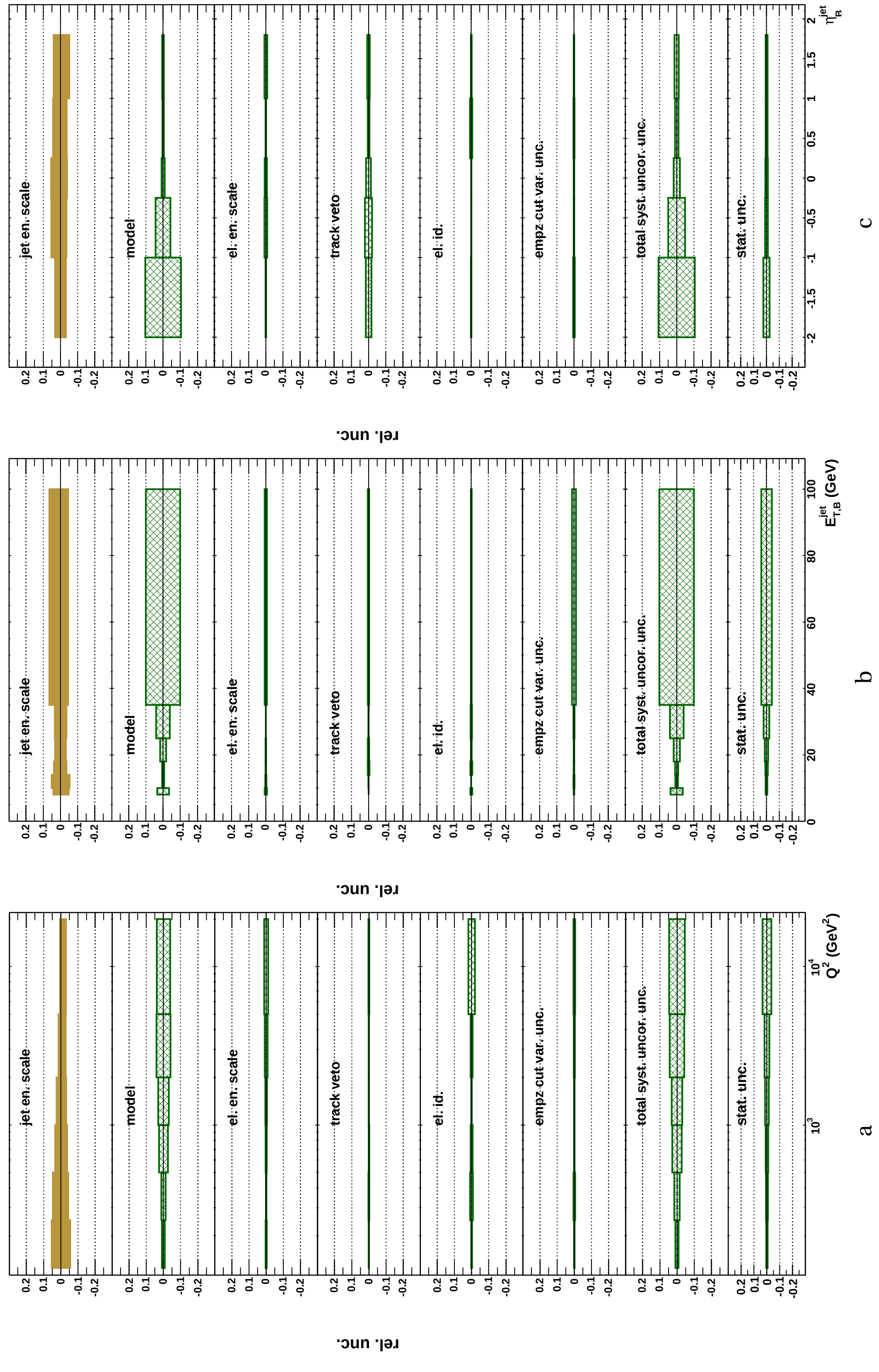


Figure 8.5.: Estimated systematic uncertainties on inclusive-jet cross sections as functions of Q^2 (a), $E_{T,B}^{jet}$ (b) and η_B^{jet} (c).

8.4. Statistical Correlations

In counting experiments, the statistical uncertainty is typically assumed to follow a Poisson distribution. However in the case of inclusive jet production, in which every jet is counted, such an assumption is incorrect because jets appear in pairs or with even higher multiplicity. Thus, the occurrence of jets is not a Poisson process. However, taking into account that the number of *events* with a given jet multiplicity accumulated during a fixed time interval is indeed Poisson distributed, it was demonstrated [168] that the probability of observing n jets in a given bin of a jet cross section[6] is given by the multi-Poisson:

$$P(n_1, n_2, n_3, \ldots) = \prod_i e^{-\lambda_i} \cdot \frac{\lambda_1^{n_i}}{n_i!}, \tag{8.4.0.5}$$

where λ_i denotes the expectation value for the number, n_i, of i-jet events and where the product runs over allowed jet multiplicities in a given reaction. By definition, the expected number of jets in a given bin is:

$$\langle n \rangle = \sum_{n=0}^{\infty} n \sum_{n_1, n_2, n_3, \ldots} P(n_1, n_2, n_3, \ldots) \tag{8.4.0.6}$$

with the number of jets, n, constrained by the requirement

$$n = n_1 + 2 \cdot n_2 + 3 \cdot n_3 + \ldots \tag{8.4.0.7}$$

Substituting the expression for $P(n_1, n_2, n_3, \ldots)$, the mean value and the variance of the number of jets can be shown to be

$$\langle n \rangle = \lambda_1 + 2 \cdot \lambda_2 + 3 \cdot \lambda_3 + \ldots \tag{8.4.0.8}$$

$$\left\langle n^2 \right\rangle - (\langle n \rangle)^2 = \lambda_1 + 4 \cdot \lambda_2 + 9 \cdot \lambda_3 + \ldots \tag{8.4.0.9}$$

Since the expectation values λ_i are not known, the number of observed events with a given jet multiplicity was used as an estimator. Thus, the statistical uncertainty in this analysis was calculated according to the following expression:

$$\sigma_{\text{stat}} = \sqrt{N_1 + 4 \cdot N_2 + 9 \cdot N_3 + \ldots}, \tag{8.4.0.10}$$

where N_i is the number of observed i-jet events.

The model for statistical uncertainties for E_T^{jet} and η^{jet} distributions is more complicated, because different jets in a single event can end up in different cross-section bins. Such assignment of jets results in a statistical correlation between entries in different bins. The probability distribution is still multi-Poissonian, however more possibilities for assignment of the jets to different cross-section bins have to be taken into account. Those are

---

[6] It is assumed that all jets in event are assigned to the same bin, as for example, in Q^2 distribution.

1. 1-,2-,3-jet events, ..., in which only *one* jet lies in a given bin. The corresponding event numbers and expectation values are $n_{1,1}, n_{1,2}, n_{1,3}, \ldots$ and $\lambda_{1,1}, \lambda_{1,2}, \lambda_{1,3}, \ldots$, respectively.
2. 2-,3-jet events, ..., in which exactly *two* jets fall in a given bin. The corresponding variables are denoted by $n_{2,2}, n_{2,3}, \ldots$ and $\lambda_{2,2}, \lambda_{2,3}, \ldots$, respectively.
3. events with larger jet multiplicity are treated analogously.

Taking these possibilities into account, the corresponding probability distribution is expressed as:

$$P\left(n_{1,1}, n_{1,2}, \ldots, n_{2,2}, n_{2,3}, \ldots\right) = \left(\prod_{i_1} e^{-\lambda_{1,i_1}} \cdot \frac{\lambda_1^{n_{1,i_1}}}{n_{1,i_1}!}\right) \cdot \left(\prod_{i_2} e^{-\lambda_{i_2}} \cdot \frac{\lambda_1^{n_{i_2}}}{n_{i_2}!}\right) \cdot \ldots. \tag{8.4.0.11}$$

Evaluating the mean number of jets and the variance, the following expressions can be obtained:

$$\langle n \rangle = \lambda_{1,1} + \lambda_{1,2} + \ldots + 2 \cdot \left(\lambda_{2,2} + \lambda_{2,3} + \ldots\right) + 3 \cdot \left(\lambda_{3,3} + \lambda_{3,4} + \ldots\right) + \ldots; \tag{8.4.0.12}$$

$$\left\langle n^2 \right\rangle - (\langle n \rangle)^2 = \lambda_{1,1} + \lambda_{1,2} + \ldots + 4 \cdot \left(\lambda_{2,2} + \lambda_{2,3} + \ldots\right) + 9 \cdot \left(\lambda_{3,3} + \lambda_{3,4} + \ldots\right) + \ldots. \tag{8.4.0.13}$$

An estimator of statistical uncertainty calculated from observed numbers of events is:

$$\sigma_{\text{stat}} = \sqrt{N_{1,1} + N_{1,2} + \ldots + 4 \cdot \left(N_{2,2} + N_{2,3} + \ldots\right) + 9 \cdot \left(N_{3,3} + N_{3,4} + \ldots\right) + \ldots}. \tag{8.4.0.14}$$

The estimate of covariance between numbers of jets in bins i and j can also be derived [168]:

$$\text{cov}_{ij} = N_{i,j}^{(2)} + \sum_{k \neq i,k} \left(N_{i,j,k}^{(3)} + N_{i,k,j}^{(3)} + N_{k,i,j}^{(3)}\right) + 2 \cdot \left(N_{i,j,j}^{(3)} + N_{i,i,j}^{(3)}\right) + \ldots, \tag{8.4.0.15}$$

where $N_{i,j}^{(2)}$ is the number of 2-jet events with jets in bins i and j; $N_{i,j,k}^{(3)}$ is the number of 3-jet events with jets in bins i, j and k, correspondingly. As follows from Eq. (8.4.0.15), only positive correlation between different bins is possible.

The resulting correlation values for different bins of the $d\sigma / dE_{T,B}^{jet}$ cross section in different regions of Q^2 are presented in Figure 8.6 and in Appendix in Tables A.6, A.7. As Figure 8.6 shows, the statistical correlation may reach up to 40%. Only those $E_{T,B}^{jet}$ bins belonging to the same Q^2 range have non-vanishing correlations. It is important to take correlations into account when a fit to the cross section is performed, especially for bins in the phase-space region with limited jet statistics, e.g. $Q^2 > 5000$ GeV^2 (see Chapter 9).

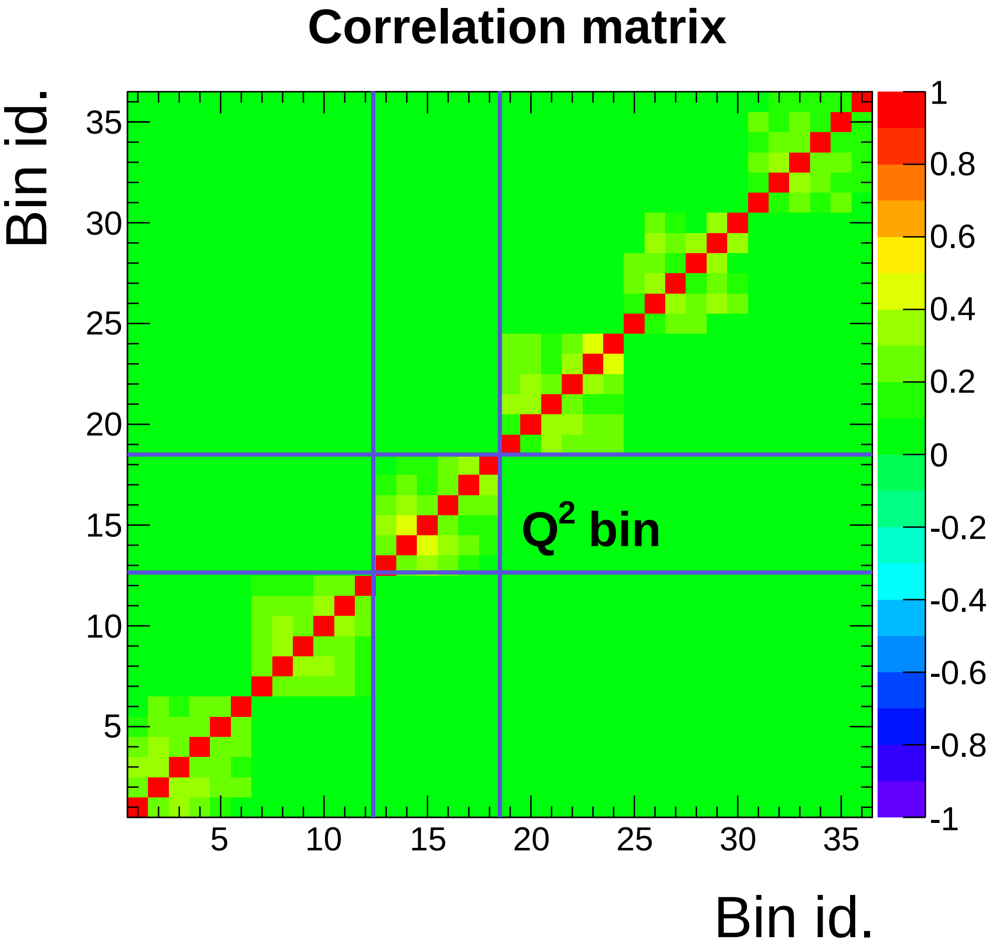


Figure 8.6.: Correlation matrix for double-differential inclusive-jet cross section as a function of $E_{T,B}^{jet}$ in different regions of Q^2

8.5. Inclusive Jets in High-Q^2 DIS

In this section the results of the inclusive-jet cross-sections measurements in neutral current deep inelastic scattering are presented. The measurements include single- and double-differential jet-production cross sections. The later were used for determination of the value of the strong coupling, $\alpha_s(M_Z)$.

The cross sections were measured in the high-Q^2 region of phase space defined by $125 < Q^2 < 20000$ GeV^2 and $0.2 < y < 0.6$. The measurements refer to each jet of hadrons reconstructed using the k_T algorithm in the longitudinally invariant inclusive mode in the Breit reference frame with transverse energies $E_{T,B}^{jet} > 8$ GeV and pseudorapidity $-1 < \eta_{lab}^{jet} < 2.5$. The total integrated luminosity of the data amounts to $\mathscr{L} = 295$ pb^{-1}.

The absolute values of cross sections, statistical and systematic uncertainties as well as values for QED and Z^0 cross sections are summarised in Appendix A.

8.5.1. Observables

The single-differential jet cross sections measured as functions of Q^2, $E_{T,B}^{jet}$ and η_{lab}^{jet} are shown in Figures 8.7, 8.8, 8.9. In addition, the inclusive-jet cross sections $d\sigma/dE_{T,B}^{jet}$ and the relative difference of the cross sections with respect to the next-to-leading-order predictions, measured in different regions of Q^2, are shown in Figures 8.10 and 8.11, respectively.

The measurements were corrected for detector effects like limited detector acceptance, resolution and efficiency as well as for higher-order QED radiation and running of the fine-structure constant, α_{em}, (see Chapter 7) and, thus, correspond to the Born-level QED process. Perturbative NLO QCD predictions calculated using the NLOJET++ program based on HERAPDF 1.5 proton PDFs sets were compared to the data. The NLOJET++ predictions refer to jets of partons, while the measurements refer to those of hadrons, therefore the theoretical predictions were corrected for hadronisation effects (see Section 8.1.1). These predictions do not include contributions from Z^0 exchange and γZ^0 interference effects, for which the calculations have been corrected as described in Section 8.1.2. The renormalisation and factorisation scales were set to $\mu_R^2 = Q^2 + E_{T,B}^{jet^2}$ and $\mu_F^2 = Q^2$ in the calculations, respectively.

The measured cross sections presented in this chapter are shown as dots in the figures. The inner error bars represent the statistical uncertainties. The statistical and systematic uncertainties not associated with the uncertainty in the absolute energy scale of the jets, added in quadrature are shown as the outer error bars. The total theoretical uncertainty and the uncertainty on the measured cross sections due to the absolute energy scale of the jets are displayed as hatched and shaded bands, respectively. The uncertainty due to jet energy scale is highly correlated across individual bins and therefore indicated separately.

8.5.2. Single-Differential Cross Sections

Inclusive-jet $d\sigma/dQ^2$ Cross Section

The inclusive-jet production cross-section $d\sigma/dQ^2$ is presented in the Figure 8.7. The cross section is a steeply falling function, decreasing more than three orders of magnitude in the measured phase space, reflecting the propagator of the exchanged boson.

The dominant sources of systematic uncertainties are those due to the jet energy scale and modelling of the parton shower. The size of the statistical uncertainty varies within the 1–3% range, depending on the value of Q^2. The total relative uncertainty on the theoretical predictions is less than 13% and typically within 5%.

Overall, the NLO QCD predictions describe the data very well in shape and normalisation within the experimental and theoretical uncertainties.

Inclusive-jet $d\sigma/dE_{T,B}^{jet}$ Cross Section

In Figure 8.8 the single-differential inclusive-jet $d\sigma/dE_{T,B}^{jet}$ cross section is presented. The cross sections decrease as the transverse energy of the jet increases, falling by more than two orders of magnitude in the measured range, complying with the dynamics of hard-scattered partons. The statistical uncertainty is less than 1% at low $E_{T,B}^{jet}$ and reaches about 4% at large values of $E_{T,B}^{jet}$. The total uncorrelated systematic uncertainty is less than 10% and typically within 5%. It is dominated by the uncertainty due to model variation for the acceptance correction. The correlated part of the systematic uncertainty amounts to 5(7)% in the regions of low (high) values of $E_{T,B}^{jet}$.

The predictions describe the measured cross sections very well. The total theoretical uncertainty, dominated by the uncertainty due to terms beyond NLO, is about 9–7% at low and high $E_{T,B}^{jet}$, respectively.

Inclusive-jet $d\sigma/\eta_B^{jet}$ Cross Section

The measured inclusive-jet single differential $d\sigma/\eta_B^{jet}$ cross section is presented in Figure 8.9. The measured range spans $-2 < \eta_B^{jet} < 2$ region. In the region $-2 < \eta_B^{jet} < 1$ the cross section rapidly increases by more than one order of magnitude, reaching a maximum value at $\eta_B^{jet} \approx 0.75$. At higher values

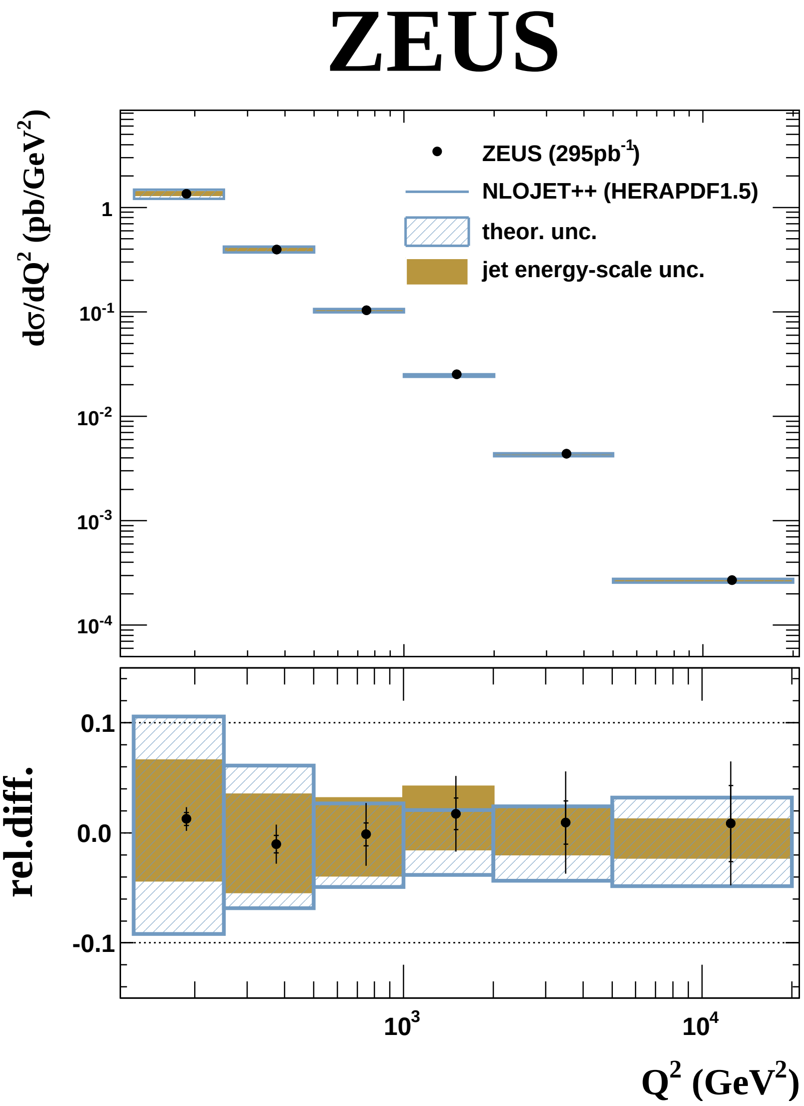


Figure 8.7.: The measured single-differential cross-section $d\sigma/dQ^2$ for inclusive-jet production in NC DIS for jets reconstructed in the Breit frame with $E^{jet}_{T,B} > 8$ GeV and $-1 < \eta^{jet}_{lab} < 2.5$, in the kinematic region given by $0.2 < y < 0.6$ and $125 < Q^2 < 20000$ GeV. The data are shown as dots with inner error bars representing the statistical uncertainties; the outer error bars show the statistical and systematic uncertainty not associated with the uncertainty in the absolute energy scale of the jets, added in quadrature; the shaded band displays the uncertainty due to the absolute energy scale of the jets. In some bins, the error bars on the data points are smaller than the marker size and are therefore not visible. The NLO QCD calculation, corrected for hadronisation effects and Z^0 exchange and using the HERAPDF 1.5 parametrisations of the proton PDFs, is also shown as a solid line and the hatched band displays the total theoretical uncertainty. The lower part of the figure shows the relative difference between the measured $d\sigma/dQ^2$ and the NLO QCD calculations.

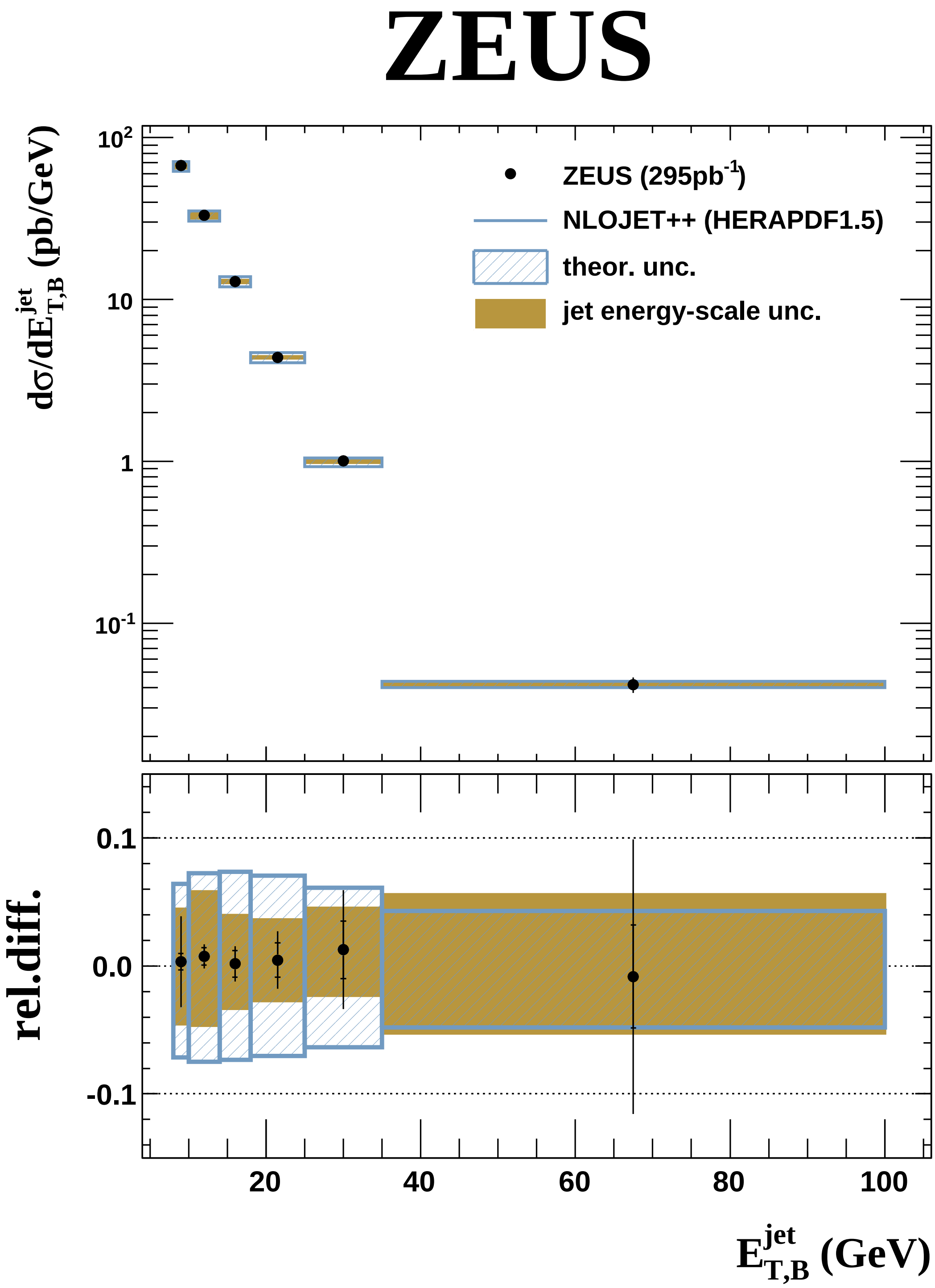


Figure 8.8.: The measured single-differential cross-section $d\sigma/dE^{jet}_{T,B}$. The other features are defined in the caption to Figure 8.7

of the jet pseudorapidity the cross section decreases. Such a behaviour is consistent with the hard scattering dynamics of partons. The cross sections have small relative statistical uncertainty, which is about 2.5%–1% in the measured range. The uncertainty due to modelling of the parton shower reaches 10% in the low η_B^{jet} region, where MC statistics was limited and less than 1% in the $-0.25 < \eta_B^{jet} < 2$ region. The pQCD prediction describes the measurements very well in shape and normalisation within the theoretical and experimental errors.

8.5.3. Double-Differential Cross Sections

In order to investigate inclusive-jet production in more detail, single-differential cross sections as functions of $E_{T,B}^{jet}$ were measured in different regions of Q^2. The obtained cross sections differ from the double-differential measurements by a constant factor equal to the Q^2 bin width. Figures 8.10, 8.11 show the measured cross sections and the relative difference with respect to theoretical predictions. The cross sections exhibit the same features as the single-differential cross-section $d\sigma/dE_{T,B}^{jet}$, however, the transverse energy spectrum becomes harder as the virtuality of exchange-boson increases.

In Figure 8.11, the NLO pQCD predictions based on different proton PDF parametrisations MSTW08 [169] and CT10 [170] are also shown. The predictions based on alternative proton PDF sets are typically above those obtained with HERAPDF1.5, except for calculations based on MSTW08 for $8 < E_{T,B}^{jet} < 25$ GeV and $Q^2 < 2000$ GeV^2. Moreover, the difference between the predictions calculated using MSTW08 and CT10 sets and nominal pQCD predictions exceeds the size of the total theoretical uncertainty and is about three times larger than the PDF uncertainty, especially in the high-$E_{T,B}^{jet}$ and high-Q^2 region, indicating high sensitivity of the jet data to the proton PDFs.

