Measurement of Jet Production in Deep-Inelastic *ep* Scattering at HERA

Dissertation

zur Erlangung des Doktorgrades
des Department Physik
der Universität Hamburg

vorgelegt von

Roman Kogler

aus Wolfsberg

Hamburg
2010

Gutachter der Dissertation: Dr. Günter Grindhammer
Prof. Dr. Robert Klanner

Gutachter der Disputation: Dr. Hannes Jung
Prof. Dr. Johannes Haller

Datum der Disputation: 20. 12. 2010

Vorsitzender des Prüfungsausschusses: Prof. Dr. Joachim Bartels

Vorsitzender des Promotionsausschusses: Prof. Dr. Joachim Bartels

Dekan der MIN Fakultät: Prof. Dr. Heinrich Graener

Leiterin des Fachbereichs Physik: Prof. Dr. Daniela Pfannkuche

Abstract

In this thesis, precision measurements of inclusive jet, dijet and trijet cross sections in neutral current (NC) deep-inelastic *ep* scattering at a centre-of-mass energy of $\sqrt{s} \simeq 319\,\text{GeV}$ are presented. The analysis is based on data collected by the H1 detector during the HERA-2 running phase in the years 2003–07, corresponding to an integrated luminosity of $351.6\,\text{pb}^{-1}$. The kinematic phase space of the measurement is defined by $150 < Q^2 < 15000\,\text{GeV}^2$ and $0.2 < y < 0.7$, where Q^2 and y are the negative four-momentum transfer squared and the inelasticity respectively. Jets are measured in the pseudorapidity range $-1.0 < \eta_{\text{lab}} < 2.5$ in the laboratory rest frame. The jet transverse momentum in the Breit frame of reference is required to be $P_{\text{T}} > 7\,\text{GeV}$ for the inclusive jet measurement and $P_{\text{T}} > 5\,\text{GeV}$ for the dijet and trijet measurements. In the case of the dijet and trijet measurements, the invariant mass of the two jets with the highest transverse momenta is required to be greater than 16 GeV.

A large part of this work is devoted to the improvement of the reconstruction of the hadronic final state. This is achieved by a separation of showers originating from electromagnetically and hadronically interacting particles in the liquid argon calorimeter of the H1 detector on a statistical basis. A novel method to calibrate the hadronic final state, which is reconstructed with an energy-flow-algorithm, is developed. This calibration is based on the probability of a shower resulting from an electromagnetically interacting particles and it is shown to improve the absolute energy scale uncertainty at high jet transverse momenta to 1%. Improvements of the resolution of the jet energy measurement with respect to the standard reconstruction of about 10% are reported.

The new calibration in combination with improvements of the reconstruction algorithms by the H1 collaboration leads to a precision measurement of inclusive jet, dijet and trijet cross sections with average uncertainties of 4%, 5.2% and 7.2%, respectively. Measurements of jet cross sections normalised to the inclusive NC cross sections are then performed, which reduce the experimental uncertainties considerably to 1.9%, 2.4% and 6.5%. The measured cross sections are found to be compatible with previously published data, but have significantly reduced experimental uncertainties. Predictions from perturbative QCD calculations in next-to-leading order are compared to the data and are found to give a good description both in terms of absolute size and shape of the measured cross sections.

Kurzfassung

In der vorliegenden Arbeit wird eine Präzissionsmessung der Wirkungsquerschnitte von inklusiver Jetproduktion, sowie von Dijet- und Trijetproduktion in tiefunelastischer *ep*-Streuung bei neutralem Strom vorgestellt. Der dazu verwendete Datensatz hat eine integrierte Luminosität von $351{,}6\,\mathrm{pb}^{-1}$ und wurde bei einer Schwerpunktsenergie von $\sqrt{s} \simeq 319\,\mathrm{GeV}$ mit dem H1 Detektor während der HERA-2 Phase in den Jahren 2003–07 aufgezeichnet. Die Analyse wird in dem Bereich $150 < Q^2 < 15000\,\mathrm{GeV}^2$ und $0{,}2 < y < 0{,}7$ vorgenommen, wobei Q^2 der negative quadrierte Viererimpulsübertrag ist und y die Inelastizität bezeichnet. Der Phasenraum der Jetmessung entspricht dem Bereich $-1{,}0 < \eta_{\mathrm{lab}} < 2{,}5$ in der Jetpseudorapidität im Laborsystem. Bei der Messung von inklusiven Jetwirkungsquerschnitten wird ein Transversalimpuls von $P_{\mathrm{T}} > 7\,\mathrm{GeV}$ gefordert. Im Fall der Dijet-, beziehungsweise der Trijetmessung, werden Jets mit $P_{\mathrm{T}} > 5\,\mathrm{GeV}$ verlangt und die invariante Masse der zwei Jets mit dem höchsten Transversalimpuls muss mindestens 16 GeV betragen.

Ein großer Teil dieser Arbeit ist der Verbesserung der Rekonstruktion des hadronischen Endzustands gewidmet. Dabei wurde eine Methode entwickelt, um elektromagnetisch und hadronisch induzierte Schauer im flüssig-Argon Kalorimeter von H1 mit statistischen Methoden zu trennen. Der hadronische Endzustand wird durch einen Algorithmus rekonstruiert, der sowohl Spuren als auch Kalorimeterdepositionen berücksichtigt. Eine neue Kalibrationsmethode, die auf den Ergebnissen der Schauerseparation beruht, wird in dieser Arbeit vorgestellt. Mit Hilfe dieser Kalibration wird eine Unsicherheit der Energiemessung von Jets von 1% bei hohem transversalem Impuls erreicht. Es wird außerdem gezeigt, dass diese Kalibration die Auflösung der Jetenergiemessung im Vergleich zu der Standardrekonstruktion um näherungsweise 10% verbessert.

Die neue Kalibration in Kombination mit Verbesserungen der Rekonstruktionsalgorithmen durch die H1 Kollaboration führt bei der durchgeführten Analyse zu einer Präzissionsmessung von inklusiven Jetwirkungsquerschnitten mit durchschnittlichen Unsicherheiten von 4%. Bei den Dijetwirkungsquerschnitten liegen diese bei 5,2% und die Trijetwirkungsquerschnitte werden mit einer Präzission von 7,2% gemessen. Als weiterführenden Schritt werden normierte Jetwirkungsquerschnitte gemessen, wobei die Normierung zu inklusiven Wirkungsquerschnitten des neutralen Stroms durchgeführt wird, was zu einer beträchtlichen Verringerung der experimentellen Unsicherheit führt. Die normierten inklusiven Jetwirkungsquerschnitte haben eine durchschnittliche Unsicherheit von 1,9%, für die normierten Dijet- und Trijetwirkungsquerschnitte werden Unsicherheiten von 2,4% und 6,5% erreicht. Die gemessenen Wirkungsquerschnitte sind mit publizierten Wirkungsquerschnitten konsistent bei bedeutend verringerter experimenteller Unsicherheit. Ein Vergleich der Daten mit störungstheoretischen QCD Berechnungen in nächst-führender Ordnung zeigt gute Übereinstimmung, sowohl in der absoluten Größe, als auch in der Form der Wirkungsquerschnitte.

Contents

Chapter 1

Introduction

The Standard Model (SM) of particle physics is a theoretical framework which describes the electromagnetic, weak and strong interactions between all known fundamental particles. The SM has been very successful so far in describing all experimentally measured processes very accurately. A large effort has been ongoing and is currently culminating with the Large Hadron Collider (LHC) to reveal the Higgs sector, which is the last crucial missing link in the verification of the SM. Different extensions of the SM are theoretically possible. They have been developed to either incorporate higher symmetries with the goal of the further unification of the fundamental forces, or to explain phenomena like neutrino masses and dark matter. Experimentally, numerous searches for physics beyond the SM have been performed for a wide range of hard scattering processes at the highest available centre-of-mass energies in e^+e^- collisions at the Large Electron Positron Collider (LEP) [LEP07], in $p\bar{p}$ collisions at the Tevatron [Dup09] and in ep collisions at the Hadron-Elektron-Ring-Anlage (HERA) [H109b]. With the advent of the LHC, pp collisions with so far unprecedented centre-of-mass energies became available, allowing for particles with masses of up to several TeV to be produced, which is much higher than previously possible. In order to claim the discovery of new phenomena at the LHC, the predictions of the SM have to be extrapolated to high energy scales, which makes a precise understanding of the SM indispensable [Man08]. This includes a precise knowledge in particular of the strong interaction, which is responsible for the largest part of the pp cross section, and the internal structure of the colliding protons, which is parametrised by Parton Distribution Functions (PDFs).

Within the SM the strong interaction is described by Quantum Chromodynamics (QCD). QCD is a non-abelian quantum field theory describing the dynamics of quarks and gluons. In the framework of QCD quarks and gluons carry colour charge, which comes in three flavours. The relative strength of interactions between coloured particles is governed by the strong coupling α_s, which decreases with increasing energy. This so called running of α_s can be predicted within QCD, but its value at some starting scale needs to be extracted from experimental data. The smallness of α_s at high energies (much larger than the mass of the proton)[1] leads to an asymptotically free theory, allowing for perturbative methods to be applicable. At small energies the value of α_s becomes large. In this energy regime QCD is believed to exhibit the property of confinement, such that coloured particles cannot be

---

[1] Throughout this work natural units are used, obtained by setting $c = \hbar = 1$.

observed as free states but are confined in colourless bound states, called hadrons. In this case perturbative methods are not applicable and therefore the structure of hadrons, most importantly the proton PDFs, need to be determined experimentally. The uncertainties from the determination of the proton PDFs and α_s translate into uncertainties of SM predictions in important discovery channels for Higgs searches and searches for new phenomena at the LHC [BD$^+$10].

In this work a precision measurement of jet production in ep collisions at high photon virtuality Q^2 and high jet transverse momenta P_T is presented. Jets are collimated sprays of particles and can be defined for coloured quarks and gluons as well as for colourless hadrons. The excellent correspondence between these definitions makes jets important tools to study the dynamics of QCD. The restriction to high Q^2 and high P_T in this analysis ensures reliable predictions within the framework of perturbative QCD (pQCD). The cross section for jet production is directly proportional to the value of α_s, making jet production an important mean for the precise determination of α_s [H110c]. Furthermore, jet production in ep collisions is sensitive to the gluon content of the proton already in the leading order, in contrast to inclusive deep-inelastic scattering (DIS). Jet cross section data can thus be used not only as an independent test of proton PDFs, but also provide further input in the determination of PDFs [ZEUS05b].

Jets consist of a number of different particles with varying energies. Some of these particles interact hadronically with the detector material, some induce electromagnetic cascades and some escape undetected. A precise measurement of jet production therefore needs an excellent understanding of the detector and its response to various particles over a wide range of energies. In this work a considerable effort is undertaken to improve the reconstruction of the hadronic final state and consequently the jet energy scale uncertainty, which is the dominant uncertainty in every jet measurement. Sources of various other experimental uncertainties are also studied in detail.

The analysis presented here is based on data collected by the H1 experiment during the HERA-2 running phase in the years 2003–2007. Inclusive jet, dijet and trijet cross sections are measured single and double-differentially. The obtained jet cross sections have a considerably reduced uncertainty compared to previous jet measurements performed by the H1 collaboration [H101a, H107]. It is the first double-differential measurement of trijet cross sections in neutral current (NC) DIS at high Q^2 at H1, which has so far been only measured at small values of Q^2 [H110b]. The measurement of normalised jet cross sections results in a further reduction of the experimental uncertainties due to partial cancellations of measurement uncertainties. An improved understanding of the hadronic final state has allowed the phase space for jet production to be enlarged and the total experimental uncertainties reduced with respect to a previous measurement of normalised jet cross sections [H110c].

This thesis is organised as follows: In chapters 2 and 3 an overview of the SM and jet production in NC DIS is given. This is followed by a description of the HERA facility and the H1 detector in chapter 4. In chapter 5 the interaction of highly energetic particles with matter and the basic concepts of calorimetry are reviewed. The chapter concludes with an overview of the Liquid Argon (LAr) calorimeter of H1 and its performance as determined from test beams. In chapter 6 a new method to separate electromagnetic from hadronic cascades in the LAr calorimeter on a statistical basis is developed. This is the foundation of an improved understanding of the jet energy scale uncertainty. The reconstruction of the hadronic final state with an energy-flow algorithm is described in chapter 7. A novel method to calibrate the resulting hadronic final state is introduced in chapter 8, which treats energy deposits with a high probability to originate from electromagnetic showers differently from those originating from hadronic showers. It is shown how this improves the uncertainty on the absolute jet energy scale, as well as the resolution of the measurement of jets. The chapter concludes with a detailed study of the achieved resolutions for the different regions in the H1 detector. In chapter 9 the data selection is described. The selection of the NC DIS sample is first explained, which ensures high trigger efficiencies together with effectively no background. This is followed by a description of the inclusive jet, dijet and trijet selections together with a comparison of the data to different Monte Carlo (MC) simulations. The concept of the jet cross section measurement is discussed in chapter 10, together with a review of the experimental uncertainties. In chapter 11 the measured jet cross sections are presented and compared to theoretical predictions as well as to previous jet measurements. Chapter 12 gives a summary of the results.

Chapter 2

Theoretical Framework

2.1 The Standard Model of Particle Physics

The fundamental building blocks of matter can be grouped into two categories: fermions with half-integer spin and bosons with integer spin. The fermions are found in two groups: strongly interacting fermions which are called quarks [GM64] and those which do not exhibit the strong interaction are called leptons. They are all spin-$\frac{1}{2}$ particles and have no internal structure within the limits of our resolution. The Standard Model of particle physics describes the interaction of the quarks and leptons through bosons, which are the force carriers of the fundamental interactions. The photon is the force carrier of the electromagnetic interaction, the $W^{\pm}$ and Z^0 mediate the weak interaction and eight gluons are the exchange particles of the strong interaction.

Quarks and leptons come in six flavours and fall into left-handed isospin doublets and right-handed isospin singlets. These come in three different families, also called generations. Quarks and leptons of one generation carry the same quantum numbers as their partners of another generation, with the exception of different masses. Nuclei, atoms and molecules are built up out of the particles of the first generation only.

The quarks carry colour charge and the six different flavours are called up, down, charm, strange, top and bottom. They are denoted by the first letter of their name and the left-handed isospin doublets are

$$\begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} c \\ s \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} t \\ b \end{pmatrix}_{\mathrm{L}}, \tag{2.1}$$

where the quarks in the top row have electric charge $q = \frac{2}{3}e$, and the ones in the bottom row have $q = -\frac{1}{3}e$, where e is the charge of the electron. The leptons do not carry colour charge and can also be arranged in left-handed doublets,

$$\begin{pmatrix} \nu_{\mathrm{e}} \\ e \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_{\mathrm{L}}, \tag{2.2}$$

where the electron (e), muon (μ) and tau (τ) have electric charge $-e$, while the corresponding neutrinos have electric charge zero and interact only through the weak interaction. No right-handed neutrinos exist in the SM, reflecting the fact that left-handed and right-handed leptons transform differently under weak interactions, a phenomenon known as parity violation.

It was realised between the late 1960's and early 1970's that the electromagnetic, weak and strong force can be described with non-abelian local gauge field theories[1]. Glashow, Weinberg and Salam recognised that the electromagnetic and weak interactions can be unified by joining a weak-isospin symmetry $SU(2)_\mathrm{L}$ with a weak hypercharge $U(1)_\mathrm{Y}$ symmetry [Gla61,Wei67,SW64]. The resulting theory is a non-abelian gauge theory which unifies Quantum Electrodynamics (QED) with a quantum field theory of the weak interaction. The Lagrangian of this electroweak theory is invariant under symmetry transformations of the common electroweak $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry group. The four resulting gauge fields, which compensate for the variations induced by gauge transformations, can be interpreted as the Goldstone bosons of the theory. However, local gauge invariance requires these bosons to be massless [GS$^+$62], which is in contradiction with the experimental evidence for large masses of the W^+, W^- and Z^0 bosons. A solution to this problem is provided by the Higgs mechanism [EB64, Hig64, GH$^+$64], which introduces a spontaneous symmetry breaking of the $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry. This mechanism leaves the photon massless, while the other three electroweak gauge bosons acquire mass. The resulting theory is invariant under local phase transformations, i.e. the $U(1)_\mathrm{em}$ symmetry of QED is recovered. The Higgs mechanism requires a scalar field which manifests itself as the Higgs boson, the only Standard Model particle which has not yet been discovered [CD10]. The Higgs field couples not only to the massive electroweak bosons, but also to itself and to the leptons and quarks, thus giving them mass. The Yukawa couplings of the Higgs boson to the quarks and leptons cannot be predicted by the Standard Model and enter as free parameters in the theory, like the mass of the Higgs boson itself.

Three kinds of colour charges are carried by the quarks, with the corresponding anti-colours carried by the anti-quarks. The colour charges are labelled red, green and blue in analogy to everyday colour theory. Rotations in colour space form a $SU(3)_\mathrm{C}$ colour group [FGM$^+$73], and all colourless SM particles are invariant under these rotations. The bosons associated to the $SU(3)_\mathrm{C}$ gauge transformations are called gluons, which carry a colour and an anticolour, and are labelled G_a^μ, where $a = 1, \ldots, 8$. These eight gluons represent the eight generators of the group and form a colour octet. All physical states are singlets with respect to $SU(3)_\mathrm{C}$, making colour a non-observable internal degree of freedom. The theory describing the interaction between quarks and gluons is called Quantum Chromodynamics (QCD).

The full Lagrangian of the Standard Model is obtained from the invariance under symmetry transformations of the $SU(3)_\mathrm{C} \otimes SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry group together with the gauge invariant terms and Yukawa interactions of the scalar Higgs field and can be found in the literature (see for example [HM84, PDG10, Qui09]).

---

[1] Including the fourth fundamental interaction, gravity, as a quantum field theory in the Standard Model has turned out to be extremely difficult. No satisfactory and commonly accepted theory has been established so far.

2.2 Quantum Chromodynamics

Within the SM the strong force is fully accounted for by the $SU(3)_{\mathrm{C}}$ invariant Lagrangian of QCD, which is given by [YM54, FGM⁺73]

$$\mathcal{L} = \sum_{j}^{n_f} \bar{q}_j (i\gamma^{\mu}\mathcal{D}_{\mu} - m_q) q^j - \frac{1}{4} F_a^{\mu\nu} F^a_{\mu\nu} , \tag{2.3}$$

with the covariant derivative

$$\mathcal{D}_{\mu} = \partial_{\mu} - i g_s T_a G^a_{\mu} , \tag{2.4}$$

where the spinor indices have been suppressed[2]. The quark and anti-quark fields with flavour j and mass m_q are labelled q_j and $\bar{q}_j$, respectively. The total number of flavours is given by n_f. The Dirac matrices are denoted by γ_{μ} and g_s is the QCD coupling. T_a are a set of linearly independent traceless 3×3 matrices, which obey the commutation relation

$$[T_a, T_b] = i f_{abc} T_c , \tag{2.5}$$

where f_{abc} are the structure constants of the $SU(3)$ algebra. Einstein convention is used, which implies a summation over repeated indices. Gauge invariance of the Lagrangian (2.3) is obtained by defining the gluon field strength tensor as

$$F^a_{\mu\nu} = \partial_{\mu} G^a_{\nu} - \partial_{\nu} G^a_{\mu} - g_s f_{abc} G^b_{\mu} G^c_{\nu} . \tag{2.6}$$

The Feynman rules of QCD can be obtained for example by varying the action integral [ES⁺96]

$$S = i \int \mathcal{L} \, \mathrm{d}^4 x . \tag{2.7}$$

For this purpose it is instructive to expand the Lagrangian density (2.3) using equations (2.4) and (2.6). The Lagrangian density can then be split into a free piece $\mathcal{L}_0$ and an interacting piece $\mathcal{L}_I$. The free Lagrangian density is given by

$$\begin{aligned} \mathcal{L}_0 = & \sum_{j}^{n_f} \bar{q}_j (i\gamma^{\mu}\partial_{\mu} - m_q) q^j \\ & - \frac{1}{4} (\partial^{\mu} G^{\nu}_a - \partial^{\nu} G^{\mu}_a)(\partial_{\mu} G^a_{\nu} - \partial_{\nu} G^a_{\mu}) \end{aligned} \tag{2.8}$$

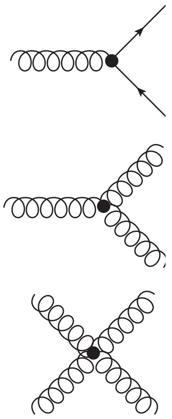
which contains the kinetic terms for the different fields. The free Lagrangian density gives rise to the quark and gluon propagators. The graphical representations of the Feynman rules for these are illustrated next to the corresponding terms. A solid line represents a quark propagator and a curly line represents a gluon propagator.

---

[2] The gauge fixing and ghost terms in the Lagrangian are ignored for the sake of simplicity.

The interaction piece $\mathcal{L}_I$ contains the colour interactions,

$$\begin{aligned}\mathcal{L}_I = g_s G^a_\mu \sum_j^{n_f} \bar{q}_j \gamma^\mu T_a q^j \\ - \frac{g_s}{2} f^{abc} (\partial^\mu G^\nu_a - \partial^\nu G^\mu_a) G^b_\mu G^c_\nu \\ - \frac{g_s^2}{4} f^{abc} f_{ade} G^\mu_b G^\nu_c G^d_\mu G^e_\nu\end{aligned} \tag{2.9}$$



with the corresponding Feynman diagrams being illustrated on the right. The first line of equation (2.9) describes the interaction between quarks and gluons. The second and third lines generate the triple and four gluon vertices. They are due to the non-Abelian character of QCD and describe self-interactions between gluons, which are a consequence of the fact that gluons themselves carry colour charge. Whereas the quark-gluon and three gluon interactions are proportional to g_s, the four gluon vertex is proportional to g_s^2. Contrary to the free propagators and the quark-gluon interaction, the gluon self coupling diagrams have no QED analogues and give rise to asymptotic freedom [GW73, Pol73] and confinement [Wil74].

2.2.1 Renormalisation

A crucial requirement for any field theory is renormalisability, which eventually allows to make finite physical predictions. Renormalisation can be understood as the replacement of the bare charge and consequently the bare coupling g_s, which appears in the Lagrangian, by a physically observable quantity. This is motivated by the fact that the coupling appearing in the lowest order Feynman graph is modified by higher order corrections, see figure 2.1. It can be shown that the modification of the coupling by the vertex correction exactly cancels the effect of self energies to all orders, which is a basic property of gauge field theories known as Ward identities [War50]. Thus, the coupling is only modified by vacuum polarisation graphs. However, contributions from these graphs involve ultraviolet (UV) divergences due to infinite loop momenta. Different possibilities exist to remove these divergences, called renormalisation schemes. The most commonly used renormalisation scheme is the modified minimal subtraction, or $\overline{\mathrm{MS}}$, scheme [BB+78]. It uses dimensional

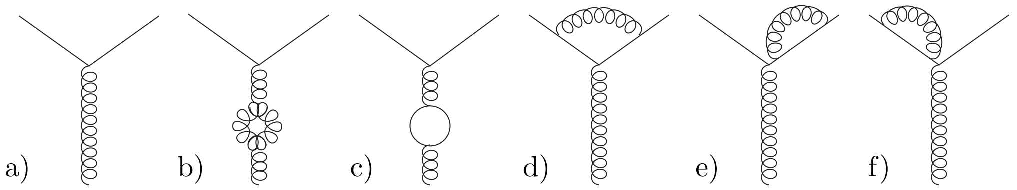


Figure 2.1: Feynman graphs of the quark-gluon interaction (a) and its first order corrections: vacuum polarisation (b), (c), vertex correction (d) and self energies (e), (f).

regularisation [tHV72] to factor out the divergences from the loop contributions. These are then absorbed in the definition of the strong coupling, which is defined in analogy to the fine structure constant of QED to be $\alpha_s = g_s^2/4\pi$. The renormalisation introduces a new mass scale μ_r in the theory. This scale defines the point at which the subtractions which remove the UV divergences are performed. The theory obtained is now free of divergent loop graphs with the trade-off of introducing the scale μ_r.