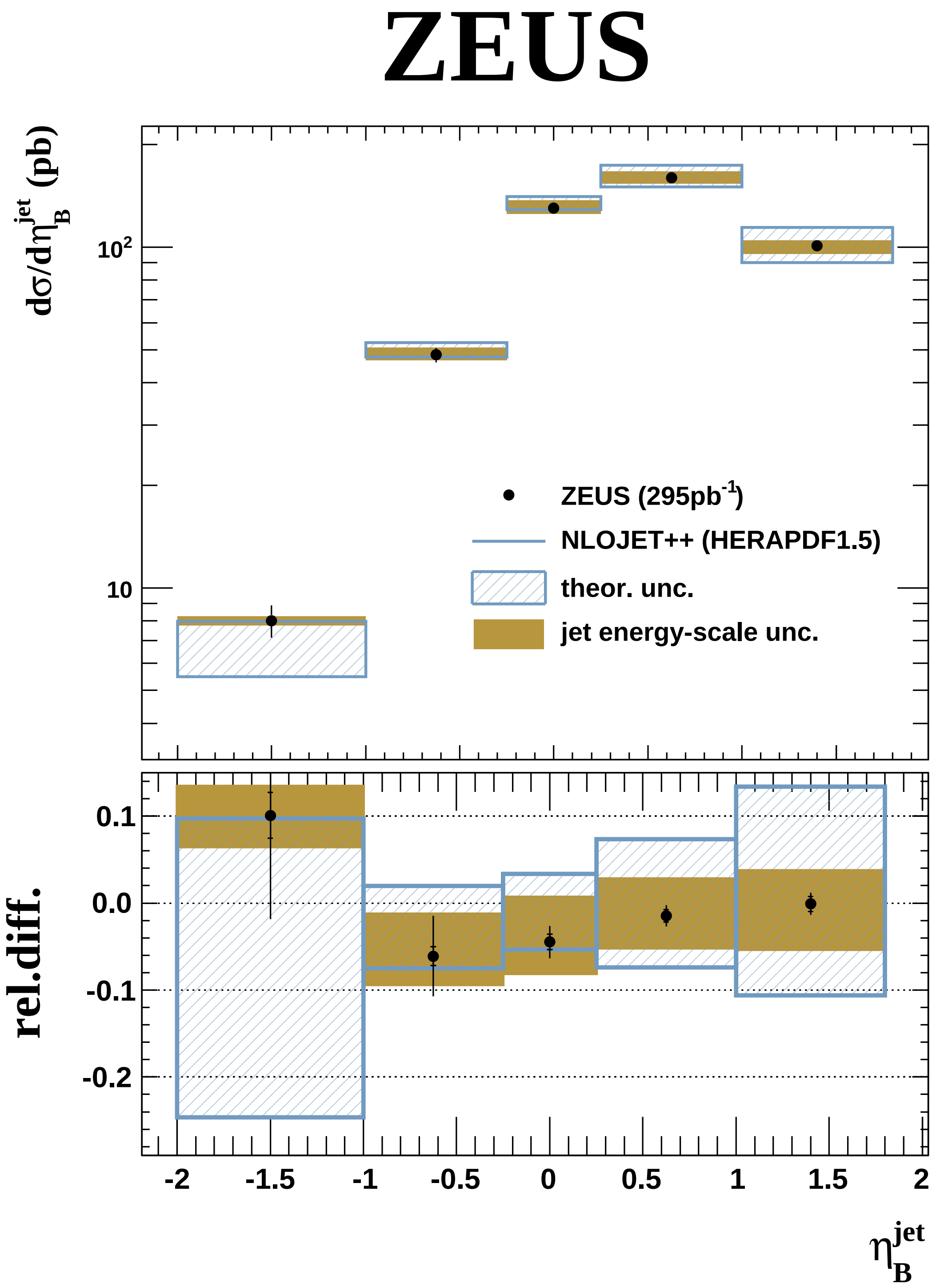


Figure 8.9.: The measured single-differential cross-section $d\sigma/d\eta_B^{jet}$. The other features are defined in the caption to Figure 8.7

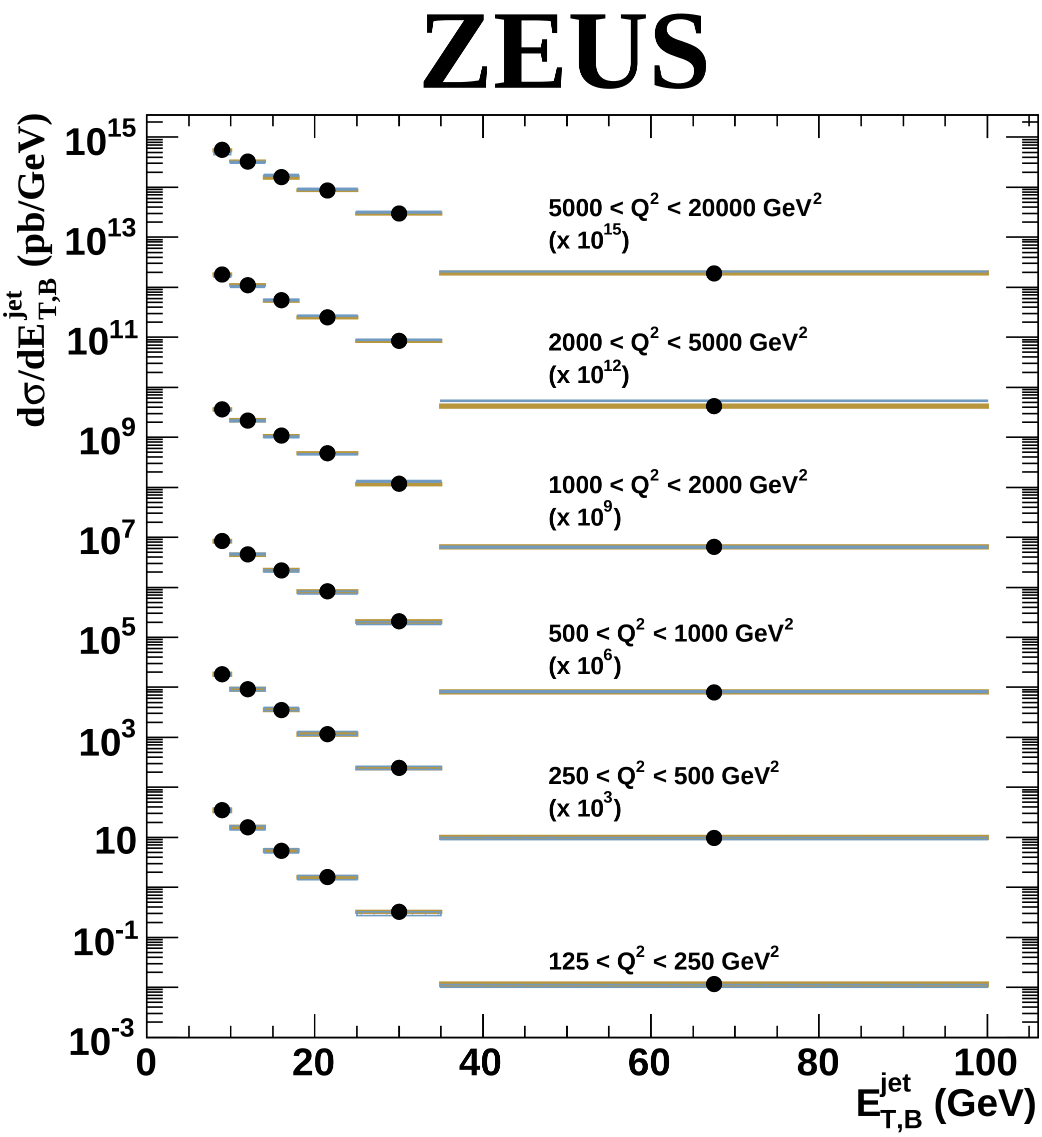


Figure 8.10.: The measured differential cross-sections $d\sigma/dE^{jet}_{T,B}$ in different regions of Q^2. To aid visibility, the series of measurements are shown with vertical offset defined by a multiplicative factor indicated in brackets. The other features are defined in the caption to Figure 8.7

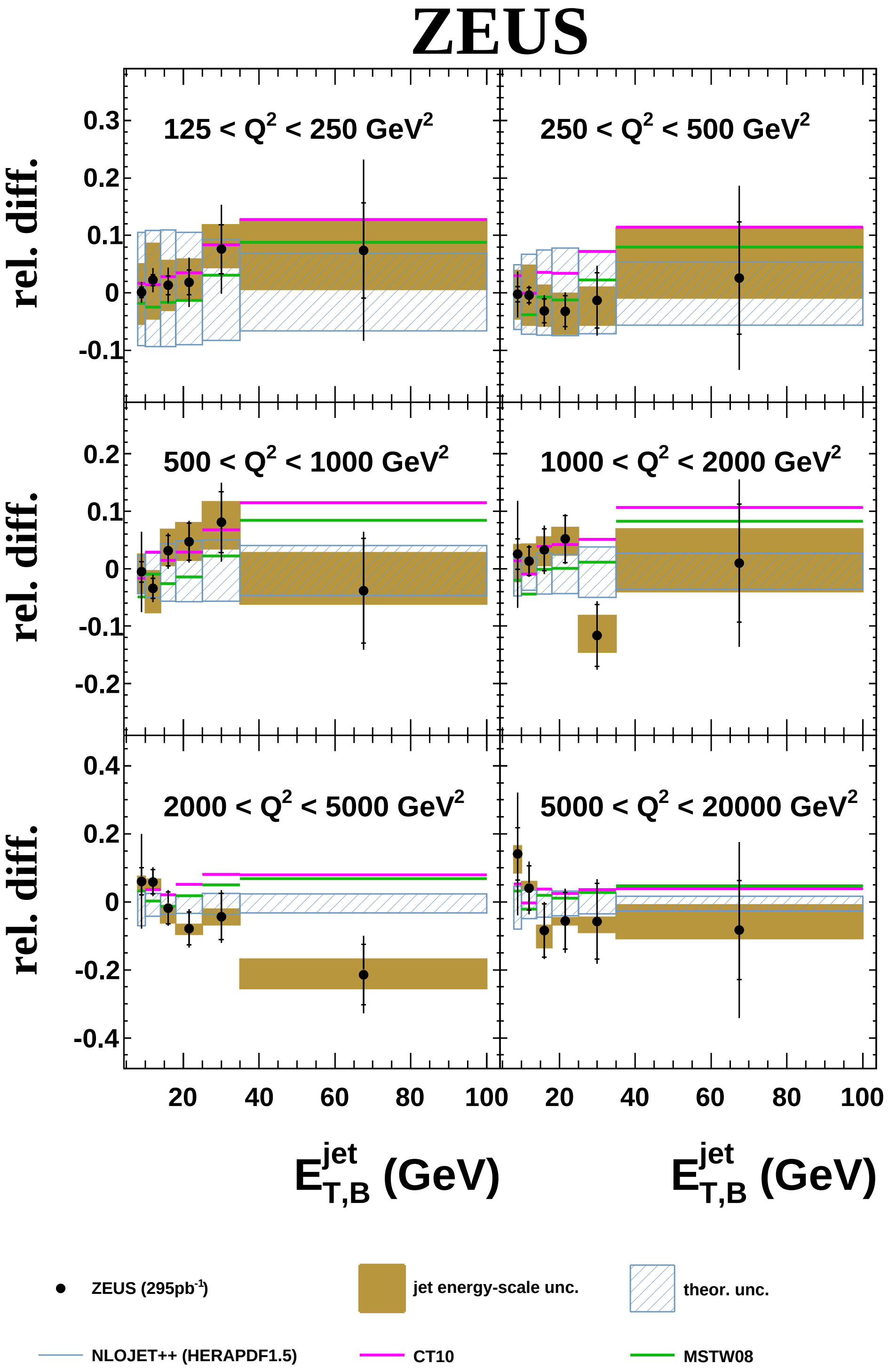


Figure 8.11.: Relative differences between the measured differential cross-sections $d\sigma/dE_{T,B}^{jet}$ presented in Figure 8.10 and the NLO QCD calculations preformed using HERAPDF1.5 pPDF parametrisations and those using MSTW08 (CT10). The other features are defined in the caption to Figure 8.7

CHAPTER 9

Results. QCD Analysis of Inclusive-Jet Data

This chapter summarises the results of the determination of the strong coupling constant at various scales from the measured inclusive-jet cross sections At the beginning, the theoretical framework of the analysis is outlined. Then the least-squares method and the notion of the 'QCD fit' are explained and the results of the $\alpha_s(M_Z)$ determination from the single- and double-differential inclusive-jet cross sections and a test of the running of the strong coupling are presented. Afterwards, the investigation of the impact of experimental and theoretical uncertainties on the extracted $\alpha_s(M_Z)$ values is described and the influence of the assumptions on proton PDFs set is emphasised.

9.1. Introduction

As demonstrated in Chapter 8, the measured inclusive-jet cross sections are very well described by the predictions based on the proton PDFs and the $\alpha_s(M_Z)$-value extracted in global fits to the inclusive DIS data from HERA. Assuming the validity of NLO pQCD predictions for inclusive-jet production in NC DIS, the accurate jet data can be exploited to constrain values of input parameters used in the calculations.

According to the factorisation ansatz, explained in Chapter 2, the perturbative QCD predictions for inclusive-jet cross sections can be expressed as a convolution of the hard-scattering matrix elements and the proton PDFs. The parton level NLO predictions, e.g. for $d\sigma/dE_{T,B}^{jet}$, have the following formal rep-

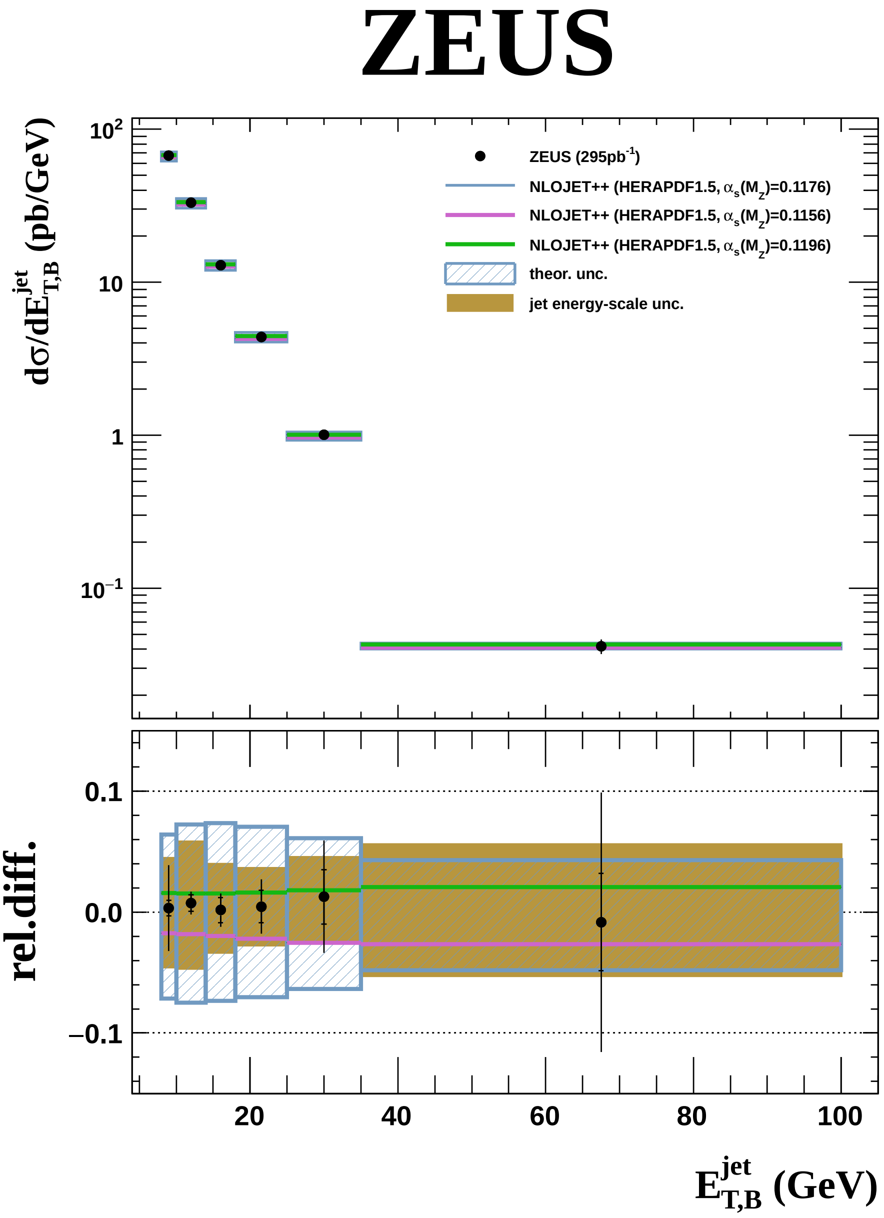


Figure 9.1.: Comparison of the measured inclusive-jet cross sections to NLO pQCD predictions calculated assuming different values of $\alpha_s(M_Z)$.

resentation:

$$\begin{aligned}\sigma\left(\mu_R,\mu_F\right)=\alpha_s{}^1\left(\mu_R\right)\cdot&\left[c_{1,i}\left(\xi,\mu_R,\mu_F\right)\overset{\xi}{\bigotimes}F_i\left(\xi,\mu_f,\mu_R\right)\right]\\+\alpha_s{}^2\left(\mu_R\right)\cdot&\left[c_{2,i}\left(\xi,\mu_R,\mu_F\right)\overset{\xi}{\bigotimes}F_i\left(\xi,\mu_f,\mu_R\right)\right],\end{aligned}\tag{9.1.0.1}$$

where $c_{n,i}$ are the perturbative coefficients for the jet-production subprocess i, e.g. QCD-Compton scattering or boson-gluon fusion at LO, and $F_i\left(\xi,\mu_f,\mu_R\right)$ are the corresponding linear combinations of the proton PDFs depending on the longitudinal fraction of the proton momentum, ξ, renormalisation and factorisation scales μ_R,μ_F and implicitly (through the DGLAP evolution equations) α_s. The convolution symbol $\overset{\xi}{\otimes}$ corresponds to the ξ-integration over the available phase space interval. Thus, jet rates in DIS are sensitive to the value of the strong coupling and the proton PDFs. The comparison of the predictions based on different proton PDFs set to the measured cross sections was presented in Chapter 8; the predictions calculated assuming different values of $\alpha_s\left(M_Z\right)$ in the matrix elements and PDFs[1] are illustrated in Figure 9.1. It can be seen that the calculations with lower (higher) value of the strong coupling provide lower (higher) jet cross sections. The inclusive-jet measurements obtained in this thesis can thus be utilised to constrain α_s and the proton PDF distributions. The details of this procedure, also called the *'QCD fit'*, are described below.

9.2. QCD Fit Setup

In the least-squares method [71], the optimal values of parameters providing the best agreement between theoretical predictions and data are estimated by minimising sums of weighted squared differences between measurements and calculations. Given theoretical predictions, $t_i\left(\mathbf{a}\right)$, for the cross-section bin i, that depend on the vector of parameters of interest **a**, the outcome of the measurement, m_i, can be represented as follows [171, 172]:

$$m_i=t_i\left(\mathbf{a}\right)+w_i\sigma_{i,\text{stat}}+\varepsilon_i\sigma_{i,\text{uncor}}+\sum_{\mu=1}^{K}b_\mu\sigma^\mu_{i,\text{cor}},\qquad i=1\ldots N_{\text{bins}},\tag{9.2.0.2}$$

where $\sigma_{i,\text{uncor}}$ and $\sigma_{i,\text{stat}}$ are estimate; of the total uncorrelated statistical and systematic uncertainties, respectively; ε_i and w_i are random normally distributed variables with mean values $\langle\varepsilon_i\rangle=\langle w_i\rangle=0$ and variance $\text{Var}\left(\varepsilon_i\right)=$

---

[1] In the HERAPDF 1.5 analysis alternative sets of PDFs were determined for different values of $\alpha_s\left(M_Z\right)$ assumed in the fitting procedure.

$\mathrm{Var}(w_i) = 1$; $\sigma^{\mu}_{i,\mathrm{cor}}$ and b_μ represent the correlated uncertainty component from source μ and the corresponding systematic shift[2], respectively. It is assumed here that the systematic shifts b_μ can also be modelled by normally distributed random variables with zero mean and unit variance. In order to take into account the proportionality of the systematic uncertainties to the central value of the measurement[3], the corresponding uncertainty components are expressed as $\sigma^{\mu}_{i,\mathrm{cor}} = \gamma_{i,\mu} t_i$, where

$$\gamma_{i,\mu} = \frac{1}{t_i}\frac{\partial t_i}{\partial b_\mu} \tag{9.2.0.3}$$

is the relative correlated systematic uncertainty quantifying the sensitivity of the measurement i to the systematic source μ. However, in practice, the relative systematic is replaced with the following approximation[4]:

$$\gamma_{i,\mu} = \frac{1}{m_i}\frac{\partial m_i}{\partial b_\mu}. \tag{9.2.0.4}$$

The statistical uncertainty $\sigma_{i,\mathrm{stat}}$ is assumed to scale with the square root of expected number of events. Correcting for the bias from systematic shifts, the statistical uncertainty is equal to

$$\sigma^2_{\mathrm{stat},i} = \delta^2_{\mathrm{stat},i} \cdot m_i \left(t_i - \sum_\mu \gamma_{i,\mu} t_i b_\mu \right), \tag{9.2.0.5}$$

where $\delta_{\mathrm{stat},i}$ is the relative statistical uncertainty determined from the observed number of events in bin i. A least-squares method is used for the QCD fit; the so-called χ^2-function is defined as [173]:

$$\chi^2 = \sum_i \frac{\left(m_i - t_i - \sum_\mu \gamma_{i,\mu} t_i b_\mu\right)^2}{\delta^2_{\mathrm{stat},i} m_i \left(t_i - \sum_\mu \gamma_{i,\mu} t_i b_\mu\right) + \left(\delta_{i,\mathrm{uncor}} t_i\right)^2} + \sum_\mu b^2_\mu, \tag{9.2.0.6}$$

with $\delta_{i,\mathrm{uncor}}$ denoting the relative uncorrelated systematic uncertainty of measurement i. The last term in this equation represents the penalty for the systematic-shift parameters. In order to take into account statistical correlations

---

[2] Parameters b_μ are also called “nuisance” parameters.

[3] Most of the systematic uncertainties in this analysis e.g. absolute normalisation uncertainty or acceptance correction uncertainty are to good approximation proportional to the central value of the measured cross section. This means that the fractional uncertainties are constant and the absolute uncertainties scale with the measurement value.

[4] This approximation is exact for the luminosity uncertainty. In case of the jet-energy-scale uncertainty the dependence of the absolute normalisation cancels out in the ratio, however the residual dependence on the shape of the cross section or the jet migrations might still be present.

between different data points the χ^2-function can be re-expressed as follows[5]:

$$\chi^2 = \sum_{ij}\left(m^i - \sum_{\mu}\Gamma^i_{\mu} b_{\mu} - t^i\right) C^{-1}_{ij}\left(m^j - \sum_{\mu}\Gamma^j_{\mu} b_{\mu} - t^j\right) + \sum_{\mu} b^2_{\mu}, \tag{9.2.0.7}$$

where C_{ij} is the element i, j of the covariance matrix for the measured data points. It is equal to the sum of covariance matrices for statistical and uncorrelated systematic-uncertainty components $C_{ij} = C_{ij}(t_i) = C^{\text{stat.}}_{ij} + C^{\text{uncor.}}_{ij}$. The covariance matrix for the uncorrelated uncertainties has a diagonal form. The quantities Γ^i_{μ} are related to the relative correlated systematic uncertainties

$$\Gamma^i_{\mu} = \gamma^i_{\mu} \cdot t^i. \tag{9.2.0.8}$$

A more elaborate description of the χ^2-function definition can be found in [174].

As was mentioned earlier, some systematic variations in this analysis resulted in asymmetric changes of the cross-section values, therefore some measurements have asymmetric uncertainties. In order to conform with the definition of the χ^2-function, the asymmetric uncertainties are converted into a symmetric form according to:

$$\sigma^{sym}_i = \frac{\sigma^+_i + \sigma^-_i}{2}, \tag{9.2.0.9}$$

with $\sigma^{+(-)}_i$ expressing the positive (negative) component of the asymmetric uncertainty[6]. This procedure is applied to correlated as well as uncorrelated components of the systematic error.

The procedure for determination of the parameters that provide the best fit to the data is based on the numerical minimisation of the χ^2 function with respect to the fit parameters. The systematic shifts b_{μ}, described above, are fitted simultaneously with the parameters **a**.

The quality of the fit can be assessed by the value of the χ^2-function at its minimum. Assuming that the measurements are sampled according to a normal distribution around the expectation value provided by the theoretical predictions and with a variance described by the covariance matrix, the χ^2

---

[5] This χ^2 definition is a combination of different representations documented in [174]. In principle, the absence of bias, possibly introduced in this approach, has to be validated in a dedicated study.

[6] This approach is ill-defined if same-sign variation of the measured cross sections is encountered. However, this was not the case in this thesis.

value can be treated as a random variable expected to follow approximately[7] a χ^2-distribution [12] with mean value equal to the number of degrees of freedom $\langle \chi^2 \rangle = n_{\mathrm{DF}}$, which in case of an n_{par} parameters fit to n_{p} points is $n_{\mathrm{DF}} = n_{\mathrm{p}} - n_{\mathrm{par}}$. Therefore, ratios $\chi^2 / n_{\mathrm{DF}} \approx 1$ indicate a good quality of the fit while larger values can indicate some inconsistency.

9.2.1. Parameters of the QCD Fit

In the QCD fit, the predictions **t** are usually treated as functions of the proton PDF parameters, **a**, and the strong coupling constant, $\alpha_s(M_Z)$. In the kinematic phase space of HERA the following parton species contribute effectively to the PDFs: the gluon (g), and five quark flavours: up (u), down (d), strange (s), charm (c), bottom (b) and corresponding anti-quarks; because of the large mass of the top quark its contribution can be neglected. The PDFs are functions of the parton longitudinal-momentum fraction, x. In principle, their shape is a calculable quantity in QCD. However, due to the non-perturbative nature of the PDFs, only predictions for a limited number of first moments

$$M_n = \int \mathrm{d}x\, x^n f(x), \qquad (n \geq 1) \tag{9.2.1.1}$$

are currently available [175], and the most precise extraction of the proton PDFs is achieved in a QCD analysis of experimental data. The usual ansatz is to describe the PDFs at a starting scale Q_0 by a suitable parametrisation. Then PDFs can be evolved to any scale by using the DGLAP equations [60–63]. Some assumptions concerning the shape of the distributions at the starting scale have to be made. In the HERAPDF approach [173], the generic functional form of the parton distribution functions is chosen to be:

$$xf(x) = A_i x^{B_i} (1-x)^{C_i} \left(1 + \epsilon_i \sqrt{x} + D_i x + E_i x^2\right). \tag{9.2.1.2}$$

In general, variables sets $(A_i, B_i, C_i, D_i, E_i, \epsilon_i) \subset \mathbf{a}$ for each individual parton type have to be considered as independent free parameters in the QCD fit, although the normalisation parameters A_i are constrained by the quark number and momentum sum rules [176], i.e. quantum numbers of total intrinsic strangeness, charm and beauty must vanish $S|uud\rangle = C|uud\rangle = B|uud\rangle = 0$ and the total parton momentum has to sum up to the proton momentum. Including $\alpha_s(M_Z)$ into the fit and applying the aforementioned constraints results in

$$2\left(\text{quarks/anti-quarks}\right) \cdot 6\left(g, u, d, s, c, b\right) \cdot 6 + 1\left(\alpha_s(M_Z)\right) - n_{\mathrm{constr}} > 60 \tag{9.2.1.3}$$

---

[7] In principle, the statistical uncertainty of the measurements follows Poisson statistics, nevertheless the normal approximation can be used, since no bins with less than 100 counts were observed.

free parameters in the fit. In addition, the contribution of different parton species cannot be distinguished in NC DIS processes, e.g. apart from their different masses, all u-type quarks behave similarly in hard ep-scattering. Usually, in the predictions for high-energy reactions, the masses of the light and sometimes even heavy quarks are ignored, resulting in identical predictions for every quark species. This leads to an ill-defined problem that requires further assumptions about the parameter space to be introduced.

Based on the above symmetry arguments for u- and d-type quarks, the number of free parameters can be reduced. For example, instead of fitting individual parton distributions for each parton type, linear combinations of PDFs can be introduced. In HERAPDF 1.0 [173], the following PDF components are treated independently:

$$xg\,(x) = A_g x^{B_g}\,(1-x)^{C_g}\,, \tag{9.2.1.4}$$

$$xu_v\,(x) = A_{u_v} x^{B_{u_v}}\,(1-x)^{C_{u_v}}\left(1+E_{u_v}x^2\right), \tag{9.2.1.5}$$

$$xd_v\,(x) = A_{d_v} x^{B_{d_v}}\,(1-x)^{C_{d_v}}\,, \tag{9.2.1.6}$$

$$x\bar{U}\,(x) = A_{\bar{U}} x^{B_{\bar{U}}}\,(1-x)^{C_{\bar{U}}}\,, \tag{9.2.1.7}$$

$$x\bar{D}\,(x) = A_{\bar{D}} x^{B_{\bar{D}}}\,(1-x)^{C_{\bar{D}}}\,, \tag{9.2.1.8}$$

where xg corresponds to the gluon, $xu\,(d)_v$ represents $u\,(d)$-valence quark distributions and $x\bar{U}\,(x)$ and $x\bar{D}\,(x)$ are u-type and d-type anti-quark distributions, respectively. At the initial scale Q_0^2, the anti-quark sea distributions are $x\bar{U}\,(x) = x\bar{u}$ and $x\bar{D}\,(x) = x\bar{d} + x\bar{s}$. Apart from parameters in Eqs.(9.2.1.4)–(9.2.1.8), all other parameters in the generic parametrisation Eq. (9.2.1.2) are set to zero by default[8] and released one at a time, when the sensitivity of the fit to the parametrisation ansatz is investigated. Further assumptions imposed on the fit parameters are:

- at the initial scale the strange sea is expressed as a constant fraction of the d-type sea $x\bar{s} = f_s x\bar{D}$ with $f_s = 0.31$ [169, 177]; the contribution from the strange quarks is assumed to be suppressed due to the larger s-quark mass;

- in order to conform with the isospin symmetry hypothesis and ensure $x\bar{u} \to x\bar{d}$ as $x \to 0$, the constraints $A_{\bar{U}} = \left(1-f_s\right)A_{\bar{D}}$ and $B_{\bar{U}} = B_{\bar{D}}$ are imposed;

- $B_{u_v} = B_{d_v}$ is assumed for the central fit, however this requirement is omitted in the study of systematic effects.

---

[8] Such a combination is called ‘central-fit‘.

These restrictions leave 10 or 11 free parameters for the proton PDFs, depending on the chosen fit configuration. The strong coupling constant, $\alpha_s(M_Z)$, can be added as an additional parameter. The measured jet data alone cannot constrain these parameter sets because jet production populates a restricted region of phase space (mostly at medium and high ξ), moreover certain types of initial-states (i.e. incoming partons) are indistinguishable in NC jet reactions, since the jet algorithm used is insensitive to the parton flavour. In HERAPDF 1.0 the combined inclusive DIS data from H1 and ZEUS were used for the proton PDF determination. In the following section, these additional data sets and the possible fit strategies are discussed.

9.2.2. Additional Data Set: Inclusive DIS Cross Sections

The inclusive DIS double-differential reduced cross-sections $\mathrm{d}\tilde{\sigma}^2/\mathrm{d}x\mathrm{d}Q^2$ measured at HERA constitute the core of all modern proton parton-density extractions [169, 173, 178–180]. The combined dataset [173] used in all these analyses comprises measurements of inclusive NC and CC DIS cross sections by the H1 and ZEUS collaborations. A summary of the data samples is provided in [173]. The kinematic phase space of the combined NC DIS cross sections is $6\cdot 10^{-7} \leq x \leq 0.65$ and $0.045 \leq Q^2 \leq 30000$ GeV^2 for values of inelasticity in the range $0.005 \leq y \leq 0.95$. The kinematic range of the CC DIS data is $1.3\cdot 10^{-2} \leq x \leq 0.40$ and $300 \leq Q^2 \leq 30000$ GeV^2 for y in the region $0.037 \leq y \leq 0.76$. It is important to note that the jet measurements provided in this work are based on a statistically independent data sample.

In practice, datasets for different reactions have different sensitivity to various PDF components. For example, NC and CC DIS cross sections are very sensitive to the quark content of the proton, because electroweak bosons couple to the quarks. In contrast, gluons do not interact directly with γ, Z^0 or $W^{\pm}$ but via $g \to q\bar{q}$ processes with one of the quarks coupling to the electroweak boson. Therefore, in the predictions for inclusive ep scattering, the gluon density is always accompanied by a factor α_s, which results in a strong correlation between the strong coupling and the gluon parton density in QCD fits to inclusive DIS data. An approximation [181] for the scaling violation of the structure function $F_2\left(x, Q^2\right)$ illustrates this property:

$$\frac{\partial F_2\left(x, Q^2\right)}{\partial \ln Q^2} \approx \sum_i^{n_f} e_{q_i}^2 \frac{\alpha_s}{\pi} xg(2x) \int_0^1 P_{q\leftarrow g}(z)\,\mathrm{d}z, \tag{9.2.2.1}$$

where $P_{q\leftarrow g}(z)$ is the $g \to q\bar{q}$ splitting function. An independent source of information constraining the magnitude of the strong coupling or the gluon density is necessary for an unbiased determination of both quantities. It was demonstrated in earlier studies [182, 183] that including jet data into the fit mitigates the problem of large correlations between α_s and $xg(x)$ and significantly improves the precision of a simultaneous $xg\,(x)$ and α_s extraction. Figure 9.2 shows Feynman diagrams for different powers of the strong coupling constant, illustrating the processes with quarks and gluons in the initial state. For example, the gluon density $xg\,(x)$ provides the dominant contribution[9] to the NC DIS ep-scattering cross section at low and medium x at high Q^2, although it appears for the first time at NLO, while jet production in the Breit frame (BF) is already sensitive to $xg\,(x)$ at leading order. The sum over all quark and anti-quark densities weighted with respective electric charges $x\Delta\,(x) = x\sum_a e_a^2\left(q_a\,(x) + \bar{q}\,(x)\right)$ is the next most important contribution to inclusive DIS and jet processes and appears in all orders of perturbation theory.

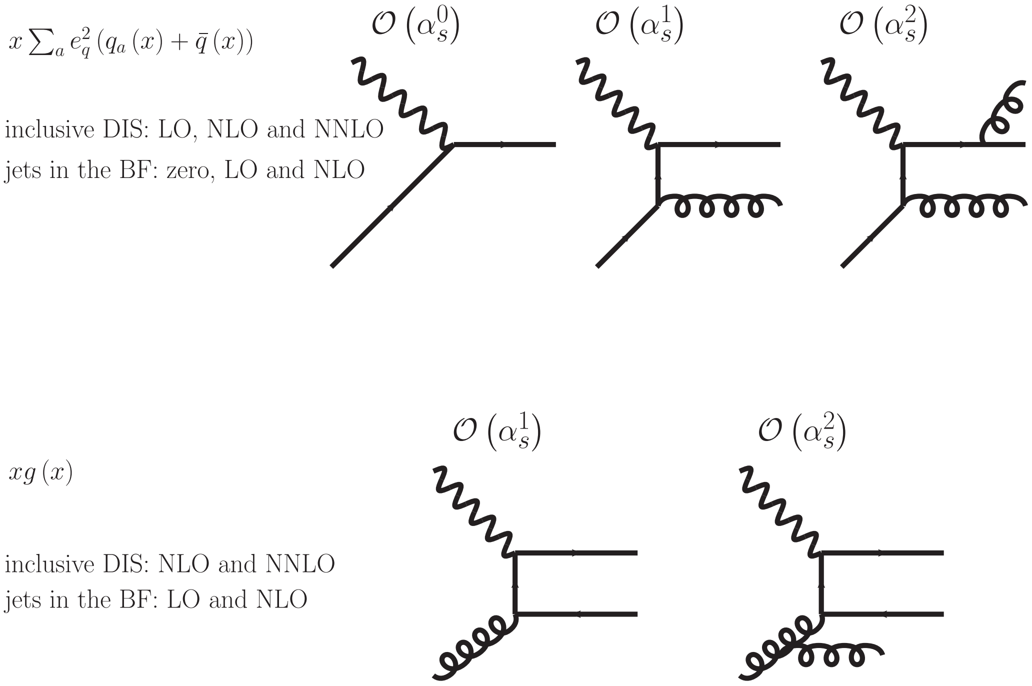


Figure 9.2.: Example Feynman diagrams illustrating the interplay of different PDF components and α_s for various DIS processes.

---

[9] Due to infrared divergences, the gluon splitting probability $\mathscr{P}_{gg}$ tends to infinity as $E_g \to 0$. This leads to a rapid proliferation of the number of soft gluons constituting the proton at low longitudinal momentum x and, in principle, leads to violation of unitarity. However, at large gluon densities, non-perturbative recombination processes have to be taken into account and unitarity is restored.

9.2.3. Fit Strategies

The simplest variant of the fit comprises the determination of the strong coupling $\alpha_s(M_Z)$ from inclusive-jet data only, assuming a fixed PDF parametrisation. A more complex strategy includes simultaneous determination of the proton PDFs and $\alpha_s(M_Z)$ from inclusive DIS and inclusive-jet data. The different fit options are discussed below.

Determination of $\alpha_s(M_Z)$
In this approach the PDFs are assumed to be fixed and the theoretical predictions for inclusive jet production are parametrised as a function of a single parameter $\alpha_s(M_Z)$. Similar analyses have been performed in a variety of environments. Examples are recent studies in ep [48, 184], pp [185, 186] and $p\bar{p}$ [187–189] collisions. This approach, however, has the following limitations:

- the dominant contribution to the jet cross sections measured in this analysis comes from the medium to high-ξ region $0.003 < \xi < 0.4$ of the phase space. In this region the knowledge of PDFs from HERA inclusive data alone is limited by the statistical precision of the measurements;

- typically, in the global PDF analyses the value of $\alpha_s(M_Z)$ is an external parameter. Thus a particular assumption of $\alpha_s(M_Z)$ may bias the extraction of the strong coupling in the α_s fit, when the proton PDFs are fixed and treated as independent quantities. For consistent determinations of the strong coupling, the value assumed for the PDF and that extracted in the fit must coincide. Recent global analyses provide collections of proton PDF parametrisations extracted assuming different values of $\alpha_s(M_Z)$. This information can be used to investigate the influence of the $\alpha_s(M_Z)$ assumption in the used PDF on the α_s results obtained in the fit and to check the consistency of the fitting procedure;

- typically, in the global PDF analyses the value of $\alpha_s(M_Z)$ is an external parameter. Thus a particular assumption of $\alpha_s(M_Z)$ may bias the extraction of the strong coupling in the α_s fit, when the proton PDFs are fixed and treated as independent quantities. For consistent determinations of the strong coupling, the value assumed for the PDF and that extracted in the fit must coincide. Recent global analyses provide collections of proton PDF parametrisations extracted assuming different values of $\alpha_s(M_Z)$. This information can be used to investigate the influence of the $\alpha_s(M_Z)$ assumption in the used PDF on the α_s results obtained in the fit and to check the consistency of the fitting procedure;

- nearly all PDF fitting groups provide an error analysis for their PDF sets. Nevertheless, differences between predictions based on different PDFs are often larger than the uncertainties suggested by the error analyses (see Section 9.3.4.2). Different prescriptions for the estimation of uncertainties attributed to proton PDFs exist. One of them, proposed by the PDF4LHC group [190], consists in taking an envelope of the predictions based on different PDF sets.

The advantage of the described variant of the fit is its simplicity. The influence of different assumptions can be easily studied because every change in the input manifests itself in the extracted $\alpha_s(M_Z)$ value.

Simultaneous determination of $\alpha_s(M_Z)$ and PDFs
The aim of this approach is to fit $\alpha_s(M_Z)$ and gluon and quark densities simultaneously. For this purpose inclusive-jet measurements are fitted together with inclusive DIS data. Such a treatment takes into account correlations between α_s and PDFs and leads to unbiased determination of both quantities.

To test various fitting options, the HERAFitter program [173, 191] was used in this thesis. The main features of the fitting code are briefly outlined in the next section.

9.2.4. HERAFitter Package and FastNLO Framework

The HERAFitter program [173, 191] is a tool developed for least-square fits of the proton PDFs. Nowadays, it provides an efficient code for various QCD analyses of high-energy collider data. In this analysis the HERAFitter program package was used for extraction of α_s as well as for simultaneous α_s + PDF fits.

The minimisation of the χ^2-function performed employing the MIGRAD algorithm from the MINUIT package [192]. The minimisation procedure requires repeated recalculation of the χ^2-function. Hence, the NLO pQCD predictions for the jet observables, when included in the fit, have to be calculated repeatedly in each iteration of the QCD fit. However, evaluation of the predictions for the jet production in an iterative procedure using the techniques described in Section 8.1 has two main limitations:

- due to the stochastic nature of the Monte Carlo approach for calculation of the phase-space integrals in the NLO calculation, the χ^2 minimisation procedure, in general, may not converge, because of fluctuations of the cross-section predictions;

- the computing time for a single iteration (i.e. evaluation of the χ^2 function for given fit parameter values) is prohibitively long.

Therefore, for the iterative fit procedure, the FastNLO framework is used for the recalculation of the NLO predictions for jet production.

The FastNLO [193–196] approach is a based on the idea of calculating the LO and NLO weights, $\mathscr{C}_{n,a,i,j,k}$, which are independent of α_s and the PDFs, on a 3d-grid in the convolution variable ξ_i and the factorisation and renormalisation scales $\mu_{F,j}, \mu_{R,k}$; index a runs over the number of contributing processes. The cross section can be obtained from a simple sum:

$$\sigma = \sum_{n,a,i,j,k} \alpha_s{}^n \left(\mu_{R,k}\right) F_a\left(\xi_i, \mu_{F,j}\right) \mathscr{C}_{n,a,i,j,k}, \tag{9.2.4.1}$$

where the magnitude of the strong coupling and PDFs can be chosen independently, without repeating the time-consuming MC integration. Since certain parton configurations lead to identical final-states, only weights for the linear combinations of PDFs $a = \left(\Delta, \Sigma, g\right)$ are stored, where $\Delta = \sum_a e_q^2 \left(q_a(\xi) + \bar{q}_a(\xi)\right)$, $\Sigma = \sum_a q_a(\xi) + \bar{q}_a(\xi)$ and $g = g(\xi)$ is the gluon PDF. The weights $\mathscr{C}_{n,a,i,j,k}$ are organised in form of a table and can be efficiently processed leading to jet cross-section calculation times $\mathscr{O}(ms)$ on a standard CPU[10], which is practical for use in the fitting procedure.

The predictions for the inclusive-jet cross sections were calculated using the NLOJET++ program with the FastNLO interface. The settings for the NLO pQCD calculations were described in detail in Section 8.1 and are briefly summarised in Table 9.1. Unless otherwise stated, these settings define the so-called "central-fit". The value of the strong coupling determined with this setup is used as a reference for studies of the sensitivity of the extracted $\alpha_s(M_Z)$ to the variation of the theory parameters discussed in Section 9.3.4.

Parameter	Default setup
proton PDF set	HERAPDF 1.5 (NLO)
renormalisation scale	$\mu_R^2 = Q^2 + E_{T,B}^{jet^2}$
factorisation scale	$\mu_F^2 = Q^2$
number of active flavours	$n_f = 5$

Table 9.1.: Summary of the theory settings used for the calculations for the α_s determination.

Within the HERAFitter program, the evolution of the parton densities according to the DGLAP-equations is performed using the QCDNUM code [72]. In the QCDNUM approach the equations are solved numerically on a $n \times m$ lattice in the x and $\mu = \mu_F$ variables. The method is based on a numerical spline interpolation of the parton densities on an equally spaced grid in $\ln(x)$

---

[10] By 'standard CPU' a single core 1.5GHz CPU is meant.

and in $\ln(\mu^2)$[11] . Thus, the convolution (see Eqn. (2.2.3.2)) is represented as a finite sum and the $t \equiv \ln \mu^2$ integration is formulated as an iterative process. The analytic form of the DGLAP-equations and a specific choice of the type of interpolation were used to recast the continuous evolution into a linear problem that admits a very efficient numerical solution. For the sake of illustration, an example for the evolution of a non-singlet combination of parton distributions is considered. Let $\mathbf{h_i} = \hat{S}\mathbf{a_i}$ denote vectors of the non-singlet PDF values on the discrete x grid determined at different factorisation scales $t_0 < ... < t_i = \ln \mu_i^2 < t_n$, where $t_0 = \ln Q_0^2$ is the initial scale and $\mathbf{a_i}$ is the corresponding vector of spline coefficients, while the matrix $\hat{S}$ contains spline-specific constants. Then, for the initial scale t_0, the discretised version of the DGLAP equations has the following form:

$$\frac{\mathrm{d}\mathbf{h_0}}{\mathrm{d}t} \equiv \frac{\mathrm{d}\hat{S}\mathbf{a_0}}{\mathrm{d}t} = \hat{W}_0\mathbf{a_0}, \tag{9.2.4.2}$$

where $\hat{W}_0$ represents the x-space PDF convolution. The subscript of matrix $\hat{W}_i$ corresponds to the grid node t_i, while the r.h.s of the DGLAP equations includes $\alpha_s(t_i)$. Moving $\hat{S}$ to the r.h.s of the equation gives:

$$\frac{\mathrm{d}\mathbf{a_0}}{\mathrm{d}t} = \hat{S}^{-1}\hat{W}_0\mathbf{a_0}. \tag{9.2.4.3}$$

Approximating the infinitesimal differential $\mathrm{d}t$ by small Δt, the equation can be solved numerically for the scale $t_1 = t_0 + \Delta t$:

$$\mathbf{a_1} = \hat{S}^{-1}\hat{W}_0\mathbf{a_0}\Delta t. \tag{9.2.4.4}$$

The solution can be iterated[12] to find PDFs for all scales $t_i > t_0$. However, in the QCDNUM method, a more precise higher-order integration scheme was proposed [72].

The NLOJET++ program only provides predictions for the single-photon exchange, therefore the NLO calculations were corrected for electroweak and non-perturbative effects using the MC models as described in Sections 8.1.1, 8.1.2, to match the definition of the cross section in the data.

9.2.5. Treatment of Systematic Uncertainties

A detailed description of systematic uncertainties on the measured jet cross sections was presented in Chapter 8. The uncertainties attributed to the inclusive DIS measurements are summarised in [173]. Besides experimental uncertainties on the measured values, additional sources appear due to that of

---

[11] An unequally spaced μ^2-grid can be set as an option.

[12] In principle, backward evolution $t_i < t_0$ is possible.

external input used in theoretical predictions. In general, the absolute magnitude of the uncertainties on the cross-section values must be accompanied by information about mutual correlations between the measured values in different cross-section bins. The correlations are taken into account in the χ^2, as described above (see Eq. 9.2.0.7). Sometimes, an estimate of the correlation of systematic uncertainties is not known because the evaluation procedure can be very difficult or even does not exist. However, information about correlations is important and some assumptions have to be made[13]. Fortunately, positive statistical correlations in inclusive-jet measurements emerging due to the fact that more than one jet can appear in the same event, are relatively easy to assess. In such a case the correlations are represented by the statistical covariance matrix $C_{ij}^{\text{stat.}}$ that was described in Section 8.4. In general, positive correlations between individual measurements result in larger uncertainty when propagated to the fit parameters. The treatment of individual systematic sources is summarised below and is motivated by the following argument:

- **correlated systematic sources:**
    - the uncertainty on the cross-section normalisation due to the luminosity measurement δ_{lumi} is fully correlated across all measurement points because according to Eq. (7.1.0.2) the integrated luminosity $\mathscr{L}$ enters the cross section as a constant common factor for all measurements;
    - the absolute jet-energy scale uncertainty δ_{JES} is treated as fully correlated. This error accounts for the precision of the jet-energy calibration. The calibration constants were determined as functions of the jet pseudorapidity in the laboratory frame. It is assumed that the extracted values deviate from the true ones in a correlated manner, because the jet-calibration technique may have a common bias for all bins.
    - the model uncertainty was treated as correlated. Such an approach is supported by the observation that the individual systematic variations performed for this source (see Section 7.2) result in correlated

---

[13] The following simple example can be considered. In general, the least-squares method for the problem with single fit-parameter e.g. $\alpha_s(M_Z)$, can be interpreted as a weighted averaging of $\alpha_s(M_Z)$ values extracted from individual cross-section bins with weights that are inversely proportional to the squared experimental uncertainty. Assume the luminosity is the only source of systematic uncertainty. The relative error due to luminosity is equal for all bins, thus, if treated as uncorrelated, the resulting uncertainty $\Delta\alpha_s(M_Z)$ scales $\propto 1/\sqrt{\text{N}_{\text{bins}}}$, because there are N_{bins} independent $\alpha_s(M_Z)$ measurements. However, in the case of a fully correlated treatment, $\Delta\alpha_s(M_Z)$ will not improve when extra measurements are included, because the cross sections are assumed to shift up or down in all bins simultaneously. Obviously, for this example, the fully correlated treatment is the correct one.

deviations of the extracted cross sections with respect to the central value;

- **uncorrelated systematic sources:**
  - the normalisation uncertainty due to the simulation of the track-veto efficiency in MC, δ_{TV}, was treated as uncorrelated by the same argument;
  - conceptually, the uncertainty due to the absolute electron-energy scale δ_{EES} has to be treated in the same way as the jet-energy scale error. However, as demonstrated in Section 8.3 the contribution of δ_{EES} to the total experimental uncertainty is typically below 1%, and the effect of correlations can be neglected. Therefore, the δ_{EES} contribution was treated as uncorrelated for simplicity;
  - the uncertainty due to electron identification δ_{eID} and due to the $(E - p_Z)$-cut variation $\delta_{\mathrm{E\text{-}Pz}}$ were also assumed to be uncorrelated.

The uncertainties from uncorrelated sources except for statistical errors were added in quadrature and included as a single uncorrelated source in the χ^2-function. The fit uncertainties of $\alpha_s(M_Z)$, resulting from the experimental uncertainties, were determined by means of the "Hesse" method[14] implemented in MINUIT [192]. A study of the effect of experimental precision of the data on $\alpha_s(M_Z)$ extraction is summarised in Section 9.3.3.

Besides experimental uncertainties, the stability of the fit with respect to theoretical assumptions has to be investigated. In this analysis, theoretical uncertainties related to the variation of renormalisation and factorisation scales were treated as correlated. A detailed discussion of the effect of different approaches for the treatment of scale variations is presented in Section 9.3.4.1 together with a discussion of the sensitivity of the fit, procedure to PDF parametrisation and implicit $\alpha_s(M_Z)$ assumptions. Table 9.2 contains an outline of the systematic-uncertainty treatment.

9.3. QCD Analysis of Inclusive-Jet data

The results of QCD fits are presented in the following. The outcome from different fit variants is described in detail along with a discussion of the stability of the fits with respect to different assumptions imposed in the $\alpha_s(M_Z)$ extraction procedure. Also the dependence of the results on the choice of the data used in the fit is emphasised.

---

[14] A detailed discussion of the "Hesse" approach can be found in [197].

systematic source	correlated	uncorrelated
	experimental	
luminosity	×	
jet energy scale	×	
MC model	×	
electron energy scale		×
electron identification		×
track-veto reweighting		×
$(E-p_Z)$-cut		×
	theoretical	
renormalisation scale	×	
factorisation scale	×	
hadronisation corrections		×
$\alpha_s(M_Z)$ assumed in pPDF	$\alpha_s(M_Z)$ =0.1184±0.0007 [198]	
parton density function	PDF error analysis	

Table 9.2.: Splitting of the systematic-uncertainty sources into correlated and uncorrelated components.

The presentation starts with the simplest variant of the fit in which the strong coupling $\alpha_s(M_Z)$ is determined from the measured inclusive-jet cross sections using proton PDFs extracted in the global fits as external parameters.

9.3.1. Fit to $d\sigma/dE_{T,B}^{jet}$ Cross Section

As a first step, $\alpha_s(M_Z)$ is determined in a fit to individual measurements of the single-differential inclusive-jet cross section as a function of $E_{T,B}^{jet}$. Since a single parameter is extracted from exactly one data point the minimum of the χ^2-function is attained when $\chi^2=0$. The systematic shifts, represented by nuisance parameters, **b** (see (9.2.0.2)), cannot be determined in this case and the only constraints, imposed by the penalty terms, ensure vanishing of the systematic shifts. The extracted values of $\alpha_s(M_Z)$ are illustrated in Figure 9.3. Overall, all α_s-values have comparable experimental uncertainty. However, the α_s-value determined from the highest-$E_{T,B}^{jet}$ point is characterised by a somewhat smaller experimental precision which is due to the larger statistical uncertainty of the data in this region of phase space. In general, α_s values from individual measurements are in very good agreement within the experimental uncertainties, which indicates the overall consistency of the measurements.

The combined $\alpha_s(M_Z)$-value determined from a simultaneous fit to all

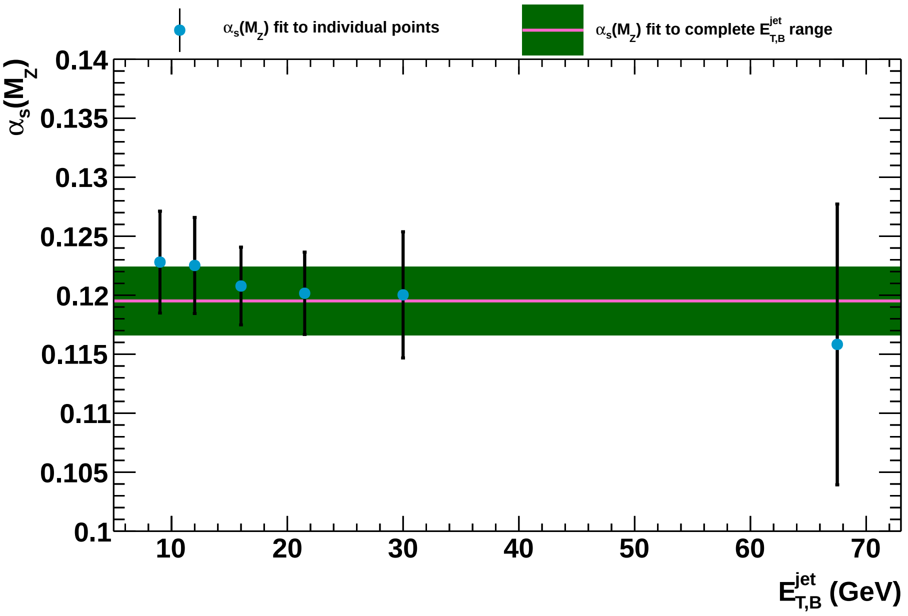


Figure 9.3.: Values of $\alpha_s(M_Z)$ extracted from the individual $d\sigma/dE_{T,B}^{jet}$ measurements. The determined values are compared to the result of a simultaneous fit to all data points in the measured $E_{T,B}^{jet}$ range (pink line and green band). The error bars for individual points indicate the size of the total experimental uncertainty, while the green band represents that for the simultaneous fit.

measured data points is

$$\alpha_s(M_Z) = 0.1195 \pm 0.0029\,(\text{exp.}) \tag{9.3.1.1}$$

as shown in Figure 9.3. The quality of the combined fit is characterised by $\chi^2/N_{\text{DF}} = 1.79/5$ which is well below 1 indicating a possible overestimation of the experimental uncertainties.

Besides the $\alpha_s(M_Z)$ extraction, the energy-scale dependence of the strong coupling was investigated. The α_s-values were determined in the QCD fit to the measured $d\sigma/dE_{T,B}^{jet}$ values. The predictions for individual $d\sigma/dE_{T,B}^{jet}$ cross section points were parametrised in terms of $\alpha_s\left(\langle E_{T,B}^{jet}\rangle\right)$ instead of $\alpha_s(M_Z)$ (see Eq. (9.1.0.1)), where $\langle E_{T,B}^{jet}\rangle$ is the average $E_{T,B}^{jet}$ of the data in a particular cross section bin. For this determination the renormalisation and factorisation scales were set to $\mu_R = E_{T,B}^{jet}$ and $\mu_F = Q^2$, respectively. Figure 9.4 shows the extracted values of α_s. The data demonstrate the running of the strong coupling over

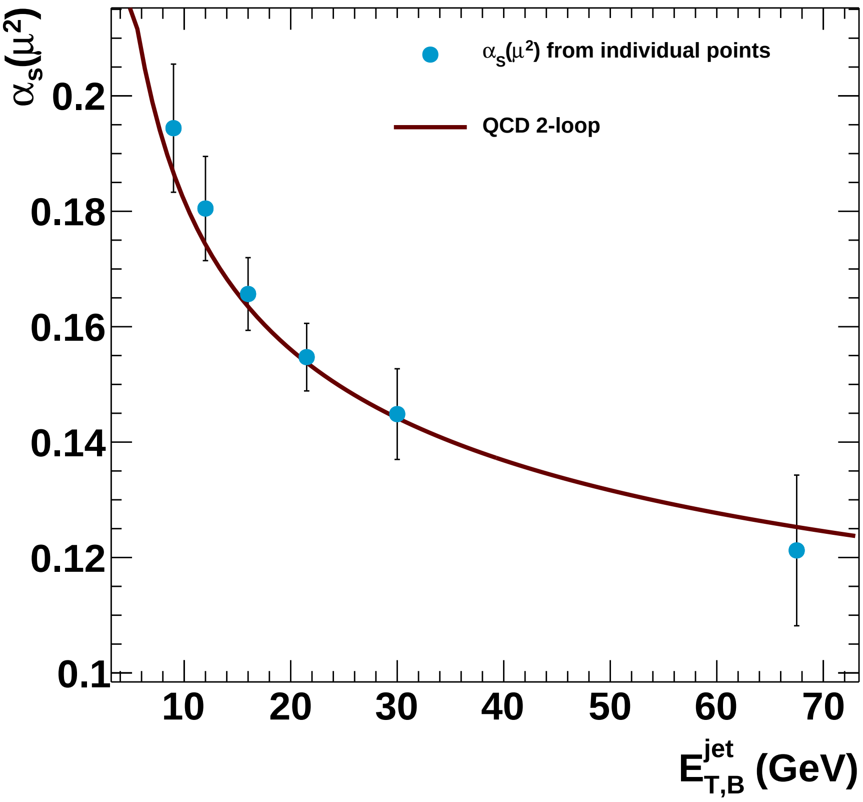


Figure 9.4.: The $\alpha_s\left(\langle E^{jet}_{T,B}\rangle\right)$ values determined from the analysis of the measured $d\sigma/dE^{jet}_{T,B}$ cross sections. The error bars represent the total experimental uncertainty. The solid line represents the renormalisation-group prediction at two-loop approximation obtained from the corresponding $\alpha_s(M_Z)$ value determined in this analysis (Eq. (9.3.1.1)).

a large range of $E^{jet}_{T,B}$. The Renormalisation Group Equations performed with two-loop accuracy [199–202] are in good agreement with the measured values. Since jet production in NC DIS naturally involves two scales, $E^{jet}_{T,B}$ and Q^2, the process scale cannot be unambiguously defined for a particular reaction. Therefore the α_s-extraction procedure described must not be regarded as a rigorous test of the running of the coupling , but rather illustrates the self-consistency of the α_s determination approach.

9.3.2. Simultaneous $\alpha_s(M_Z)$-fit to Double-Differential Cross Sections

As a next step, the sensitivity of the measured double-differential cross sections to the value of the strong coupling was investigated. Similarly to the single-differential cross-section case considered above, α_s-values were determined from individual measurements in bins of $E_{T,B}^{jet}$ and in different regions of Q^2. The results of the extraction are summarised in Figure 9.5 and Table 9.3.

Q^2-range	$\alpha_s(M_Z) \pm$ exp. unc.	N_{DF}	χ^2/N_{DF}	Prob $\chi^2 > \chi^2_{\mathrm{obs}}$
$125 < Q^2 < 250\,\mathrm{GeV}^2$	0.1206 ± 0.0034	5	0.18	0.97
$250 < Q^2 < 500\,\mathrm{GeV}^2$	0.1189 ± 0.0036	5	0.60	0.7
$500 < Q^2 < 1000\,\mathrm{GeV}^2$	0.1224 ± 0.0038	5	0.56	0.73
$1000 < Q^2 < 2000\,\mathrm{GeV}^2$	0.1242 ± 0.0039	5	0.61	0.69
$2000 < Q^2 < 5000\,\mathrm{GeV}^2$	0.1248 ± 0.0046	5	1.44	0.17
$5000 < Q^2 < 20000\,\mathrm{GeV}^2$	0.1246 ± 0.0073	5	0.58	0.72

Table 9.3.: Values of $\alpha_s(M_Z)$ obtained in a fit to the measured double-differential inclusive-jet cross sections. The experimental uncertainty includes statistical as well as correlated and uncorrelated systematic uncertainties.

Several observations can be made:

- The χ^2/N_{DF} of the fits to $d\sigma/dE_{T,B}^{jet}$ in different regions of Q^2 varies between 0.18 and 1.44, but is typically about 0.6. In general, such values can be considered as a sign of a reasonable fit quality.

- The minimal value of χ^2/N_{DF} arise from the lowest Q^2 bin, which, in principle, is characterised by the smallest total experimental uncertainty. This could indicate an overestimation of experimental errors in this region.

- All values of $\alpha_s(M_Z)$ determined in different bins of Q^2 are in good agreement within experimental uncertainties.

- The α_s values extracted in the four regions in the virtuality range $125 < Q^2 < 2000\,\mathrm{GeV}^2$ have comparable experimental precision. The values determined from the measurements in the two last Q^2 bins have larger uncertainty, which is due to the statistical precision of the data in this region.

The value of the strong coupling extracted from the simultaneous fit to all points of the measured double-differential cross section is

$$\alpha_s(M_Z) = 0.1216 \pm 0.0026\,(\mathrm{exp.})\,. \tag{9.3.2.1}$$

Uncertainty source	$\delta\alpha_s$	$\delta\alpha_s$ uncor. treatment
Statistical	0.3%	—
Tot. uncor. syst.	0.4%	—
JES	1.7%	0.3%
Model	0.3%	0.2%
Luminosity	1.2%	0.2%

Table 9.4.: Contribution of different sources of experimental uncertainty to the total experimental uncertainty on $\alpha_s(M_Z)$ determined from the fit to double-differential cross sections. The considered sources of experimental error on the $\alpha_s(M_Z)$ value are statistical, combined uncorrelated systematic uncertainty, jet-energy scale, model and luminosity uncertainties. The second column contains the breakdown of uncertainty contributions attributed to different sources and treated as described in Section 9.2.5. The last column presents the contribution of correlated errors sources, when treated as uncorrelated.