However, even though μ_r is an arbitrary parameter, the theory remains predictive as long as any physical quantity calculated does not depend on the particular choice of μ_r. Consider a physical observable $\Gamma(p_i, \alpha_s)$ which is a function of all external momenta p_i and the renormalised strong coupling α_s. For simplicity, assume that $\Gamma(p_i, \alpha_s)$ is measured at high enough energy such that the quark masses can be neglected. The requirement that $\Gamma(p_i, \alpha_s)$ is independent of μ_r leads to the Renormalisation Group Equation (RGE),

$$\left(\mu_\mathrm{r}\frac{\partial}{\partial\mu_\mathrm{r}} + \mu_\mathrm{r}\frac{\partial\alpha_s}{\partial\mu_\mathrm{r}}\frac{\partial}{\partial\alpha_s} + \gamma_\Gamma(\alpha_s)\right)\Gamma(p_i, \alpha_s) = 0\,. \tag{2.10}$$

The term $\gamma_\Gamma(\alpha_s)$ is an anomalous dimension, which vanishes for dimensionless quantities [Wei73]. With the definition of the β function,

$$\beta(\alpha_s) = \mu_\mathrm{r}\frac{\partial\alpha_s}{\partial\mu_\mathrm{r}} \tag{2.11}$$

the RGE equation can be written as

$$\left(\mu_\mathrm{r}\frac{\partial}{\partial\mu_\mathrm{r}} + \beta(\alpha_s)\frac{\partial}{\partial\alpha_s} + \gamma_\Gamma(\alpha_s)\right)\Gamma(p_i, \alpha_s) = 0\,. \tag{2.12}$$

For this equation to hold a change in the renormalisation scale μ_r has to be compensated by a change of α_s. This is expressed by the dependence of α_s on the renormalisation scale μ_r, such that the renormalised coupling becomes a running coupling $\alpha_s = \alpha_s(\mu_\mathrm{r})$.

2.2.2 The Running Coupling

To compute the scale dependence of α_s on μ_r, the β function (2.11) can be expanded in powers of α_s,

$$\beta(\alpha_s) = -\alpha_s \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi}\right)^{(n+1)}\,. \tag{2.13}$$

The coefficients β_n are known up to four loops:

$$\beta_0 = 11 - \frac{2}{3}n_f \tag{2.14}$$

$$\beta_1 = 102 - \frac{38}{3}n_f \tag{2.15}$$

$$\beta_2 = \frac{2857}{2} - \frac{5033}{18}n_f + \frac{325}{54}n_f^2 \tag{2.16}$$

$$\begin{aligned}\beta_3 = {} & \frac{149753}{6} + 3564\zeta_3 - \left(\frac{1078361}{162} + \frac{6508}{27}\zeta_3\right)n_f \\ & + \left(\frac{50065}{162} + \frac{6472}{81}\zeta_3\right)n_f^2 + \frac{1093}{729}n_f^3\,,\end{aligned} \tag{2.17}$$

where ζ_3 is the third constant of the Riemann zeta-function, and n_f is the number of active light flavours. The coefficients β_0 [GW73, Pol73] and β_1 [Cas74] are independent of the renormalisation scheme, whereas β_2 [TV$^+$80] and β_3 [RV$^+$97] are calculated in the $\overline{\mathrm{MS}}$ scheme.

In order to obtain the running coupling at four loops, it is helpful to define the dimensionless variable

$$L = \ln\left(\frac{\mu_\mathrm{r}^2}{\Lambda^2}\right) \tag{2.18}$$

with Λ denoting the asymptotic scale parameter. The running coupling is obtained by integrating equation (2.13) and iteratively solving for α_s, which yields [CK$^+$97]

$$\begin{aligned}\frac{\alpha_s}{4\pi} &= \frac{1}{\beta_0 L} - \frac{b_1 \ln L}{(\beta_0 L)^2} + \frac{1}{(\beta_0 L)^3}\left[b_1^2(\ln^2 L - \ln L - 1) + b_2\right] \\ &+ \frac{1}{(\beta_0 L)^4}\left[b_1^3\left(-\ln^3 L + \frac{5}{2}\ln^2 L + 2\ln L - \frac{1}{2}\right) - 3b_1 b_2 \ln L + \frac{b_3}{2}\right],\end{aligned} \tag{2.19}$$

where terms of $\mathcal{O}(1/L^5)$ have been neglected, and the notation $b_n = \beta_n/\beta_0$ is used. The evolution of the strong coupling $\alpha_s(\mu_\mathrm{r})$ is thus fully determined once the scale Λ and the renormalisation scheme have been fixed. All β_n are positive for $n_f < 6$ and the leading order term β_0 is positive for $n_f \leq 16$, which is far larger than the $n_f = 6$ in the Standard Model. Thus the strength of α_s decreases when the scale at which it is probed increases, $\alpha_s(\mu_\mathrm{r}) \to 0$ for $\mu_\mathrm{r} \to \infty$. This behaviour is known as asymptotic freedom and implies that at high enough energies, or equivalently short enough distances, quarks and gluons may be considered to be quasi-free. Asymptotic freedom thus ensures convergence of perturbative calculations at high scales, where the interactions between quarks and gluons can be treated as perturbations from quasi-free propagation. As μ_r approaches Λ, α_s becomes large and in the limit $\mu_\mathrm{r} \to \Lambda$, $\alpha_s \to \infty$. This leads to a property of QCD which is called confinement, which states that free coloured objects are not observable. A mathematical proof of confinement is difficult, since the perturbative expansion breaks down in the small energy regime and non-perturbative methods need to be applied. However, lattice QCD suggests that the non-linearities of the $SU(3)_\mathrm{C}$ gauge field are indeed responsible for the confinement of quarks [Lüs03].

Consequently, Λ takes the meaning of a reference scale above which α_s is small enough that perturbation theory is applicable. The absolute value of Λ cannot be predicted within the Standard Model and has to be determined from experiment. However, in calculations above two loops Λ depends on the renormalisation scheme used, since a constant of integration is absorbed in its definition (here: $\Lambda = \Lambda_{\overline{\mathrm{MS}}}$). Therefore α_s is often expressed at a well known reference scale, which is typically the mass of the Z^0 boson $\alpha_s(M_Z)$. The corresponding value of Λ can be calculated by iteratively solving equation (2.19). The value of the strong coupling can then be obtained at any scale μ_r. For the one loop solution,

$$\alpha_s(\mu_\mathrm{r}) = \frac{4\pi}{\beta_0 \ln(\mu_\mathrm{r}^2/\Lambda^2)}\,, \tag{2.20}$$

the dependence on Λ can be expressed analytically. Calculating Λ from a previously measured value of $\alpha_s(\mu_\mathrm{r} = M_Z)$, and inserting the result in equation (2.20), one

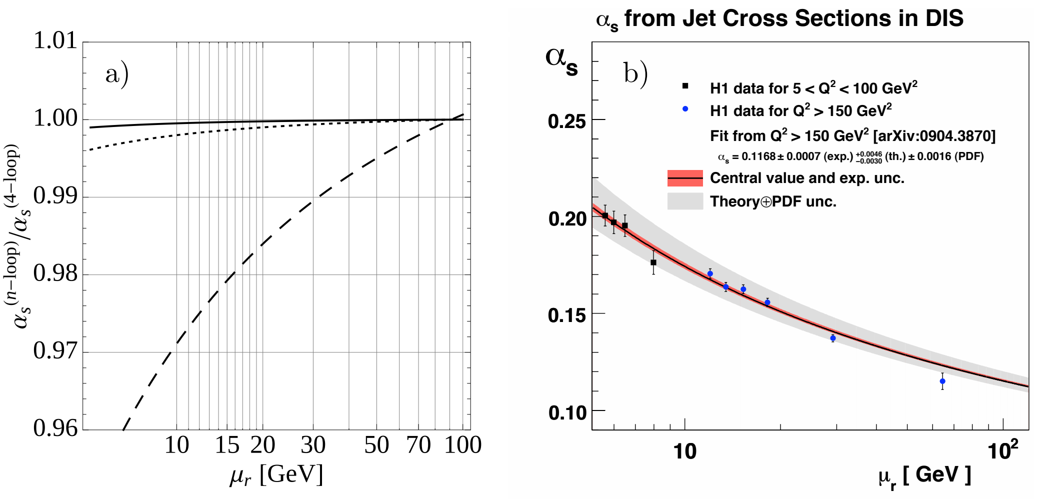


Figure 2.2: a) Ratio of the running of the strong coupling in one loop (dashed), two loops (dotted) and three loops (solid) to the four-loop solution (2.19). The values $n_f = 5$ and $\alpha_s(M_Z) = 0.118$ are used in the calculation, which has been performed with the package RunDec [CK$^+$00]. b) Running of the strong coupling as measured by H1 [H110c], compared to the two loop solution [H110b].

obtains

$$\alpha_s(\mu_\mathrm{r}) = \frac{\alpha_s(M_Z)}{1 + \frac{\beta_0}{4\pi}\alpha_s(M_Z)\ln(\mu_\mathrm{r}^2/M_Z^2)}\,. \tag{2.21}$$

In figure 2.2a the running of the strong coupling in the n-loop approximation, with $n = 1, 2, 3$, is compared to the exact four loop solution. The asymptotic scale parameter Λ is determined from equation (2.19) including terms up to $\mathcal{O}(1/L^n)$, with the condition $\alpha_s(M_Z) = 0.118$. Hence, all curves cross exactly at $\mu_\mathrm{r} = M_Z$. α_s is evolved using $n_f = 5$ down to 5 GeV, which is just above the threshold of the b-quark. It can be seen that the one loop result is far off the exact solution, whereas the two loop solution already provides a very useful approximation. The three loop result differs from the four loop result by less than 0.2% over the full energy range considered.

The running of the strong coupling is shown in figure 2.2b, where α_s values extracted from jet cross sections in deep-inelastic ep scattering at different values of μ_r and are compared to the two loop solution. The uncertainty on the theoretical prediction comes from missing higher orders in the next-to-leading order (NLO) calculation, which introduce a much larger uncertainty than missing orders in the α_s evolution. Dijet production in DIS is a prominent tool to extract the strong coupling because the cross sections are proportional to α_s already in leading order (LO). Recent determinations of the strong coupling using jets are compared to the world average from Bethke [Bet09] in figure 2.3. The uncertainties of the extracted values of $\alpha_s(M_Z)$ are divided into theoretical and experimental uncertainties. It can be observed that the latest determinations of the strong coupling by the H1 collaboration are dominated by theoretical uncertainties, which are mostly due to missing higher order terms in the perturbative NLO calculations. It can also be seen that these missing terms are more important at small scales (H1, low Q^2) than at high scales (H1, high Q^2). The measurements by the Zeus collaboration [ZEUS07b, ZEUS10b] are

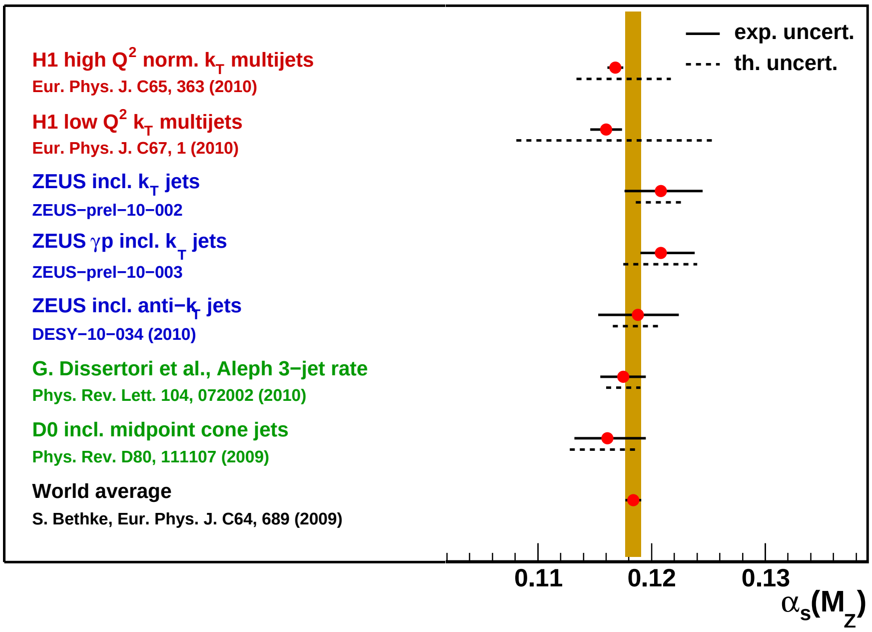


Figure 2.3: Recent determinations of the strong coupling $\alpha_s(M_Z)$ from jet cross sections compared to the world average [Kog10].

optimised for the smallest total uncertainty and are performed at a higher scale than the H1 determination of $\alpha_s(M_Z)$, which results in a smaller theoretical, but larger experimental uncertainties. The extraction of $\alpha_s(M_Z)$ from the 3-jet rate in e^+e^- annihilation [DG$^+$10b] is performed at next-to-next-to-leading order (NNLO), including terms $\mathcal{O}(\alpha_s^4)$, which results in a considerable smaller theoretical uncertainty.

2.2.3 Perturbative QCD

The small size of the strong coupling $\alpha_s(\mu_\mathrm{r})$ at large values of the renormalisation scale μ_r allows the calculation of an observable $\mathcal{R}$ to be expanded in terms of $\alpha_s(\mu_\mathrm{r})$. $\mathcal{R}$ can then be written in terms of a perturbative series,

$$\mathcal{R} = \sum_{n=0}^{\infty} c_n(\mu_\mathrm{r})\, \alpha_s(\mu_\mathrm{r})^n \,, \tag{2.22}$$

where $c_n(\mu_\mathrm{r})$ are coefficients that can be calculated by the evaluation of Feynman diagrams. This approach is called perturbative QCD (pQCD) and is valid at $\mu_\mathrm{r} \gg \Lambda$, where terms of order $\alpha_s(\mu_\mathrm{r})^{n+1}$ are assumed to be less important than preceding terms of order $\alpha_s(\mu_\mathrm{r})^n$. Differentiating N terms of the perturbative series (2.22) by μ_r, one obtains

$$\frac{\mathrm{d}}{\mathrm{d}\ln\mu_\mathrm{r}^2} \sum_{n=0}^{N} c_n(\mu_\mathrm{r})\, \alpha_s(\mu_\mathrm{r})^n \sim \mathcal{O}\left(\alpha_s(\mu_\mathrm{r})^{N+1}\right) \,, \tag{2.23}$$

and as $\alpha_s(\mu_\mathrm{r}) < 1$ for sufficiently large μ_r, the observable $\mathcal{R}$ will depend less on the choice of μ_r the more terms are included in the perturbative series [Sop97].

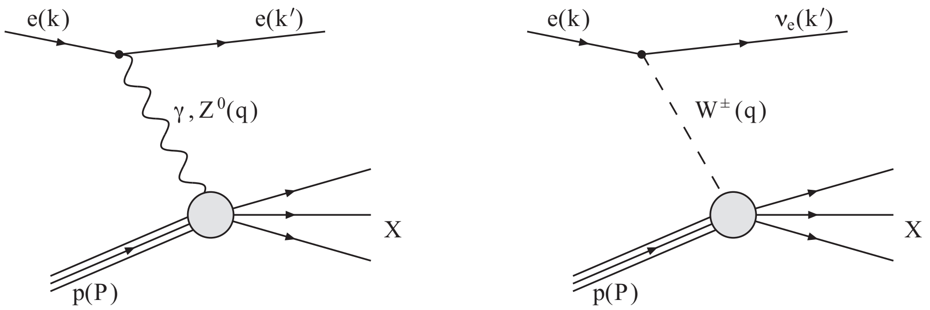


Figure 2.4: Kinematics of $ep \rightarrow \ell X$ scattering. The four-momenta associated to the particles are given in brackets. The exchanged boson may be a photon (γ) or Z^0 in NC or a $W^{\pm}$ in CC interactions.

2.3 Lepton-Proton Scattering

The scattering of either charged (e, μ) or neutral (ν_{e}, ν_{μ}) leptons from protons has been indispensable in the development and subsequent testing of QCD and the structure of the proton. A diagrammatic representation of the scattering of electrons[3] from protons, $ep \rightarrow \ell X$, is shown in figure 2.4. The electron interacts with the proton either via the exchange of a virtual photon or Z^0, which is termed neutral current (NC), or via the exchange of a $W^{\pm}$, which is called charged current (CC) interaction. The squared transferred momentum is given by the negative square of the invariant mass of the exchanged virtual boson, $Q^2 = -q^2$. Deep-inelastic scattering (DIS) is characterised by high momentum transfer, $Q^2 \gg m_p^2$, and large inelasticity, $M_X \gg m_p$, where m_p is the mass of the proton and M_X is the invariant mass of the hadronic final state X.

2.3.1 Kinematics of Deep-Inelastic Scattering

The centre-of-mass energy in ep scattering is given by the square root of the Mandelstam variable

$$s = (k + P)^2 \tag{2.24}$$

where k and P are the four-vectors of the incoming electron and proton, respectively. Assuming massless particles, the centre-of-mass energy at HERA can be well approximated by $\sqrt{s} \approx 4E_{\mathrm{e}}^0 E_{\mathrm{p}}^0$, where E_{e}^0 and E_{p}^0 are the energies of the incoming electron and proton, respectively. The virtuality Q^2 can be calculated from the four-vectors of the incoming and outgoing leptons,

$$Q^2 = -(k - k')^2\,. \tag{2.25}$$

In the limit $Q^2 \rightarrow 0$ the exchanged boson becomes on-mass shell and quasi-real photons are scattered off the proton. This process is termed $\gamma^* p$ scattering or photoproduction. The invariant mass of the final hadronic system is denoted by W and is given by

$$W^2 = (q + P)^2\,. \tag{2.26}$$

---

[3] Unless otherwise stated, the term electron will be used generically for electron and positron throughout this work.

The Bjorken scaling variable x is given can be interpreted in LO as the fraction of the momentum of the incoming proton taken by the struck quark and is given by

$$x = \frac{Q^2}{2P \cdot q} = \frac{Q^2}{W^2 + Q^2 - m_p^2} \tag{2.27}$$

and since $Q^2 \geq 0$ and W^2 cannot be smaller than m_p^2, $x \leq 1$. The lower limit on x is obtained for small Q^2 and large W^2, such that as $s \to \infty$, $x \to 0$. The Bjorken scaling variable is thus restricted to values $0 \leq x \leq 1$. The inelasticity y, which provides a measure of the amount of energy lost by the lepton in the proton's rest frame, is defined as

$$y = \frac{P \cdot q}{P \cdot k} \, . \tag{2.28}$$

The inelasticity is limited to values $0 \leq y \leq 1$, which can be easily seen by its definition in the laboratory rest frame,

$$y = 1 - \frac{E_\mathrm{e}}{E_\mathrm{e}^0} \sin^2 \frac{\theta_\mathrm{e}}{2} \, , \tag{2.29}$$

where E_e and θ_e are the energy and the polar angle of the scattered electron, respectively. At fixed centre-of-mass energy only two of the variables x, y and Q^2 are independent and ignoring masses, they are related through

$$Q^2 = sxy \, . \tag{2.30}$$

The virtuality Q^2 takes a special role, since it can be associated with the resolving power of the interaction. If the virtual boson is considered as a probe of the proton, it has a spatial resolution, related to its wavelength, of

$$\Delta[\mathrm{fm}] \approx \frac{\hbar c}{Q} \approx \frac{0.2}{\sqrt{Q^2[\mathrm{GeV}^2]}} \, , \tag{2.31}$$

which corresponds to a maximum resolution of $\Delta = 10^{-18}$ m at the HERA centre-of-mass energy of 319 GeV.

2.3.2 The Neutral Current DIS Cross Section

The double-differential cross section for NC DIS scattering can be written as a contraction of a leptonic tensor $L_{\mu\nu}$ and a hadronic tensor $W_{\mu\nu}$. Whereas $L_{\mu\nu}$ is fully defined by the electron momenta, the definition of $W_{\mu\nu}$ is more complicated and involves a summation over all possible hadronic final states. Since $W_{\mu\nu}$ cannot be calculated, the NC DIS cross section is written in its most general form in terms of the structure functions of the proton, F_1, F_2, xF_3 and F_L. For unpolarised *ep* scattering it is given by [DCS04]

$$\frac{\mathrm{d}^2\sigma_{\mathrm{NC}}}{\mathrm{d}x\mathrm{d}Q^2} = \frac{2\pi\alpha_{\mathrm{em}}}{xQ^4} \left[Y_+ F_2(x,Q^2) - y^2 F_L(x,Q^2) \mp Y_- xF_3(x,Q^2) \right] \, , \tag{2.32}$$

where the short-hand notation $Y_\pm = 1 \pm (1-y)^2$ has been used, and α_{em} is the electromagnetic coupling. The structure functions F_i have to be determined experimentally. They are related via $F_L = F_2 - 2xF_1$. F_1 is proportional to the transversely

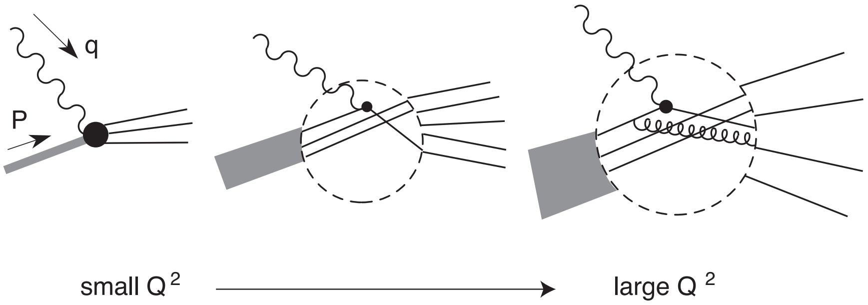


Figure 2.5: Resolution of the photon probe as function of Q^2, from [Pic95].

polarised component of the structure function. The contribution of the longitudinal structure function F_L can be neglected in a large part of the phase space, but becomes relevant at high y. F_3 is the parity violating structure function which includes effects from Z^0 exchange and γ/Z^0 interference and needs to be taken into account only at high Q^2 ($Q^2 \gtrsim M_Z^2$). Therefore, the structure of the proton is over a large part of the phase space driven by the structure function F_2.

The Naïve Quark Parton Model

The Quark Parton Model (QPM) was developed by Feynman in attempt to explain the scaling behaviour of F_2 predicted by Bjorken [Bjo69]. Bjorken scaling anticipates F_2 to be independent of Q^2 and depend only on x at high enough Q^2. First DIS experiments at the Stanford Linear Accelerator Center (SLAC) confirmed this behaviour at $x \sim 0.25$ [BC⁺69, BF⁺69, FK72]. Feynman explained this behaviour by the proton being made out of pointlike constituents which he called partons [Fey69]. The inelastic electron-proton cross section can then be approximated by an incoherent sum of elastic lepton-parton scatters via the exchange of virtual vector bosons. It turns out that the partons interacting directly with the lepton are spin-$\frac{1}{2}$ particles, which could later be identified with the quarks. In the QPM the structure function F_2 can then be written as

$$F_2(x, Q^2) = \sum_i e_i^2 \, x \, f_i(x), \tag{2.33}$$

where the function $f_i(x)$ is the Parton Distribution Function (PDF) of the proton. The sum runs over all quarks and anti-quarks in the proton, which have electrical charge e_i. In the QPM, which corresponds to the Born level of DIS, $xf_i(x)$ gives the probability to find a parton i carrying the momentum fraction x in the proton. $f_i(x)$ is independent of α_s and any scale and directly proportional to F_2. The QPM predicts $F_L(Q^2, x) = 0$, which is a consequence of the scattering from spin-$\frac{1}{2}$ partons in the absence of QCD radiation. It follows that $F_2 = 2xF_1$, which is known as the Callan-Gross relationship [CG69].