This is consistent with the extractions from the individual Q^2 bins. The experimental uncertainty was obtained by propagating the uncertainties listed in the previous section. The obtained fitted nuisance parameters b_μ for the correlated systematic uncertainties are $b_{JES} = -0.04$, $b_{lumi} = -0.2$ and $b_{model} = -0.25$ for the absolute jet-energy calibration, the luminosity and model error, respectively. Since the parameters b_μ are introduced with a negative sign and $\Gamma_{JES(lumi)} > 0$ in Eq. (9.2.0.7), it indicates that the α_s-fit prefers the 'down-shift' of the measured cross sections. Nevertheless, the uncertainty on b_μ was approximately of the same magnitude as the parameters itself, indicating that the obtained values are consistent with $b_\mu = 0$ hypothesis. It is important to note that the extracted α_s values are very sensitive to the overall normalisation of the data, because, as was shown in Figure 9.1, the absolute cross-section value depends monotonically on $\alpha_s(M_Z)$.

9.3.3. Experimental Uncertainties on $\alpha_s(M_Z)$

The influence of systematic uncertainties on the extracted $\alpha_s(M_Z)$ value was investigated and is summarised in Table 9.4. The contribution from individual sources of uncertainty were were estimated setting all errors except of the one for the studied source to zero in the α_s fit.

The dominant contribution to the total experimental uncertainty on α_s arises from the absolute jet-energy scale and luminosity errors on the jet cross sections and amounts to about 1%. The other contributions due to limited statistics and due to uncorrelated sources of systematic uncertainties are smaller.

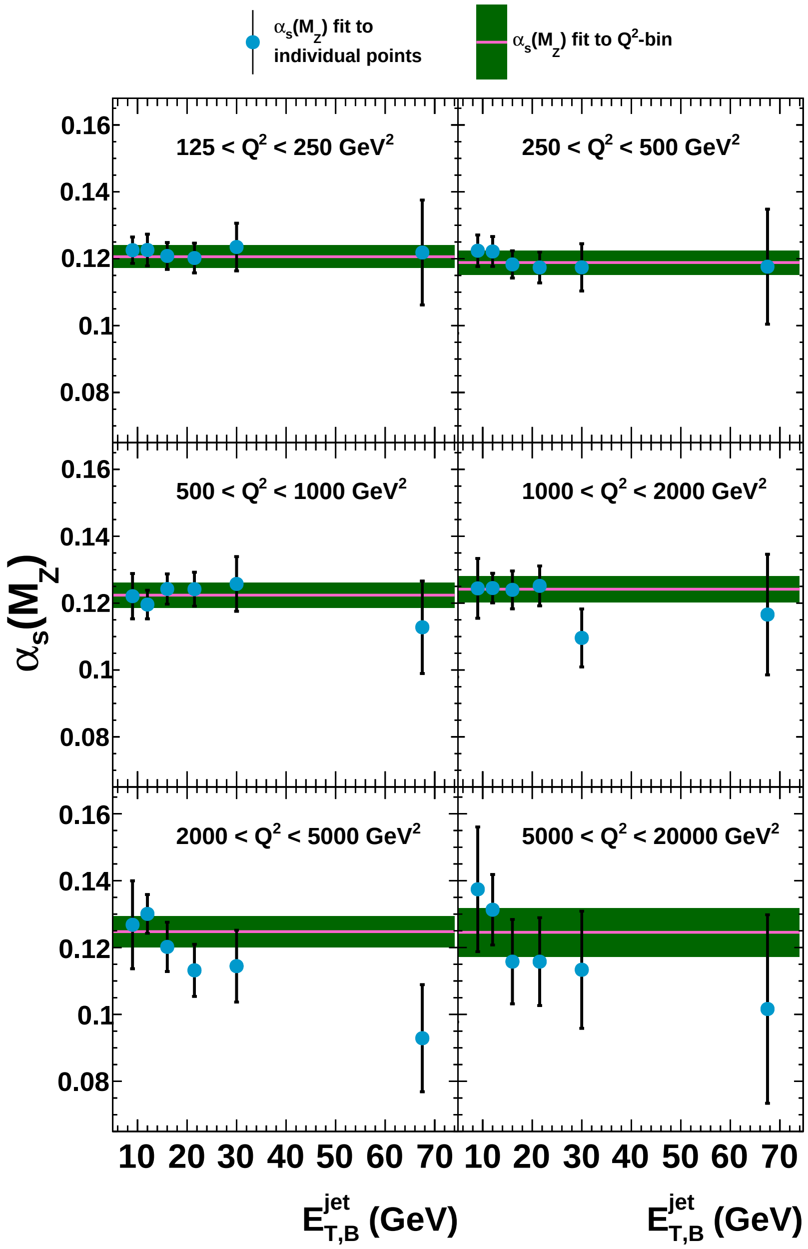


Figure 9.5.: Values of $\alpha_s(M_Z)$ extracted from the individual measurements of the $d\sigma/dE_{T,B}^{jet}$ cross sections in different regions of Q^2. The determined values are compared to the result of a simultaneous fit to all data points in the separate Q^2 ranges (pink line and green band). The error bars for individual points indicate the size of the total experimental uncertainty, while the green bands represents that for the simultaneous fits.

For comparison, the last column in Table 9.4 illustrates the effect of treating jet-energy scale, model and luminosity errors as uncorrelated. As can be seen, the difference between two approaches to the treatment of the model uncertainty is marginal. Moreover, the uncorrelated approach, which is totally inconsistent for the luminosity, results in an underestimated magnitude of experimental uncertainties.

9.3.4. Theoretical Uncertainties on $\alpha_s(M_Z)$

Additional sources of uncertainty on α_s arise from the precision of theoretical calculations. In this section the results of a study of the sensitivity of the $\alpha_s(M_Z)$ fits to external theoretical assumptions is presented. In general, uncertainties attributed to assumptions are difficult to quantify or even impossible to define due to their non-stochastic nature. Approaches used for the treatment of particular error sources are typically tailored for each problem and the statistical interpretation of the resulting variations of the results if often doubtful. In this section the errors due to the truncation of the perturbative series, limited precision of the proton PDFs and hadronisation corrections are described in detail.

9.3.4.1. Uncertainties due to Missing Higher Orders

A natural approach that can be used to estimate the uncertainties attributed to theoretical assumptions is to repeat the parameter determination under different assumptions and treat the difference between the results as an uncertainty on the extracted parameter value. If the value of $\alpha_s(M_Z)$ obtained under different theoretical assumption is denoted by α_s^* and the central value is α_s^0, then the uncertainty is:

$$\Delta\alpha_s = \alpha_s^* - \alpha_s^0. \tag{9.3.4.1}$$

Such an approach is called the “offset method” [197]. In principle, it takes into account non-linear effects in the error propagation. When the underlying assumptions can be parametrised by a single continuous parameter, this method can be used for linear error propagation employing a numerical approximation of the first derivative:

$$\frac{\partial\alpha_s}{\partial a} \approx \frac{\Delta\alpha_s}{\Delta a}, \tag{9.3.4.2}$$

where Δa is a 1σ change of the parameter in question.

By default, the offset method was used in this analysis for the determination of the uncertainty attributed to the truncation of the perturbative expansion for the jet cross sections. As described in Chapter 2, one way to estimate the size of the missing higher-order terms is to measure the change of the pQCD predictions with respect to the variation of the renormalisation and factorisation scales. For this purpose the $\alpha_s(M_Z)$ fit was repeated using theoretical predictions evaluated with modified definitions of the renormalisation and factorisation scales. The commonly accepted definition of a '1σ confidence interval' for a scale variation uncertainty [203] was adopted in this work. Two variants of the scale variation were considered:

- **Independent variation of μ_R and μ_F.** The theoretical predictions for the double-differential jet cross sections were recalculated with the scaled μ_R and μ_F parameters:

  $$\mu_R \to 2\mu_R; \qquad \mu_R \to 1/2\mu_R; \tag{9.3.4.3}$$

  $$\mu_F \to 2\mu_F; \qquad \mu_F \to 1/2\mu_F; \tag{9.3.4.4}$$

  performing each variation one at a time. In general, the fits were characterised by larger values of χ^2/N_{DF}. Only the fit with $\mu_F \to 2\mu_F$ had smaller value of χ^2/N_{DF} than the central $\alpha_s(M_Z)$-fit. The resulting variation of $\alpha_s(M_Z)$ was:

  $$\alpha_s(M_Z) = 0.1216\,^{+0.0057}_{-0.0050}\,(\mu_R)\,^{+0.0003}_{-0.0003}\,(\mu_F) \tag{9.3.4.5}$$

- **Simultaneous variation of μ_R and μ_F.** Because the analytic expressions for the renormalisation and factorisation scales are related, alternatively both scales were varied up and down simultaneously. The obtained variations are:

  $$\alpha_s(M_Z) = 0.1216\,^{+0.0065}_{-0.0053}\,(\mathrm{scales}) \tag{9.3.4.6}$$

  Such variant results in a more conservative estimate of the uncertainty.

It can be seen that the fits are much more sensitive to the renormalisation-scale variation and result in an asymmetric uncertainty on the $\alpha_s(M_Z)$ value. The comparison of the fit results for different scale variations is illustrated in Figure 9.6 and denoted as δ_1.

Other approaches for the treatment of the scale-variation uncertainty were also investigated. As an alternative, theoretical uncertainties were included into the covariance matrix assuming that errors attributed to the scale variations are completely uncorrelated across the measured points. In this case the total error attributed to the theoretical and experimental uncertainties was propagated to the uncertainty on $\alpha_s(M_Z)$ value using the Hesse method. The uncertainty contribution on α_s from the scale variations was calculated from the total error Δ_{tot} according to:

$$\Delta_{\mathrm{scales}} = \sqrt{(\Delta_{\mathrm{tot}})^2 - \left(\Delta_{\mathrm{exp}}\right)^2} \tag{9.3.4.7}$$

Another approach was also tested. The uncertainties due to missing higher-order contributions were included into the χ^2 using the nuisance parameters, similar to jet-energy scale or luminosity uncertainty. In this case the systematic shifts attributed to the theoretical uncertainty are treated as correlated across different bins and propagated to the result using the Hesse method. Overall, fits with Hesse treatment of the scale-variation uncertainties resulted in substantial reduction of the $\alpha_s(M_Z)$ error in comparison to the offset method. The comparison of different approaches is provided in Table 9.5.

Treatment approach	$\Delta\alpha_s$
Offset	$^{+0.0065}_{-0.0053}$
Hesse (nuissance)	0.0026
Hesse (uncorrelated)	0.0015

Table 9.5.: Comparison of the uncertainty on the extracted $\alpha_s(M_Z)$ value due to missing higher orders obtained using different error-treatment approaches. The uncertainties due to variations in the renormalisation and factorisation scales are combined in a single contribution.

The reduction of the $\alpha_s(M_Z)$ uncertainty can be explained as follows. For simplicity, only the Q^2 dependence of the μ_R-variation uncertainties is considered. As was demonstrated in Section 8.2 (see Figure 8.3), the dominant component of the theoretical uncertainty due to renormalisation scale variation decreases with increasing Q^2. Therefore, in contrast to the offset method, in the the Hesse approach, the measurement point at high virtualities has higher sensitivity to the value of the nuisance parameter b_{μ_R}, which effectively can be interpreted as assigning larger weight to the high-Q^2 measurements. This, in turn, leads to the reduction of the uncertainty on the extracted $\alpha_s(M_Z)$ value.

In the case of using nuisance parameters to carry out the fit, the numerical value for the systematic shifts $b_{\mu_R} = \ln \mu_R/\mu_R^0$, $b_{\mu_F} = \ln \mu_F/\mu_F^0$, where μ_R^0 and μ_F^0 denote the default scale, represents the rescaling factor that has to be applied to the renormalisation or factorisation parameters in order to obtain a better fit. This treatment assumes full correlation of missing higher-order contributions across the phase space, which is, in principle, difficult to justify. As demonstrated, simultaneous variation of scales results in a more conservative estimate of uncertainty on α_s, therefore the errors on theoretical predictions estimated from simultaneous rescaling of μ_R and μ_F were used in the fit and a single parameter $b = b_{\text{scale}}$ was introduced instead of b_{μ_R} and b_{μ_F}. The systematic shift obtained from the fit was $b_{\text{scale}} = 0.11 \pm 0.34$. This indicates that within the uncertainty on b_{scale} an $\mu_R = \sqrt{Q^2 + E_{T,B}^{jet\,2}}$ and $\mu_F = Q$ cannot be distinguished from the 'optimal' scale choice.

Fits with alternative definitions of the renormalisation and factorisation scales were also performed. Scale choices such as $\mu_R = E^{jet}_{T,B}$, $\mu_R = Q$ and $\mu_F = E^{jet}_{T,B}$, $\mu_F = \sqrt{Q^2 + E^{jet^2}_{T,B}}$ were investigated. In general, alternative fits had comparable χ^2/N_{DF}, but resulted in lower values of $\alpha_s(M_Z)$. Only the fit with $\mu_R = \mu_F = \sqrt{Q^2 + E^{jet^2}_{T,B}}$ gave rise to a somewhat larger value of the strong coupling. The discrepancy between the results of different fit variants was smaller than the scale-variation uncertainty estimated using the offset method and therefore was assumed to be covered by the scale-variation uncertainty.

9.3.4.2. Sensitivity to the Proton PDF Sets

The proton PDFs have limited precision, therefore the stability of the $\alpha_s(M_Z)$-fits with respect to PDF assumptions has to be investigated. Several effects contributing to the proton PDF uncertainty can be identified. All of them can be approximately attributed to one of the following categories:

- **Limited PDFs precision due to data uncertainty.** Because PDFs are extracted in QCD fits to data, the uncertainty on the measurements propagates to the error on the proton PDFs, which in turn affects the precision of the $\alpha_s(M_Z)$ determination.

- **QCD-fit setup differences.** As explained above, a procedure involving simultaneous PDF fits of all individual parton species is a not well defined problem and additional assumptions have to be imposed in order to obtain reasonable fit results. However, the choice of the fit procedure is usually ambiguous. Typically, different fitting groups use partially different datasets to constrain PDF parameters[15]. Furthermore, they can have different choices of PDF parametrisation at the starting evolution scale, distinct approaches to the treatment of heavy quarks, alternative χ^2-functions and/or error definitions, etc. This leads to different results for PDFs obtained by different groups.

- **PDF and $\alpha_s(M_Z)$ correlation.** In many high-energy processes, the strong coupling and the proton PDFs are coupled, i.e. the predictions are proportional to the product $\alpha_s \cdot f_i(x, \mu)$. In particular, the gluon component is always accompanied by a factor α_s in the matrix elements, therefore, a value of $\alpha_s(M_Z)$ has to be assumed in the PDF fit, which introduces an

---

[15] In particular, the exact values for the Q^2-cut used to remove the data from kinematic phase space where higher-twist effects might have significance contribution rely on the model assumptions and therefore may differ in different analyses.

implicit dependence of the results on $\alpha_s(M_Z)$. In order to enable propagation of this type of uncertainty into pQCD predictions, different PDF fitting groups provide series of PDFs extracted assuming different values of $\alpha_s(M_Z)$.

In the HERAPDF 1.5 set, which was used as a default in the calculations of the inclusive-jet cross sections, 34 systematic variations were performed to estimate the PDF uncertainty and were classified into three groups roughly corresponding to the categories listed above.

The uncertainty due to the experimental precision of the data is propagated to the PDF parameters and is represented in the covariance matrix form. In order to simplify propagation of the uncertainties attributed to the proton PDFs, the matrix is transformed to diagonal form with eigenvectors representing the uncertainty on the corresponding linear combinations of the PDF parameters; $+1\sigma$ parameter variation along the eigenvectors translates into a 68% confidence interval for the pQCD predictions around their central values according to the formula [204]

$$\delta^{\pm}\mathbf{t} = \sqrt{\sum_{i=1}^{N_{eig}} \left(\delta_i^{\pm}\mathbf{t}\right)^2}, \tag{9.3.4.8}$$

where the positive (negative) uncertainty, $\delta^{\pm}\mathbf{t}$, is calculated from the quadratic sum of positive (negative) variations of the predictions with respect to those based on the central PDF, $\delta_i^{\pm}\mathbf{t}^2$. The summation runs over eigenvectors, the number of which is equal to the number of free PDF parameters[16]. When propagated to the strong coupling, PDF eigenvector variation results in an asymmetric uncertainty on the α_s value. A symmetric estimate of $\delta\alpha_s$ was defined as:

$$\delta\alpha_s = \sqrt{\sum_{i=1}^{N_{eig}} \frac{\left(\delta_i^{+}\alpha_s - \delta_i^{-}\alpha_s\right)^2}{2}}. \tag{9.3.4.9}$$

It provides a reasonable average of the asymmetric uncertainties and can be considered as a modification of the offset method. In order to estimate the uncertainty on $\alpha_s(M_Z)$, the α_s fits were repeated for every individual up-and down-variation of the PDF eigenvectors and compared to the central value given in Eq. 9.3.2.1. The results of these fits are summarised in Figure 9.6 and denoted by δ_2. As can be seen, the result of the variations of the PDF eigenvectors is typically very small, except for single upward variation of the parameter 3. Unfortunately, since the eigenvectors represent the linear combination of the original PDF parameters, the nature of the parameter, resulting in the largest variation of $\alpha_s(M_Z)$, is obscured.

---

[16] In total, the HERAPDF 1.5 set provides 10 up and 10 down eigenvector variations.

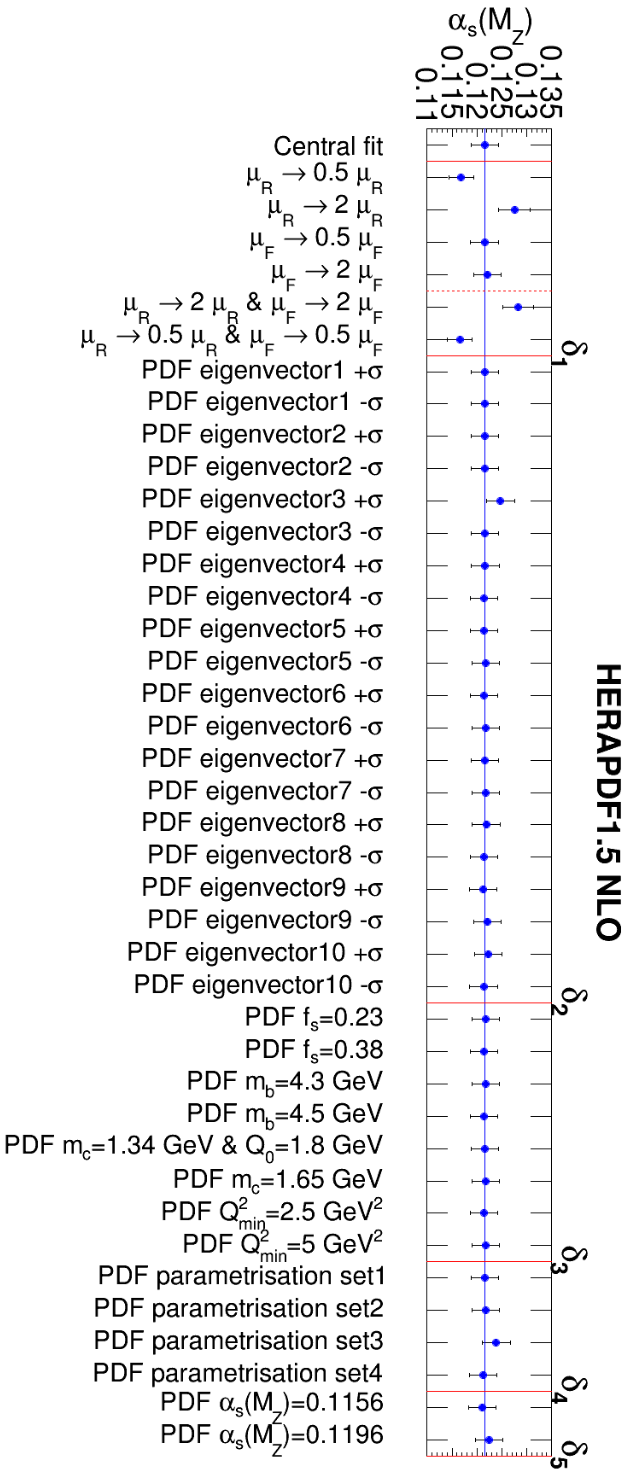


Figure 9.6.: The results of $\alpha_s(M_Z)$ extraction from double-differential cross sections. The individual systematic variations are combined into groups $\delta_1 - \delta_5$ (see text).

PDF set	$\alpha_s(M_Z)$-value	χ^2/n_{DF}
HERAPDF 1.5	$0.1216^{+0.0029}_{-0.0018}$	0.626
NNPDF23	0.1221	1.00
MSTW08	0.1232	0.884
CT10	0.1240	0.831

Table 9.6.: The results of the $\alpha_s(M_Z)$ extractions based on different PDF sets. For comparison the uncertainty estimated using the HERAPDF analysis is presented.

The model uncertainty on the PDFs, attributed to the variation of the assumptions such as the fraction of the strange quarks in the d-type sea PDFs, mass of the charm and bottom quarks, etc. were also treated using the offset approach. A detailed description of the nature of individual variations is presented in [167]. The uncertainty on $\alpha_s(M_Z)$ was estimated by taking the difference between the variation and the central value, and then adding in quadrature all positive (negative) differences to obtain the positive (negative) model error (see Eq. (9.3.4.8)). Corresponding results of the α_s fits are combined into the δ_3 group in Figure 9.6. The effect of these variations was found to be negligible.

Variations of the functional form of the PDF parameterisation (see [167]) were also treated using the offset-method. To form the parametrisation envelope, the largest positive (negative) difference between the variation and the central value is taken as the positive (negative) parametrisation error. The outcome of the resulting variations of $\alpha_s(M_Z)$ are illustrated in Figure 9.6 as δ_4. The only significant variation of the $\alpha_s(M_Z)$ value resulted from the alternative gluon PDF parametrisation Eq. (9.2.1.2), where $\varepsilon_g, E_g \neq 0$.

The uncertainty attributed to the assumption imposed on the magnitude of the strong coupling in the PDF extraction procedure was estimated from two PDF sets with $\alpha_s(M_Z)$ =0.1156 and $\alpha_s(M_Z)$ =0.1196. The difference between the two results, (see Figure9.6 (δ_5)), was scaled to the current uncertainty on the world average [198].

In order to obtain the total error on the $\alpha_s(M_Z)$-value due to the proton PDF, all mentioned sources were added in quadrature, which resulted in the final value:

$$\alpha_s(M_Z) = 0.1216\,^{+0.0029}_{-0.0018}\,(\mathrm{PDF})\,. \tag{9.3.4.10}$$

Fits to $\alpha_s(M_Z)$ with alternative predictions based on PDF sets provided by different fitting-groups were performed in order to investigate the stability of the $\alpha_s(M_Z)$ determination procedure. The considered set of PDFs comprises

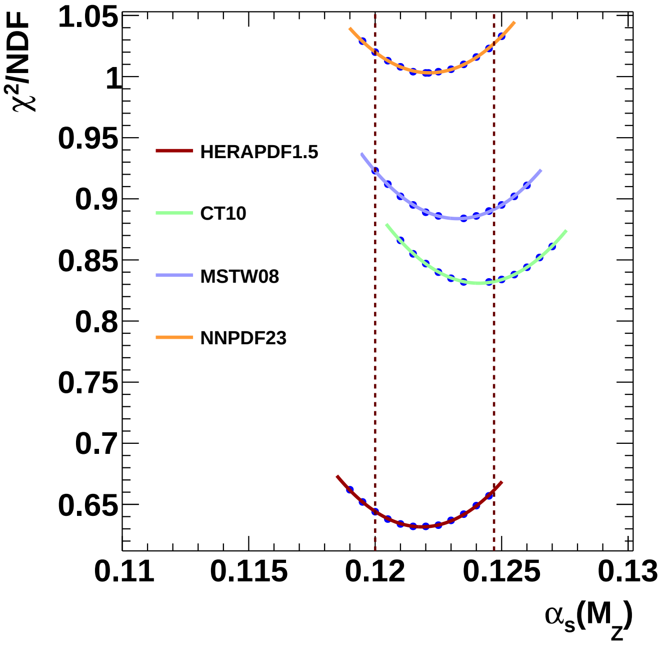


Figure 9.7.: The results of the scan of $\chi^2(\alpha_s(M_Z))$ as a function of $\alpha_s(M_Z)$ obtained using different PDF sets. The vertical dashed lines represent the PDF uncertainty from the HERAPDF 1.5 error analysis described in the text.

the extractions by the MSTW [169], NNPDF [205] and CT [170] groups. The obtained results were compared to the reference fit based on HERAPDF 1.5, described above. All fits were performed using the so-called 'central' PDF set determined with the $\alpha_s(M_Z)$ value recommended in the corresponding PDF-fit. Analysing the obtained results, several observations can be made:

- All fits with alternative PDFs have larger χ^2 than the reference one, however in all cases the obtained values of χ^2/n_{DF} were $\lesssim 1$. Figure 9.7 illustrates the scan of the χ^2 as a function of $\alpha_s(M_Z)$ for different PDF sets. One possible explanation for the difference in the fit quality is the difference in the gluon distribution of different PDF sets, especially in the high-x region, relevant for inclusive-jet production (see Figure 9.8).

- The obtained values of the strong coupling were larger than that extracted using HERAPDF 1.5, but the difference was always within the uncertainty attributed to the HERAPDF 1.5 set. A summary of the fit results based on

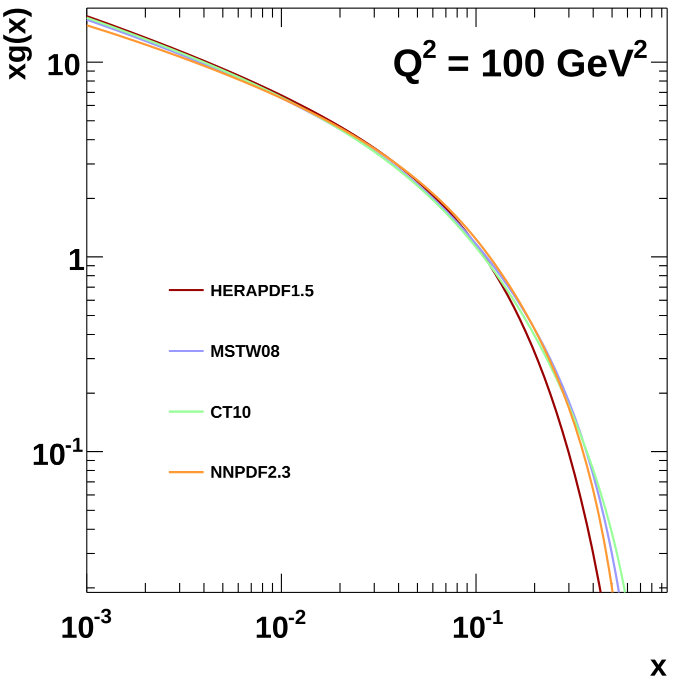


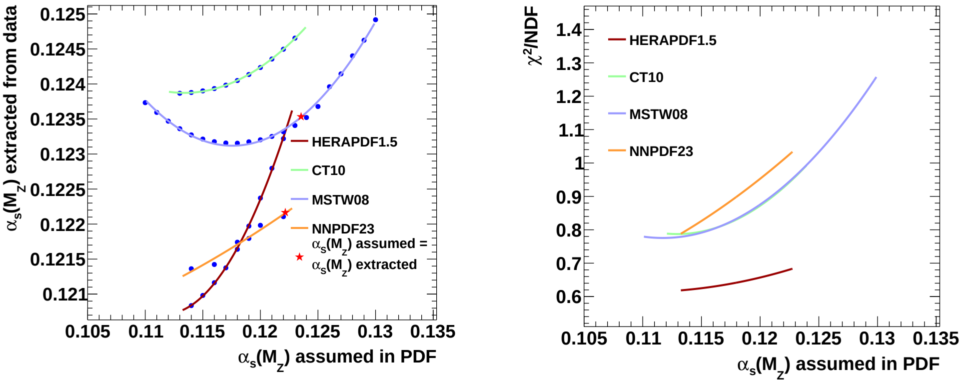
Figure 9.8.: Comparison of the gluon distribution in the high-x region for different PDF sets at Q^2 =100 GeV^2.

different PDF sets is compiled in Table 9.6

In accord with the PDF4LHC recommendations, in recent analyses [74] an additional uncertainty was attributed to the difference between results obtained with different PDF sets. In contrast to that, the checks, described above, reveal the consistency of the considered PDF sets in the phase space of the inclusive-jet measurement and support the adequacy of the HERAPDF 1.5 error analysis. Therefore no additional uncertainties due to PDFs were assigned to the $\alpha_s(M_Z)$-value determined in this thesis.

9.3.4.3. Sensitivity to the $\alpha_s(M_Z)$ Assumption

As mentioned above, the strong coupling enters pQCD predictions for hadron-induced processes through the matrix elements and the DGLAP splitting functions $P(z) = \sum_{i=1}^{N} {\alpha_s}^i \cdot P_i(z)$ (see Chapter 2). Therefore, in order to determine



a

b

Figure 9.9.: The dependence of the fitted value of $\alpha_s(M_Z)$ extracted from inclusive-jet data using PDF sets with different assumed $\alpha_s(M_Z)$-values (a). Quality of the fit as a function of $\alpha_s(M_Z)$-value assumed in the PDFs (b).

PDFs in global fits to the data, a value of $\alpha_s(M_Z)$ has to be either taken from other analyses or fitted together with PDFs. Due to the coupled nature of perturbative hard coefficients and PDFs ($c(\alpha_s) \cdot f(x,\mu)$), the decrease of one quantity leads to the increase of the other to maintain a constant cross section. Therefore the question of the influence of the $\alpha_s(M_Z)$ assumption in the PDF extraction procedure has to be addressed. Figure 9.9(a) shows the results of the $\alpha_s(M_Z)$ fits to the double-differential inclusive-jet cross sections with PDF sets from several fitting groups, extracted under different assumptions on the magnitude of the strong coupling. It can be seen that, apart from MSTW08 sets, available for a wide range of assumed $\alpha_s(M_Z)$ values, the extracted $\alpha_s(M_Z)$ increases monotonically with increasing $\alpha_s(M_Z)$ assumed in the PDFs. Only for MSTW08 and NNPDF23 do the input and output values of $\alpha_s(M_Z)$ coincide at relatively large value of α_s. Qualitatively, the correlation of the input and output values can be explained as follows: in the global fit, PDFs decrease as the assumed value of the strong coupling increases, therefore to compensate the reduction of $f(x,\mu)$ in the fit to inclusive-jet data, the determined $\alpha_s(M_Z)$ value increases.