The QCD Improved Parton Model

Lowest order QCD corrections to the DIS process include gluon emission either from the initial or final quark and gluon splittings into quark-antiquark pairs. Taking

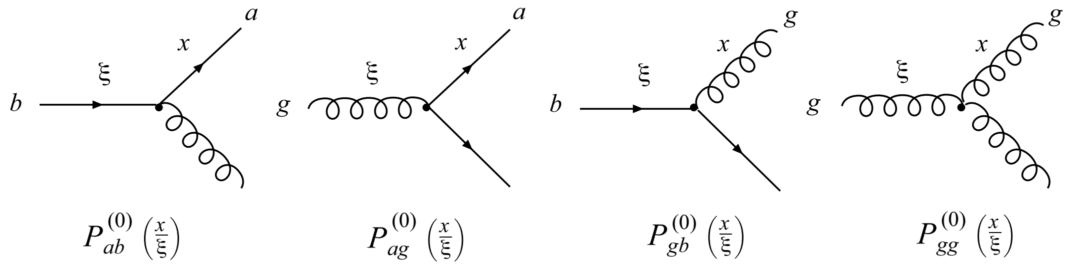


Figure 2.6: The LO DGLAP splitting functions $P^{(0)}_{ab}(\frac{x}{\xi})$, which take the meaning of the probability that an parton a with momentum fraction x emerges from an incident parton b with momentum fraction ξ.

these corrections into account, the parton densities acquire a scale dependence. This can be interpreted as an effect of the resolution power at which the proton is probed (2.31). At very small values of momentum transfer, or correspondingly large wavelength, the virtual photon cannot resolve the proton which behaves as a single object. At intermediate scales, the virtual photon scatters off the constituent quarks. By increasing the energy, the photon may resolve the scattered quark into a quark and a gluon. The initial parton with momentum fraction ξ can thus be resolved into two partons with smaller momentum fractions $x < \xi$ and $\xi - x$. This is illustrated in figure 2.5.
The Dokshitzer, Gribov, Lipatov, Altarelli, Parisi (DGLAP) formalism [Dok77, GL72, AP77] describes the evolution of the parton densities as the scale μ_f changes, where μ_f is the factorisation scale. Formally, μ_f is introduced through the factorisation theorem, as described in section 3.1. The DGLAP equations are coupled integro-differential equations which can be written in terms of splitting functions $P_{ij}(\frac{x}{\xi})$,

$$\frac{\partial}{\partial \ln \mu_\mathrm{f}^2}\begin{pmatrix} q_i(x,\mu_\mathrm{f}^2) \\ g(x,\mu_\mathrm{f}^2)\end{pmatrix} = \frac{\alpha_s(\mu_\mathrm{r})}{2\pi}\sum_j \int_x^1 \frac{\mathrm{d}\xi}{\xi}\begin{pmatrix} P_{q_iq_j}(\frac{x}{\xi}) & P_{q_ig}(\frac{x}{\xi}) \\ P_{gq_j}(\frac{x}{\xi}) & P_{gg}(\frac{x}{\xi})\end{pmatrix}\begin{pmatrix} q_j(\xi,\mu_\mathrm{f}^2) \\ g(\xi,\mu_\mathrm{f}^2)\end{pmatrix}, \tag{2.34}$$

where the argument $\alpha_s(\mu_\mathrm{r})$ in the parton distributions and splitting functions has been dropped. The functions $q_i(x,\mu_\mathrm{f}^2)$ and $g(x,\mu_\mathrm{f}^2)$ label the quark and gluon parton distributions, respectively. The splitting functions are expanded as power series in $\alpha_s(\mu_\mathrm{r})$,

$$P_{q_iq_j} = \delta_{ij}P^{(0)}_{qq}\left(\frac{x}{\xi}\right) + \frac{\alpha_s(\mu_\mathrm{r})}{2\pi}P^{(1)}_{q_iq_j}\left(\frac{x}{\xi}\right) + \dots \tag{2.35}$$

$$P_{ab} = P^{(0)}_{ab}\left(\frac{x}{\xi}\right) + \frac{\alpha_s(\mu_\mathrm{r})}{2\pi}P^{(1)}_{ab}\left(\frac{x}{\xi}\right) + \dots \quad , \tag{2.36}$$

where the indices ab correspond to q_ig, gq_i and gg. In LO the splitting functions can be directly calculated from the QCD vertices and take the meaning of an emission probability: $P^{(0)}_{ab}(\frac{x}{\xi})$ is the probability that a parton a with momentum fraction x emerges from an incident parton b with momentum fraction ξ, as shown in figure 2.6. The splitting functions are known up to NNLO, i.e. $\mathcal{O}(\alpha_s^2)$ [VM+04, MV+04].
The prescription that the momentum fraction of the incoming parton ξ is always larger than that of the outgoing parton x, leads to a strong ordering in the virtualities

$k_{i,\mathrm{t}}$ of the radiated partons, $\mu_\mathrm{f} \gg k_{i,\mathrm{t}} \gg k_{i-1,\mathrm{t}} \gg \ldots \gg Q_0$, where Q_0 is some cut-off scale below which the soft non-perturbative contributions are absorbed in the renormalised parton densities. These can be connected to the structure functions F_i with appropriate renormalisation and factorisation scheme dependent coefficient functions.

Although the DGLAP formalism gives a prescription of how the renormalised parton densities evolve with μ_f, perturbative methods do not allow the calculation of the parton densities at the starting scale Q_0. This information is usually obtained by parameterising the parton densities at the starting scale and fitting the free parameters to structure function data at some higher scale Q. An example of such a fit at NLO to the combined H1 and ZEUS HERA-1 dataset [AA+10a] is shown in figure 2.7. The data are very precise, with typical uncertainties around 1–2%. Scaling violations are apparent in the data and well described by the DGLAP formalism over a large range in Q^2 and x. Since the photon does not couple directly to the gluon, the parameters of the gluon PDF can only be obtained through the scaling violations. Furthermore, the shape of the gluon PDF is strongly correlated with the value of $\alpha_s(M_Z)$, because in the expressions for the structure functions the gluon PDF only appears convoluted with the strong coupling $\alpha_s(\mu_\mathrm{r})$. Thus, an increase in $\alpha_s(M_Z)$ can be compensated by a more steeply rising gluon distribution and vice versa. The value of the strong coupling is therefore usually kept fixed for PDF extractions [CTEQ08, H109c]. Including jet data at high Q^2 in the fit can help to disentangle the strong coupling from the gluon PDF. In this way the correlation between the gluon PDF and $\alpha_s(M_Z)$ can be taken into account, which leads to a determination of the gluon PDF with a simultaneous measurement of $\alpha_s(M_Z)$ [H101a, ZEUS05b, MS+09].

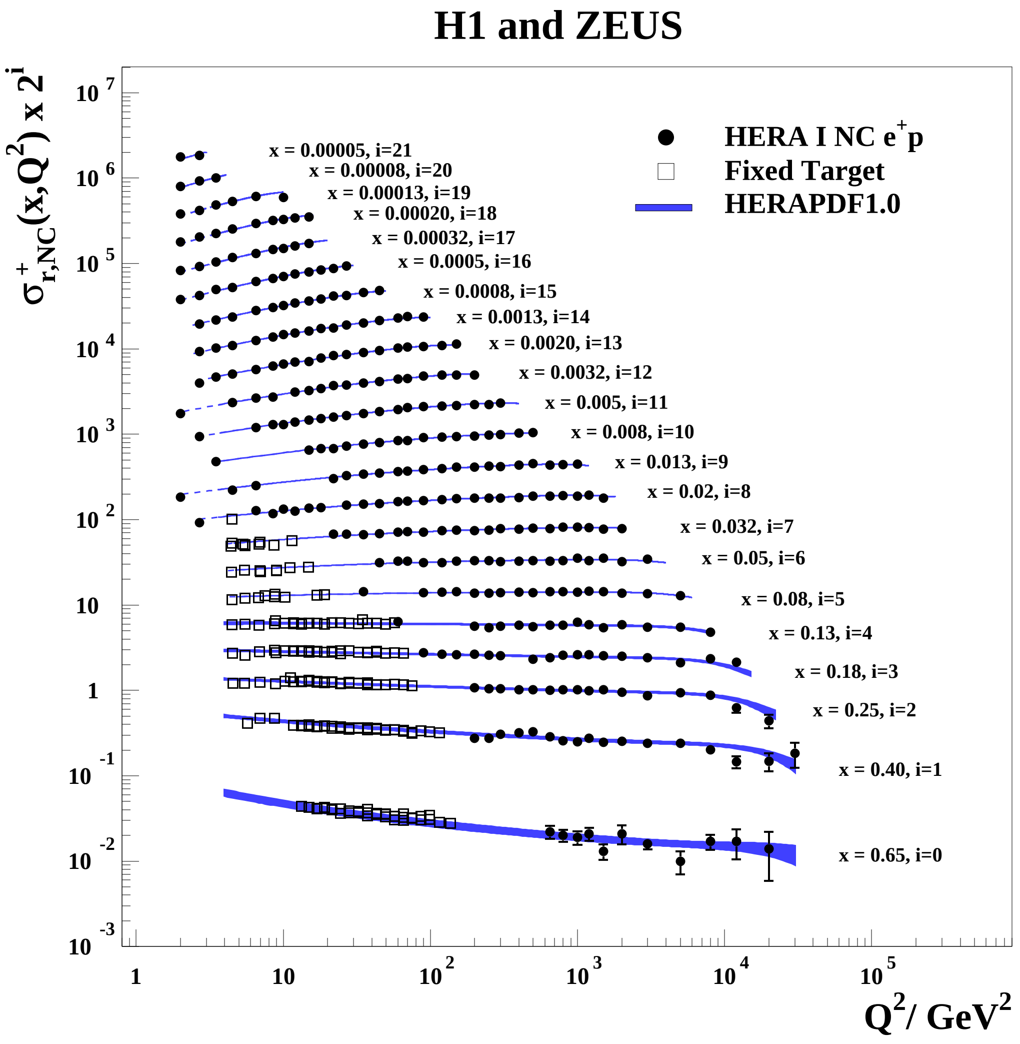


Figure 2.7: H1 and ZEUS combined structure function data from e^+p NC DIS scattering as function of Q^2 for different values of x. The blue bands correspond to NLO DGLAP calculations using the HERAPDF1.0 [AA$^+$10a]. Also shown are data from the fixed target experiments BCDMS [BCDMS89] and NMC [NMC97], which were not used for the extraction of the HERAPDF1.0.

Chapter 3

Jet Production in DIS

QCD is formulated as a theory of the interaction of quarks and gluons. These partons carry colour charge and are confined and therefore not directly observable. If two or more coloured partons with sufficient energy are separated, long distance effects, for example as described by colour strings, cause the formation of colourless hadrons. These can be observed and appear in detectors as collimated sprays of particles, which are called jets.

Apart from having the advantage of being strongly correlated to the original parton, jet definitions obeying certain rules make pQCD calculations possible in the first place. In this chapter an overview of the production of jets in the context of pQCD is given. Requirements on jet algorithms are reviewed and different jet definitions are introduced. The models that describe the formation of hadron jets out of the underlying quarks and gluons are described. An overview of the used fixed-order pQCD calculations and MC event generators is given.

3.1 The Factorisation Theorem

The applicability of the perturbative series (2.22) is only ensured for processes where no hadrons are present in the initial state, like in e^+e^--annihilation. In DIS, the presence of a bound state of coloured partons in the initial state inevitably leads to non-perturbative long-range effects with characteristic scales $\mu < \Lambda$. In hadron induced processes the applicability of pQCD is ensured by the factorisation theorem [LS78, EG+79], which states that a cross section $\sigma(x, Q^2)$ can be written as convolution of the hard process, which is calculable in pQCD, with a soft part accounting for the long-range effects. The factorisation theorem in DIS takes the form [CS+88]

$$\sigma(x, Q^2) = \sum_{i=q,\bar{q},g} \int_x^1 \frac{\mathrm{d}\xi}{\xi}\, f_i\left(\xi, \mu_\mathrm{f}^2, \alpha_s(\mu_\mathrm{r})\right)\, \hat{\sigma}_i \left(\frac{x}{\xi}, \frac{Q^2}{\mu_\mathrm{r}^2}, \frac{\mu_\mathrm{r}^2}{\mu_\mathrm{f}^2}, \alpha_s(\mu_\mathrm{r})\right) \tag{3.1}$$

and is shown schematically in figure 3.1. The sum in equation (3.1) runs over all partons in the proton, where the function $f_i(\xi, \mu_\mathrm{f}^2, \alpha_s)$ is the Parton Distribution Function (PDF) for parton i of the proton. The partonic cross section, which depends on the incoming parton and its momentum fraction, is denoted by $\hat{\sigma}_i$. In LO,

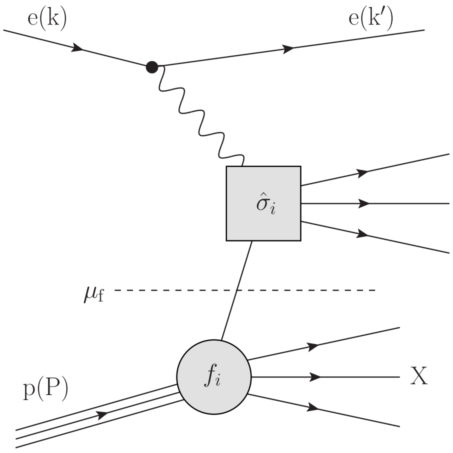


Figure 3.1: Schematic representation of the factorisation theorem in DIS. The factorisation scale $\mu_{\rm f}$ separates soft contributions, which are absorbed in the parton density function f_i, from hard contributions $\hat{\sigma}_i$, calculable in pQCD.

$f_i(\xi, \mu_{\rm f}^2, \alpha_s)\,{\rm d}\xi$ is independent of $\mu_{\rm f}$ and α_s and gives the probability to find a parton i carrying the momentum fraction between ξ and $\xi + {\rm d}\xi$ in the proton. In NLO this intuitive picture is more intricate. The PDFs depend on the factorisation and renormalisation scheme and the chosen scales and cannot be readily interpreted as physical probability distributions anymore. The long-range behaviour of the strong force is absorbed in the PDF, which leaves the partonic cross section $\hat{\sigma}_i$ calculable in pQCD. The factorisation theorem thus allows the expansion of $\hat{\sigma}_i$ in a perturbative series similar to (2.22), with the compromise of the introduction of a new arbitrary scale $\mu_{\rm f}$. This factorisation scale regulates the classification of processes between perturbative and non-perturbative QCD. Similar to the renormalisation scale $\mu_{\rm r}$, the dependence on $\mu_{\rm f}$ becomes weaker the more terms are included in the perturbative expansion. In inclusive DIS the natural choice for the scale $\mu_{\rm f}$ is Q, since it is the only hard scale available. The factorisation property of the cross section is a fundamental property of QCD and can be proven to all orders in perturbation theory [CS$^+$88].

3.2 The Jet Function

Although infrared sensitivity arising from the initial state has been factored into the PDF and UV divergences are removed by renormalisation, the partonic cross section $\hat{\sigma}_i$ in equation (3.1) is not free from singularities for all observables. The singularities arising can be grouped in two classes, infrared and collinear divergences. Infrared divergences originate from massless partons, which are allowed to emit a parton with zero momentum and still remain on-shell. Poles in the partonic cross section at $p^\mu = 0$ then lead to infrared divergences. Collinear divergences arise when a parton with momentum p^μ emits a parton collinear to it, with momentum zp^μ, where $0 < z < 1$. Integration over momenta leads to divergences close to zp^μ [Ste78]. For inclusive observables, like the total $ep \to eX$ cross section these infrared and collinear singularities from real emissions are cancelled by virtual corrections to all orders in perturbation theory. In order for these cancellations to work for more

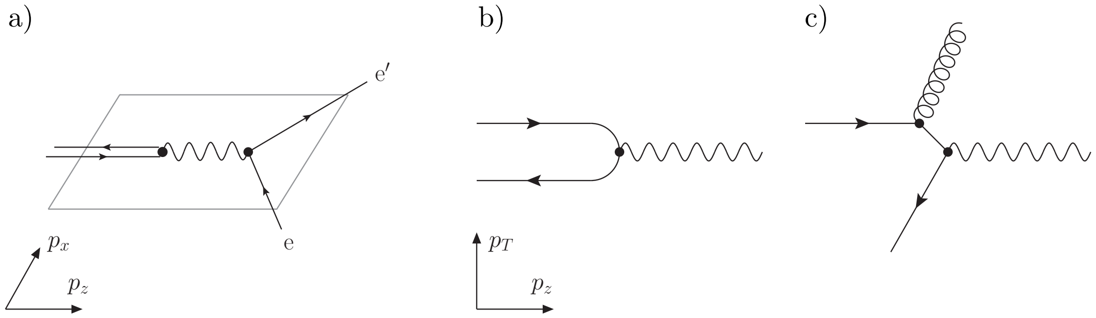


Figure 3.2: a) Definition of the Breit frame. b) Born ($\mathcal{O}(\alpha_s^0)$) and QCD Compton process ($\mathcal{O}(\alpha_s)$) (c) of deep-inelastic scattering in the Breit frame.

exclusive quantities like jet production, the defined cross section has to be infrared and collinear safe [SS+95]. To meet these requirements the n-parton cross section has to be convoluted with a jet function $\mathcal{J}_n(p_1, \ldots, p_n)$, which has the momenta p_i of all final state partons as arguments. Infrared safety means that the observable must not change if a final state consists of one or more partons with zero momentum or if these partons are omitted entirely. In terms of the jet function, this requirement can be expressed as

$$\mathcal{J}_{n+1}(p_1^\mu, \ldots, p_n^\mu, p_{n+1}^\mu = 0) = \mathcal{J}_n(p_1^\mu, \ldots, p_n^\mu)\,. \tag{3.2}$$

Collinear safety requires

$$\mathcal{J}_{n+1}(p_1^\mu, \ldots, zp_n^\mu, (1-z)p_{n+1}^\mu) = \mathcal{J}_n(p_1^\mu, \ldots, p_n^\mu) \tag{3.3}$$

which means that the observable should not differ between a final state in which two particles are collinear and the same final state where the two particles are replaced by one.

Since infrared and collinear divergences originate from soft, non-perturbative dynamics, using a jet definition fulfilling equations (3.2) and (3.3) ensures that the resulting jets are insensitive to the long-range behaviour of QCD and the n-jet cross section becomes calculable in pQCD. These cross sections are defined for coloured partons and need to be corrected for hadronisation effects before a comparison to data can be made.

3.3 The Breit Frame

Besides being collinear and infrared safe, a jet algorithm also has to provide a factorisable jet cross section. This means that the factorisation theorem (3.1) is not violated by the jet definition, and initial state singularities are absorbed in the definition of the PDF. There exists a class of reference frames which ensure factorisation, one of them being the Breit frame [Web93]. The Breit frame is defined through the relation [Fey72]

$$2\,x\,\vec{p} + \vec{q} = 0\,, \tag{3.4}$$

where $\vec{p}$ and $\vec{q}$ are the momenta of the incoming proton and the exchanged virtual boson, respectively. The positive z-direction is chosen to be the direction of the incoming proton. The x-z plane is given by the plane spanned by the momentum vectors of the incoming and outgoing scattered electron, such that the azimuthal angle of the scattered electron in the Breit frame $\phi_{\mathrm{e}}^{\mathrm{Breit}} = 0$ (see figure 3.2a).

In the Breit frame the exchanged boson and the struck quark collide head-on. The boson is purely space-like with $q = (0, 0, 0, -Q)$. In the Born process of DIS the incoming quark has momentum $p_z = Q/2$ and is back-scattered with momentum $p_z = -Q/2$ ("brick-wall frame"). In processes involving the strong interaction ($\mathcal{O}(\alpha_s)$) final state partons acquire transverse momentum in the Breit frame, see figures 3.2b and c. Hence, in the Breit frame the hardness of the QCD interaction is reflected by the jet transverse momentum. Using the Breit frame as reference frame for a jet analysis brings the additional experimental advantage that Born processes can be distinguished from higher order QCD processes simply by the requirement of significant transverse momentum.

3.4 Jet Algorithms

A set of rules which identify jets and merge the constituents to obtain the jet four-vector is called a jet algorithm or jet finder. The exact implementation of an algorithm incorporating the requirements described above is non-trivial, especially if the jet algorithm should be infrared and collinear safe in all orders of perturbation theory. For the following discussion, the word particle will be used to generically refer to the input objects of a jet finder, and can mean partons, stable particles or protojets from a previous combination step.

A class of jet finders which fulfils the requirements of infrared and collinear safety in all orders is defined through the distance measures d_{ij} and d_i between the input particles,

$$d_{ij} = \min(k_{\mathrm{Ti}}^{2p}, k_{\mathrm{Tj}}^{2p}) \frac{\Delta_{ij}^2}{R^2} \, , \tag{3.5}$$

$$d_i = k_{\mathrm{Ti}}^{2p} \tag{3.6}$$

with $\Delta_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$. The variables k_{Ti}, ϕ_i and η_i are the transverse momentum, azimuthal angle and pseudorapidity[1] of particle i, respectively. The free parameter R is a distance measure comparable to the radius of cone algorithms. The parameter p governs the importance of the geometrical distance with respect to the 'distance' in transverse momenta of the particles. The longitudinally invariant k_{T}-algorithm [CD+93, ES93] is obtained by setting $p = 1$, the Cambridge/Aachen algorithm [DL+97, Wob00] uses $p = 0$, and the choice $p = -1$ yields the anti-k_{T} algorithm [CS+08]. These jet algorithms are clustering algorithms which define jets by successive recombination of the input particles. The iterative procedure which results in a list of jets works as follows:

1. For all pairs of input particles calculate the distance d_{ij} and for each particle calculate its distance to the beam d_i.

---

[1] The pseudorapidity is defined as $\eta = -\ln(\tan\theta/2)$, where θ is the polar angle of the particle. For massless particles it is identical to the rapidity.

2. Find the minimum d_{min} of all d_{ij} and d_i. If d_{min} is a d_{ij} then merge particles i and j into a single particle. If d_{min} is a d_i, remove it from the list and add it to the list of jets.

3. Repeat from step one until no particles are left.

In step two there is a certain degree of freedom on how to merge two particles into one. In this work the P_{T}-recombination scheme [HW+90] is employed. For massless input particles, it defines the total transverse momentum as sum of transverse momenta and the azimuthal angle and pseudorapidity as weighted sums,

$$k_{\mathrm{Tij}} = k_{\mathrm{Ti}} + k_{\mathrm{Tj}}\,, \tag{3.7}$$

$$\phi_{ij} = (\phi_i k_{\mathrm{Ti}} + \phi_j k_{\mathrm{Tj}})/k_{\mathrm{Tij}}\,, \tag{3.8}$$

$$\eta_{ij} = (\eta_i k_{\mathrm{Ti}} + \eta_j k_{\mathrm{Tj}})/k_{\mathrm{Tij}}\,, \tag{3.9}$$

which results in massless jets. In the case of massive input particles, their energies are scaled to be equal to the magnitude of their three-momenta [BC+03].

A different class of jet algorithms are cone algorithms, which define jets based on momentum flow within a cone. Cone algorithms are mostly employed by hadron-hadron collider experiments due to the predictable geometrical shape of the resulting jets. Due to limited computing power often seeded cone algorithms were used [CDF08, D008], which however turned out not to be infrared and collinear safe to all orders in perturbation theory. With the new SISCone algorithm [SS07] a fast implementation of a seedless cone algorithm became available which is infrared and collinear safe to all orders.

In a recent publication the ZEUS collaboration compared the k_{T}, anti-k_{T} and SISCone algorithms for jet production in NC DIS [ZEUS10b]. The inclusive jet cross sections obtained with the three algorithms are very similar, as shown in figure 3.3a and b. They differ by less than 3.6% over most of the phase space, only at high P_{T} the differences between the k_{T} algorithm and the other two become around 5%. The observed ratios are very well described by calculations including $\mathcal{O}(\alpha_s^3)$ terms. In figures 3.3c and d the hadronisation corrections (defined in section 10.6.2) together with their uncertainties are shown for the three different jet finders. Jets defined with the k_{T} and anti-k_{T} algorithms have the smallest hadronisation corrections, meaning that they are relatively insensitive to the non-perturbative fragmentation of quarks and gluons into hadrons. The hadronisation corrections obtained for the k_{T} algorithm show overall the smallest uncertainties.

Due to these considerations the k_{T} algorithm with $R = 1$ is used in this work. Unless otherwise stated, jets are found and clustered in the Breit frame.

3.5 Jet Cross Sections in Perturbative QCD

At high enough scales the small size of the strong coupling allows for the perturbative expansion of jet observables in terms of α_s. The leading order (LO) contributions of these perturbative series can already explain specific jet topologies and give a basic understanding of the jet observables. However, reliable predictions can only be obtained by the inclusion of higher orders. In the following an overview of the

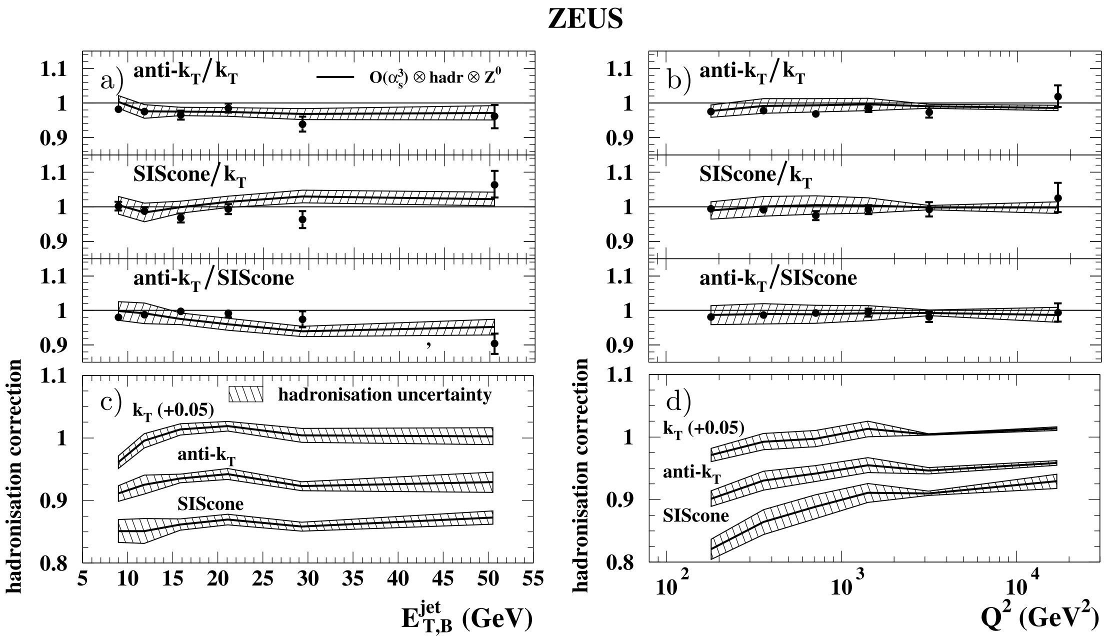


Figure 3.3: Ratios of inclusive jet cross sections obtained with different jet algorithms as function of the jet transverse momentum in the Breit frame (a) and Q^2 (b). Also shown are calculations for these ratios including $\mathcal{O}(\alpha_s^3)$ terms. Hadronisation corrections for the different jet algorithms as function of P_T^{jet} (c) and Q^2 (d) are shown below, from [ZEUS10b].

leading order contributions to dijet and trijet production is given. This is followed by a summary of calculations used in this work which include higher order corrections.

3.5.1 Dijet and Trijet Production in Leading Order

The two leading order processes for dijet production in DIS are the QCD Compton (QCDC) and Boson Gluon Fusion (BGF) processes. Both are 2 → 2 scatterings of order $\mathcal{O}(\alpha_s)$, where QCDC is the $\gamma^* q \to gq$ process and BGF is the $\gamma^* g \to q\bar{q}$ process. The relevant Feynman diagrams are shown in figure 3.4. Besides the kinematic variables Q^2 and x, three more independent variables are needed to fully characterise the kinematics of dijet production. Usually, these are chosen to be z, x_p and φ [PR80]. The angle φ is the relative azimuth between the electron and parton scattering planes in the hadronic centre-of-mass frame or any frame related to it with a longitudinal boost. The variables z and x_p are given by

$$z = \frac{p_a \cdot p_1}{q \cdot p_a}\,, \qquad x_p = \frac{Q^2}{2p_a \cdot q}\,, \tag{3.10}$$

where p_a is the momentum of the incoming parton, p_1 and p_2 are the momenta of the outgoing partons, and q is the momentum of the virtual boson as shown in figure 3.4. The variable z is related to the scattering angle in the parton-parton centre-of-mass frame and is in the range $0 \le z \le 1$. The variable x_p can be interpreted as the momentum fraction of the proton taken by the parton interacting with the virtual boson for which $x \le x_p \le 1$ holds. In LO, the centre-of-mass energy squared of the virtual boson and incoming parton, $\hat{s} = (p_a + q)^2$, coincides with the squared

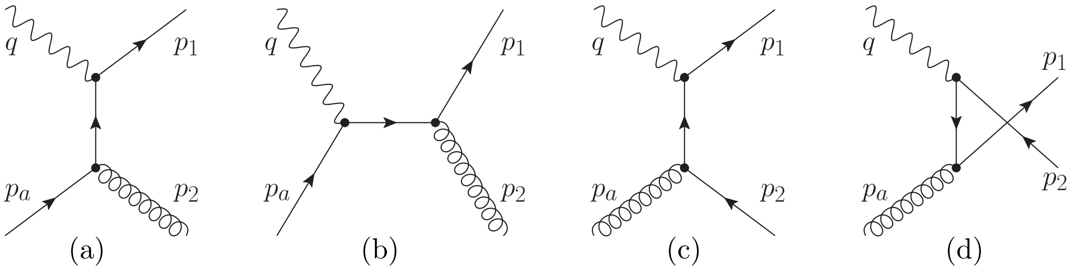


Figure 3.4: LO Feynman diagrams for 2-jet production in NC DIS. The QCDC (a), (b) and BGF diagrams of order $\mathcal{O}(\alpha_s)$.

invariant mass of the outgoing partons, $M_{12}^2 = (p_1 + p_2)^2$. Denoting by ξ the fraction of the proton's momentum carried by the initial state parton, p_a can be written as $p_a = \xi P$. This leads to

$$\xi = x \left(1 + \frac{M_{12}^2}{Q^2}\right) . \tag{3.11}$$

It follows from the definition of x (2.27) that $\xi = x/x_p$. In LO, where the centre-of-mass energy of the partonic final state coincides with the invariant mass of the dijet system, the variable ξ coincides with the momentum fraction of the parton to be used in the evaluation of the PDF of the proton, $f_i(x = \xi, \mu_\mathrm{f}^2, \alpha_s)$. With the inclusion of higher order corrections, the invariant mass of the dijet system becomes smaller than the centre-of-mass energy of the partonic final state. As a consequence $\xi \geq x$, which can also be seen from its definition (3.11). The variable ξ then takes the meaning of the lower bound for the integration over all momentum fractions when convoluting the PDFs with the partonic matrix elements $\hat{\sigma}_i$.

Integrating over φ, the partonic cross section $\hat{\sigma}$ for the QCDC and BGF processes can be written as [PR80]

$$\hat{\sigma}_\mathrm{QCDC} \propto \frac{1 + x_p^2 z^2}{(1-z)(1-x_p)} , \tag{3.12}$$

$$\hat{\sigma}_\mathrm{BGF} \propto \frac{\left(z^2 + (1-z)^2\right)\left(x_p^2 + (1-x_p)^2\right)}{z(1-z)} . \tag{3.13}$$

These expressions are singular for $x_p \to 1$, $z \to 1$ or $z \to 0$, which are related to either collinear parton configurations or one of the final partons being soft. A detailed study shows that the singular regions in the jet phase space appear if the jets have either vanishing transverse momentum or infinite pseudorapidities. The requirement of non-zero transverse momentum of the outgoing partons avoids the singular regions during the integration over x_p and z.

In leading order, trijet production can be considered as the radiation of an extra gluon from either the incoming or one of the outgoing partons in dijet production. Thus, it includes $\mathcal{O}(\alpha_s^2)$ terms and is the lowest order of jet production in NC DIS where the triple-gluon vertex directly contributes. Some of the diagrams contributing to $\gamma^* q \to qgg$ and $\gamma^* g \to q\bar{q}g$ scattering are shown in figure 3.5. In these processes, the centre-of-mass energy squared is equal to the invariant mass squared of the three outgoing partons, $M_{123}^2 = (p_1 + p_2 + p_3)^2$. The fraction of the proton's momentum carried by the initial state parton is thus given by a similar expression

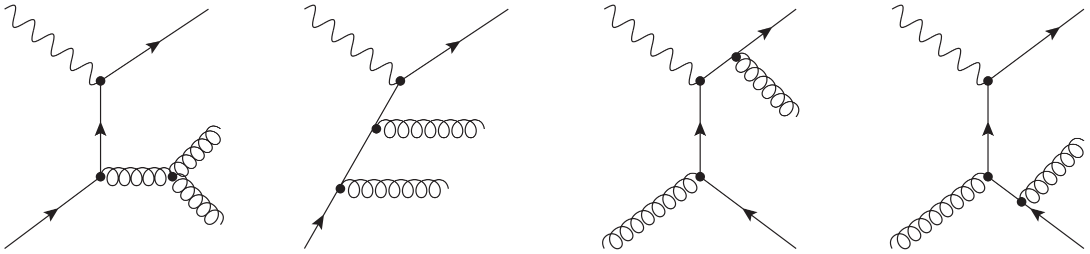


Figure 3.5: Some of the LO Feynman diagrams contributing to trijet production in NC DIS.

as equation (3.11), by replacing M_{12} by M_{123}. Similar to the dijet case, infrared and collinear singularities can be avoided by requiring non-vanishing transverse momenta of all three outgoing partons.

3.5.2 Jet Production Beyond Leading Order

Although perturbative calculations in LO can already explain specific event topologies and give an estimate for the magnitude of a cross section, reliable predictions need at least the inclusion of terms proportional to the next higher order in the strong coupling. Since dijet production is in LO proportional to $\mathcal{O}(\alpha_s)$, the NLO contributions include terms of $\mathcal{O}(\alpha_s^2)$. The NLO contributions for trijet production include terms of $\mathcal{O}(\alpha_s^3)$.

The contributions to the NLO calculations can be grouped in two classes: real and virtual corrections. Real corrections include the emission of one additional gluon with respect to the LO diagrams. For example, the LO diagrams for trijet production, some of which are shown in figure 3.5, are the NLO real corrections for dijet production. These contributions introduce divergences due to collinear and soft parton configurations. Virtual corrections, on the other hand, do not change the number of emitted partons. They include internal loops, which introduce infrared (IR) and UV divergences due to vanishing and infinite loop momenta. The UV divergences are removed by renormalising the fields, gauge parameters and masses in the Lagrangian, similar to the renormalisation of the strong coupling. The renormalisation results in counter terms, which render the physical Green's function finite. These counter terms exhibit an explicit logarithmic dependence on the renormalisation scale μ_r, which cancels the first term of the expansion of the running coupling (2.19), thus reducing the scale dependence of the NLO calculation. The IR divergences from the virtual contributions cancel order-to-order in the perturbative expansion with the divergences from the real contributions according to the Kinoshita-Lee-Nauenberg theorem [Kin62,LN64]. The requirements for these cancellations to work are infrared and collinear safety, which is ensured by the jet function. Additional divergences arise from configurations where the parton from the proton, undergoing the hard scattering, radiates a soft or collinear secondary parton which cannot be resolved. Divergences of this sort are called initial state singularities and can be absorbed in the definition of the proton PDFs. This introduces a scale dependence of the coefficient functions on the factorisation scale μ_f, which partially cancels the scale dependence of the PDF. Consequently, already the NLO corrections substantially reduce the scale dependence of perturbative calculations and lead to reliable predictions.

In practice, the jet phase space is usually more restricted than only requiring non-zero jet transverse momenta. For example, a restriction to a specific region of jet pseudorapidity due to limited detector acceptances is unavoidable in experimental analyses. This makes analytic calculations almost impossible. Numerical methods allow the calculation of arbitrary infrared safe observables through MC integration. These computer codes can be classified in two categories. Firstly, fixed-order pQCD calculations which calculate cross sections on first principles. These programs require only the proton PDF and physical parameters of the SM (i.e. couplings, masses,...) as input parameters. The resulting cross sections are parton level cross sections, which in general cannot be directly compared to measurements since the defined observables may be sensitive to hadronisation processes. However, for jet observables it is possible to estimate the hadronisation corrections, making fixed-order calculations a powerful tool to test the validity of pQCD and extract SM parameters. The second class of available MC codes are full-scale event generators, which include terms to all orders in α_s through parton showers or subsequent gluon emission from colour dipoles. These approximations neglect virtual contributions such that probabilities are added, instead of the summation of probability amplitudes. In this framework it thus becomes possible to model long-range effects like hadronisation with appropriate approximations. Often QED corrections are included in MC event generators such that calculated observables can be directly compared to data. An overview of the event generators used in this work is given in section 3.6.

3.5.3 Fixed Order Calculations

Calculations beyond the Born level quickly become very involved due to the large number of contributing Feynman diagrams. In order to perform numerical integrations the divergences have to be separated such that cancellations can either be achieved analytically or through the construction of a non-trivial local counter-term. The latter method is implemented in the program NLOJet++ [NT98, NT01] and is outlined in the following.

In NC DIS jet production the n-jet cross section is of the order $\mathcal{O}(\alpha_s^{n-1})$. The NLO corrections thus involve $\mathcal{O}(\alpha_s^{n})$ terms, which include $n+1$ parton configurations, which are the real emissions, and n-parton configurations with one internal loop, which are the virtual corrections. In NLO the partonic n-jet cross section can be written symbolically as

$$\hat{\sigma}_n = \int\limits_{n} \mathrm{d}\hat{\sigma}_n^{\mathrm{B}} + \int\limits_{n+1} \mathrm{d}\hat{\sigma}_{n+1}^{\mathrm{R}} + \int\limits_{n} \mathrm{d}\hat{\sigma}_n^{\mathrm{V}} \,, \tag{3.14}$$

where the superscripts B, R and V respectively denote the Born, real and virtual contributions, and the integrations are performed over all outgoing parton momenta. The exclusive partonic cross sections $\mathrm{d}\hat{\sigma}_n$ are infrared and collinear safe because of a convolution with the jet function $\mathcal{J}_n$,

$$\mathrm{d}\hat{\sigma}_n = \mathrm{d}\Phi^{(n)} |\mathcal{M}_n(p_1, \ldots, p_n)|^2 \mathcal{J}_n(p_1, \ldots, p_n) \,, \tag{3.15}$$

where $\mathrm{d}\Phi^{(n)}$ and $\mathcal{M}_n$ denote the phase space and matrix elements to produce the n-parton final state. The last two integrals in equation (3.14) are separately divergent.

The general method of removing the singularities is based on counter terms $\mathrm{d}\hat{\sigma}^{A}_{n+1}$, which can be constructed to have the same singular behaviour as $\mathrm{d}\hat{\sigma}^{\mathrm{R}}_{n+1}$ but with opposite sign. The n-jet cross section then becomes

$$\hat{\sigma}_n = \int\limits_{n} \mathrm{d}\hat{\sigma}^{\mathrm{B}}_{n} + \int\limits_{n+1} \left[\mathrm{d}\hat{\sigma}^{\mathrm{R}}_{n+1} - \mathrm{d}\hat{\sigma}^{\mathrm{A}}_{n+1}\right] + \int\limits_{n+1} \mathrm{d}\hat{\sigma}^{\mathrm{A}}_{n+1} + \int\limits_{n} \mathrm{d}\hat{\sigma}^{\mathrm{V}}_{n} \,, \tag{3.16}$$

where the first two integrals are finite and numerically calculable. The last two integrals have different phase space integrations, but by construction the counter-term can be split into a one parton and an n-parton phase space. The contributions to the one parton subspace are termed dipoles. These terms have the same singular behaviour as the virtual contributions but with opposite sign, which renders the last two integrals numerically integrable. The cancellations rely on infrared and collinear safe observables which fulfil the requirements of the jet function as described in section 3.2. This method is called dipole subtraction method and was developed by Catani and Seymour [CS97]. The difficulty lies in the construction of the counter terms, but once they have been constructed they are universal. The counter terms have been calculated for 2- and 3-jet production in NC DIS at NLO. They are implemented in the program NLOJet++. At present no NNLO calculations for jet production in DIS are available.

3.5.4 Higher Order Corrections

Calculations beyond NLO become increasingly demanding due to the large number of diagrams which need to be considered. Nevertheless, there are regions in phase space where higher-order terms become non-negligible and cannot be ignored. Instead of including all terms of a fixed order in α_s, it is possible to take enhanced terms approximately into account to all orders. This leads to parton showers, which make it possible to generate events at parton level down to an almost non-perturbative scale Q_0, which is typically taken to be of the order of 1 GeV.

In DIS the quark struck by the virtual boson can emit additional partons before and after the hard interaction. The emerging cascade of partons is termed initial or final state parton shower, respectively. In a parton shower a propagating parton can split into a pair of partons with momentum fractions z and $(1-z)$. The probability of such a splitting to happen can be expressed in terms of the Sudakov form factor [Sud56]. This form factor encodes the probability for a parton not to split between some initial maximum virtuality and some minimal value of it. The Sudakov form factor depends on the DGLAP splitting functions $P_{ab}(z)$ from section 2.3.2 and the PDFs. A DIS event including initial and final state parton showers is depicted in figure 3.6. Viewing the interaction in ‘chronological’ order, the initial state parton shower is initiated by a parton close to mass-shell from the incoming proton. At each splitting one parton becomes increasingly off-shell with space-like virtuality $k_i^2 < 0$ and the other parton is on-shell or has a time-like virtuality $m^2 > 0$. At each splitting the parton with space-like virtuality continues towards the hard interaction, such that either a gluon or quark with large space-like virtuality enters the hard matrix element. The outgoing partons from the hard interaction are close to on mass-shell or have time-like virtualities. This leaves all partons left in the event with

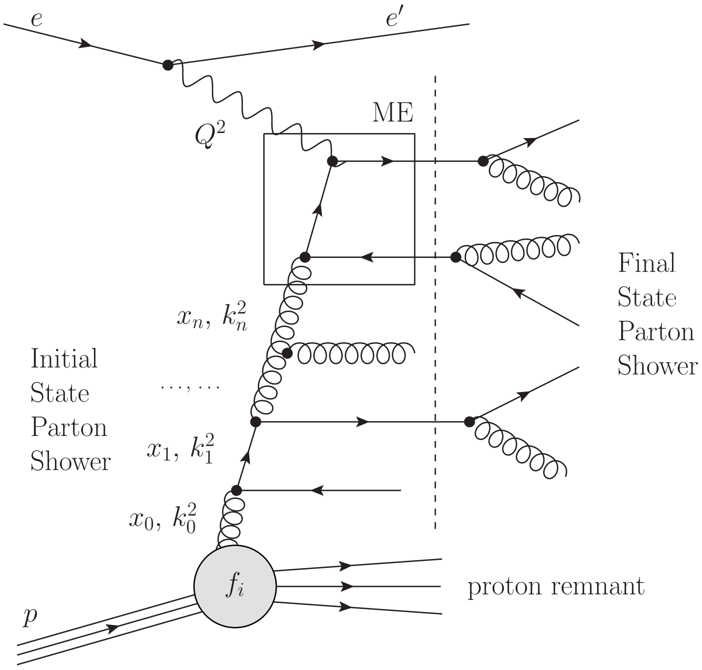


Figure 3.6: Higher order corrections to jet production in NC DIS approximated by the parton shower approach. The subdiagram where the virtual boson couples to the outgoing quark is the matrix element (ME) of the hard interaction. It defines the separation between the initial and final state parton shower, indicated as dashed line. A space-like initial state parton i is described by its virtuality k_i^2 and its fractional momentum of the proton x_i.

either time-like virtualities or being close to mass-shell. In the final state parton shower all partons which are off mass-shell decrease their masses by branching into daughter partons with decreasing off-shell masses and decreasing opening angles. The branchings continue until all partons are essentially on mass-shell.

It can be shown that these initial and final state parton showers obey the solutions of the DGLAP equations (2.34) [ES+96]. In the previous chapter the DGLAP equations described the evolution of the PDFs when probed at a given scale. In the parton shower context, these evolution equations instead represent the momentum fraction distribution of the produced partons. The space-like virtualities k_i^2 of the partons approaching the hard matrix element are strongly ordered, $k_0^2 \ll \dots \ll k_n^2 \ll Q^2$. This is equivalent to a rearrangement of the perturbative expansion in terms of $\alpha_s^m \ln Q^2$. Other logarithmic terms appearing in the expansion, like terms of the form $\alpha_s^m \ln \frac{1}{x}$, are neglected, such that the resummation only includes single logarithmic terms (leading log approximation).

One of the shortcomings of the parton shower approach is the neglect of interference effects between the initial state and final state parton emissions and that it is not gauge invariant. While the parton shower is a good approximation for the emission of soft partons or partons close to the direction of the emitting parton, the emission of hard partons at large angles is not well accounted for [IE+97].

In principle it is possible to attach parton showers to any fixed-order calculation. However, already beyond LO this task becomes very involved since care has to be taken that configurations produced by the hard matrix element are not double counted. At present, only one MC event generator exists for lepton-proton scattering which matches NLO matrix elements with parton showers [Tol10]. Unfortunately, this `MC@NLO` program has been developed for heavy flavour photoproduction and is thus not applicable in this analysis. The MC event generators used in this work have LO matrix elements matched with parton showers.

3.6 Event Generators

MC event generators are computer programs which attempt to fully simulate high energy collisions of particles. These programs comprise a collection of different QCD motivated, phenomenological models used to simulate the different stages of particle collisions. These programs start out from a given initial state and obtain events by the random sampling of possible processes and final states. Through the summation of many events in a certain region of phase space a prediction for the cross section may be obtained. In this analysis the importance of MC event generators lies in the possibility to combine them with detector simulations in order to investigate the detector performance and obtain corrections for the recorded data.

A MC program starts out by calculating the hard interaction in pQCD from first principles typically only in LO. In the next step the available phase space is filled with additional parton radiation. This can either be modelled with DGLAP type parton showers or through gluon radiation according to the colour dipole model. The initial state parton branching and/or gluon emission is performed backwards in time such that the hard interaction is not modified by the parton radiation [BS88]. The simulation of initial state radiation is terminated at a cut-off scale Q_0, which is close to Λ and regularises the parton splittings. At this point the parton propagator is taken from the proton PDF. The final state parton shower is simulated until another cut-off scale is reached. At this stage the parton virtualities are too small for perturbative methods to be applicable any longer. Before all produced partons enter the hadronisation stage corrections to the parton momenta are required, such that energy and momentum conservation is ensured. This is an artefact from having neglected interference terms between the initial and final state parton showers. Different models exist which describe the formation of mesons and hadrons from coloured partons, with the two most commonly used ones being the Lund string model [AG+83] and the cluster fragmentation model [Web84]. The final result of the simulation of a high energy particle collision is a list of stable hadrons together with their four-vectors. This hadronic final state is subject to the detector simulation to model the finite resolutions, efficiencies and acceptances of the experimental apparatus.

The various approximations and models implemented in MC programs introduce a number of free parameters which can be tuned to available data. Many parameters are connected to the hadronisation stage, which is assumed to be universal and independent of the specific model used to obtain the partonic final state. However, some parameters are sensitive to the specific implementation of the parton shower and hadronisation and thus need to be adjusted for every MC program individually.

3.6.1 MC Programs Used

A number of MC programs is used in this analysis for the simulation of signal and background events, the determination of detector efficiencies, detector corrections and corrections for electroweak (EW) effects. In the following the utilised MC programs are briefly described.