As can be seen in Figure 9.9(b), the quality of the fit deteriorates with increasing assumed value of the strong coupling and the fits based on PDFs with the smallest $\alpha_s(M_Z)$ value yield the smallest values of χ^2/n_{DF}. However, as can be seen in Figure 9.9(a) the correlation between assumed and extracted values of $\alpha_s(M_Z)$ is very mild. For example, in case of NNPDF23, which has an approximately linear dependence, the resulting difference in the extracted $\alpha_s(M_Z)$ value of about 1% corresponds to the extreme variation of the assumed value of the strong coupling in range $\alpha_s(M_Z) \in (0.113; 0.123)$, which in turn corresponds to the variation of χ^2/n_{DF} between 0.8 and 1.04. Taking into account that $n_{\mathrm{DF}} = 35$, this corresponds to the χ^2 variation between χ^2_{min} and $\chi^2_{\mathrm{min}} + 8$, which is significantly larger than that, used in the 1σ definition for the $\alpha_s(M_Z)$ uncertainty. Summarising, the resulting values of α_s for the fits with minimal value of χ^2/n_{DF} are consistent with the central value within the experimental uncertainty.

9.3.4.4. Uncertainties due to Hadronisation Corrections

The uncertainty on α_s due to the that on the hadronisation corrections applied to the NLO pQCD predictions were estimated by error propagation utilising the Hesse method. The uncertainty on the correction factors was defined as a half-difference[17] between the predictions obtained using the ARIADNE or LEPTO models (see Section 8.1.1).

For simplicity, the fits for α_s were performed by adding the uncertainty on the hadronisation correction factors quadratically to the uncorrelated systematic uncertainty[18]. The effect of the additional source of error on the extracted $\alpha_s(M_Z)$ value was determined using Eq. (9.3.4.7), where in this case, the Δ_{tot} represented the combined uncertainty on α_s due to hadronisation corrections and experimental errors.

The uncertainties on the predictions due to hadronisation correction factors was $\lesssim 2\%$ but typically less than 0.5% in most of the cross-section bins. The α_s fits were characterised by comparable quality of the fit and resulted in $\alpha_s(M_Z)$ very similar to the central value. The uncertainty on $\alpha_s(M_Z)$ due hadronisation corrections was smaller than 0.1% and therefore was neglected.

9.3.5. Summary and Conclusion on $\alpha_s(M_Z)$ fits with Fixed PDFs

A value of the strong coupling constant has been determined from the double-differential inclusive-jet cross sections measured as functions of $E^{jet}_{T,B}$ in different regions of Q^2. The extraction refers to a wide interval of hard scales corresponding to $19 < \mu_R < 150$ GeV. The sensitivity of the extraction procedure to theoretical assumptions e.g. the choice of the PDFs or renormalisation or fac-

---

[17] An average of the LEPTO and ARIADNE correction factors was used to correct parton-level predictions for hadronisation effects.

[18] In principle, the uncertainty on the hadronisation corrections has to be treated as correlated; however the difference between the two approaches was found to be negligible.

torisation scales was investigated in detail. The extracted value of $\alpha_s(M_Z)$ is:

$$\alpha_s(M_Z) = 0.1216 \pm 0.0026\,(\text{exp.})^{+0.0066}_{-0.0053}\,(\text{scales})^{+0.0029}_{-0.0018}\,(\text{PDF})\,, \tag{9.3.5.1}$$

where individual uncertainties attributed to various sources are indicated separately.

The dominant source of experimental uncertainty is due to the precision of the calibration of the absolute jet-energy-scale, while the theoretical uncertainty is dominated by missing higher orders in the pQCD predictions. Different approaches to the treatment of experimental and theoretical uncertainties were investigated. It was demonstrated that the results of the fit can significantly differ when the systematic uncertainties are treated as correlated or uncorrelated.

To check the consistency of the data, the $\alpha_s(M_Z)$-values were determined from individual cross-section bins and were found to be in good agreement within the estimated experimental uncertainties.

The running of the strong coupling was also illustrated. The predictions of the Renormalisation Group Equation calculated with two-loop precision were found to be in good agreement with the observed evolution of α_s over a wide range of scales.

9.3.6. Comparison with Other $\alpha_s(M_Z)$ Determinations at HERA

QCD has become the accepted theory of the strong interaction. If one ignores the masses of quarks in the QCD Lagrangian (2.2.0.2)[19] the success of the theory with a single free parameter ($\alpha_s = \frac{g_s^2}{4\pi}$) is striking. Nevertheless, precision tests of QCD are ongoing. The value of the strong coupling, extracted from a wide variety of experimental data, can be regarded as a figure of merit unifying diverse approaches for testing QCD. This section contains a brief summary of α_s determinations from jet measurements at HERA and LHC.

Previous determinations of the strong coupling at HERA include extractions from NC DIS inclusive-jet data at high-Q^2 [17, 206], low-Q^2 [207] as well as from photoproduction data [27]. The extractions from inclusive-jet data in NC DIS performed at HERA before 2007 were combined, yielding a HERA average [208]. More recent determinations include extractions from normalised inclusive-jet [18] and multijet [184] cross sections as well as from the ratio of three-to-two jet rate, $R_{3/2}$, [149] and photoproduction [48]. A preliminary analysis of the data set very similar to the one presented in the thesis was reported in [209].

---

[19] Such an approximation is valid at sufficiently high energy $m_q \ll E$.

Overall, the result obtained in this work, is in good agreement with previous analyses performed in different kinematic regions at ZEUS and with early extractions by H1. When compared to HERA I results [17,27,206], this determination features increased statistical precision, especially in the high-$E_{T,B}^{jet}$ and high-Q^2 region and has experimental precision comparable to HERA (2007) average. The presented extraction is also consistent within experimental uncertainties with more recent determinations from $R_{3/2}$ [149] and photoproduction [48] at ZEUS.

Comparing to [149] this determination benefits from the fact that an inclusive analysis gains larger event sample, due to excluding the requirement that both jets have to be within the detector acceptance[20]. On the other hand, in the ratio $R_{3/2}$ significant amount of systematic effects cancel.

In comparison to the photoproduction result [48] both determinations have very similar experimental uncertainties. Besides most the same sources of theoretical uncertainty the photoproduction analysis includes an additional source attributed to the knowledge of the photon PDF. Nevertheless, the photoproduction result has smaller theoretical uncertainty when compared to the values obtained in this thesis with the offset method[21]. In both cases the dominant source of theoretical uncertainty is the one due to terms beyond NLO, however, better precision of the analysis [48] can be explained by the existence of a unique hard scale, $E_{T,lab}^{jet}$, and typically larger jet transverse energy $\left(E_{T,lab}^{jet} > 21\mathrm{GeV}\right)$.

This $\alpha_s(M_Z)$ determination is consistent with the preliminary result, obtained in an independent inclusive-jet study [209], which was also based on HERA II data but performed within somewhat different phase space. That determination has slightly larger experimental uncertainty, while the theoretical error of that analysis benefits from using the reduced dataset with lower sensitivity to the variations in pQCD predictions and has to be compared to the uncertainty treatment, performed using the Hesse method, where the varying sensitivity of the data to pQCD predictions was taken into account.

Normalised inclusive-jet [18] and multijet [184] cross sections from H1 experiment are characterised by small experimental uncertainty and are typically lower than that from ZEUS. High experimental precision in these determinations was achieved by utilising normalised jet cross sections in which systematic effects cancel. Besides that different sensitivity of various jet data sets to pQCD predictions (e.g. NLO pQCD predictions for trijet cross sections are $\mathcal{O}\left({\alpha_s}^3\right)$) was exploited.

The extractions of $\alpha_s(M_Z)$ from HERA are complementary to the determi-

---

[20] At leading order in pQCD an NC DIS event always has exactly two partons (see Section 2.1.1).

[21] In the photoproduction analysis an offset method was used.

nations from jet cross sections in pp collisions at the LHC. When compared to CMS results [185, 186, 210], the value of $\alpha_s(M_Z)$ obtained in this work features similar theoretical uncertainty and somewhat larger experimental uncertainty. The CMS values are systematically lower then that presented in this thesis.

All discussed determinations including world average [12] are summarised in Figure 9.10.

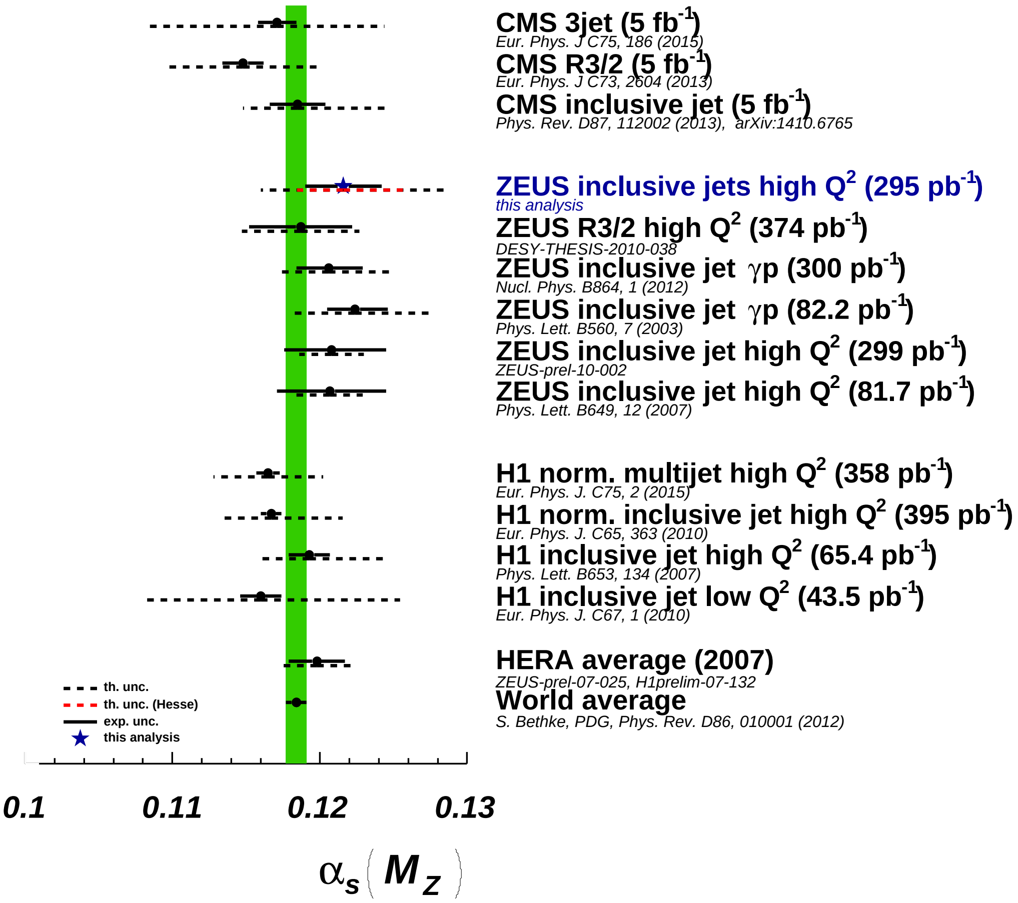


Figure 9.10.: Comparison of the values of the strong coupling constant determined at HERA by H1 and ZEUS collaborations as well as those from CMS. The green band indicates the uncertainty on the world average α_s value. Solid and dashed lines correspond to combined experimental and theoretical uncertainties, respectively.

Summary and Outlook

In this thesis, single-differential and double-differential inclusive-jet cross sections in neutral current deep inelastic scattering have been measured with the ZEUS detector at HERA. Jets were reconstructed with the k_T-clustering algorithm in the longitudinally invariant inclusive mode in the Breit frame with $E_{T,B}^{jet} > 8$ GeV and $-1 < \eta_{lab}^{jet} < 2.5$. The measurement was carried out in the kinematic range $125 < Q^2 < 20000$ GeV^2 and $0.2 < y < 0.6$. An extensive study of the systematic uncertainties has been performed. In particular, multidimensional reweighting of the MC spectra made possible a reliable estimate of the corresponding uncertainty. When compared to other similar analyses from ZEUS, these measurements make use of a large statistics sample collected during the HERA II running phase. The total integrated luminosity of the data amounts to 295 pb^{-1}, which is almost three time large than in the previously published measurement [49]. The NLO pQCD predictions based on different pPDF sets were compared with the measured jet cross sections. In general, theoretical calculations provide a good description of the data within the experimental and theoretical uncertainties. The measurements demonstrated high sensitivity to the proton parton distributions used in the predictions.

Several potential directions can be considered for further improvement of the precision of the jet cross-section measurements. For instance, as was mentioned, the jets with low transverse energy, $3 < E_{T,lab}^{jet} < 10$ GeV, are characterised by larger energy-scale uncertainty than those with $E_{T,lab}^{jet} > 10$ GeV. In order to improve the precision of the jet-energy measurement, especially in the low-$E_{T,lab}^{jet}$ region, the combined information from tracking and calorimeter systems can be used. In particular, an improvement of the jet energy-scale uncertainty is expected in the measurements based on the so-called ZUFOS clusters [211]. Moreover, as was shown, the systematic uncertainty due to the variation of MC models is substantial in this analysis, therefore further improvement of the results can be achieved when more advanced MC models, like those based on higher-order perturbative QCD predictions, are used. Furthermore, an improvement of the precision can be achieved in future if these measurements will be combined with those from the H1 experiment.

From the measured data, a value of the strong coupling at the scale of the Z^0-boson mass has been determined and the running of α_s over a wide range of scales was demonstrated. A detailed analysis of the uncertainties on $\alpha_s(M_Z)$-values was carried out and the importance of the treatment of systematic errors was emphasised. The extracted $\alpha_s(M_Z)$ value is:

$$\alpha_s(M_Z) = 0.1216 \pm 0.0026\,(\text{exp.})^{+0.0066}_{-0.0053}\,(\text{scales})^{+0.0029}_{-0.0018}\,(\text{PDF})\,. \qquad (9.3.6.1)$$

The uncertainty is dominated by the scale uncertainty of the predictions which is an estimate of the effect of missing higher-order terms beyond NLO, therefore in order to take advantage of the full precision of the data, higher-order calculations are vital.

These data contribute to the ultimate legacy of HERA. Further steps in the improvement of the $\alpha_s(M_Z)$ precision can be made if the HERA data were to be used simultaneously with the jet measurements from other high-energy colliders such as the LHC and TEVATRON. Such an extensive data sample will have a strong impact on the precision of future $\alpha_s(M_Z)$ extractions as well as on the proton PDF.

Looking into the far future, the ongoing studies for the lepton-nucleon collider LHeC demonstrate the possibility of accessing an unprecedented kinematic range for lepton-nucleon scattering. LHeC would open a new window to unexplored regions of the electroweak and strong sectors of the Standard Model, where jet studies will play a very significant role.

APPENDIX A

Cross Section Tables

Q^2 bin (GeV^2)	$\mathrm{d}\sigma/\mathrm{d}Q^2$ (pb/GeV^2)	δ_{stat} (%)	$\delta_{\mathrm{syst}}^{\mathrm{uncor}}$ (%)	δ_{JES} (%)	$\mathscr{C}_{\mathrm{hadr}}$	$\mathscr{C}_{Z^0}$	$\mathscr{C}_{\mathrm{QED}}$
125 . . . 250	1.353	0.6	1.1	5.4	0.95	1.00	0.97
250 . . . 500	0.394	0.8	1.8	4.5	0.95	1.00	0.95
500 . . . 1000	0.104	1.0	2.9	3.5	0.95	1.01	0.94
1000 . . . 2000	0.0252	1.4	3.4	2.8	0.94	1.04	0.94
2000 . . . 5000	0.00437	1.9	4.6	2.0	0.94	1.09	0.93
5000 . . . 20000	0.00027	3.4	5.6	1.7	0.94	1.20	0.92

Table A.1.: Inclusive-jet cross-section $\mathrm{d}\sigma/\mathrm{d}Q^2$

$E_{T,B}^{jet}$ bin (GeV)	$\mathrm{d}\sigma/\mathrm{d}E_{T,B}^{jet}$ (pb/GeV)	δ_{stat} (%)	$\delta_{\mathrm{syst}}^{\mathrm{uncor}}$ (%)	δ_{JES} (%)	$\mathscr{C}_{\mathrm{hadr}}$	$\mathscr{C}_{Z^0}$	$\mathscr{C}_{\mathrm{QED}}$
8 . . . 10	67.16	0.7	3.6	4.5	0.92	1.01	0.96
10 . . . 14	33.13	0.7	0.9	5.2	0.96	1.01	0.95
14 . . . 18	12.94	1.1	1.4	3.6	0.98	1.01	0.95
18 . . . 25	4.398	1.3	2.3	3.2	0.98	1.02	0.96
25 . . . 35	1.004	2.2	4.6	3.4	0.98	1.03	0.95
35 . . . 100	0.0417	4.1	10.8	5.5	0.97	1.04	0.94

Table A.2.: Inclusive-jet cross-section $\mathrm{d}\sigma/\mathrm{d}E_{T,B}^{jet}$

η_B^{jet} bin	$\mathrm{d}\sigma/\mathrm{d}\eta_B^{jet}$ (pb)	δ_{stat} (%)	$\delta_{\mathrm{syst}}^{\mathrm{uncor}}$ (%)	δ_{JES} (%)	$\mathscr{C}_{\mathrm{hadr}}$	$\mathscr{C}_{Z^0}$	$\mathscr{C}_{\mathrm{QED}}$
-2.00 … -1.00	8.006	2.4	10.8	3.2	0.79	1.09	0.93
-1.00 … -0.25	48.28	1.1	4.9	4.4	0.81	1.03	0.94
-0.25 … 0.25	130.2	0.9	1.9	4.6	0.90	1.02	0.95
0.25 … 1.00	159.6	0.7	1.2	4.0	0.99	1.01	0.96
1.00 … 1.80	100.8	0.8	1.3	4.5	1.00	1.01	0.95

Table A.3.: Inclusive-jet cross-section $\mathrm{d}\sigma/\mathrm{d}\eta_B^{jet}$

$E_{T,B}^{jet}$ bin (GeV)		δ_{stat} (%)	$\delta_{\text{syst}}^{\text{uncor}}$ (%)	δ_{JES} (%)	$\mathcal{C}_{\text{hadr}}$	$\mathcal{C}_{Z^0}$	$\mathcal{C}_{\text{QED}}$
$125 < Q^2 < 250$ GeV^2							
8 … 10	34.60	0.9	1.8	5.2	0.93	1.00	0.97
10 … 14	15.81	1.0	2.1	6.4	0.96	1.00	0.97
14 … 18	5.427	1.6	3.0	4.3	0.97	1.00	0.97
18 … 25	1.592	2.1	4.2	3.5	0.97	1.01	0.97
25 … 35	0.322	3.9	7.2	3.5	0.96	1.01	0.98
35 … 100	0.0115	7.7	14.7	5.7	0.95	1.01	0.96
$250 < Q^2 < 500$ GeV^2							
8 … 10	18.19	1.3	4.0	4.3	0.93	1.00	0.95
10 … 14	9.186	1.4	1.7	5.2	0.96	1.01	0.95
14 … 18	3.541	2.1	2.8	3.7	0.98	1.01	0.95
18 … 25	1.161	2.8	3.4	3.7	0.99	1.00	0.95
25 … 35	0.243	4.9	6.2	3.3	0.97	1.00	0.95
35 … 100	0.00969	9.5	15.6	6.0	0.97	1.00	0.92
$500 < Q^2 < 1000$ GeV^2							
8 … 10	8.367	1.8	7.1	3.4	0.90	1.01	0.93
10 … 14	4.524	1.8	2.5	3.8	0.96	1.01	0.94
14 … 18	2.195	2.6	3.0	3.1	0.99	1.01	0.94
18 … 25	0.824	3.1	3.5	3.1	0.99	1.00	0.96
25 … 35	0.209	4.9	6.4	3.8	0.98	1.03	0.95
35 … 100	0.00795	9.5	10.7	4.7	0.98	1.04	0.94

Table A.4.: Inclusive-jet cross-section $\mathrm{d}\sigma/\mathrm{d}E_{T,B}^{jet}$ in different regions of Q^2

$E_{T,B}^{jet}$ bin (GeV)		δ_{stat} (%)	$\delta_{\text{syst}}^{\text{uncor}}$ (%)	δ_{JES} (%)	$\mathscr{C}_{\text{hadr}}$	$\mathscr{C}_{Z^0}$	$\mathscr{C}_{\text{QED}}$
$1000 < Q^2 < 2000$ GeV^2							
8 ... 10	3.657	2.6	9.1	3.1	0.90	1.04	0.94
10 ... 14	2.172	2.4	2.8	2.5	0.95	1.03	0.93
14 ... 18	1.070	3.5	4.1	2.4	0.96	1.03	0.94
18 ... 25	0.484	3.9	4.1	2.3	1.00	1.05	0.95
25 ... 35	0.117	6.1	6.8	3.6	0.99	1.05	0.92
35 ... 100	0.00641	10.2	14.5	5.4	0.98	1.06	0.94
$2000 < Q^2 < 5000$ GeV^2							
8 ... 10	1.801	3.7	13.2	2.1	0.91	1.08	0.93
10 ... 14	1.111	3.4	4.0	1.5	0.93	1.09	0.93
14 ... 18	0.552	4.8	5.3	2.3	0.98	1.08	0.92
18 ... 25	0.251	5.2	6.2	1.6	0.99	1.11	0.92
25 ... 35	0.0853	7.1	8.1	2.4	0.99	1.10	0.92
35 ... 100	0.00421	11.2	14.6	5.5	0.97	1.16	0.99
$5000 < Q^2 < 20000$ GeV^2							
8 ... 10	0.555	6.8	15.9	3.5	0.87	1.19	0.94
10 ... 14	0.329	6.3	7.4	1.2	0.93	1.20	0.93
14 ... 18	0.159	8.6	9.2	3.7	0.97	1.18	0.93
18 ... 25	0.0866	8.8	10.0	1.1	1.01	1.23	0.88
25 ... 35	0.03002	11.8	13.2	2.4	0.96	1.23	0.96
35 ... 100	0.00190	15.9	28.3	5.4	1.01	1.18	0.88

Table A.5.: Inclusive-jet cross-section $d\sigma/dE_{T,B}^{jet}$ in different regions of Q^2. Continuation of Table A.4.

		8 . . . 10	10 . . . 14	14 . . . 18	18 . . . 25	25 . . . 35	35 . . . 100
$125 < Q^2 < 250$ GeV^2	8 . . . 10	1.00	0.22	0.33	0.30	0.18	0.09
	10 . . . 14	0.22	1.00	0.37	0.38	0.29	0.21
	14 . . . 18	0.33	0.37	1.00	0.27	0.26	0.16
	18 . . . 25	0.30	0.38	0.27	1.00	0.25	0.24
	25 . . . 35	0.18	0.29	0.26	0.25	1.00	0.24
	35 . . . 100	0.09	0.21	0.16	0.24	0.24	1.00
$250 < Q^2 < 500$ GeV^2	8 . . . 10	1.00	0.21	0.29	0.28	0.23	0.10
	10 . . . 14	0.21	1.00	0.34	0.34	0.27	0.10
	14 . . . 18	0.29	0.34	1.00	0.25	0.27	0.10
	18 . . . 25	0.28	0.34	0.25	1.00	0.35	0.25
	25 . . . 35	0.23	0.27	0.27	0.35	1.00	0.20
	35 . . . 100	0.10	0.10	0.10	0.25	0.20	1.00
$500 < Q^2 < 1000$ GeV^2	8 . . . 10	1.00	0.21	0.34	0.24	0.19	0.07
	10 . . . 14	0.21	1.00	0.41	0.37	0.23	0.13
	14 . . . 18	0.34	0.41	1.00	0.24	0.16	0.13
	18 . . . 25	0.24	0.37	0.24	1.00	0.25	0.22
	25 . . . 35	0.19	0.23	0.16	0.25	1.00	0.30
	35 . . . 100	0.07	0.13	0.13	0.22	0.30	1.00

Table A.6.: Statistical correlations matrix for the double-differential inclusive-jet cross section as a function of $E_{T,B}^{jet}$ in different regions of Q^2

		8...10	10...14	14...18	18...25	25...35	35...100
$1000 < Q^2 < 2000\ \mathrm{GeV}^2$	8...10	1.00	0.17	0.33	0.26	0.22	0.30
	10...14	0.17	1.00	0.38	0.37	0.25	0.26
	14...18	0.33	0.38	1.00	0.23	0.11	0.15
	18...25	0.26	0.37	0.23	1.00	0.30	0.26
	25...35	0.22	0.25	0.11	0.30	1.00	0.45
	35...100	0.30	0.26	0.15	0.26	0.45	1.00
$2000 < Q^2 < 5000\ \mathrm{GeV}^2$	8...10	1.00	0.19	0.26	0.20	0.06	0.04
	10...14	0.19	1.00	0.38	0.24	0.36	0.24
	14...18	0.26	0.38	1.00	0.19	0.26	0.12
	18...25	0.20	0.24	0.19	1.00	0.36	0.08
	25...35	0.06	0.36	0.26	0.36	1.00	0.32
	35...100	0.04	0.24	0.12	0.08	0.32	1.00
$5000 < Q^2 < 20000\ \mathrm{GeV}^2$	8...10	1.00	0.13	0.27	0.15	0.28	0.06
	10...14	0.13	1.00	0.34	0.21	0.15	0.13
	14...18	0.27	0.34	1.00	0.23	0.28	0.19
	18...25	0.15	0.21	0.23	1.00	0.19	0.13
	25...35	0.28	0.15	0.28	0.19	1.00	0.19
	35...100	0.06	0.13	0.19	0.13	0.19	1.00

Table A.7.: Statistical correlations matrix for the double-differential inclusive-jet cross section as a function of $E_{T,B}^{jet}$ in different regions of Q^2. Continuation of Table A.6.

APPENDIX B

Running of the Strong Coupling

In this section the values of the strong coupling extracted as a function of the hard scale are provided.

$\langle E^{jet}_{T,B}\rangle$ (GeV)	α_s	δ_{exp}
8.86	0.195182	0.011196
11.5	0.182313	0.009220
15.45	0.166910	0.006396
20.25	0.156607	0.00600
28.06	0.146688	0.008063
39.76	0.132095	0.015568

Table B.1.: The α_s values determined in each $\langle E^{jet}_{T,B}\rangle$ from the analysis of the measured $d\sigma/dE^{jet}_{T,B}$ cross sections.

APPENDIX C

Corrections and Reweightings

In this section the supplementary materials to the Chapter 6 are provided.

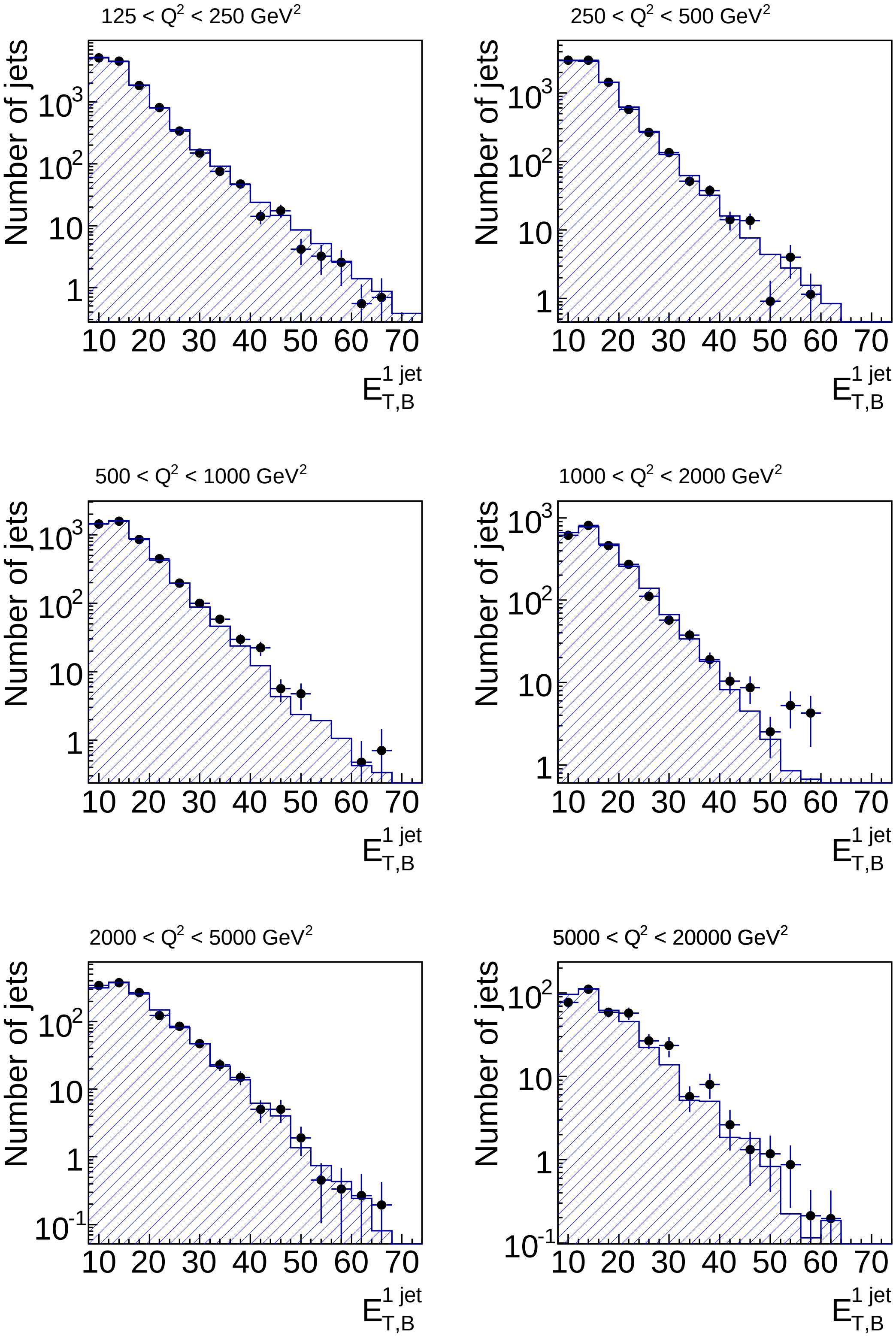


Figure C.1.: Comparison of transverse energy of the leading-jet spectra in data and ARIADNE after MC reweighting for events containing exactly two jets.

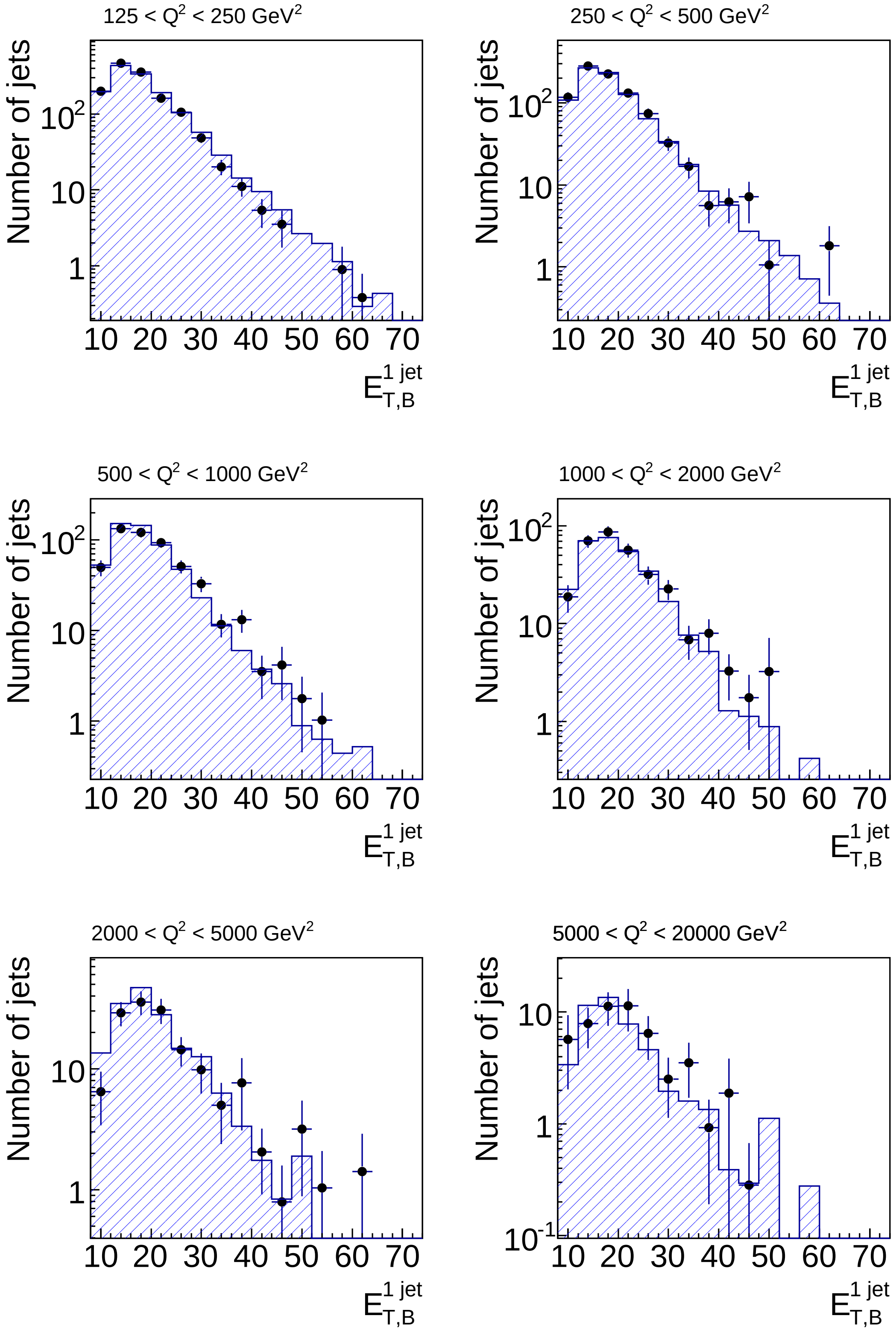


Figure C.2.: Comparison of transverse energy of the leading-jet spectra in data and ARIADNE after MC reweighting for events containing exactly three or more jets.

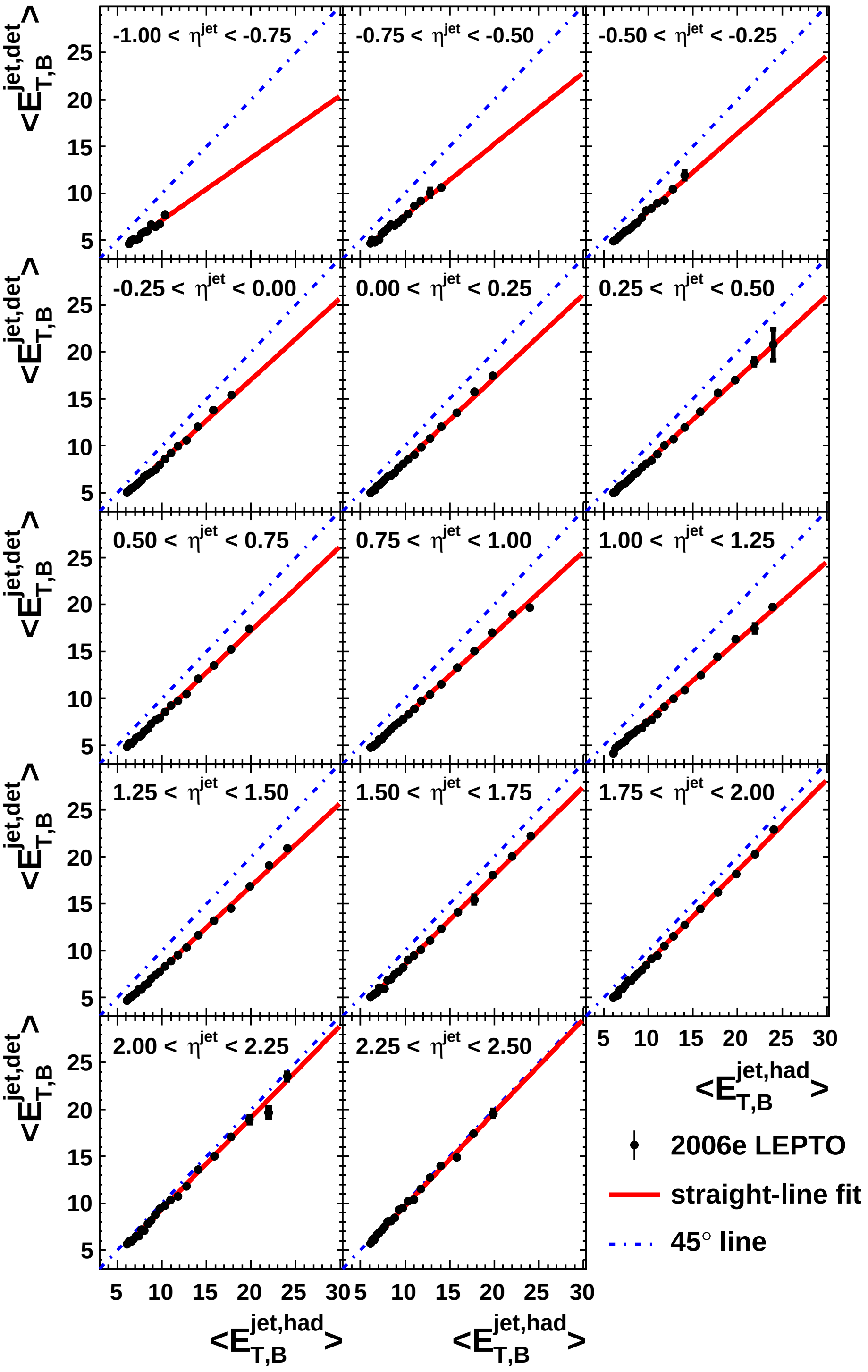


Figure C.3.: The average measured detector-level jet transverse energy $\left\langle E_{T,B}^{jet,det}\right\rangle$ as a function of $\left\langle E_{T,B}^{jet,had}\right\rangle$ and the corresponding straight-line fits in different regions of η_{lab}^{jet} for the data-taking period 2006 e^{-} in the LEPTO MC sample.

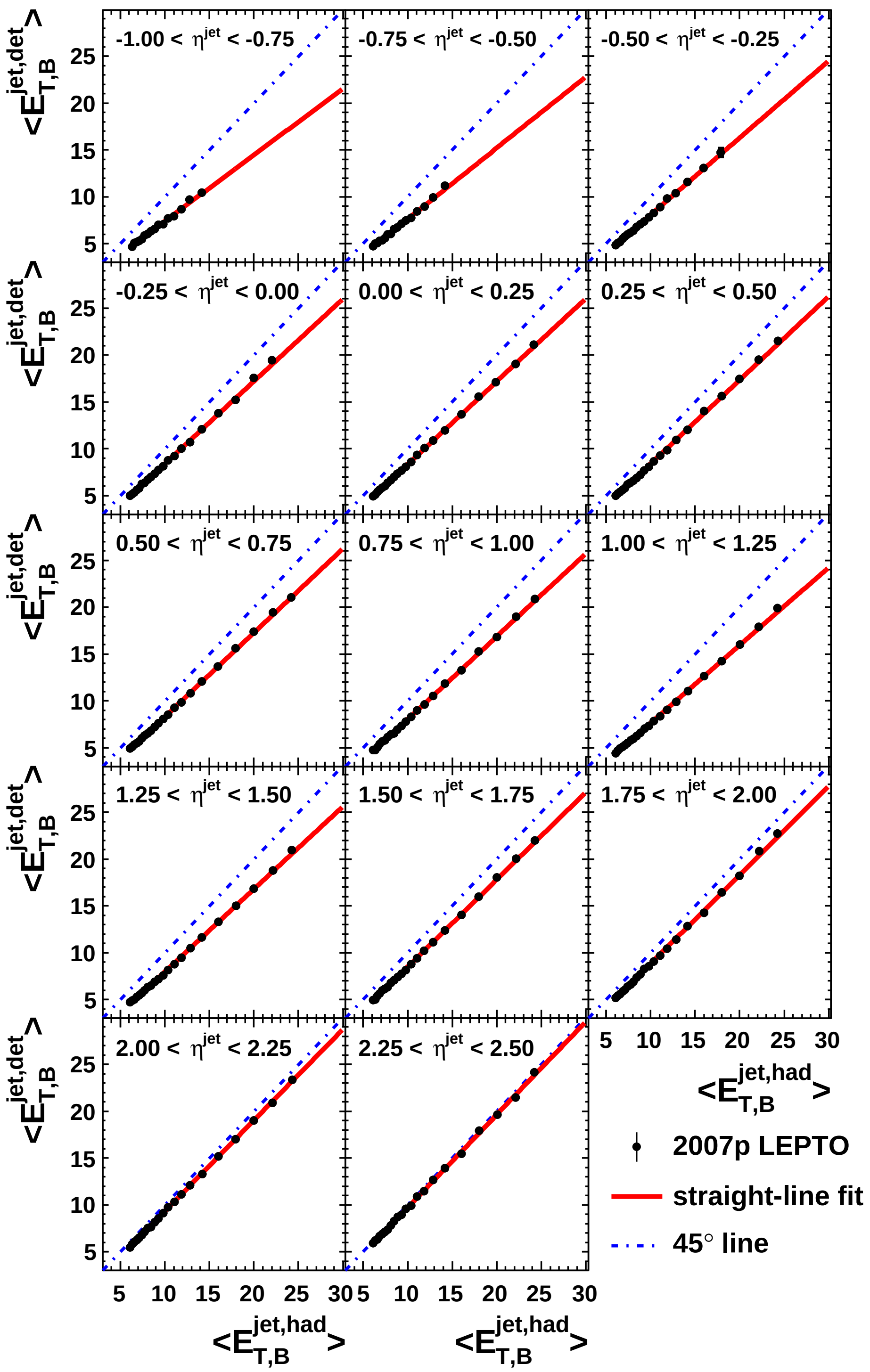


Figure C.4.: The average measured detector-level jet transverse energy $\left\langle E_{T,B}^{jet,det}\right\rangle$ as a function of $\left\langle E_{T,B}^{jet,had}\right\rangle$ and the corresponding straight-line fits in different regions of η_{lab}^{jet} for the data-taking period 2007 e^+ in the LEPTO MC sample.

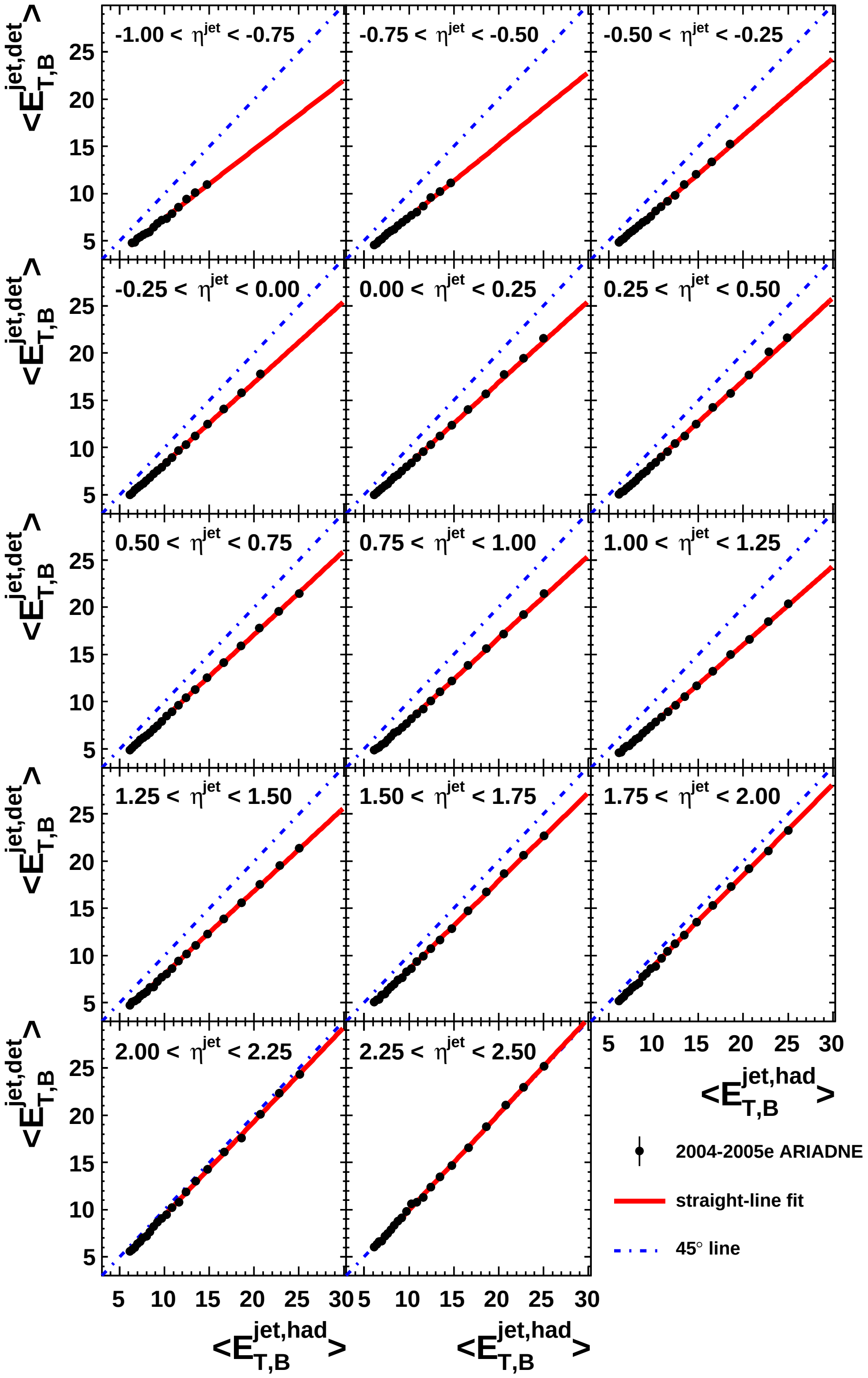


Figure C.5.: The average measured detector-level jet transverse energy $\left\langle E_{T,B}^{jet,det} \right\rangle$ as a function of $\left\langle E_{T,B}^{jet,had} \right\rangle$ and the corresponding straight-line fits in different regions of η_{lab}^{jet} for the data-taking period 2004–2005 e^- in the ARIADNE MC sample.

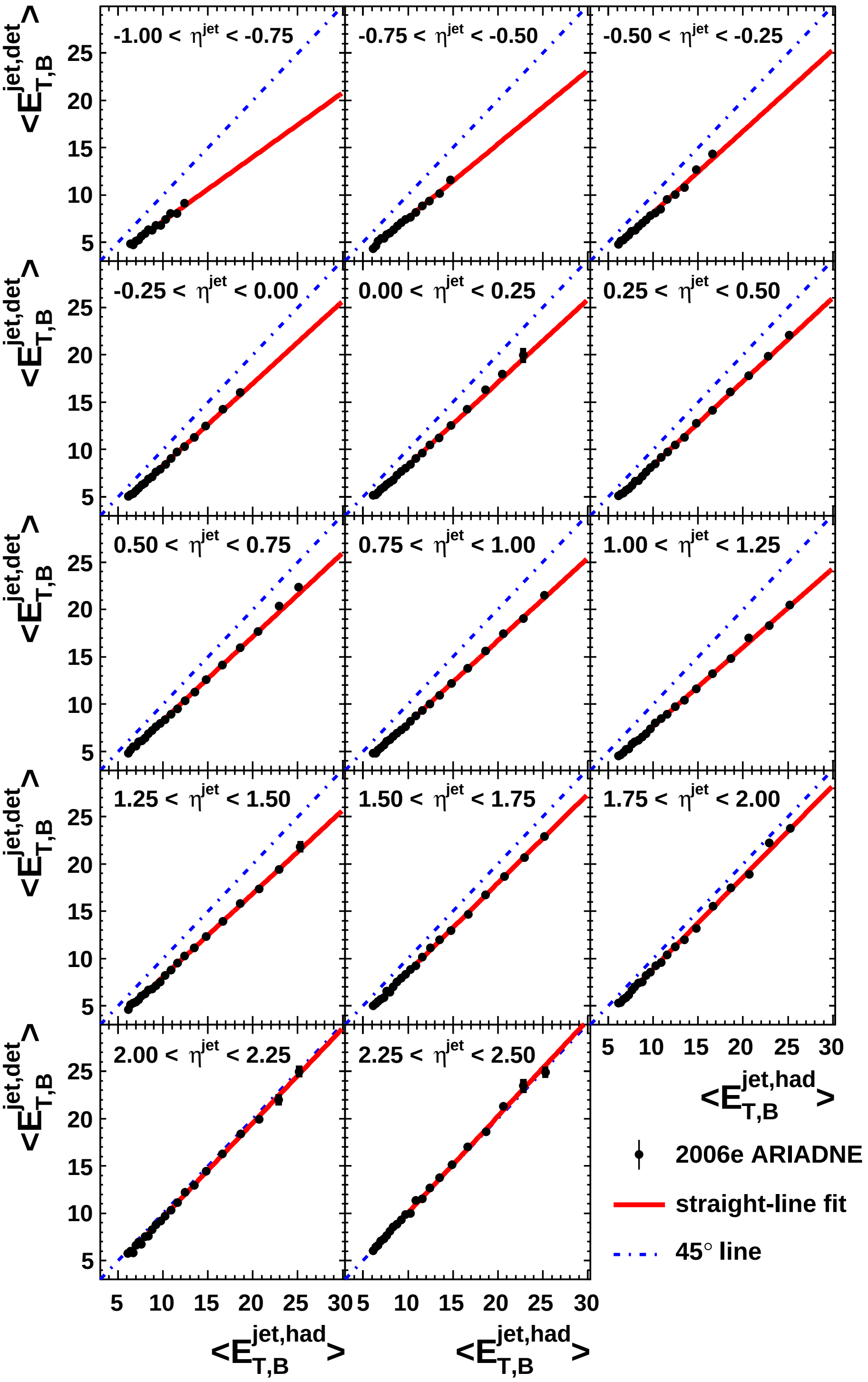


Figure C.6.: The average measured detector-level jet transverse energy $\left\langle E_{T,B}^{jet,det}\right\rangle$ as a function of $\left\langle E_{T,B}^{jet,had}\right\rangle$ and the corresponding straight-line fits in different regions of η_{lab}^{jet} for the data-taking period 2006 e^- in the ARIADNE MC sample.

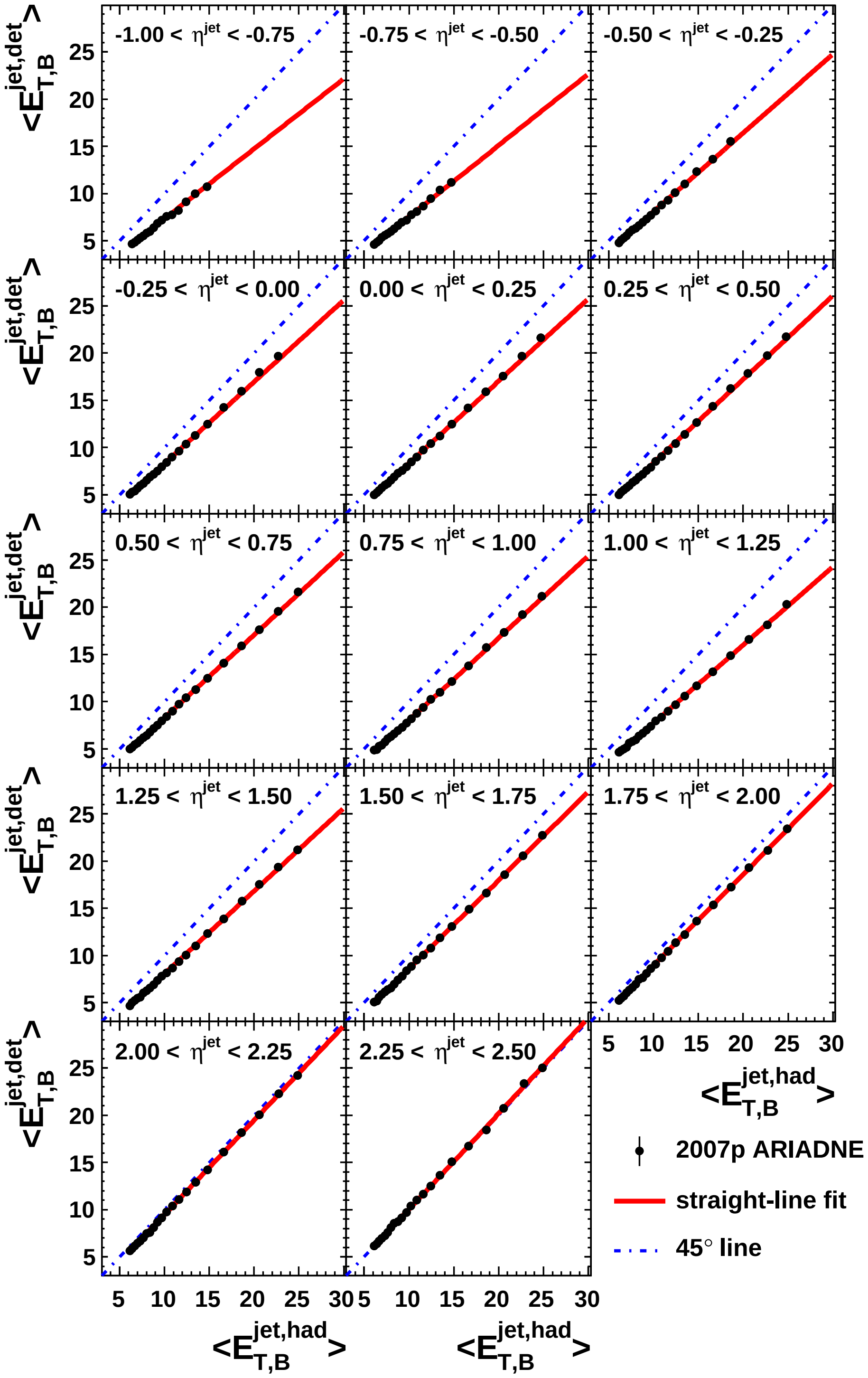


Figure C.7.: The average measured detector-level jet transverse energy $\left\langle E^{jet,det}_{T,B} \right\rangle$ as a function of $\left\langle E^{jet,had}_{T,B} \right\rangle$ and the corresponding straight-line fits in different regions of η^{jet}_{lab} for the data-taking period 2007 e^+ in the ARIADNE MC sample.

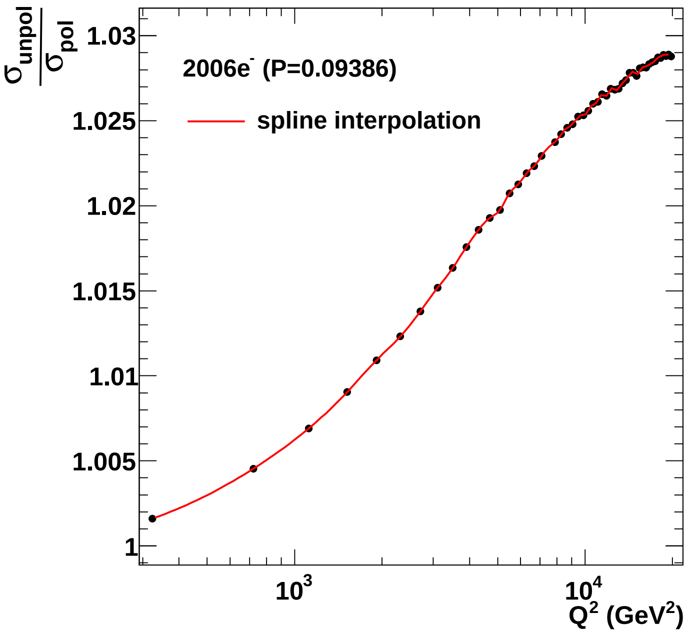


Figure C.8.: The polarisation reweighting factors for the 2006 e^- data taking period determined using the HECTOR program. The red curve represents the spline interpolation.

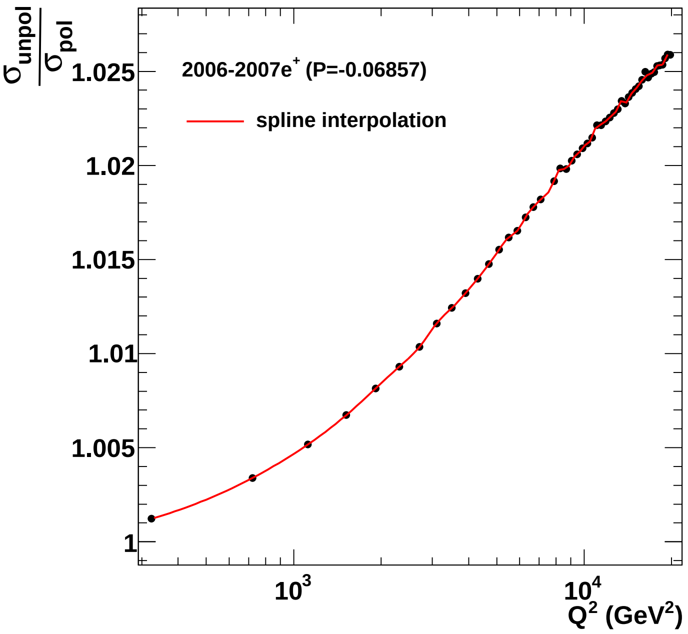


Figure C.9.: The polarisation reweighting factors for the 2006–2007 e^+ data taking period determined using the HECTOR program. The red curve represents the spline interpolation.

APPENDIX D

MC Validation

This section contains pull distributions for individual single-differential cross-section measurements as functions of Q^2, $E_{T,B}^{jet}$ and η_B^{jet} obtained using LEPTO and ARIADNE simulation.

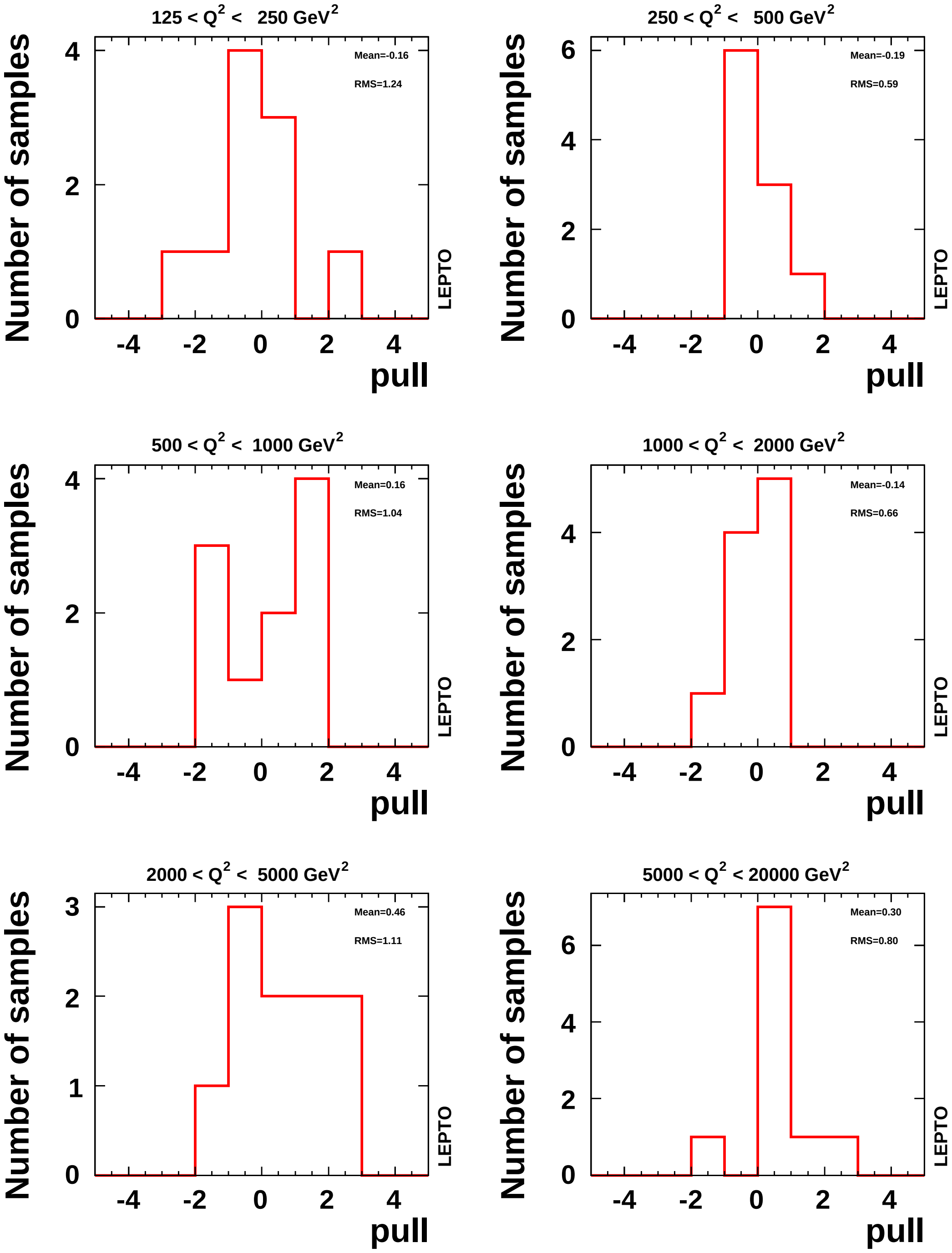


Figure D.1.: Pull distributions for individual $d\sigma/dQ^2$ cross-section bins in LEPTO MC.