Rapgap

RAPGAP [Jun95] uses LO matrix elements matched with DGLAP parton showers (MEPS) for the simulation of DIS events. The hard matrix elements include BGF and QCDC processes for γ^* and Z^0 exchange. In all these respects RAPGAP is similar to LEPTO, which is described below. In RAPGAP the hadronisation in is modelled with the Lund string fragmentation as implemented in PYTHIA. QED radiation from the lepton line, virtual corrections to the leptonic vertex and vacuum polarisation are included through the HERACLES program [KS+92]. In this work RAPGAP is used for the derivation of detector and radiative corrections, as well as for the estimation of hadronisation corrections.

Lepto

The LEPTO [IE+97] generator includes all electroweak processes in LO for lepton-proton scattering. LEPTO is a LO matrix elements and parton shower (MEPS) program, which uses HERACLES for the inclusion of $\mathcal{O}(\alpha_{\text{em}})$ processes.
Since NLO calculations do not include Z^0 exchange, LEPTO is used to estimate its effect on jet cross sections and derive EW correction factors for the NLO calculations.

Djangoh

In its original version the package DJANGO [CS+94] has been developed as an interface between LEPTO and HERACLES. In its present version DJANGOH includes an implementation of the Colour Dipole Model (CDM) [Gus86] by providing an interface to ARIADNE [Lön92].
The CDM can be seen as an alternative formulation of parton showers. Instead of quark and gluon splittings, parton radiation is modelled through gluon emission from colour dipoles. Since the first dipole in an event is spanned between the struck quark and the proton remnant, only QCDC processes and higher order corrections can be simulated by the CDM. A matching procedure has been implemented in ARIADNE which matches the BGF processes with the CDM [Lön95], which leads to an improved description of trijet events.
DJANGOH is used in the CDM mode as an alternative to the MEPS approach implemented in RAPGAP. Changes in the measured jet cross sections due to differences in the derived corrections from DJANGOH and RAPGAP are assigned as model uncertainty.

Pythia

$\gamma^* p$ scattering with $Q^2 \approx 0$, which is also called photoproduction, is one of the sources of background in this analysis. In order to estimate the amount of background, the Pythia program [SM$^+$06] is used for the simulation of light and heavy quark production in direct and resolved photoproduction.

Compton

The quasi-real QED Compton (QEDC) process $ep \rightarrow e\gamma X$ is simulated with higher precision in the Compton generator [CC$^+$91, CK92] than in Heracles. To get a reliable estimate of the background from QEDC processes these events are removed from Djangoh and Rapgap and are instead simulated with the Compton program.

Grape

The background from lepton-pair production, $ep \rightarrow e\,\ell^+\ell^- X$, is estimated with the Grape program [Abe01], which includes electron, muon and tau production in electron-proton scattering.

Chapter 4

Experimental Setup

The HERA collider is the first and so far only high energy electron-proton collider in the world. It is located at the research laboratory Deutsches Elektronen Synchrotron (DESY) in Hamburg and was in operation from 1992 to 2007. During this period the two multi-purpose detectors H1 and Zeus recorded roughly 500 pb^{-1} of data each.

In this chapter a brief overview of the HERA *ep* collider is given, followed by a review of the H1 detector. The focus of the detector description will be on the detector components used in this analysis.

4.1 HERA

The electron-proton storage ring HERA [W^{+}81] is situated in a tunnel between 10 to 20 m below the surface with approximately 6.4 km circumference. It is composed of four straight sections of 360 m connected with arcs with a radius of 797 m (figure 4.1).

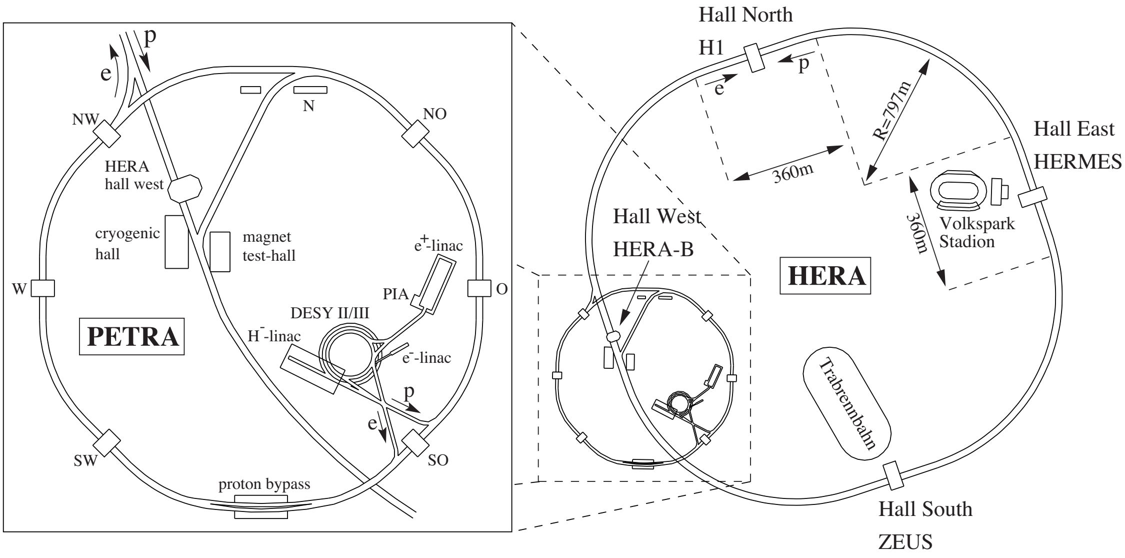


Figure 4.1: Overview of the HERA storage ring with the two collider experiments H1 and Zeus and the two fixed-target experiments Hermes and Hera-b (right). The preaccelerator chain is enlarged and shown on the left.

	HERA-1		HERA-2	
	electrons	protons	electrons	protons
beam energy E [GeV]	27.5	920	27.6	920
current I [mA]	45	100	58	140
particles per bunch N ($\times 10^{10}$)	3.5	7.3	4.0	10.3
number of beam bunches	189	180	189	180
beam sizes $\sigma_x \times \sigma_y$ [μm]	190×50		112×30	
synchrotron radiation at IP [kW]	6.9		25	
luminosity [cm^{-2}s^{-1}]	$1.4 \cdot 10^{31}$		$7 \cdot 10^{31}$	

Table 4.1: Achieved parameters of the HERA storage ring, before and after the luminosity upgrade. HERA I denotes the running period between 1998 and 2000, based on [Hol08, Sch98].

The two interaction points in hall North and South were equipped by the two multi-purpose detectors H1 and ZEUS, respectively. Both experiments started data taking in 1992 and were dismantled between 2007 and 2009. At the experimental halls East and West the fixed-target experiments HERMES and HERA-B were installed. The HERMES experiment used the polarised electron beam on a gas target and was in operation from 1995 to 2007, HERA-B made use of the proton beam halo on a wire target between 1998 and 2003.

Several stages of pre-acceleration were necessary before electrons and protons could be injected into HERA. Different linear accelerators and synchrotrons as well as the storage ring Positron-Elektron-Ring-Anlage (PETRA) were used to accelerate electrons and protons to their injection energies of 14 and 40 GeV. After injection the beams were stored in separate bunches with a distance of 28.8 m between successive bunches. This resulted in a bunch crossing time of 96 ns or a collision rate of 10.4 MHz. The final acceleration to the nominal electron and proton beam energies of $E_e = 27.6$ and $E_p = 920$ GeV[1] was performed in HERA, resulting in a center-of-mass energy of $\sqrt{s} \approx 320$ GeV. In order to keep the high momentum proton beam on a circular orbit, high magnetic fields ($B \approx 4.5$ T) were required. These were realised with superconducting magnets operated at a temperature of 4.2 K. Whereas the maximal proton energy was limited by the attainable magnetic field, the limiting factor for the electron beam energy were losses due to synchrotron radiation [Sch85].

An important characteristic of a particle collider is its luminosity, which is given for a two ring collider by [PDG10, p. 277]

$$\mathcal{L} = \frac{N_p N_e n_b f_0}{4\pi\sigma_x\sigma_y}. \tag{4.1}$$

in units of cm^{-2}s^{-1}. Here n_b is the number of colliding bunches, N_p and N_e denote

---

[1] Until 1998 the proton beam energy was 820 GeV.

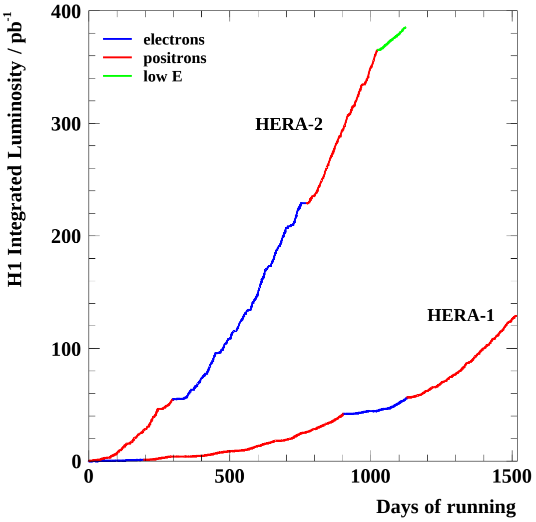


Figure 4.2: Integrated luminosity collected by the H1 experiment as a function of running time for HERA-1 and HERA-2.

the single bunch intensities, and f_0 is the revolution frequency. The transverse dimensions of the beams are characterised by σ_x and σ_y. The integrated luminosity

$$L = \int \mathcal{L}\,\mathrm{d}t \tag{4.2}$$

is directly proportional to the expected number of events N for a specific process

$$N = L\,\sigma, \tag{4.3}$$

where σ is the interaction cross section. Hence, an increase of the luminosity means an increase of the event yield for the particle physics experiments. In the years 2001 and 2002 an upgrade of HERA and the experiments was performed with the aim of increasing the luminosity by a factor of 5. It was neither possible to significantly increase the proton bunch population N_p due to limitations in the pre-accelerator chain, nor was it feasible to increase the electron current $I_e = N_e\,n_b\,f_0\,e$ due to the limited power of the accelerator cavities [Sei99]. However, from equation 4.1 it can be seen that a reduction of the beam spot sizes also leads to an increased luminosity. This was achieved by strong super conducting focusing magnets which were installed close to the interaction regions inside the experiments. Table 4.1 summarises the main parameters of the HERA storage ring before and after the upgrade.

The luminosity upgrade caused some problems with beam-induced backgrounds which made data taking only possible after two more years of machine studies and improvements. Between the years 2004 and 2007 the H1 experiment collected data corresponding to an integrated luminosity of roughly 400 pb^{-1} (figure 4.2). When HERA was decommissioned in 2007 most machine parameters had reached their design values [Hol08].

4.2 The H1 Detector

The H1 detector was a multi-purpose particle physics detector with an angular coverage of most of the solid angle around the Interaction Point (IP). It was located in the experimental hall North and had roughly the dimensions 12 m×10 m×15 m (length×width×height) with a weight of approximately 2800 tons. Figure 4.3 shows a schematic drawing of the detector.

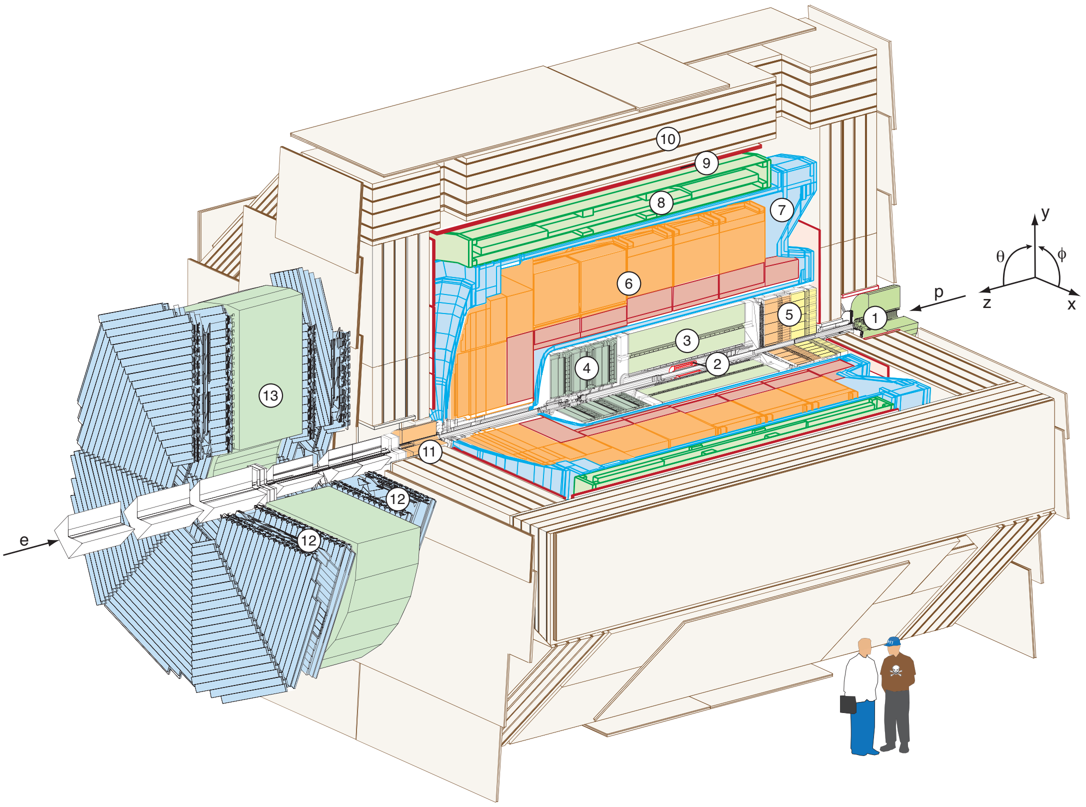
The adopted reference frame is a right-handed coordinate system with the $+z$ direction along the proton beam direction and the $+x$- direction pointing to the centre of the HERA ring (cf. figure 4.3). The nominal IP is chosen as origin of the coordinate system. The detector had a cylindrical symmetry around the beam pipe. The natural spherical coordinate system with (r, θ, ϕ) is chosen such that the electron direction is at $\theta = \pi$ (the $-z$-direction). This is referred to as the backward direction, whereas the proton direction ($\theta = 0$) is called forward direction. Because of the asymmetric beam energies the centre-of-mass is boosted along the forward direction which is why the H1 detector was considerably more massive and highly segmented there.

The H1 detector consisted of multiple subdetectors which were arranged around the IP in several layers. The innermost part, the beampipe, was surrounded by the central and forward track detectors, consisting of silicon track detectors, drift and proportional chambers. In the central and forward region the tracking detectors were surrounded by a cryostat vessel containing the LAr calorimeter. The backward region was equipped with the Backward Proportional Chamber (BPC) and Spaghetti Calorimeter (SpaCal). The LAr calorimeter and SpaCal comprised of electromagnetic and hadronic sections. A superconducting coil, providing an axial magnetic field of 1.15 T, surrounded the cryostat vessel. The iron return yoke of the magnet was instrumented with streamer tube detectors and was used to measure muons and small fractions of hadronic showers leaking out of the LAr calorimeter. Additional chambers inside and outside of the iron yoke were installed to improve muon identification. For an independent measurement of highly energetic muons in the forward direction the Forward Muon Detector (FMD) was installed. It comprised a toroidal magnet with 1.6 T, sandwiched between drift chambers. Between the FMD and the cryostat vessel the PLUG scintillation detector enclosed the forward beam pipe hole of the LAr calorimeter. Further along the proton direction systems to tag intact scattered protons were installed. In the electron direction the Electron Tagger (ET) was used to detect scattered electrons under very small angles. Also in this direction the Photon Detector (PD) was installed to measure photons from Bethe-Heitler processes to determine the luminosity.

In the following an overview of the main subdetectors used in this analysis is given. A detailed description of the H1 detector and its subsystems can be found elsewhere [H197a, H197b].

4.2.1 Track Detectors

In particle physics measured charged particle trajectories are called tracks. Although there exists a vast variety of techniques how to measure such tracks, all track detectors are based on the same physical principle. Charged particles traversing a medium ionise its atoms and thus create free charges. Track detectors aim to amplify, collect and measure these free charges in order to reconstruct the trajectories of particles. A well measured track provides multiple information about the particle's nature: the spatial measurement shows the path the particle followed, through



① Beam pipe and beam magnets
② Silicon tracking detector
③ Central tracking detector
④ Forward tracking detector
⑤ Spacal calorimeter (em and had)
⑥ Liquid Argon calorimeter (em and had)
⑦ Liquid Argon cryostat
⑧ Superconductiong coil
⑨ Muon chambers
⑩ Instrumented iron (streamer tube detectors)
⑪ Plug calorimeter
⑫ Forward muon detector
⑬ Muon toroid magnet

Figure 4.3: Schematic drawing of the $\mathrm{H1}$ detector with its main components. The $\mathrm{H1}$ coordinate system is shown on the right. The proton (electron) beam is entering from the right (left).

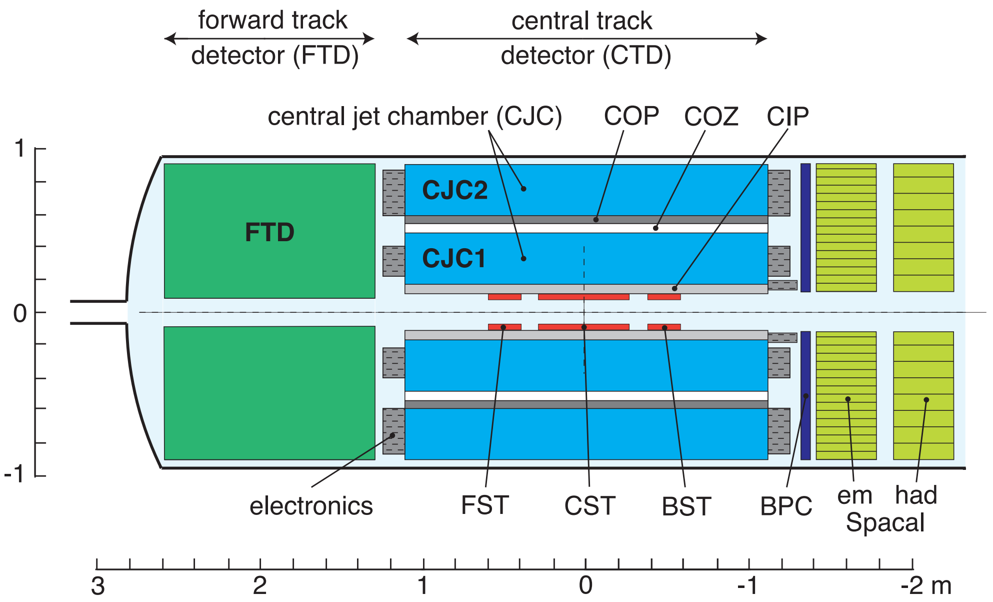


Figure 4.4: Cross-sectional view of the H1 tracking system.

the track's curvature in a magnetic field the sign of the particle's charge and the particle's momentum can be deducted, and through its specific energy loss in the medium the particle may be identified.
Since every tracking detector has its strengths and weaknesses in large particle physics experiments usually combinations of different techniques are used to acquire the needed precision.
In H1 three different tracking detector technologies were used: drift chambers, multi-wire proportional chambers and silicon strip detectors. Figure 4.4 provides an overview of the configuration of the various track detectors.

The Central Track Detector

The two main components of the Central Track Detector (CTD) were two large concentric drift chambers, the Central Jet Chambers (CJCs) 1 and 2. Two thin proportional chambers, the Central Inner Proportional Chamber (CIP) and Central Outer Proportional Chamber (COP) surrounded the CJC1 and were used mainly for triggering purposes. An additional thin drift chamber, the Central Outer z-Chamber (COZ) was used for a precise measurement of the z-coordinate and in close proximity to the beam pipe three silicon detectors (forward, central and backward silicon trackers, FST, CST and BST) provided an exact vertex determination. Figure 4.5 shows an $r\phi$ view of the CTD.
The Central Jet Chambers [H197b] covered an angular range of $15° < \theta < 165°$ in the polar angle with full coverage in the azimuth. The CJC1 consisted of 30 drift cells with 24 sense (anode) wires each, while the CJC2 comprised 60 drift cells with 32 sense wires each. The sense wires stretched parallel to the beam pipe and magnetic field, whereas the drift cells were inclined by about 30° with respect to the radial direction. This tilt was approximately equal to the Lorentz angle and was introduced such that in the presence of the 1.15 T magnetic field

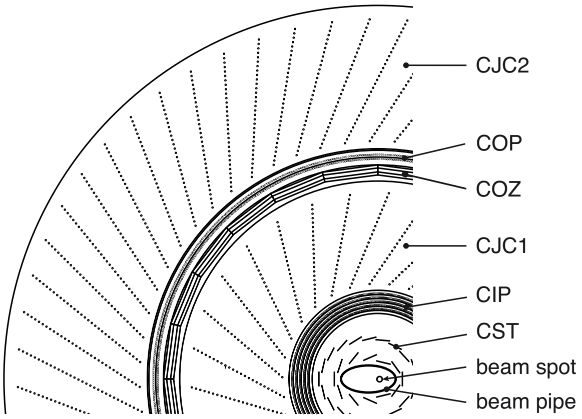


Figure 4.5: The central tracking system in the $r\phi$ plane.

the ionisation electrons from a stiff high momentum track drifted approximately perpendicular to the radial direction. This tilt had the additional advantages of resulting in an optimal track resolution and in resolving the usual left/right drift chamber ambiguity. Besides, every stiff track crossed the sense wire plane at least once in the CJC1 and CJC2. From the match at the crossing the passing time of a particle could be deducted to a precision of 0.5 ns which allowed an easy separation of tracks from different bunch crossings. From the drift time measured by the anode wires a single hit resolution in the $r\phi$ plane of $\sigma_{r\phi} \sim 170\,\mu\text{m}$ was obtained. By measuring the signal propagation time on both wire ends the z-coordinate could be measured with accuracies of about 1% of the wire length, resulting in a resolution $\sigma_z \sim 2.2\,\text{cm}$.

The Central Outer z-Chamber surrounded the CJC1 and had a polar angle coverage of $25° < \theta < 156°$. Its sense wires were strung perpendicular to the beam axis resulting in a drift direction along z-direction. Thus it measured the z-coordinates with a better accuracy than that obtained by charge division. The track elements it delivered had a typical z-resolution of $\sigma_z \sim 300\,\mu\text{m}$. These could be linked to the track segments obtained from the jet chambers to acquire an improved accuracy of the track's longitudinal momentum component.

The Central Inner Proportional Chamber was revised during the HERA upgrade in the years 2000–2002. The new design (CIP2k) incorporated a five-layer multiwire proportional chamber with cathode pad readout. The CIP was mounted on the inside of the CJC1. The five concentric layers provided space points which defined a track's direction. Through backward extrapolation a determination of the z-position of the event vertex could be made available after $2.3\,\mu\text{s}$ with a resolution of approximately 16 cm [Urb04]. This information was used for a z-vertex trigger which rejected high multiplicity background events with an efficiency close to 100% [BB+08].

The Central Outer Proportional Chamber was a two layer multiwire proportional chamber which was located between the CJC1 and CJC2. It had a short response time and could be used in combination with the CIP for a fast z-vertex determination. After the HERA upgrade it was of less importance due to the improved CIP performance.

The Silicon Track Detectors [Lis05] consisted of three individual components. The Central Silicon Tracker (CST) was the innermost tracking detector with a distance to the nominal IP of 5–10 cm. It was designed to provide a precise vertex determination as well as allowing the identification of displaced secondary vertices from heavy flavour particles with decay lengths of a few hundred micrometers. The CTD consisted of two layers of silicon sensors which were arranged around the elliptical beam pipe such that particle trajectories originating at the IP were perpendicular to the silicon ladders. The achieved point resolutions were 12 µm in $r\phi$ and 22 µm in the z-direction, which translated to an impact parameter resolution of 37 µm for high momentum tracks with an angular coverage of $30° < \theta < 150°$ [PB+00]. The Backward Silicon Tracker (BST) and Forward Silicon Tracker (FST) [HL00, Nož03] were installed to extend the angular coverage of the CST into the backward and forward direction. The BST covered the polar angle between 165° and 176°, whereas the FST measured in the angular region of 7°-19°.