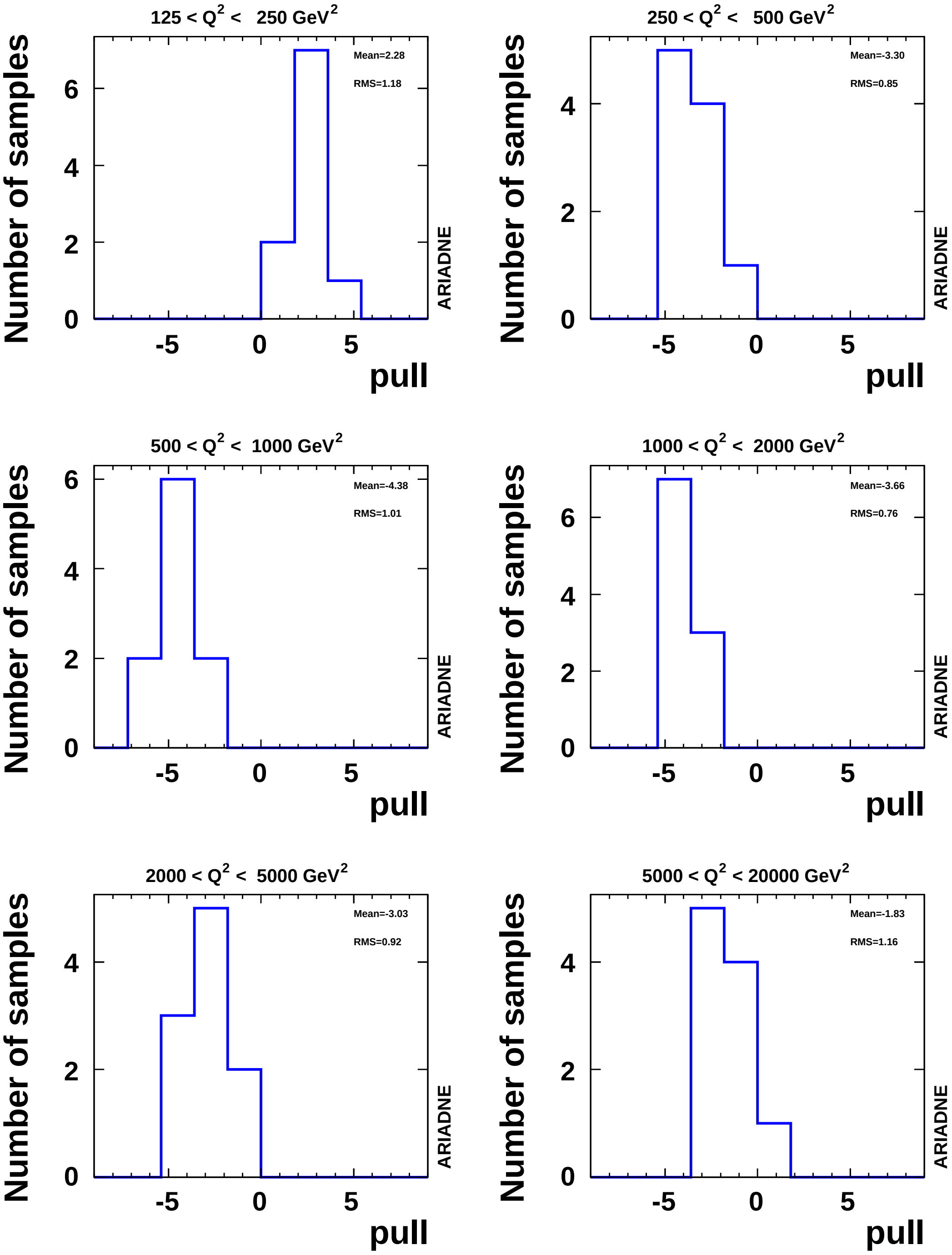


Figure D.2.: Pulls distributions for individual $d\sigma/dQ^2$ cross-section bins in ARIADNE MC.

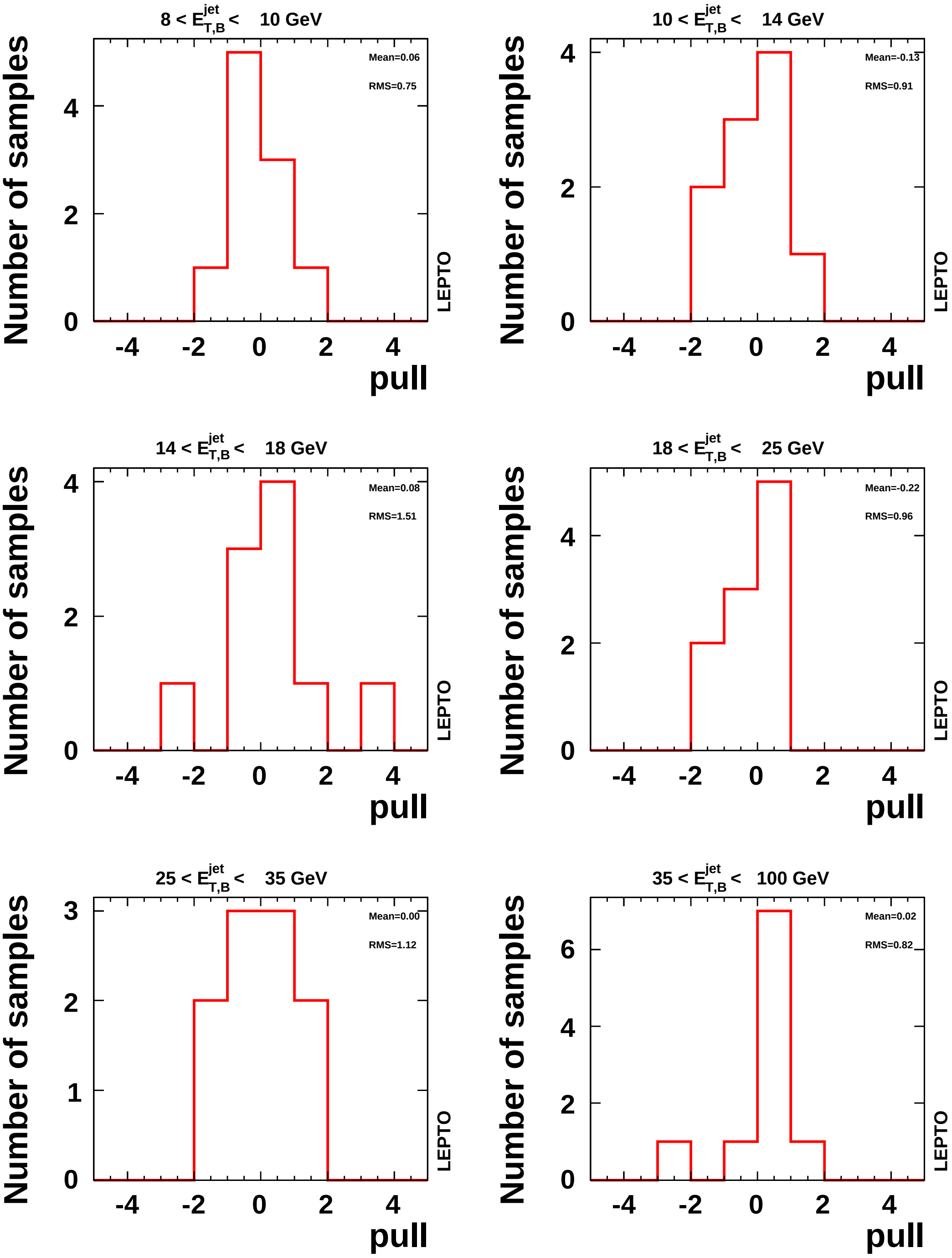


Figure D.3.: Pull distributions for individual $d\sigma/dQ^2$ cross section-bins in LEPTO MC.

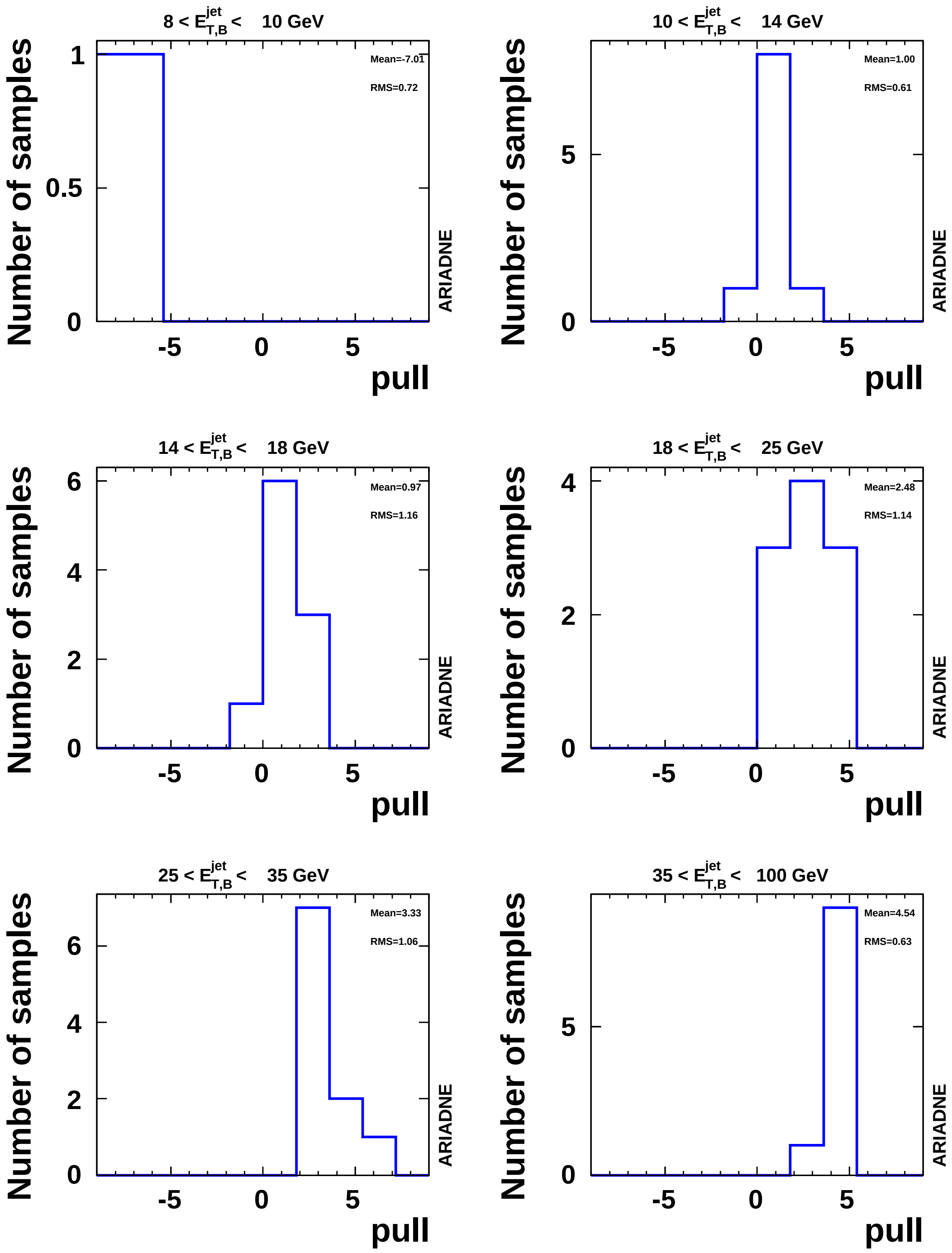


Figure D.4.: Pulls distributions for individual $d\sigma/dE_{T,B}^{jet}$ cross-section bins in ARIADNE MC.

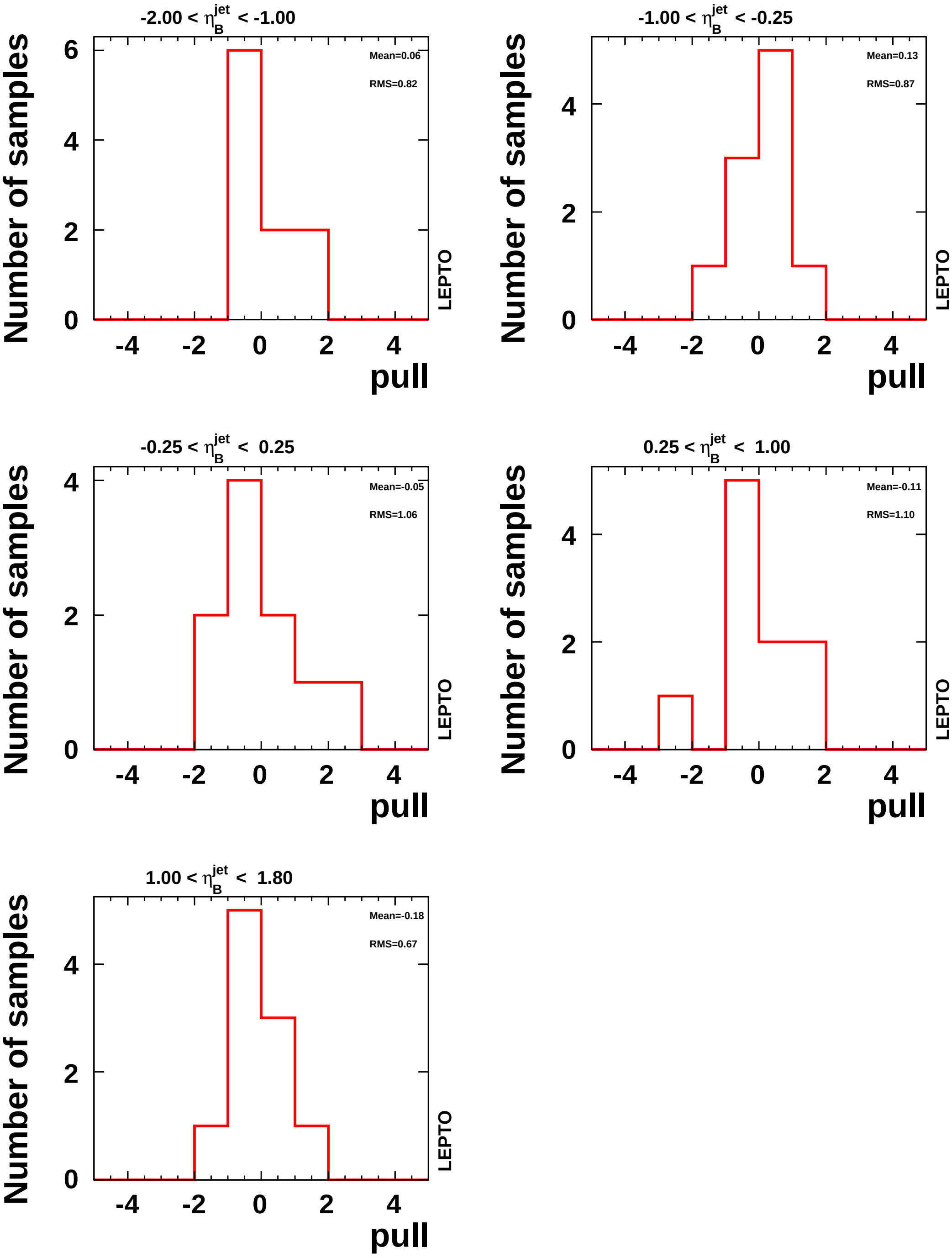


Figure D.5.: Pull distributions for individual $d\sigma/\eta_B^{jet}$ cross-section bins in LEPTO MC.

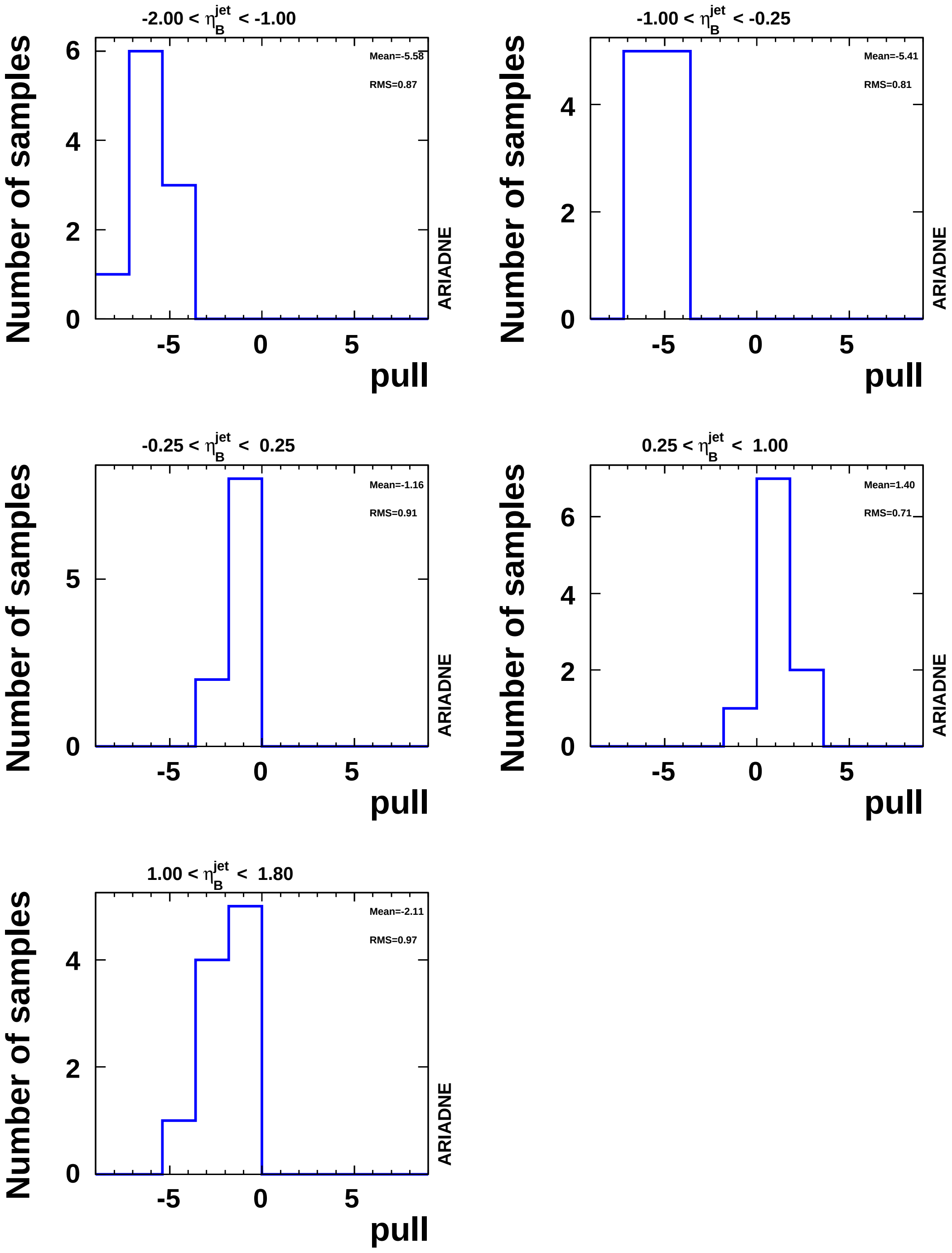


Figure D.6.: Pulls distributions for individual $d\sigma/\eta_B^{jet}$ cross-section bins in ARIADNE MC.

List of Figures

List of Tables

Bibliography

[1] E. Rutherford, *The scattering of α and β particles by matter and the structure of the atom*, *Phil. Mag.* **21** (1911) 669.

[2] J. Chadwick, *Possible existence of a neutron*, *Nature* **129** (1932) 312.

[3] R. Frisch and O. Stern, *über die magnetische ablenkung von wasserstoffmolekulen und das magnetische moment des protons. i (magnetic deflection of hydrogen molecules and the magnetic moment of the proton.*, *Z. Phys.* **85** (1933) 4.

[4] R. Bacher, *Note on the magnetic moment of the nitrogen nucleus*, *Phys. Rev.* **43** (1933) 1001.

[5] L. Alvarez and F. Bloch, *A quantitative determination of the neutron moment in absolute nuclear magnetons*, *Phys. Rev.* **57** (1940) 111.

[6] R. Hofstadter, *Electron scattering and nuclear structure. A collection of reprints with and introduction.* New York, Benjamin, 1963.

[7] W.K.H. Panofsky, *Electromagnetic interactions: Low Q^2 electrodynamics: Elastic and inelastic electron (and muon) scattering*, *Proceedings 14th International Conference of High-Energy Physics*, J.Prentki and J. Steinberger (eds.), p. 23. Vienna, Austria (1968).

[8] E.D. Bloom et al., *High-energy inelastic ep scattering at 6-degrees and 10-degrees*, *Phys. Rev. Lett.* **23** (1969) 930.

[9] M. Briedenbach et al., *High-energy inelastic ep scattering at 6-degrees and 10-degrees*, *Phys. Rev. Lett.* **23** (1969) 935.

[10] R.E. Taylor, *Deep inelastic scattering: The early years*, *Rev. Mod. Phys.* **63** (1991) 573.

[11] H.W. Kendall, *Deep inelastic scattering: Experiments on the proton and the observation*, *Rev. Mod. Phys.* **63** (1991) 597.

[12] K.A. Olive et al. (Particle Data Group), *Statistics*, *Chin. Phys.* **C 38** (2014) 090001.

[13] H1 collaboration, C. Adloff et al., *Measurement and QCD analysis of jet cross-sections in deep inelastic positron–proton collisions at* $\sqrt{s} = 300 GeV$, *Eur. Phys. J.* **C 19** (2001) 289, [`arXiv:hep-ex/0010054`].

[14] ZEUS collaboration, S. Chekanov et al., *Inclusive jet cross sections in the Breit frame in neutral current deep inelastic scattering at HERA and determination of* α_s, *Phys. Lett.* **B 547** (2002) 164, [`arXiv:hep-ex/0208037`].

[15] ZEUS collaboration, S. Chekanov et al., *Study of the azimuthal asymmetry of jets in neutral current deep inelastic scattering at HERA*, *Phys. Lett.* **B 551** (2003) 226, [`arXiv:hep-ex/0210064`].

[16] ZEUS collaboration, S. Chekanov et al., *Inclusive-jet and dijet cross-sections in deep inelastic scattering at HERA*, *Nucl. Phys.* **B 765** (2007) 1, [`arXiv:hep-ex/0608048`].

[17] ZEUS collaboration, S. Chekanov et al., *Jet-radius dependence of inclusive-jet cross-sections in deep inelastic scattering at HERA*, *Phys. Lett.* **B 649** (2007) 12, [`arXiv:hep-ex/0701039`].

[18] H1 collaboration, F.D. Aaron et al., *Jet production in* ep *collisions at high* Q^2 *and determination of* α_s, *Eur. Phys. J.* **C 65** (2010) 363, [`arXiv:0904.3870`].

[19] H1 collaboration, F.D. Aaron et al., *Jet production in* ep *collisions at low* Q^2 *and determination of* α_s, *Eur. Phys. J.* **C 67** (2010) 1, [`arXiv:0911.5678`].

[20] ZEUS collaboration, J. Breitweg et al., *Measurement of dijet production in neutral current deep inelastic scattering at high* Q^2 *and determination of* α_s, *Phys. Lett.* **B 507** (2001) 70, [`arXiv:hep-ex/0102042`].

[21] ZEUS collaboration, S. Chekanov et al., *Dijet production in neutral current deep inelastic scattering at HERA*, *Eur. Phys. J.* **C 23** (2002) 13, [`arXiv:hep-ex/0109029`].

[22] H1 collaboration, C. Adloff et al., *Three-jet production in deep inelastic scattering at HERA*, *Phys. Lett.* **B 515** (2001) 17, [`arXiv:hep-ex/0106078`].

[23] ZEUS collaboration, S. Chekanov et al., *Multijet production in neutral current deep inelastic scattering at HERA and determination of* α_s, *Eur. Phys. J.* **C 44** (2005) 183, [`arXiv:hep-ex/0502007`].

[24] ZEUS collaboration, H. Abramowicz et al., *Angular correlations in three-jet events in* ep *collisions at HERA*, *Phys. Rev.* **D 85** (2012) 052008, [`arXiv:0808.3783`].

[25] ZEUS collaboration, S. Chekanov et al., *Jet production in charged current deep inelastic* e^+p *scattering at HERA*, *Eur. Phys. J.* **C 31** (2003) 149, [`arXiv:hep-ex/0306018`].

[26] ZEUS collaboration, S. Chekanov et al., *Multi-jet cross-sections in charged current* $e^{\pm}p$ *scattering at HERA*, *Phys. Rev.* **D 78** (2008) 032004, [`arXiv:0802.3955`].

[27] ZEUS collaboration, S. Chekanov et al., *Scaling violations and determination of* α_s *from jet production in* γp *interactions at HERA*, *Phys. Lett.* **B 560** (2003) 7, [`arXiv:hep-ex/0212064`].

[28] H1 collaboration, C. Adloff et al., *Measurement of inclusive jet cross-sections in photoproduction at HERA*, *Eur. Phys. J.* **C 29** (2003) 497, [`arXiv:hep-ex/0302034`].

[29] ZEUS collaboration, J. Breitweg et al., *Measurement of dijet photoproduction at high transverse energies at HERA*, *Eur. Phys. J.* **C 11** (1999) 35, [`arXiv:hep-ex/9905046`].

[30] ZEUS collaboration, S. Chekanov et al., *Dijet photoproduction at HERA and the structure of the photon*, *Eur. Phys. J.* **C 23** (2002) 615, [`arXiv:hep-ex/0112029`].

[31] ZEUS collaboration, S. Chekanov et al., *High-mass dijet cross sections in photoproduction at HERA*, *Phys. Lett.* **B 531** (2002) 9, [`arXiv:hep-ex/0112030`].

[32] H1 collaboration, C. Adloff et al., *Measurement of dijet cross-sections in photoproduction at HERA*, *Eur. Phys. J.* **C 25** (2002) 497, [`arXiv:hep-ex/0201006`].

[33] H1 collaboration, A. Aktas et al., *Photoproduction of dijets with high transverse momenta at HERA*, *Phys. Lett.* **B 639** (2006) 21, [`arXiv:hep-ex/0201006`].

[34] ZEUS collaboration, S. Chekanov et al., *High-e_t dijet photoproduction at HERA*, *Phys. Rev.* **D 76** (2007) 072011, [`arXiv:0706.3809`].

[35] ZEUS collaboration, J. Breitweg et al., *Measurement of three-jet distributions in photoproduction at HERA*, *Phys. Lett.* **B 443** (1998) 394, [`arXiv:hep-ex/9810046`].

[36] ZEUS collaboration, S. Chekanov et al., *Three- and four-jet final states in photoproduction at HERA*, *Nucl. Phys.* **B 792** (2008) 1, [`arXiv:0707.3749`].

[37] ZEUS collaboration, S. Chekanov et al., *An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment*, *Eur. Phys. J.* **C 42** (2005) 1, [`arXiv:hep-ph/0503274`].

[38] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B 716** (2012) 1, [`arXiv:1207.7214`].

[39] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 gev with the CMS experiment at the LHC*, *Phys. Lett.* **B 716** (2012) 30, [`arXiv:1207.7235`].

[40] J.D. Bjorken, *Asymptotic sum rules at infinite momentum*, *Phys. Rev.* **179** (1969) 1547.

[41] R.P. Feynman, *Very high-energy collisions of hadrons*, *Phys. Rev. Lett.* **23** (1969) 1415.

[42] R.P. Feynman, *Photon-hadron interactions*. 1973.

[43] H1 and ZEUS collaboration, H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, *submitted to Eur. Phys. J* (2015), [`arXiv:1506.0604`].

[44] J.E. Huth et al., *Toward a standardization of jet definitions*, *Research Directions for the Decade. Proceedings of Summer Study on High Energy Physics, 1990*, E.L. Berger (ed.), p. 134. World Scientific (1992). Also in preprint FERMILAB-CONF-90-249-E.

[45] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k(t) jet clustering algorithm*, *JHEP* **0804** (2008) 063, [`arXiv:0802.1189`].

[46] Yu.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **9708** (1997) 001, [`arXiv:hep-ph/9707323`].

[47] S. Catani, Yu.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant k_T clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B 406** (1993) 187.

[48] ZEUS collaboration, H. Abramowicz et al., *Inclusive-jet photoproduction at HERA and determination of α_s*, *Nucl. Phys.* **B 864** (2012) 1, [`arXiv:1205.6153`].

[49] ZEUS collaboration, H. Abramowicz et al., *Inclusive-jet cross sections in NC DIS at HERA and a comparison of the k_T, anti-k_T and SIScone jet algorithms*, *Phys. Lett.* **B 691** (2010) 127, [`arXiv:1003.2923`].

[50] R.P. Feynman, *Photon-Hadron Interactions*, in Frontiers in Physics. W.A. Benjamin Inc., Reading, MA, 1972.

[51] K.H. Streng, T.F. Walsh and P.M. Zerwas, *Quark and gluon jets in the Breit frame of lepton-nucleon scattering*, *Z. Phys.* **C 2** (1979) 237.

[52] C. Yang and R.L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (1954) 191.

[53] L.D. Faddeev and V.N. Popov, *Feynman diagrams for the yang-mills field*, *Phys. Lett.* **B 25** (1967) 29.

[54] B.S. DeWitt, *Theory of radiative corrections for non-abelian gauge fields*, *Phys. Rev. Lett.* **12** (1964) 742.

[55] J.M. Henn and J.C. Plefka, *Scattering Amplitudes in Gauge Theories. Lecture Notes in Physics 883.* Springer-Verlag, 2014.

[56] B.L. Ioffe, V.S Fadin and L.N. Lipatov, *Quantum chromodynamics: Perturbative and nonperturbative aspects.* Wiley-VCH, 2010.

[57] Bardeen, William A. and Buras, A.J. and Duke, D.W. and Muta, T., *Deep inelastic scattering beyond the leading order in asymptotically free gauge theories*, *Phys. Rev.* **D 18** (1978) 3998.

[58] K.A. Olive et al. (Particle Data Group), *Quantum chromodynamics*, *Chin. Phys.* **C 38** (2014) 090001.

[59] J. Collins et al., *Factorization of hard processes in qcd, Adv. Ser. Direct. High Energy Phys.* **5** (1988) 1, [`arXiv:hep-ph/0409313`].

[60] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language, Nucl. Phys.* **B 126** (1977) 298.

[61] V.N. Gribov and L.N. Lipatov, *e^+e^- pair annihilation and deep inelastic ep scattering in perturbation theory, Sov. J. Nucl. Phys.* **15** (1972) 438.

[62] Yu.L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e^+e^- pair annihilation by perturbation theory in quantum choromodynamics, Sov. Phys. JETP* **46** (1977) 641.

[63] I.I. Balitsky and L.N. Lipatov, *The pomeranchuk singularity in quantum chromodynamics, Sov. J. Nucl. Phys.* **28** (1978) 822.

[64] A. Vogt, S. Moch, and J.A.M. Vermaseren, *The three-loop splitting functions in qcd: The singlet case, Nucl. Phys.* **B 691** (2004) 129, [`arXiv:hep-ph/0404111`].

[65] A. Vogt, S. Moch, and J.A.M. Vermaseren, *The three-loop splitting functions in qcd: The nonsinglet case, Nucl. Phys.* **B 688** (2004) 101, [`arXiv:hep-ph/0403192`].

[66] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, *The pomeranchuk singularity in nonabelian gauge theories, Sov. Phys. JETP* **45** (1977) 199.

[67] M. Ciafaloni, *Coherence effects in initial jets at small Q^2, Nucl. Phys.* **B 296** (1988) 49.

[68] S. Catani, F. Fiorani and G. Marchesini, *Small x behavior of initial state radiation in perturbative qcd, Nucl. Phys.* **B 336** (1990) 18.

[69] S. Catani, F. Fiorani and G. Marchesini, *Qcd coherence in initial state radiation, Phys. Lett.* **B 234** (1990) 339.

[70] A.D. Martin, *Bfkl / ccfm phenomenology*, p. 156. Rome, Italy (1996).

[71] O. Behnke, K. Krninger, T. Schrner-Sadenius and G. Schott, *Data analysis in high energy physics*. Wiley-VCH, 2013.

[72] M. Botje, *Qcdnum: Fast qcd evolution and convolution, Comput. Phys. Commun.* **182** (2011) 490, [`arXiv:1005.1481`].

[73] P.M. Stevenson, *Resolution of the renormalization scheme ambiguity in perturbative qcd, Phys. Lett.* **B100** (1981) 61.

[74] D. Britzger, *Regularized Unfolding of Jet Cross Sections in Deep-Inelastic ep Scattering at HERA and Determination of the Strong Coupling,* Ph.D. Thesis, Hamburg University, Hamburg, Germany, Report DESY-THESIS-2013-045, 2013.

[75] T. Kinoshita, *Mass singularities of feynman amplitudes, J. Math. Phys.* **3** (1962) 650.

[76] T.D. Lee and M. Nauenberg, *Degenerate systems and mass singularities, Phys. Rev.* **133** (1964) B 1549.

[77] S. Catani and M.H. Seymour, *A general algorithm for calculating jet cross-sections in nlo qcd, Nucl. Phys.* **B 485** (1997) 291, [`arXiv:hep-ph/9605323`].

[78] A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover and J. Pires, *Second order qcd corrections to jet production at hadron colliders: the all-gluon contribution, Phys. Rev. Lett.* **110** (2013) 162003, [`arXiv:1301.7310`].

[79] S. Frixione, P. Nason and C. Oleari, *Matching nlo qcd computations with parton shower simulations: the powheg method, JHEP* **0711** (2007) 070, [`arXiv:0709.2092`].

[80] S. Frixione and B.R. Webber, *Matching nlo qcd computations and parton shower simulations, JHEP* **0206** (2002) 029, [`arXiv:hep-ph/0204244`].

[81] G. Gustafson and U. Pettersson, *Dipole formulation of QCD cascades, Nucl. Phys.* **B 306** (1988) 746.

[82] B. Andersson, P. Dahlkvist and G. Gustafson, *An infrared stable multiplicity measure on qcd parton states, Phys. Lett.* **B 214** (1988) 604.

[83] B. Andersson, G. Gustafson, A. Nilsson and C. Sjogren, *Fluctuations and anomalous dimensions in qcd cascades.*

[84] G. Gustafson, *Multiplicity distributions in qcd cascades, Nucl. Phys.* **B 392** (1993) 251.

[85] B. Andersson, G. Gustafson, L. Lonnblad and U. Pettersson, *Coherence effects in deep inelastic scattering.*

[86] G. Gustafson, *Dual description of a confined color field, Phys. Lett.* **B 175** (1986) 453.

[87] B. Andersson, G. Gustafson and G. Ingelman and T. Sjostrand, *Parton fragmentation and string dynamics, Phys. Rept.* **97** (1983) 31.

[88] B.R. Webber, *A qcd model for jet fragmentation including soft gluon interference, Nucl. Phys.* **B 238** (1984) 492.

[89] R.D. Field, and S.Wolfram, *A qcd model for e+ e- annihilation, Nucl. Phys.* **B 213** (1983) 65.

[90] D. Amati and G. Veneziano, *Preconfinement as a property of perturbative qcd, Phys. Lett.* **B 83** (1979) 87.

[91] A. Kwiatkowski, H. Spiesberger and H.-J. Möhring, HERACLES*: An event generator for ep interactions at HERA energies including radiative processes, Comp. Phys. Comm.* **69** (1992) 155. Also in *Proc. Workshop Physics at HERA*, eds. W. Buchmüller and G. Ingelman, (DESY, Hamburg, 1991).

[92] K. Charchula, G.A. Schuler and H. Spiesberger, *Combined QED and QCD radiative effects in deep inelastic lepton-proton scattering: The Monte Carlo generator DJANGO6, Comp. Phys. Comm.* **81** (1994) 381.

[93] Ingelman, G. and Edin, A. and Rathsman, J., *Lepto 6.5: A monte carlo generator for deep inelastic lepton-nucleon scattering, Comput. Phys. Commun.* **101** (1997) 108, [`arXiv:hep-ph/9605286`].

[94] L. Lönnblad, ARIADNE *version 4 – A program for simulation of QCD cascades implementing the colour dipole model, Comp. Phys. Comm.* **71** (1992) 15.

[95] L. Lonnblad, *Rapidity gaps and other final state properties in the color dipole model for deep inelastic scattering.*

[96] H.L. Lai et al., *Global QCD analysis and the CTEQ parton distributions, Phys. Rev.* **D 51** (1995) 4763.

[97] T. Sjöstrand and M. Bengtsson, *The Lund Monte Carlo for jet fragmentation and e^+e^- physics:* JETSET *version 6.3: An update, Comp. Phys. Comm.* **43** (1987) 367.

[98] U. Schneekloth (ed.), *The HERA luminosity upgrade,* Preprint HERA-98-05, DESY, 1998, `http://www.desy.de/~ahluwali/HERA-98-05/`.

[99] A.A. Sokolov and I.M. Ternov, *On polarization and spin effects in the theory of synchrotron radiation, Sov. Phys. Dokl.* **8** (1964) 1203.

[100] V.N. Baier and V.A. Khoze, *Determination of the transverse polarization of high-energy electrons*, *Sov. J. Nucl. Phys.* **9** (1969) 238.

[101] D.P. Barber et al., *The first achievement of longitudinal spin polarization in a high-energy electron storage ring*, *Phys. Lett.* **B 343** (1995) 436.

[102] ZEUS collaboration, U. Holm (ed.), *The ZEUS Detector*. Status Report (unpublished), DESY (1993). `http://www-zeus.desy.de/bluebook/bluebook.html`.

[103] A. Polini et al., *The design and performance of the ZEUS microvertex detector*, *Nucl. Instrum. Meth.* **A 581** (2007) 656, [`arXiv:0708.3011`].

[104] ZEUS collaboration, *A Microvertex Detector for ZEUS*, Proposal DESY PRC 97/01, DESY, 1997. Also as internal ZEUS Note ZEUS-97-006.

[105] B. Foster et al., *The design and construction of the ZEUS central tracking detector*, *Nucl. Instrum. Meth.* **A 338** (1994) 254.

[106] ZEUS collaboration, UK group, C.B. Brooks et al., *Development of the ZEUS central tracking detector*, *Nucl. Instrum. Meth.* **A 283** (1989) 477.

[107] N. Harnew et al., *Vertex triggering using time difference measurements in the ZEUS central tracking detector*, *Nucl. Instrum. Meth.* **A 279** (1989) 290.

[108] B. Foster et al., *The performance of the ZEUS central tracking detector* z*-by-timing electronics in a transputer based data acquisition system*, *Nucl. Phys. Proc. Suppl.* **B 32** (1993) 181.

[109] 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.

[110] M. Derrick et al., *Design and construction of the ZEUS barrel calorimeter*, *Nucl. Instrum. Meth.* **A 309** (1991) 77.

[111] A. Andresen et al., *Construction and beam test of the ZEUS forward and rear calorimeter*, *Nucl. Instrum. Meth.* **A 309** (1991) 101.

[112] A. Caldwell et al., *Design and implementation of a high-precision readout system for the ZEUS calorimeter*, *Nucl. Instrum. Meth.* **A 321** (1992) 356.

[113] A. Bernstein et al., *Beam tests of the ZEUS barrel calorimeter*, *Nucl. Instrum. Meth.* **A 336** (1993) 23.

[114] I. Kudla et al., *Test of a prototype of the ZEUS backing calorimeter, Nucl. Instrum. Meth.* **A 300** (1991) 480.

[115] H. Bethe, W. Heitler, *On the stopping of fast particles and on the creation of positive electrons, Proc. Roy. Soc. Lond.* **A 146** (1934) 83.

[116] M. Helbich et al., *The spectrometer system for measuring ZEUS luminosity at HERA,* 2005, [`arXiv:physics/0512153`]. Submitted to Nucl. Inst. Meth.

[117] J. Andruszków et al., *First measurement of HERA luminosity by ZEUS lumi monitor,* Preprint DESY-92-066, DESY, 1992.

[118] ZEUS collaboration, M. Derrick et al., *Measurement of the total and partial photon proton cross sections at 180 gev center of mass energy, Z. Phys.* **C 63** (1994) 391.

[119] J. Andruszków et al., *Luminosity measurement in the ZEUS experiment, Acta Phys. Pol.* **B 32** (2001) 2025.

[120] M. Beckmann et al., *The longitudinal polarimeter at HERA, Nucl. Instrum. Meth.* **A 479** (2002) 334.

[121] D.P. Barber et al., *The HERA polarimeter and the first observation of electron spin polarization at HERA, Nucl. Instrum. Meth.* **A 329** (1993) 79.

[122] W.H. Smith, K. Tokushuku, L.W. Wiggers, *The ZEUS trigger system, Contributed to 10th International Conference on Computing in High Energy Physics (CHEP 92).* Annecy, France (21-25 Sept. 1992).

[123] ZEUS collaboration, R. Carlin et al., *The trigger of ZEUS, a flexible system for a high bunch crossing rate collider, Nucl. Instrum. Meth.* **A 379** (1996) 542.

[124] R. Carlin, W.H. Smith, K. Tokushuku, L. Wiggers, *Experience with the ZEUS trigger system,* Vol. 44, p. 430. (1995), doi:10.1016/0920-5632(95)00565-X.

[125] W.H. Smith et al., *The ZEUS calorimeter first level trigger, Nucl. Instrum. Meth.* **A 355** (1995) 278.

[126] G.P. Heath et al., *The ZEUS first level tracking trigger, Nucl. Instrum. Meth.* **A 315** (1992) 431.

[127] P.D Allfrey et al., *The design and performance of the ZEUS global tracking trigger, Nucl. Instrum. Meth.* **A 580** (2007) 1257.

[128] G. Abbiendi, *Global Tracking of Muons in the Barrel and Rear Region* (unpublished). ZEUS-99-063, internal ZEUS Note, 1999.

[129] F. Chlebana, *Description of the ZEUS Global Second Level Trigger in 1994* (unpublished). ZEUS-94-102, internal ZEUS Note, 1994.

[130] H.A.J.R Uijterwaal, *The global second level trigger for ZEUS, RX-1392 (AMSTERDAM)* (1992).

[131] D.C. Bailey et al., *The ZEUS third level trigger hardware architecture and the control system, DESY-92-150E* (1992).

[132] S. Bhadra, M. Crombie, D.P. Kirkby and R.S. Orr, *The ZEUS third level trigger system, Comput. Phys. Commun.* **57** (1989) 321.

[133] J.C. Hart, *Experience with ADAMO in ZEUS, In Proceedings, Data structures for particle physics experiments.* (1990).

[134] R. Brun et al., GEANT3, Technical Report CERN-DD/EE/84-1, CERN, 1987.

[135] G.F. Hartner, *VCTRAK Briefing: Program and Math* (unpublished). ZEUS-98-058, internal ZEUS Note, 1998.

[136] G.F. Hartner et al., *VCTRAK(3.07/04): Offline Output Information* (unpublished). ZEUS-97-064, internal ZEUS Note, 1997.

[137] R.E. Klaman, *A new approach to linear filtering and prediction problems, Journal of Basic Engineering* **82** (1960) 35.

[138] A. Spiridonov, *Mathematical framework for fast and rigorous track fit for the ZEUS detector*, 2008, [`0812.1245`].

[139] R. Frühwirth and A. Strandlie, *Track and vertex reconstruction: From classical to adaptive methods, Rev. Mod. Phys.* **82** (2010) 1419.

[140] J. Grosse-Knetter, *Energy Correction for Islands* (unpublished). ZEUS-97-039, internal ZEUS Note, 1997.

[141] H. Abramowicz, A. Caldwell and R. Sinkus, *Neural network based electron identification in the ZEUS calorimeter, Nucl. Instrum. Meth.* **A 365** (1995) 508.

[142] ZEUS collaboration, J. Breitweg et al., *Measurement of high-Q^2 neutral current e^+p deep inelastic scattering cross sections at HERA, Eur. Phys. J.* **C 11** (1999) 427, [`arXiv:hep-ex/9905032`].

[143] B. Straub, *The EM Electron Finder* (unpublished), `http://www-zeus.desy.de/ZEUS_ONLY/zeus_notes/notes01.html`.

[144] H. Abramowicz, A. Caldwell, R. Sinkus, *Neural Network Based Electron Identification in the ZEUS Calorimeter* (unpublished). DESY-95-054.

[145] 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.

[146] U. Bassler and G. Bernardi, *Structure function measurements and kinematic reconstruction at HERA, Nucl. Instrum. Meth.* **A 426** (1999) 583.

[147] F. Jacquet and A. Blondel, *Detection of the charged current event – Method II, Proceedings of the Study for an ep Facility for Europe,* U. Amaldi (ed.), p. 391. Hamburg, Germany (1979). Also in preprint DESY 79/48.

[148] S. Bentvelsen, J. Engelen and P. Kooijman, *Reconstruction of* (x, Q^2) *and extraction of structure functions in neutral current scattering at HERA, Proc. Workshop on Physics at HERA,* W. Buchmüller and G. Ingelman (eds.), Vol. 1, p. 23. Hamburg, Germany, DESY (1992).

[149] J. Behr, *Jets at High Q^2 at HERA and Test Beam Measurements with the EUDET Pixel Telescope,* Ph.D. Thesis, Hamburg University, Hamburg, Germany, Report DESY-THESIS-2010-038, 2010.

[150] M. Moritz, *Measurement of the High Q^2 Neutral Current DIS Cross Section at HERA,* Ph.D. Thesis, Universität Hamburg, Report DESY-THESIS-02-009, 2001.

[151] G. Marchesini et al., HERWIG*: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons, Comp. Phys. Comm.* **67** (1992) 465.

[152] G. Corcella et al., HERWIG 6*: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes), JHEP* **01** (2001) 010, [`arXiv:hep-ph/0011363`].

[153] T. Sjöstrand, *High-energy physics event generation with* PYTHIA 5.7 *and* JETSET 7.4, *Comp. Phys. Comm.* **82** (1994) 74.

[154] H. Perrey, *Jets at Low Q^2 at HERA and Radiation Damage Studies for Silicon Sensors for the XFEL,* Ph.D. Thesis, Hamburg University, Hamburg, Germany, Report DESY-THESIS-2011-021, 2011.

[155] F. Januschek, *Measurement of e^+p Neutral Current Deep Inelastic Scattering with a Longitudinally Polarised Positron Beam and X-Ray Radiation Damage for Silicon Sensors*, Ph.D. Thesis, Hamburg University, Hamburg, Germany, Report DESY-THESIS-2012-012, 2010.

[156] J.R. Silvia, *Measurement of the High-Q^2 Neutral Current DIS Cross Sections with the ZEUS Detector at HERA*, Ph.D. Thesis, Imperial College - London, Hamburg, Germany, Report DESY-THESIS-2003-022, 2010.

[157] K. Oliver, J. Ferrando, R. Devenish, *A Minimum Bias Z vertex Distribution for 2005–2007 ep Interactions at ZEUS* (unpublished). ZEUS-07-008, internal ZEUS-note, 2007.

[158] ZEUS collaboration (ed.), *The CTD FLT event classification.* (unpublished). `http://www-zeus.desy.de/components/glft/yycontents/trkv.html`.

[159] ZEUS collaboration, H. Abramowicz et al., *Inclusive dijet cross sections in neutral current deep inelastic scattering at HERA*, *Eur. Phys. J.* **C 70** (2010) 965, [`arXiv:1010.6167`].

[160] T.P. Stewart, *Measurement of High-Q^2 Neutral Current Cross-Sections with Longitudinally Polarised Positrons with the ZEUS Detector*, Ph.D. Thesis, University of Toronto, Toronto, Canada, Report DESY-THESIS-2012-028, 2012.

[161] G. Cowan, *A survey of unfolding methods for particle physics*, *Conf. Proc.* **Conference on Advanced Statistical Techniques in Particle Physics** (2002) 248.

[162] A. Arbuzov et al., HECTOR 1.00*: A program for the calculation of QED, QCD and electroweak corrections to ep and $\ell^{\pm}N$ deep inelastic neutral and charged current scattering*, *Comp. Phys. Comm.* **94** (1996) 128, [`arXiv:hep-ph/9511434`].

[163] K. Nagano, *HECTOR interfaced to BASES* (unpublished), `http://www-zeus.desy.de/~nagano/ZEUS_ONLY/nc0304/hector.html`.

[164] I. Makarenko, *Multijet production in γp interactions and at High Q^2 at HERA (thesis in preparation)*, Ph.D. Thesis, Hamburg University, Hamburg, Germany.

[165] Z. Nagy and Z. Trocsanyi, *Next-to-leading order calculation of four jet observables in electron positron annihilation*, *Phys. Rev.* **D 59** (1999) 014020, [`arXiv:hep-ph/9806317`].

[166] Z. Nagy and Z. Trocsanyi, *Multijet cross-sections in deep inelastic scattering at next-to-leading order*, *Phys. Rev. Lett.* **87** (2001) 082001, [`arXiv:hep-ph/0104315`].

[167] H1 and ZEUS collaboration, F.D. Aaron et al., *HERAPDF1.5 (Preliminary)* (unpublished), 2010, `https://www.desy.de/h1zeus/combined_results/index.php?do=proton_structure`.

[168] O. González, J. Terrón, *Study of the statistical fluctuations and correlations in the measurements of inclusive jet cross sections* (unpublished). 2002.

[169] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C 63** (2009) 189, [`arXiv:0901.0002`].

[170] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, *Phys. Rev.* **D 89** (2014) 033009, [`arXiv:1302.6246`].

[171] D. Stump et al., *Uncertainties of predictions from parton distribution functions. 1. the lagrange multiplier method*, *Phys. Rev.* **D 65** (2001) 014012, [`arXiv:hep-ph/0101051`].

[172] M. Botje, *Error estimates on parton density distributions*, *J. Phys.* **G 28** (2002) 779, [`arXiv:hep-ph/0110123`].

[173] H1 and ZEUS collaboration, F.D. Aaron et al., *Combined measurement and qcd analysis of the inclusive $e^{\pm}p$ scattering cross sections at HERA*, *JHEP* **1001** (2010) 109, [`arXiv:0911.0884`].

[174] H1 and ZEUS collaboration (eds.), *HERAFitter. PDF Fitting package.* Program Manual (unpublished), DESY (2013). `http://wiki-zeuthen.desy.de/HERAFitter/HERAFitter`.

[175] Ph. Hagler, *Hadron structure from lattice quantum chromodynamics*, *Phys. Rept.* **490** (2010) 49, [`arXiv:0912.5483`].

[176] F.J. Yndurain, *The Theory of Quark and Gluon Interactions.* Springer, 2006. doi:10.1007/3-540-33210-3.

[177] P.M. Nadolsky et al., *Implications of cteq global analysis for collider observables*, *Phys. Rev.* **D 78** (2008) 013004, [`arXiv:0802.0007`].

[178] H. Lai et al., *New parton distributions for collider physics*, *Phys. Rev.* **D 82** (2010) 074024, [`arXiv:1007.2241`].

[179] S. Alekhin, J. Blumlein and S. Moch, *Parton distribution functions and benchmark cross sections at NNLO*, *Phys. Rev.* **D 86** (2012) 054009, [`arXiv:1202.2281`].

[180] R.D. Ball et al., *Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO*, *Nucl. Phys.* **B 855** (2012) 153, [`arXiv:1107.2652`].

[181] K. Prytz, *Approximate determination of the gluon density at low x from the f_2 scaling violations*, *Phys. Lett.* **B 311** (1993) 286.

[182] S. Chekanov et al., *An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment*, *Eur. Phys. J.* **C 42** (2005) 1, [`arXiv:hep-ph/0503274`].

[183] H1 and ZEUS collaboration, F.D. Aaron et al., *NLO QCD analysis of inclusive, charm and jet data from HERA (HERAPDF1.7)* (unpublished), 2011, `https://www.desy.de/h1zeus/combined_results/index.php?do=proton_structure`.

[184] H1 collaboration, V. Andreev et al., *Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s*, *Eur. Phys. J.* **C 65** (2015) 65, [`arXiv:1406.4709`].

[185] CMS collaboration V. Khachatryan et al., *Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range*, *Eur. Phys. J.* **C 75** (2015) 186, [`arXiv:1412.1633`].

[186] CMS collaboration, S. Chatrchyan et al., *Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s}$ = 7 TeV and first determination of the strong coupling constant in the TeV range*, *Eur. Phys. J.* **C 73** (2013) 2604, [`arXiv:1304.7498`].

[187] CDF collaboration, T. Affolder et al., *Measurement of the strong coupling constant from inclusive jet production at the TEVATRON $\bar{p}p$ collider*, *Phys. Rev. Lett.* **88** (2002) 042001, [`arXiv:hep-ex/0108034`].

[188] D0 collaboration, V.M. Abazov et al., *Measurement of angular correlations of jets at $\sqrt{s}$ = 1.96 tev and determination of the strong coupling at high momentum transfers*, *Phys. Lett.* **B 718** (2012) 56, [`arXiv:1207.4957`].

[189] D0 collaboration, V.M. Abazov et al., *Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV*, *Phys. Rev.* **D 80** (2009) 111107, [`arXiv:0911.2710`].

[190] M. Botje et al., *The PDF4LHC working group interim recommendations*, 2011, [`1101.0538`].

[191] H1 collaboration, F.D. Aaron et al., *A precision measurement of the inclusive ep scattering cross section at HERA*, *Eur. Phys. J.* **C 64** (2009) 561, [`arXiv:0904.3513`].

[192] F. James and M. Roos, *Minuit: A system for function minimization and analysis of the parameter errors and correlations*, *Comput. Phys. Commun.* **10** (1975) 343.

[193] M. Wobisch, *Measurement and QCD analysis of jet cross sections in deep inelastic scattering positron-proton cpllisions at $\sqrt{s} = 300$ GeV*, Ph.D. Thesis, RWTH Aachen, Aachen, Germany, Report DESY-THESIS-2000-049, 2000.

[194] T. Kluge, K. Rabbertz and M. Wobisch, *FastNLO: Fast pQCD calculations for PDF fits*, *Conf. Proc. C0604201*, p. 483. (2006). Also in preprint hep-ex/0609285. [`arXiv:hep-ex/0609285`].

[195] M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, F. Stober, *Theory-data comparisons for jet measurements in hadron-induced processes*, Preprint DESY-11-150 and FERMILAB-PUB-11-418-PPD, 2011.

[196] D. Britzger, K. Rabbertz, F. Stober and M. Wobisch, *New features in version 2 of the FastNLO project*, *Conf. Proc. C1203261*, p. 217. (2012). Also in preprint 1208.3641. [`arXiv:1208.3641`].

[197] R. Devenish, A. Cooper-Sarkar, *Deep Inelastic Scattering.* Oxford University Press, 2011.

[198] S. Bethke, *World summary of α_s (2012)*, *Nucl. Phys. Proc. Suppl.* **234** (2013) 229, [`arXiv:1210.0325`].

[199] D.J. Gross, F. Wilczek, *Ultraviolet behavior of nonabelian gauge theories*, *Phys. Rev. Lett.* **30** (1973) 1343.

[200] H.D. Politzer, *Reliable perturbative results for strong interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.

[201] D.J. Gross, F. Wilczek, *Asymptotically free gauge theories. 1*, *Phys. Rev.* **D 8** (1973) 3633.

[202] H.D. Politzer, *Asymptotic freedom: An approach to strong interactions*, *Phys. Rept.* **14** (1974) 129.

[203] D. Soper, *Basics of qcd perturbation theory,* 1997, [`arXiv:hep-ph/9702203`].

[204] J.M. Campbell, J.W. Huston and W.J. Stirling, *Hard interactions of quarks and gluons: A primer for LHC physics, Rept. Prog. Phys.* **70** (2007) 89, [`arXiv:hep-ph/0611148`].

[205] NNPDF collaboration, R.D. Ball et al., *Parton distributions with LHC data, Nucl. Phys.* **B 867** (2013) 244, [`arXiv:1207.1303`].

[206] H1 collaboration, A. Aktas et al., *Measurement of inclusive jet production in deep-inelastic scattering at high Q^2 and determination of the strong coupling, Phys. Lett.* **B 653** (2007) 134, [`arXiv:0706.3722`].

[207] H1 collaboration, F.D. Aaron et al., *Jet production in ep collisions at low Q^2 and determination of α_s, Eur. Phys. J.* **C 67** (2010) 1, [`arXiv:0911.5678`].

[208] H1 collaboration, ZEUS collaboration, *Determination of the strong coupling $\alpha_s(M_Z)$ by Combining Inclusive Jet Cross Sections in Deep-Inelastic-Scattering at HERA* (unpublished). ZEUS-07-125, H1prelim-07-132, 2007.

[209] ZEUS collaboration, *Inclusive-jet production in NC DIS with HERA II* (unpublished). ZEUS-10-002, 2010.

[210] CMS collaboration, S. Chatrchyan et al., *Measurements of differential jet cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the CMS detector, Phys. Rev.* **D 87** (2013) 112002, [`arXiv:1212.6660`].

[211] N. Tuning, *ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction* (unpublished). ZEUS-01-021, internal ZEUS-note, 2001.

Acknowledgement

This work would be impossible without contribution of numerous people to which I am deeply indebted.

I would like to express my gratitude to supervisors **PD Dr. Olaf Behnke** and **Prof. Dr. Brian Foster** for their dedication to this analysis and kind support during the tough moments. I am especially pleased to thank them for the critical review of the thesis manuscript and numerous indispensable suggestions.

I thank **Prof. Dr. Erika Garutti**, **Prof. Dr. Ludwig Mathey** and **Prof. Dr. Gudrid Moortgat-Pick** for agreeing to be the members of the examination committee.

I wish to express my appreciation to **Dr. Daniel Britzger**, **Dr. Ringaile Plačakytė**, **Dr. Voica Radescu** and **Dr. Katarzyna Wichmann** for helping to resolve various technical issues experienced within the course of my studies with FastNLO and HERAFitter. Additionaly, I express the utmost appreciation to my officemate **Inna Makarenko**. This thesis significantly owes its existence to the close collaboration with you.

I would like to thank **Dr. Ingrid-Maria Gregor** for inviting me to DESY pixel telescope group. I am grateful to **Dr. Claus Kleinwort**, **Dr. Rainer Mankel**, **Dr. Hanno Perrey** and **Dr. Igor Rubinskiy** for their help and encouragement during my involvement in the EUTelescope project.

I cordially thank to my first teachers **Dr. Volodymyr Aushev**, who provided the possibility to join the world-class research at the DESY, as well as **Prof. Dr. Cluadia Glasman** and **Prof. Dr. Juan Terron**, who introduced me to the exciting area of jet physics.

The importance of the feedback from other researchers from the **ZEUS collaboration** cannot be overestimated. I thank all in the **ZEUS group** who expressed interest in this analysis.

The research would not be such an interesting endeavour without friends, therefore I am pleased to express my warmest acknowledgement to all former colleagues from the **ZEUS-Kyiv group**. It was great to share happy moments with you.

Finally, my wholehearted gratitude goes to my loving family and especially to my dear wife for their unconditional love, patience and support.

Statement of authorship

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbs verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, 8 March 2016