The Backward Proportional Chamber was situated in front of the SpaCal and was used to improve the position measurement of particles going into the backward direction as well as to discriminate between charged and neutral particles. It consisted of six wire layers with three different orientations and measured the polar angle with a precision of $\sigma_\theta = 0.5\,\mathrm{mrad}$.

The Forward Track Detector

The Forward Track Detector (FTD) was situated in front of the CTD and extended the tracking coverage in the polar angle to $5° < \theta < 25°$. The FTD consisted of three supermodules, each containing five drift chambers (four in the case of the most forward supermodule) with different wire geometries. The planar drift chambers had wires strung parallel and the radial ones had wires strung in the radial direction. In both cases the wires were perpendicular to the beam pipe. The FTD achieved single hit resolutions of typically 210 µm in $r\phi$ with radial resolutions of approximately $\sigma_r \sim 3\,\mathrm{cm}$. The momentum resolution was between $\sigma_p/p^2 \sim 0.1\text{–}0.02\,\mathrm{GeV}^{-1}$ depending on the track length and on the track's polar angle. The track finding efficiency was approximately 70% in each supermodule.

4.2.2 Calorimeters

In particle physics a calorimeter[2] is a detector which measures the energy of incident particles. The principle relies on absorbing the particle's kinetic energy inside a given detector volume and converting it into a measurable signal. In H1 four different types of calorimeters were used. The Liquid Argon (LAr) calorimeter, the Spaghetti Calorimeter (SpaCal), the PLUG and the Tail Catcher (TC).

The LAr Calorimeter

The LAr calorimeter had an angular coverage of $4° < \theta < 154°$ in the polar angle and full coverage in the azimuth. It was designed to provide a clear identification

---

[2] Derived from the Latin word *calor*, meaning heat.

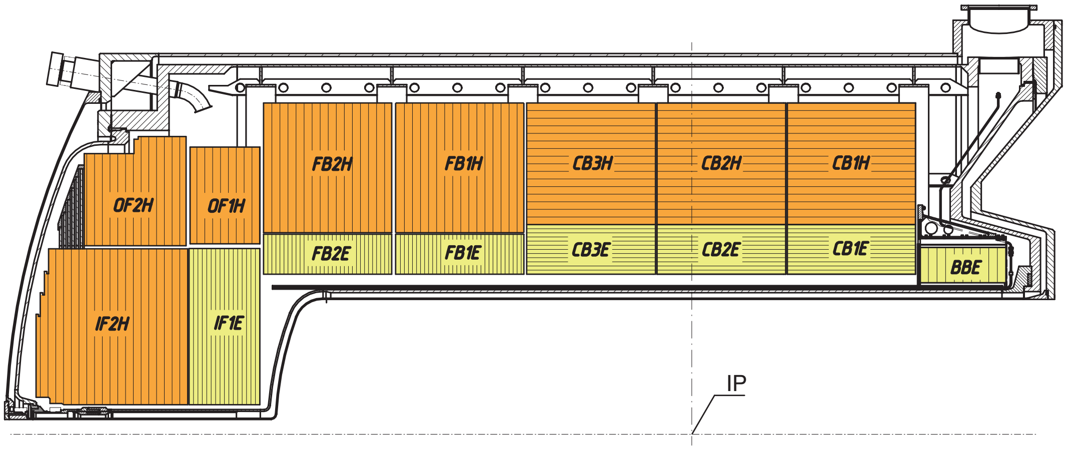


Figure 4.6: Side view of the LAr calorimeter inside the cryostats, showing the different calorimeter stacks. The electromagnetic and hadronic calorimeters are depicted in different colours. IP denotes the nominal interaction point.

and measurement of electrons together with a good hadronic measurement. In order to achieve this, the amount of dead material seen by particles originating at the IP was kept as small as possible. The calorimeter was situated inside a single cryostat and, to keep the amount of dead material as small as possible, the superconducting solenoid was surrounding the LAr vessel (cf. figure 4.3). The calorimeter consisted of two parts: an electromagnetic calorimeter with lead absorber plates and a hadronic calorimeter with stainless steel absorber plates. It was segmented into eight calorimeter wheels: the Inner Forward (IF), Outer Forward (OF), Forward Barrel (FB) 1 and 2, Central Barrel (CB) 1,2 and 3 and the Backward Barrel (BBE). An E or H behind the name denoted the electromagnetic or hadronic calorimeter, respectively (see figure 4.6). The LAr calorimeter is of crucial importance for this analysis, since it is used to trigger events, identify and measure the scattered electrons as well as a large fraction of the jet's and the hadronic final state's energy. A large part of this work has been dedicated to improve the calorimetric measurement by improving the separation between electromagnetic and hadronic showers and developing a new calibration method based on this. In the next chapter an introduction to calorimetry will be given followed by a detailed description of the LAr calorimeter of H1.

The "Spaghetti" Calorimeter

The so-called Spaghetti Calorimeter (SpaCal) was a lead/scintillating fibre calorimeter which was installed in the backward direction behind the CTD (cf. figure 4.4). It covered the polar angular range of $153° < \theta < 174°$ and was designed to precisely measure electrons scattered into the backward region. This required a good electron/pion separation, an excellent electromagnetic energy and spatial resolution as well as a fast response time [H197c].

The SpaCal was a sampling calorimeter consisting of an inner electromagnetic and an outer hadronic section (see figure 4.7). Both sections had lead as absorber and plastic scintillating fibres as active material. The name SpaCal derived from the long and thin fibres ('spaghettis') which were aligned parallel to the beam direction and

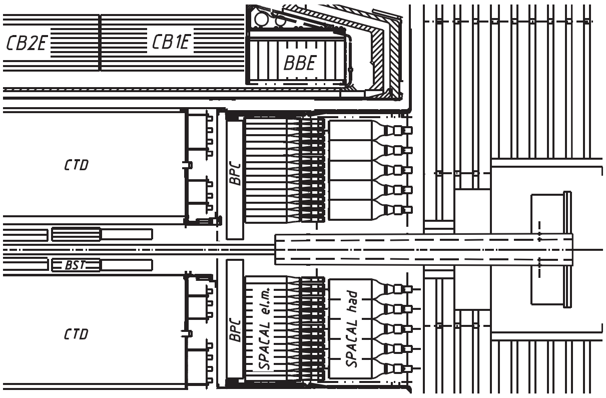


Figure 4.7: Technical drawing of the backward part of the H1 detector. Shown is the SpaCal with the BPC directly in front of the electromagnetic part. Surrounding the beam pipe is the focusing magnet from the HERA upgrade.

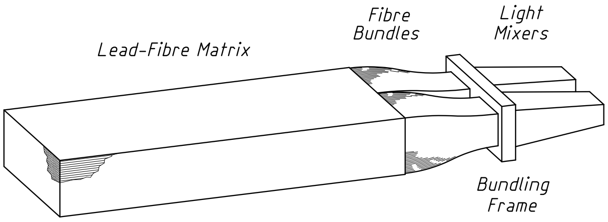


Figure 4.8: Drawing of a module from the electromagnetic SpaCal. Each module contained 2340 fibres.

embedded in a lead matrix, as shown in figure 4.8. Particles interacting with the lead absorber produced particle showers which caused scintillations in the plastic fibres. The scintillation light was conducted by the fibres and collected by photomultiplier tubes (PMTs) which converted the light into electric signals and amplified them. Since the SpaCal was situated in the strong magnetic field of the main solenoid with 1.15 T, the PMTs' gain was two orders of magnitude smaller than with no magnetic field applied [JL$^+$94]. Therefore, pulsed LEDs were used to monitor the stability of the PMTs over time.

The electromagnetic section of the SpaCal consisted of fibres with a diameter of 0.5 mm and had a lead/fibre ratio of 2.3:1. The active volume was 25 cm long, corresponding to 28 radiation length. The small fibre diameter allowed for a high sampling frequency and resulted in an excellent electromagnetic energy resolution of

$$\frac{\sigma_{\mathrm{em}}}{E} \approx \frac{7\%}{\sqrt{E(\mathrm{GeV})}} \oplus 1\%\,,$$

which was obtained from test beams at CERN and DESY [H196b]. The high granularity also ensured a position resolution of a few millimetres which corresponded to the required angular resolution of 1–2 mrad. The time resolution was found to be better than 0.4 ns which made the SpaCal sufficiently fast for triggering purposes.

The hadronic section of the SpaCal contained fibres with a diameter of 1 mm and the lead/fibre ratio was 3.4:1. The SpaCal's total active length was equal to two interaction lengths. The hadronic energy response has been obtained with test beams of pions with an energy of 1–7 GeV. An energy resolution of $\sigma_{\mathrm{had}}/E \sim 38\%$

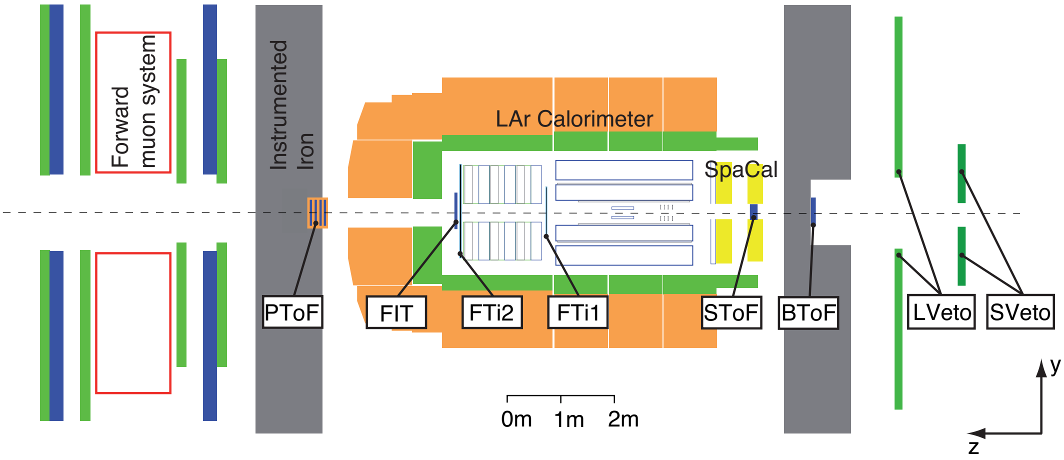


Figure 4.9: Schematic arrangement of the Time-of-Flight detectors.

within a depth of one interaction length and $\sigma_{\mathrm{had}}/E \sim 29\%$ for a total depth of two interaction length has been found [H196a].

The PLUG Calorimeter

The PLUG calorimeter was situated inside the iron return yoke around the forward part of the beam pipe. Between the IP and its position were 2-5 interaction length, therefore the requirements on its energy resolution were only moderate. However, it had to have a good time resolution of 1 ns to separate primary bunch crossings from satellite background. This was realised by the use of copper absorber plates in combination with sensitive layers of plastic scintillator tiles, which were coupled to wavelength shifting fibres for the light collection. PMTs were used to amplify the signals, and they were monitored analogous to the SpaCal PMTs.

The Tail Catcher

The iron return yoke of the main solenoid surrounded all major components of H1. It was interleaved with slits which housed limited streamer tube detectors. They served to detect hadronic activity leaking out of the LAr calorimeter as well as to detect penetrating minimum ionising particles.

4.2.3 Time-of-Flight Counters

The Time-of-Flight (ToF) system was an assembly of plastic scintillation detectors used for a fast identification of beam-induced background. The main sources of such background events were interactions with residual gas molecules in the beam pipe and stray protons interacting with the walls of the beam pipe (beam-gas and beam-wall events). The ToF system consisted of several components which were placed at various distances from the IP (cf. figure 4.9). In the forward direction at $z = 5\,\mathrm{m}$ the Plug detector was used (PToF). At 2.7 m, 2.5 m and 1.3 m the FIT, FTi2 and FTi1 (Forward Inner ToF) were installed to reject background at small forward angles. The SpaCal ToF (SToF) was positioned at the inside of the SpaCal.

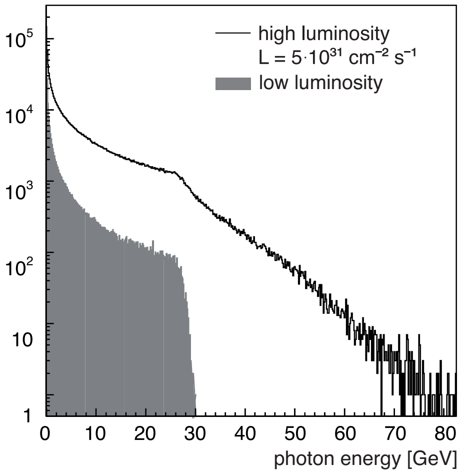


Figure 4.10: Bremsstrahlung spectrum for low and high luminosities. At the nominal HERA-2 luminosity the spectrum is distorted and shows a long tail towards large photon energies coming from pile-up. This effect is corrected for in the luminosity determination [AB⁺02].

In the backward direction, the backward ToF (BToF) at −4 m was followed by the large and small veto walls (LVeto and SVeto) at −6.5 m and −8.1 m. The various ToF components were calibrated such that particles originating from the nominal IP lead to measured time differences close to zero. Larger time differences were due to particles not coming from *ep* collisions and this information was used to reject such events. The time resolution of the ToF system of 1 ns lead to a definition of the primary vertex region in the z-direction of approximately ±30 cm around the nominal IP.

4.2.4 Luminosity System

Since the luminosity is the constant of proportionality between the event rate and the cross section for a given process (4.3), a precise knowledge of the luminosity is of crucial importance for any cross section measurement. At H1 the luminosity is determined from the rate of Bethe-Heitler events $ep \rightarrow ep\gamma$ by the measurement of the bremsstrahlung photons with the Photon Detector (PD). The detection of scattered electrons at small angles with the Electron Tagger (ET) has been used for the purpose of cross-checks and calibration. The total cross section of the Bethe-Heitler process is known with high accuracy. The main background comes from bremsstrahlung from the residual gas in the beam pipe, $eA \rightarrow eA\gamma$, but can be subtracted using electron pilot bunches. The resulting luminosity can then be calculated via [H197a]

$$L = \frac{R_{tot} - (I_{tot}/I_0)R_0}{\sigma_{vis}}. \tag{4.4}$$

Here R_{tot} is the total rate of bremsstrahlung events, R_0 is the rate in the electron pilot bunches and I_{tot} and I_0 are the corresponding electron beam currents. σ_{vis} is the visible part of the bremsstrahlung cross section, taking the detector acceptance and trigger efficiency into account.

The PD used at H1 was a Čherenkov sampling calorimeter with tungsten absorbers and quartz fibres as active medium [AB^+02]. In order to reduce the annual dose of synchrotron radiation on the apparatus, two radiation length of beryllium were placed in front of the detector. Due to the high luminosity after the HERA upgrade pile-up of bremsstrahlung photons in one bunch crossing was unavoidable and distorted the energy spectrum as shown in figure 4.10. In addition to measuring the photon energy spectrum, the impact point position was measured for the determination of the detector acceptance as well as for an online monitoring of the HERA electron beam tilt. The luminosity was measured online at 10 second intervals.
A precise determination of the luminosity is performed off-line by correcting for the pile-up, the detector acceptance, the effect of the beryllium shielding and the $eA \to eA\gamma$ bremsstrahlung background.

4.2.5 Trigger System

At HERA the time between two successive bunch crossings was 96 ns, which corresponded to a frequency of 10.4 MHz. However, not every collision lead to particles detectable in the H1 detector. The usual rate of recordable events reached the order of tens of kHz, which was still three orders of magnitude higher than the frequency at which data could be written out. The overwhelming amount of these events were background, which was mainly originating from electron- and proton-gas collisions, cosmic and beam halo muons and synchrotron radiation. The requirements to the trigger were to efficiently reject background events such that ep collisions could be recorded with a frequency of ~ 50 Hz, which was the limit of the output bandwidth. To achieve this, H1 used a pipe lined four level trigger system. Each of these levels (L1-L4) had a consecutively increasing precision at the cost of longer dead time, where the first level was effectively dead-time free and the level four had a latency of $\sim$100 ms.

The First Trigger Level (L1)

The first trigger level decision was based on the information from 256 trigger elements. These were passed from the various subdetectors into dead time free pipelines to the central trigger logic, which combined them to 128 raw subtriggers. Many of these were defined for the selection of specific physics processes. They were complemented by some subtriggers defined as monitor triggers, which controlled the performance of the various subdetectors and physics triggers. The central trigger logic decided to keep an event ("L1keep"), if at least one subtrigger was set. Some subtriggers had a prescale factor of n, which meant that only one out of n-times the subtrigger was set, corresponding to an actual subtrigger. In the case of no set subtrigger, the data in the pipelines were overwritten and no dead time was generated. If an event was kept, all pipelines were stopped and dead time started to accumulate. The output rate of the first level trigger was about 1 kHz.

The Second Trigger Level (L2)

The second trigger level consisted of two independent trigger systems, the topological trigger (L2TT) and the neural network trigger (L2NN). On this level the trigger

information from the first level was validated or disproved within 20 μs.
The L2TT trigger decision was based on topological event signatures derived from subdetector signals. It generated up to 16 trigger elements, which helped to reduce the L1 trigger rates with a good background reduction and large efficiencies.
The L2NN trigger was a set of 13 neural networks on parallel computers [KD$^+$01]. After the luminosity upgrade the high rates from the two main neutral and charged current subtriggers (S67 and S77) became critical, consuming too much bandwidth and preventing other triggers from running efficiently. The L2NN trigger offered an attractive solution and it was trained with samples from charged current, neutral current, multi-lepton events and events with isolated electrons in the final state. The working point was chosen at an efficiency of 99.89% with a 75% rejection of background events [Pla06, p57].

The Third Trigger Level (L3)

At level three the Fast Track Trigger (FTT) was commissioned in 2005. This was a trigger system which mainly used track based information to identify heavy quark decays. The latency time of the L3 was about 100 μs, which made it possible to use time-optimised routines to reconstruct event properties and decay particle resonances from L2 FTT tracks [BSC$^+$01, Sch06].

Level Four and Five Reprocessing

On the last trigger level (L45 for level four and five) a full event reconstruction and classification was performed. The input rate was approximately 50 Hz, and the full reconstruction started as soon as the complete event information from all subdetectors was available. At this stage the pipelines were started again, terminating the detector dead time. The reconstruction was handled by a dedicated PC farm, allowing an asynchronous processing which was necessary since complicated events (e.g. high particle multiplicities) required more computing time. After the reconstruction the events were classified in different physics categories. Events belonging to no category were rejected, except for a fraction of 1% which was kept for monitoring purposes. Additionally, some physics categories like high rate soft scale processes were downscaled, and only a fraction of these events was kept. Events that met all requirements were written to disk at a rate of approximately 10 Hz. The reconstructed events were stored together with the complete raw detector information as Production Output Tapes (POTs). A more compact format with reconstructed quantities only, the Data Summary Tapes (DSTs), are the starting point for physics analyses.

4.3 Detector Simulation

For the simulation of the H1 detector the program H1SIM was developed by the H1 collaboration. It is based on GEANT3 [BB$^+$87] for the implementation of the detector geometry, interaction of particles in matter and the tracking and hit management. The detector simulation has been verified with various test beam measurements with prototype detectors. After the assembly of the full detector it was constantly monitored and refined with *ep* data. The simulation of electromagnetic and hadronic showers is described in the next chapter. After the detector simulation MC events are on the same basis as genuine *ep* data and are subject to the same reconstruction program H1Rec.

Chapter 5

Calorimetry

Particles and their properties can only be measured through their interaction with matter. In a multi-purpose detector usually track detectors reside close to the interaction point, measuring the momenta of charged particles through ionisation. However, neutral particles cannot be measured in this way, and the precision of the measurement of the track momentum is proportional to the track's curvature, $\sigma_{p_\mathrm{T}}/p_\mathrm{T} \sim p_\mathrm{T}$, degrading linearly with momentum. Additionally, the track finding efficiency is never perfect and in dense environments with high particle multiplicities it is not possible to assure a reconstruction of all particle trajectories.

In contrast to track detectors where the particles' momenta are measured essentially non-destructively, calorimeters completely absorb incident particles. Through stopping processes they gradually lose energy which can then be converted into measurable signals. Ideally these are proportional to the incident particle's energy. Taking into account stochastic fluctuations, charged and neutral particles can be measured with a precision $\sigma_E/E \sim 1/\sqrt{E}$, improving with the square root of the incident energy. Hence calorimeters become important and more precise than track detectors at high energies. They are of significant importance in dense environments like jet production and crucial for a full event reconstruction.

There are numerous processes by which particles traversing dense matter lose energy and eventually get absorbed. Most of them originate from the electromagnetic and the strong force, weak interactions play only a minor role. After the first interaction the produced secondary particles can interact with the material's constituents again, producing cascades of particles. These cascades are called showers and the different developments of electromagnetic and hadronic showers lead to different calorimeter responses.

5.1 Electromagnetic Showers

Electromagnetic showers are induced by electrons, positrons and photons in dense materials with high atomic number. They are characterised by their compactness and high energy density. High energetic electrons and positrons lose their energy dominantly by ionisation and bremsstrahlung and may also excite the medium's atoms or molecules. Subordinate processes are the production of energetic knock-on electrons (δ-rays), the emission of Čerenkov light and induced nuclear reactions. The

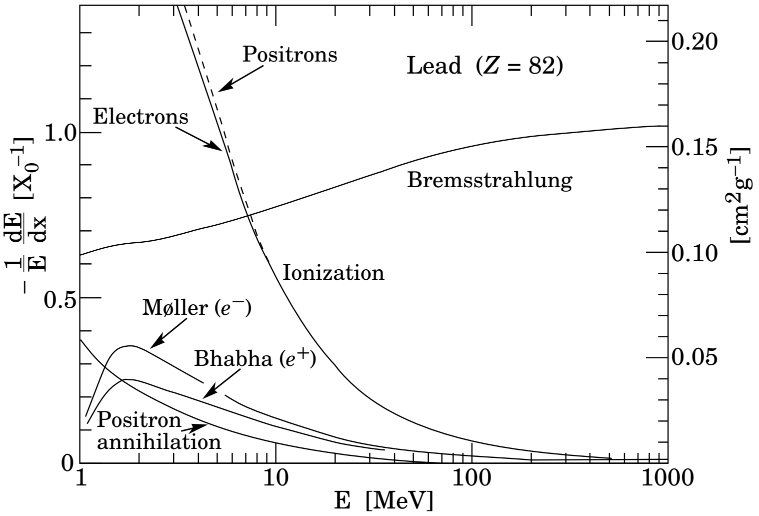


Figure 5.1: Fractional energy loss per radiation length of electrons and positrons in lead [PDG10].

energy loss of photons is dominated by the photoelectric effect, Rayleigh and Compton scattering and electron-positron pair production, depending on the photon's energy.

5.1.1 Interactions of Charged Particles

Charged particles traversing matter interact with the Coulomb field generated by the atoms. In figure 5.1 the relative energy losses for electrons and positrons in lead for different processes are shown. At low energies ionisation is the most prominent effect, but also Møller and Bhabha scattering contribute. Positrons with small energies may annihilate with the medium's electrons resulting in relatively low energetic photons with an isotropic spatial distribution.

In order to describe the dominant energy loss by ionisation and excitations the usual Bethe-Bloch equation has to be modified for electrons and positrons. Screening effects and a different energy-transfer probability due to the equal mass of the incident and the target particle have to be taken into account. Additionally, bremsstrahlung processes give a significant contribution to the energy loss below the critical energy E_c (defined below) and have to be considered. A precise treatment for the energy loss of electrons due to ionisation and excitation gives [GS08]

$$\begin{aligned} -\left.\frac{dE}{dx}\right|_{\mathrm{ion}} = 4\pi N_{\mathrm{A}} r_{\mathrm{e}} m_{\mathrm{e}} c^2 \cdot \frac{Z}{A} \cdot \frac{1}{\beta^2} \Bigg[\ln \frac{\gamma m_{\mathrm{e}} \beta \sqrt{\gamma - 1}}{\sqrt{2} I} + \frac{1}{2}(1-\beta^2) \\ - \frac{2\gamma - 1}{2\gamma^2} \ln 2 + \frac{1}{16} \left(\frac{\gamma - 1}{\gamma} \right)^2 \Bigg] \end{aligned} \tag{5.1}$$

which is given in units of MeV g^{-1} cm^2 with the constant of proportionality

$$4\pi N_{\mathrm{A}} r_{\mathrm{e}} m_{\mathrm{e}} c^2 = 0.3071 \text{ MeV g}^{-1} \text{ cm}^2 . \tag{5.2}$$

N_{A} is Avogadro's number, Z and A are the atomic number and weight of the absorber, the velocity $\beta = v/c$, the Lorentz factor $\gamma = (1-\beta^2)^{-1/2}$ and the electron

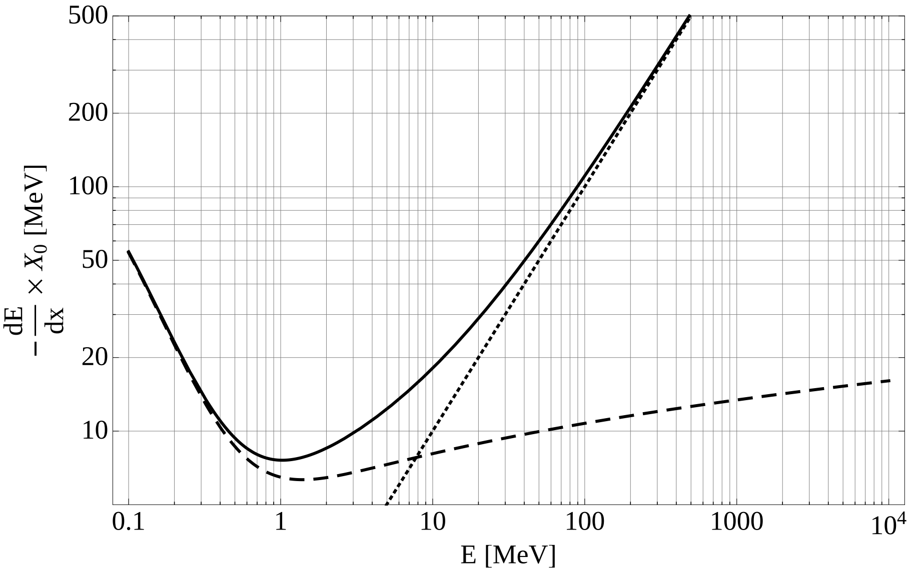


Figure 5.2: Energy loss for electrons in lead as function of the electron energy. Shown are the energy loss due to ionisation and excitations, eq. (5.1) (dashed) and the energy loss due to bremsstrahlung, eq. (5.3) (dotted). The total energy loss is shown as solid line.

mass and the classical eletron radius are given by m_e and r_e, respectively. The mean excitation energy I can be approximated by $I = 16\,Z^{0.9}\,\mathrm{eV}$.
At energies above tens of MeV the energy loss of electrons and positrons is dominated by bremsstrahlung. For this process the absolute energy loss per traversed length is proportional to the particle's energy and can be approximated by

$$-\left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_\mathrm{brems} = \frac{E}{X_0}. \tag{5.3}$$

This equation is true at very high energies, where energy losses due to ionisation can be neglected. It is only an approximation at the regime of energies below 100 MeV. Equation (5.3) defines the radiation length X_0 which includes the material's properties and gives a characteristic length of electromagnetic showers. One radiation length is the distance after which a highly energetic particle lost $1 - 1/e$ of its energy due the emission of bremsstrahlung photons. Taking the Coulomb field generated by the target's electrons and screening effects into account, the Particle Data Group [PDG10] gives the following approximation for the radiation length,

$$X_0 = \frac{716.4 \cdot A[\mathrm{g/mol}]}{Z(Z+1)\ln(287/\sqrt{Z})}\ \mathrm{g/cm^2}. \tag{5.4}$$

The energy at which the energy losses due to ionisation and bremsstrahlung are equal is called the critical energy E_c, which is defined through

$$-\left.\frac{\mathrm{d}E}{\mathrm{d}x}(E_c)\right|_\mathrm{ion} = -\left.\frac{\mathrm{d}E}{\mathrm{d}x}(E_c)\right|_\mathrm{brems}. \tag{5.5}$$

For solids and liquids the critical energy can be approximated by

$$E_c = \frac{610\,\mathrm{MeV}}{Z + 1.24} \tag{5.6}$$

with Z being the atomic number. This approximation is shown to hold for all chemical elements within $\sim 2.2\%$ [PDG10].

The radiation length as well as the critical energy scale as the square of the mass of the incident particle. Already for the next lightest particle, the muon ($m_\mu = 106\,\mathrm{MeV/c^2}$), the critical energy is $(m_\mu/m_e)^2 \approx 4 \cdot 10^4$ times larger than for electrons. Hence, bremsstrahlung plays only a minor role for the energy losses of heavy particles. Since bremsstrahlung photons are the source of electromagnetic showers only electrons and positrons give rise to them.

The energy loss of electrons in lead is shown in figure 5.2, where also the contributions from ionisation and excitations, equation (5.1), and the emission of bremsstrahlung photons, equation (5.3), are shown. In the low energy domain ($\beta\gamma \lesssim 2$) the energy loss decreases like $1/\beta^2$ and reaches a minimum around 1 MeV ($\beta\gamma \sim 2$). For high energies the energy loss due to ionisation increases logarithmically (relativistic rise), whereas the bremsstrahlung process shows a linear dependence on the electron energy and dominates the energy loss at energies above 100 MeV ($\beta\gamma \gtrsim 200$).

5.1.2 Interactions of Photons

Photons traversing a dense medium can interact either with the shell electrons of the atoms or at higher energies with the Coulomb field of the nuclei. In figure 5.3 the total cross section of incident photons on lead as function of the photon energy is shown.

At low energies the **photoelectric effect** ($\sigma_{p.e.}$) has the largest cross section. At this process the photon is absorbed by an atom leaving it in an excited state. The atom falls back to its ground state by the emission of Auger electrons or X-rays,

$$\gamma + \mathrm{atom} \rightarrow \mathrm{atom}^* \rightarrow \mathrm{atom}^+ + e^-. \tag{5.7}$$

The photoelectric cross section falls of with E^{-3} and rapidly loses its importance as the energy increases. The steps in $\sigma_{p.e.}$ originate from photon energies just above the binding energy of the shell electrons, in figure 5.3 the M, L and K edges are nicely visible. In lead the photoelectric effect is dominating below energies of 600 keV, whereas for iron inelastic scattering already starts to dominate above 100 keV.

Also at low energies **Rayleigh scattering** ($\sigma_{\mathrm{Rayleigh}}$) is of importance. This is a coherent process where the photon scatters off the atomic electrons,

$$\gamma + e^- \rightarrow \gamma + e^-. \tag{5.8}$$

In this process the photon does not lose energy so Rayleigh scattering only influences the spatial distribution of electromagnetic showers but does not contribute to the signal generation in calorimeters.

The incoherent equivalent to Rayleigh scattering is **Compton scattering** ($\sigma_{\mathrm{Compton}}$). In this process an incident photon scatters off an electron and transfers part of its energy to the electron. In electromagnetic showers typically about half the energy is

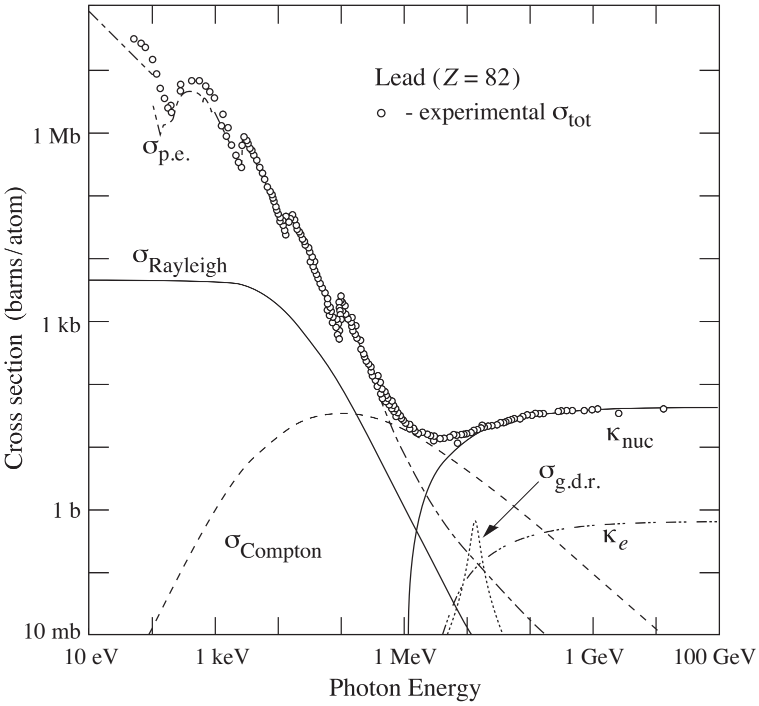


Figure 5.3: Total cross section of photons in lead as a function of the photon energy [PDG10]. The contributions from different processes are described in the text.

deposited by this process. Incident photons above 1 MeV undergo several Compton scatterings before their energy is small enough to be absorbed by the photoelectric effect. In the first scattering the struck electron has a significantly higher probability of following the direction of the initial photon. After several scatterings this preference disappears and most of the photoelectrons produced by Compton scattering are isotropically distributed.

At photon energies larger than twice the rest mass of an electron, $E_\gamma > 2m_e$, **pair production** (κ_{nuc} and κ_e) quickly becomes the dominating process. A photon with enough energy may create in the Coulomb field of a charged particle an electron-positron pair,

$$\gamma + \text{atom} \rightarrow e^- + e^+ + \text{atom}. \tag{5.9}$$

Pair production in the field of the nucleus (κ_{nuc}) has typically a much higher cross section than in the field of the atomic electrons (κ_e). The cross section for pair production rises strongly for energies above 1 MeV and reaches a plateau at energies $E_\gamma > 1\,\text{GeV}$. The asymptotic cross section for photon interactions is related to X_0 via

$$\sigma_{\text{pair}}(E \rightarrow \infty) = \frac{7}{9}\frac{A}{N_{\text{A}} X_0} \tag{5.10}$$

with N_{A} being Avogadro's number and X_0 in units of g cm^{-2}. This implies that the mean free path of very high-energy photons is $\frac{7}{9}X_0$ [Tsa74].

At energies between 5-20 MeV **photonuclear interactions** may occur, with a maximum of the cross section at the giant dipole resonance ($\sigma_{\text{g.d.r.}}$), where the photon energy is about equal to the nucleon binding energy. These reactions include the vaporisation of nucleons and photon induced nuclear fission. However, the total cross

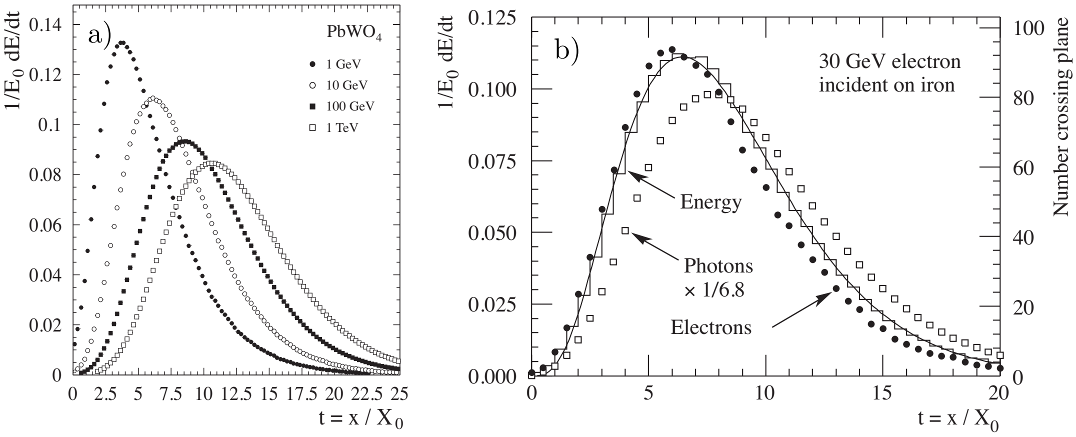


Figure 5.4: a) Simulated longitudinal shower profiles in PbWO$_4$ as a function of material thickness for incident electrons of energies 1 GeV, 10 GeV, 100 GeV and 1 TeV [FG03]. b) Simulation of a shower induced by 30 GeV incident on iron, from [PDG10]. The histogram shows the fractional energy deposition per radiation length and the curve is the parameterisation from equation 5.11. The circles (squares) show the number of electrons (photons) with energy greater than 1.5 MeV crossing planes at intervals of $X_0/2$ (scale on the right).

section is small compared to pair production, never exceeding 0.7 b in lead [Kos02] and amounting to less than 1% of the total cross section in the relevant energy region.

5.1.3 Electromagnetic Shower Development

The interplay of the processes described above is the basis of the development of electromagnetic cascades. High energy electrons or positrons entering a dense medium will radiate bremsstrahlung photons. The majority of these photons is very soft and will be absorbed through Compton scattering and the photoelectric effect. Photons with sufficient energy will create electron-positron pairs which again will radiate bremsstrahlung photons. This multiplication of shower particles reaches its maximum once the average energy of the shower particles is comparable to the critical energy E_c, i.e. the ionisation cross section is similar to the one for bremsstrahlung. The longitudinal distribution of the energy deposition in electromagnetic cascades can be parameterised by [LS75]

$$\frac{\mathrm{d}E}{\mathrm{d}t} = E_0\, b\, \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}\,, \tag{5.11}$$

with the dimensionless variable

$$t = \frac{x}{X_0}\,, \tag{5.12}$$

which is the distance normalised to radiation length and a convenient measure for longitudinal shower developments. In the above formula E_0 is the incident energy and a and b are parameters which can be obtained from measurements. For heavy absorbers (from iron to lead) $b \approx 0.5$ [GS08]. With this parameterisation, the maximum of the shower development is reached at

$$t_{\mathrm{max}} = \frac{a-1}{b} = \ln\left(\frac{E_0}{E_c}\right) + C_{\gamma\mathrm{e}} \tag{5.13}$$

where $C_{\gamma\mathrm{e}} = 0.5$ for photon induced showers and $C_{\gamma\mathrm{e}} = -0.5$ for showers caused by electrons or positrons. Equation (5.13) can then be used to obtain the energy dependent parameter a. The shower maximum has a logarithmic dependence on the incident energy (compare figure 5.4a) and occurs at greater depth for photons. Figure 5.4b shows the longitudinal development of an electromagnetic shower in iron induced by 30 GeV electrons. At small depth the shower is dominated by high energy electrons losing their energy by radiating photons which can convert into e^+e^- pairs (shower particle mulitiplication). Beyond the shower maximum the number of electrons decreases quicker than the number of photons which still increases for about two radiation length. The tail of the shower at large depth is dominated by photons. The number of electrons in an electromagnetic shower is typically two orders of magnitude larger than the number of positrons. However, positrons originate from pair production and are usually more energetic than electrons, carrying about one quarter of the shower energy [Wig00].
The lateral development of electromagnetic showers can be described in an approximately material independent way with the Molière radius,

$$R_{\mathrm{M}} = \frac{E_s}{E_c} X_0 \,, \tag{5.14}$$

where the scale energy E_s is defined by $m_e c^2 \sqrt{4\pi/\alpha}$ or 21.2 MeV. On average 90% of an electromagnetic shower are contained in a cylinder of radius R_{M} around the shower axis, almost independently of the incident energy. This defines the dimensionless scale variable for the development of the radial shower

$$u = \frac{r}{R_{\mathrm{M}}}, \tag{5.15}$$

where r denotes the radial distance to the shower axis. The lateral width of an electromagnetic shower increases with increasing shower depth. The lateral shower distriubution can approximately be described by [GR⁺90]

$$f(u, E, t) = \frac{2uR(E,t)^2}{(u^2 + R(E,t)^2)^2} \,, \tag{5.16}$$

where the parameter $R(E, t)$ describes the logarithmic behaviour of the lateral width as function of energy as well as the increasing radial extention of the shower with increasing shower depth. The mean value of the approximately log-normal distributed parameter $R(E, t)$ is given by

$$\langle R(E,t) \rangle = R_1 + (R_2 - R_3 \ln E)\, t. \tag{5.17}$$

The parameters R_1, R_2 and R_3 can be obtained from experimental data or by comparison to detailed simulations of electromagentic cascades, see for example [Rud92]. In figure 5.5 (left) the radial development of an electromagnetic shower in lead is shown at various shower depths, following equation (5.16). The radial shower profile shows a central core whose spread is dominated by multiple scattering of relatively high energy shower particles. This central core disappears beyond the shower maximum, where the radial width is dominated by isotropic processes such as Compton scattering, the photoelectric effect and photons from e^+e^- annihilation

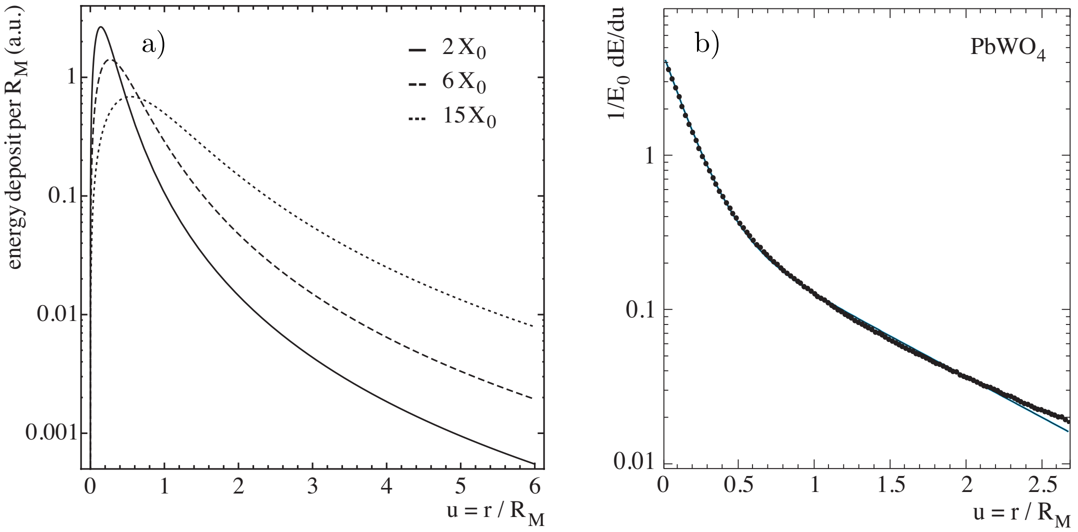


Figure 5.5: *Left:* Radial distribution of the energy deposited by 1 GeV electrons in lead at various depths. Parameters of equation (5.17) were obtained from [Rud92]. *Right:* Simulated radial shower profile as a function of the distance from the shower axis for incident electrons with an energy of 1 GeV on PbWO$_4$, from [FG03]. The line is a fit by formula (5.18).

processes, called the shower halo. The parameterisation (5.16) agrees well with experimental results, see for example the measuremens from Yuda [Yud69] and the description of electromagnetic showers obtained by Peters [Pet92]. An extension of the parameterisation (5.16) is used for the simulation of electromagnetic showers in H1 [GP00].

The radial shower distribution integrated over the full shower length can be approximated by two Gaussians [AI$^+$77], one expressing the shower core and the second one the halo,

$$\frac{\mathrm{d}E}{\mathrm{d}u} = A_1 e^{-u/b_1} + A_2 e^{-u/b_2}. \tag{5.18}$$

An example for an integrated radial shower distribution obtained from a simulation is shown in figure 5.5, together with a fit to equation 5.18. Equally good fits can be achieved with experimental data, e.g. [AC$^+$05a].

5.2 Hadronic Showers

Hadronic showers are considerably more complicated than electromagnetic ones, owing to the large number of possible processes involving the strong interaction. Unlike in the electromagnetic case, ab-initio calculations are currently not possible, since they would require many-body calculations in the context of QCD in a region where the strong coupling is large and perturbative calculations break down. However, models exist which can describe experimental data even in the soft regime, making it possible to build Monte Carlo simulations starting with hadron-nucleon scattering and developing hadronic cascades based on theoretical considerations [BB$^+$06]. Other models are driven as much as possible from experimental data using parameterisations to extrapolate over the full range of hadronic shower energies [AA$^+$03].

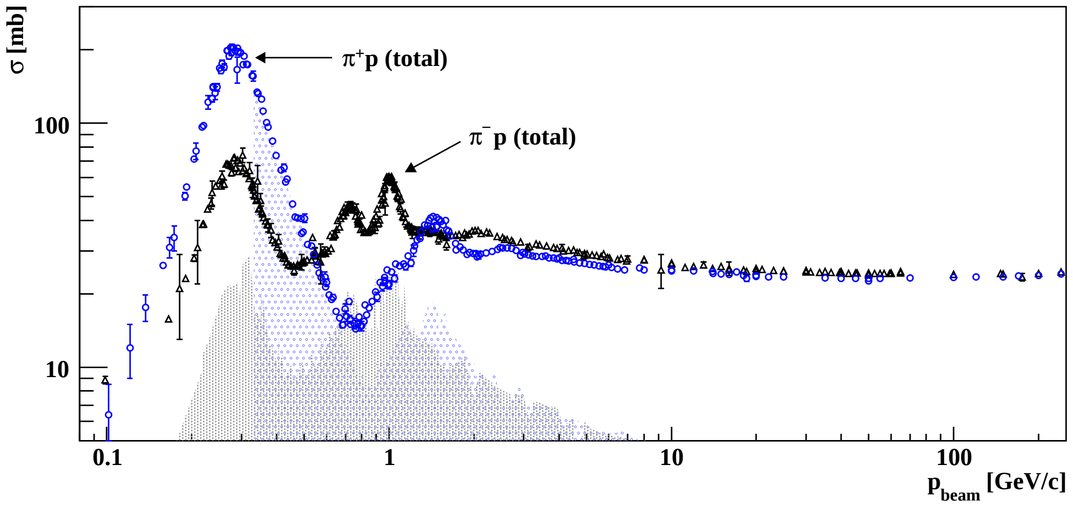


Figure 5.6: π^+p and π^-p cross sections, the elastic contributions are shown as shaded areas. In the region $p_{\text{beam}} < 0.32\,\text{GeV/c}$ no data on elastic π^+p scattering are available (data from [PDG10]).

Even though the particle multiplication in hadronic showers is conceptually similar to the electromagnetic case, there are fundamental differences, which have far-reaching consequences for calorimetry. Firstly, hadronic showers develop on a different length scale than electromagnetic ones due to the differences between the nuclear and electromagnetic cross sections. This makes hadronic showers longitudinally and laterally less compact than their electromagnetic counterparts. Secondly, while in electromagnetic showers the shower products are always electrons, positrons and photons which can be converted into measureable signals, in hadronic showers some fraction of the deposited energy is fundamentally non-detectable. This invisible energy is governed by large fluctuations with the implication of more stringent limitations on the resolutions of hadronic calorimeters.

In the following a brief overview of the processes involved in the development of hadronic showers is given, with an emphasis on the implications for calorimetry.

5.2.1 Hadron-Nucleon Scattering

For incident hadron energies above $\sim$10 GeV the constituents of the target nucleus can be considered to be quasi free. Therfore, the first step of particle production by the interaction of an incident hadron with a nucleus can be understood by looking at hadron-nucleon scattering. Figure 5.6 shows the total and elastic cross sections of pion-proton scattering. Elastic scattering contributes predominantly at small energies, where it is dominated by several direct resonances, the most prominent one being the $\Delta(1232)$. At high energies inelastic reactions dominate the cross section which is approximately energy-independent above 10 GeV. The bulk of the hadrons produced in these collisions are pions, where because of isospin conservation, the average number of neutral pions is approximately

$$\langle n_{\pi^0} \rangle = \frac{\langle n_{\pi^+} \rangle + \langle n_{\pi^-} \rangle}{2}. \qquad (5.19)$$

These π^0s decay electromagnetically, $\pi^0 \to \gamma\gamma$ with a branching ratio of 98.8% and $\pi^0 \to e^+e^-\gamma$ with a 1.2% probability. With lower multiplicities also kaons, etas, protons, neutrons and other hadrons are produced in these high energy collissions. The average charged particle multiplicity $\langle n_{\mathrm{ch}} \rangle$ per interaction varies logarithmically with energy [KM+89] and can be described by

$$\langle n_{\mathrm{ch}} \rangle = a + b \ln s + c(\ln s)^2, \tag{5.20}$$

where $\sqrt{s}$ is the available centre-of-mass energy, and values for a, b and c can be found in the literature. The average transverse momentum of the secondary particles is approximately energy-independent,

$$\langle p_{\mathrm{T}} \rangle \approx 0.35\,\mathrm{GeV/c}. \tag{5.21}$$

About half of the incident particle's energy is transferred to the produced particles. The remaining energy is carried away by a particle with the projectile's baryon number (leading particle effect). Hence, if the incident particle is a baryon, another baryon will emerge from the nuclear interaction carrying a large fraction of the energy. This suppresses the production of pions for baryon-nucleon scattering with respect to meson-nucleon scattering.

5.2.2 Hadron-Nucleus Scattering

High energy hadron-nucleus scattering can be considered as a sequence of mulitple hadron-nucleon scatterings. It can be described in two steps. In the first step the incident hadron starts an intra-nuclear cascade by interacting with one of the quasi-free nucleons. If the transferred energy is large enough, the struck and secondary particles may interact with the remaining nucleons. The ones with sufficient energy will escape from the nucleus, whereas others will distribute their energy among the remaining nucleons. The next step is the de-excitation of the intermediate nucleus. This will either happen through evaporation of nucleons until the excitation energy is less than the binding energy and/or through the emission of γ-rays. Heavy nuclei like Uranium may even fission.

In the intra-nuclear cascade the probability of producing a free proton is approximately equal to the ratio of Z/A of the target nucleus, thus releasing nearly as many protons as neutrons. The average charged particle multiplicity exhibits a weak dependence on the atomic number A. Also the rapidity distribution from the produced secondaries is slightly different than in hadron-nucleon scattering, see for example the review by Leroy and Rancoita [LR00].

In the evaporation stage more neutrons than protons are emitted on average from heavy nuclei because of the larger Coulomb barrier (∼12 MeV in lead compared to ∼5 MeV in iron) which prevents protons from being emitted from an excited heavy nucleus. The produced evaporation protons and neutrons have rather small kinetic energies of the order of a few MeV. The evaporation protons lose their energy through spallation and ionisation of the medium. The evaporation neutrons undergo reactions of the type $(\mathrm{n}, x\mathrm{n})$ until their energy is too low for these processes. In each reaction the evaporation neutrons release some amount of binding energy of the target nuclei. A few nanoseconds after the inital interaction the only free neutrons in the medium are evaporation neutrons. The multiplicities of these can be large, with about 20 neutrons per GeV of incident energy in lead and up to 60 neutrons per GeV in ^{238}U [Wig00, p76]. The kinetic energy spectrum of the evaporation neutrons can be described by a Boltzmann-Maxwell distribution,

$$\frac{\mathrm{d}N}{\mathrm{d}E} \sim \sqrt{E}\, e^{-E/T}, \tag{5.22}$$

where the energy E is given in MeV. A typical temperature of $T \sim 2\,\mathrm{MeV}$ yields an average neutron energy of about $3\,\mathrm{MeV}$ [Wig87].

The nucleons from evaporation are emitted approximately isotropically, while the particles from the initial cascade have a dominating momentum component in the direction of the incident hadron. This leaves the residual target nucleus with a net recoil with a kinetic energy of the order of m/M, where m and M are the total mass of the emitted nucleons and the residual nucleus, respectively.

5.2.3 Hadronic Shower Development

The characteristic length which governs the development of a hadronic shower is the interaction length λ, which is defined as the mean distance after which a hadron loses $(1 - 1/e)$ of its energy due to inelastic collisions. It is given by

$$\lambda = \frac{A}{N_\mathrm{A}\,\rho\,\sigma_\mathrm{inel}} \approx 35\,A^{1/3}\,\mathrm{g/cm^2} \tag{5.23}$$

with A and ρ being the atomic mass and density of the material, N_A is Avogadro's number, and σ_inel denotes the inelastic cross section. Conventionally, the inelastic proton cross section at high energy is taken. In this regime it is approximately energy independent, similar to the inelastic pion cross section (cf. figure 5.6). In the strict sense σ_inel depends on the energy and on the type of incident particle, where both dependencies are usually neglected.

The interaction length λ can be interpreted as the mean free path of a high energy hadron incident on a block of matter. After travelling a distance λ it will on average interact strongly with one of the nuclei of the medium. In the inelastic collisions mostly pions are produced, and about one third of them are π^0s. They decay electromagnetically and thus introduce electromagnetic subshowers. The remaining hadrons will continue to produce electromagnetic subshowers through inelastic collisions until the average hadron energy is below the pion production threshold. If a hadronic shower comprises n generations, the average energy fraction transferred to electromagnetic subshowers f_em follows a power law with n in the exponent. Assuming that n increases with the energy of the incident hadron, f_em will increase with increasing energy as well. This dependence is non-trivial, since in every inelastic collision the particle multiplicity of secondaries varies logarithmically with the energy of the incident particle (5.20). Gabriel *et al.* [GG+94] studied this dependence and gave an empirical formula

$$f_\mathrm{em}(E) = 1 - \left(\frac{E}{E_0}\right)^{(k-1)}, \tag{5.24}$$

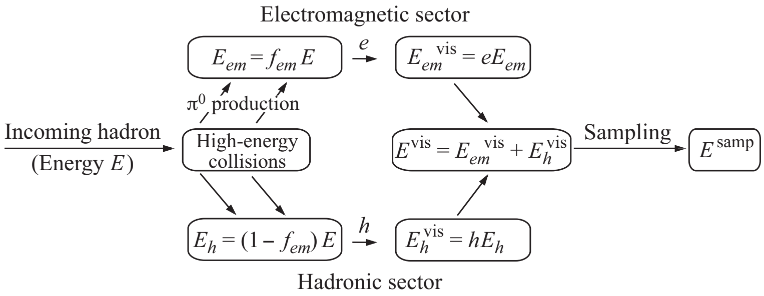


Figure 5.7: Schematic view of the energy flow in a hadronic cascade (from [Gro07]). A detailed explanation is given in the text.

where E is the energy of the incident hadron. Equation (5.24) is valid for energies above approximately 5 GeV and the free parameters E_0 and k are connected to the average π^0 fraction and the average multiplicity. Approximate values are $k \approx 0.83$, and for incident pions $E_0 \approx 1$ GeV, whereas for protons $E_0 \approx 2.6$ GeV [Gro07].

A schematic view of the energy flow in a hadronic cascade is shown in figure 5.7. Let E denote the energy of the incident hadron. After a cascade of nuclear interactions the energy of the electromagnetic subshowers will be $f_{\mathrm{em}}E$ and the energy of the hadronic sector $(1 - f_{\mathrm{em}})E$[1]. The conversion efficiency of energy into a measureable signal is denoted by e for the electromagnetic sector and by h for the hadronic sector. In contrast to the energy from the electromagnetic sector, which generates almost completely a measureable signal, a substantial fraction of the energy from the hadronic sector remains undetectable. This undetectable energy is termed invisible energy and is related to energy dissipating into the recoil of the target nuclei and nuclear binding energy, which is transformed to the kinetic energy of the evaporation neutrons. The measureable signal from the hadronic sector originates mostly from ionisation processes, where charged secondaries ionise the atoms of the medium. The invisible energy fraction of a hadronic cascade amounts usually to about 25-45%, where the absolute number and the sharing between the nuclear binding energy and the kinetic energy of the evaporation neutrons depend on the absorber. The difference between the detection efficiencies, the ratio e/h, is the degree of compensation of a calorimeter. Only if the energy of the evaporation neutrons can be recovered, meaning that the invisible energy fraction is negligible to a good approximation, the ratio e/h will be unity, which is termed compensation. From the sum of all measureable signals E^{vis}, a fraction is sampled resulting in a measureable energy E^{samp}.

The average longitudinal development of many hadronic showers can be parameterised by two exponentials [BHK+81, GR+90],

$$\frac{\mathrm{d}E}{\mathrm{d}x} = N \left(w t^{a-1} e^{-bt} + (1 - w) s^{c-1} e^{-ds} \right) \tag{5.25}$$

with the dimensionless variables

$$s = \frac{x}{\lambda} \tag{5.26}$$

---

[1] Some energy of the hadronic sector will contribute to the electromagnetic part of the shower through γ-rays from nuclear decays, amounting to 5-10% of the total deposited energy, which is neglected in this discussion.

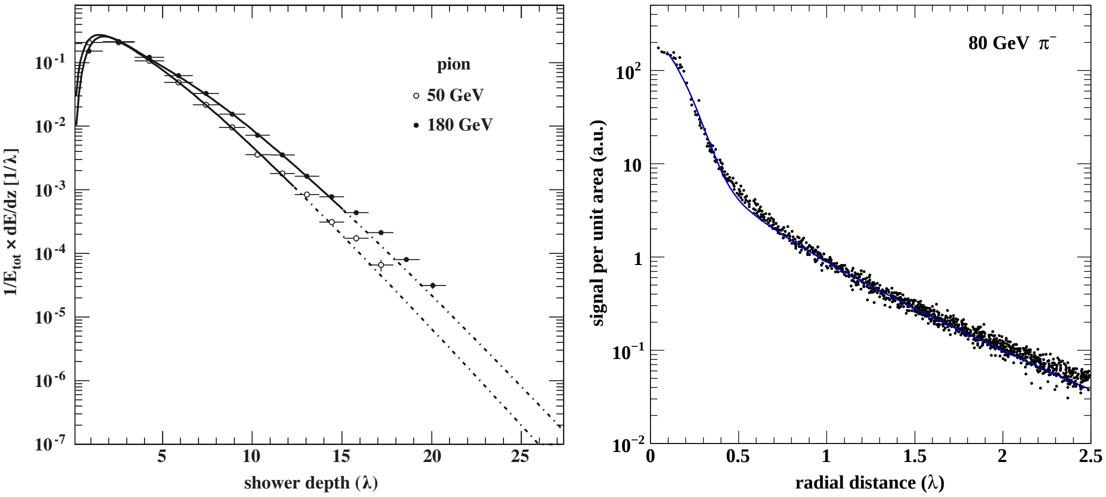


Figure 5.8: *Left:* Longitudinal development of a hadronic shower induced by pions measured in an iron-scintillator tile calorimeter. Shown is the fractional energy deposit per interaction length as a function of the shower depth. Data points from Adragna *et al.* [AA$^+$10b], the lines are fits to equation (5.25). *Right:* Lateral hadronic shower development of pions, measured with a lead/scintillating fiber calorimeter. Shown is the deposited energy per unit area as function of the distance to the shower axis [AB$^+$92]. The line is a fit to equation (5.29)

and t as defined in equation (5.12). In equation (5.25) the first exponential scales with X_0 and describes the electromagnetic sector, whereas the second exponential scales with λ and describes the hadronic sector. N, w, a, b, c and d are free parameters which need to be adjusted to experimental data. Figure 5.8 shows the longitudinal profile of hadronic showers measured by an iron-scintillator tile calorimeter of the ATLAS type. The lines are fits to equation (5.25), which has been slightly modified to account for the position of the shower start [AA$^+$10b]. The shower maximum occurs at depth $s_{\max}$, which is approximately given by

$$s_{\max} = 0.2 \ln E + 0.7. \tag{5.27}$$

A fraction of 95% of the energy of a hadronic shower is contained within a cylinder of length

$$L_{95\%} \simeq s_{\max} + 2.5\lambda E^{0.13}, \tag{5.28}$$

where in equations (5.27) and (5.28) $s_{\max}$ and $L_{95\%}$ are given in units of interaction lengths and the energy E is given in GeV.

The radial extent of a hadronic shower is mainly due to the transverse momentum of the secondary particles (5.21). Although the radial development does not scale with λ, 95% of the shower energy is contained within a cylinder of radius λ around the shower axis. Radial profiles of hadronic showers show a prominent core with a surrounding halo which can be seen figure 5.8. The central core originates from the electromagnetic component, whereas the halo with its much slower decreasing intensity is caused by the hadronic component. Radial profiles can be sufficiently described by a sum of a Gaussian with an exponential

$$\frac{\mathrm{d}E}{\mathrm{d}r} = Ae^{-(r/a)^2} + Be^{-(r/b)}, \tag{5.29}$$

representing the electromagnetic and hadronic shower contributions, respectively. The parameters A, B, a and b need to be adjusted to experimental data, see for example [AB+92, AC+08].

5.3 Sampling Calorimeters

A sampling calorimeter consists of layers of two different types. The passive layers consist of a high Z material used for the absorption of showers. Sandwiched between the passive layers are active layers for the signal generation and collection. The characteristic parameters defining such a calorimeter are the thicknesses of the active and passive layers d_a and d_p and the sampling fraction,

$$f_{\text{samp}} = \frac{E_{\text{mip}}(\text{active})}{E_{\text{mip}}(\text{active}) + E_{\text{mip}}(\text{passive})}. \tag{5.30}$$

Here $E_{\text{mip}}(\text{active})$ and $E_{\text{mip}}(\text{passive})$ denote the energy deposited by an idealized minimum-ionizing particle in the active and passive part of the calorimeter, respectively.

5.3.1 Electromagnetic Sampling Calorimeters

The energy response of an electromagnetic calorimeter is based on the principle that electromagnetic showers deposit their energy mainly through ionisation and excitation processes (see section 5.1). These processes generate free electrons which can be converted into a measureable signal, which is directly proportional to the energy of the incident particle(s). In sampling calorimeters only a part of this signal is measured, proportional to the total shower track length in the active layers. This track length fluctuates from event to event due to the stochastic nature of the shower development. These fluctuations are called sampling fluctuations and represent the most important constraint on the energy resolution of electromagntetic sampling calorimeters. The overall resolution can be parameterised by

$$\frac{\sigma(E)}{E} = \frac{a_{\text{samp}}}{\sqrt{E}} \oplus \frac{b_{\text{noise}}}{E} \oplus c \tag{5.31}$$

where $\oplus$ indicates a quadratic sum. The first term is due to the sampling fluctuations and can be approximated by [Wig00]

$$a_{\text{samp}} = 2.7\% \sqrt{d_a / f_{\text{samp}}}\,. \tag{5.32}$$

Equation (5.32) indicates that the resolution can be improved either by reducing the thickness of the active layers and thus sampling the shower more often with the same amount of active material (fixed sampling fraction), or by increasing the sampling fraction by increasing the amount of active material. In practice too small values of d_a with fixed f_{samp} are rarely feasible and quite costly. The second approach means that the effective Molière radius and radiation length become large, and showers become less compact. In practice a compromise must be found. Typical values for a_{samp} lie between 5–20%. It should be noted that equation (5.32) is only valid for calorimeters with typical design values of d_a in the range between 0.1–1.0X_0 and sampling fractions of the order 1–10%.

The second term in equation (5.31), b_{noise}/E, comes from the electronic contribution of the readout chain. It increases with decreasing energy of the incident particles and may become dominant at energies below a few GeV. It can be decreased by increasing the sampling fraction since the higher the signal from the active medium, the higher the signal-to-noise ratio. A different way to decrease the electronic noise contribution for ionisation type calorimeters is to reduce the detector capacitance C_D which is, however, experimentally challenging.

The last term in equation (5.31), c, is independent of the incident particle's energy. It is due to non-uniformities, the readout system, detector aging, radiation damage, etc. It starts to dominate at very high energies, usually around 50 GeV.

5.3.2 Hadronic Sampling Calorimeters

Hadronic showers consist of two parts with very different behaviour: an electromagnetic and a hadronic part (section 5.2.3). The signal generation for the electromagnetic shower component is identical to the one for purely electromagnetic showers, and therefore the sampling fluctuations discussed in the previous section apply to this component as well.

The energy of the hadronic shower component is carried by mesons, spallation protons, evaporation neutrons and recoil target nuclei. The calorimeter response to it, denoted by h, is energy independent for sufficiently large energies. Because of the invisible energy h is usually smaller than the electromagnetic response e,

$$e/h > 1 \tag{5.33}$$

and calorimeters for which this relation holds are called non-compensating.

However, since the electromagnetic content of hadronic showers increases with energy, equation (5.24), the calorimeter response to incident hadrons is energy dependent. This is expressed through the calorimeter's response to pions, denoted by π, which can be written as

$$\pi = f_{\text{em}} \cdot e + (1 - f_{\text{em}}) \cdot h\,. \tag{5.34}$$

This leads to a relationship between the e/h and e/π ratios,

$$e/\pi = \frac{e/h}{1 - f_{\text{em}}(1 - e/h)} \tag{5.35}$$

which shows that the signal non-linearity for hadrons is determined by the calorimeter's e/h value. The e/π ratio for different values of e/h is shown in figure 5.9. As expected, the non-linear response is worse for larger deviations from 1. In the limit of high energies the e/π ratio approaches 1 because of the increasing electromagnetic shower component.

In hadronic showers event-to-event fluctuations of the electromagnetic component are large and show a non-Gaussian behaviour. Especially fluctuations in the first inelastic interaction can lead to drastically different shower developments. Additionally, fluctuations in the invisible energy in the hadronic component result in

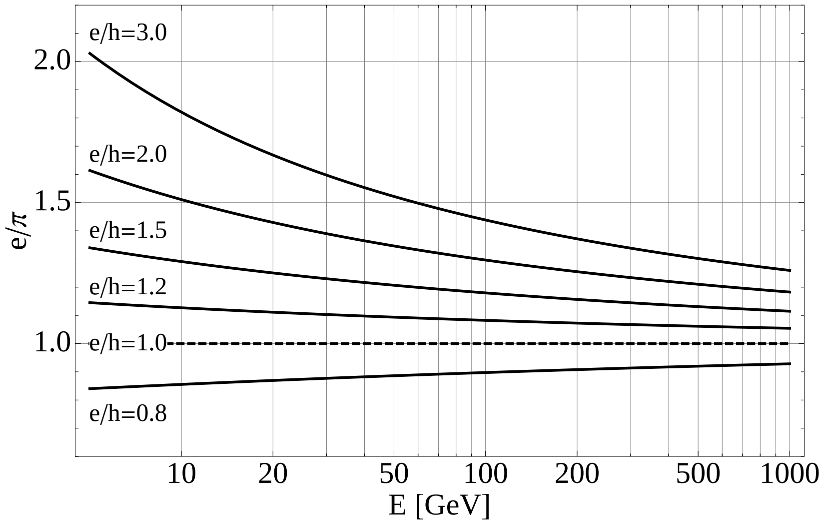


Figure 5.9: Relationship between the measured e/π ratio and the intrinsic value e/h.

asymmetric tails in the calorimeter response to incident hadrons. As a consequence the energy resolution does not scale with $E^{-1/2}$ and the resolution is significantly worse than for electromagnetic showers. Groom studied the resolution in detail and gives the following expression [Gro07]

$$\frac{\sigma(E)}{E} = \sqrt{\frac{(\pi/e)}{E}}\sigma_{\text{samp}} \oplus \sqrt{\frac{f_{\text{em}}}{E}}\sigma_e \oplus \sqrt{\frac{(1-f_{\text{em}})(h/e)}{E}}\sigma_h \oplus (1-h/e)\sigma_{\text{em}}(E)\,. \quad (5.36)$$

Here σ_{samp} is the variance of the sampled signal and σ_e and σ_h are the variances of fluctuations in the electromagnetic and the hadronic shower components, respectively. The variable σ_{em} denotes the variance of the size of electromagnetic component which exhibits a weak energy dependence.

The first term in equation (5.36) is the sampling contribution, which is multiplied by the inverse of the e/π ratio to account for the fact that the contribution is proportional to the sampled visible signal with mean $(\pi/e)E$ and not to the incident energy E. The second and third term originate from the intrinsic fluctuations of the electromagnetic and hadronic components of the shower. The last term in equation (5.36) is proportional to $(1-h/e)$ and has a weak energy dependence due to the fluctuations in f_{em}. Fits to presently available experimental data cannot resolve this energy dependence since it only affects the resolution at very high energies ($E > 400\,\text{GeV}$) and is thus compatible with a constant [Wig00]. Usually the resolution of hadron calorimeters is approximated by the form

$$\frac{\sigma(E)}{E} = \frac{a_{\text{samp}}}{\sqrt{E}} \oplus \mathrm{c}\,, \quad (5.37)$$

where the first term summarises the first three terms of equation (5.36) by neglecting the intrinsic energy dependence of f_{em} and the constant term resembles the last term. The noise contribution from equation (5.31) plays a minor role and is often left out, since the resolution is dominated by the $E^{-1/2}$ term. Typical values for a_{samp} in LAr calorimeters are of the order 50–80% with e/h in the range 1.4–1.8. Compensating calorimeters ($e/h = 1$) like the ZEUS depleted uranium/scintillator calorimeter have reported resolutions of $35\%/\sqrt{E}$ [BB+93].

5.4 The H1 Liquid Argon Calorimeter

The calorimeter of H1 was a longitudinally and laterally highly segmented lead/LAr and stainless steel/LAr sampling calorimeter [H193a,H197a]. It had good hermicity and homogeneity and was in stable operation for 15 years. It was non-compensating, but its fine granularity allowed to achieve compensation with an offline weighting method. In the following a description of the calorimeter and the weighting algorithm will be given, followed by a review of its performance from test beam measurements.

5.4.1 Calorimeter Layout

The H1 LAr calorimeter consisted of an electromagnetic and a hadronic section, which were placed in a large single cryostat (see figure 4.6). One important requirement on the cryostat was to minimise the dead material in front of the calorimeter while withstanding a pressure of 3 bar and supporting the calorimeter wheels. This was achieved with a two-shell construction separated by vacuum and insulation. The total amount of dead material from the tracker and cryostat in front of the electromagnetic calorimeter amounted from only $0.5\,X_0$ in the central region up to about $2\,X_0$ in the forward region ($\theta < 30°$) [H193a].
The calorimeter was organised in eight calorimeter wheels (section 4.2.2) which were subdivided into eight octants. They were self-supporting in order to minimise the amount of dead volumes (cracks) due the support structure. The wheels were separated by a narrow gap of the order of 10 mm with some additional space between the FB1/FB2 and FB2/OF for cabling. Independent readout cells were placed in these z-cracks to achieve a uniform sampling ratio in the crack region.
The calorimeter was not compensating. To overcome the resulting disadvantages a software weighting technique for hadronic showers has been developed. For this method to work the localisation of electromagnetic deposits in hadronic showers is crucial. This means that a developing shower needs to be sampled at least three to four times in the longitudinal direction, and the size of the readout cells is not to exceed $2\,R_{\mathrm{M}}$. This has been achieved by a high segmentation of the H1 LAr calorimeter, where the adopted segmentation in the longitudinal and radial direction is shown in figure 5.10. Since at HERA the centre-of-mass system was boosted along the proton direction, the calorimeter was significantly more segmented in the forward direction. Overall the calorimeter consisted of 45000 readout cells with about 31000 in the electromagnetic and 14000 in the hadronic part.

Electromagnetic Calorimeter (EMC)

The electromagnetic calorimeter consisted of lead absorber plates with a thickness of 2.4 mm and liquid argon gaps of 2.35 mm. The absorber plates together with copper-cladded G10[2] served alternatingly as high voltage electrodes and read-out boards. With exception of the IFE, the octants of the electromagnetic stacks were separated by φ-cracks filled with LAr, 2 mm steel plates or cabling. The effective size

---

[2] G10 is a material composed of glass woven fabric, impregnated with an epoxy resin binder. It is widely used for insulating structural parts in electrical and mechanical engineering.

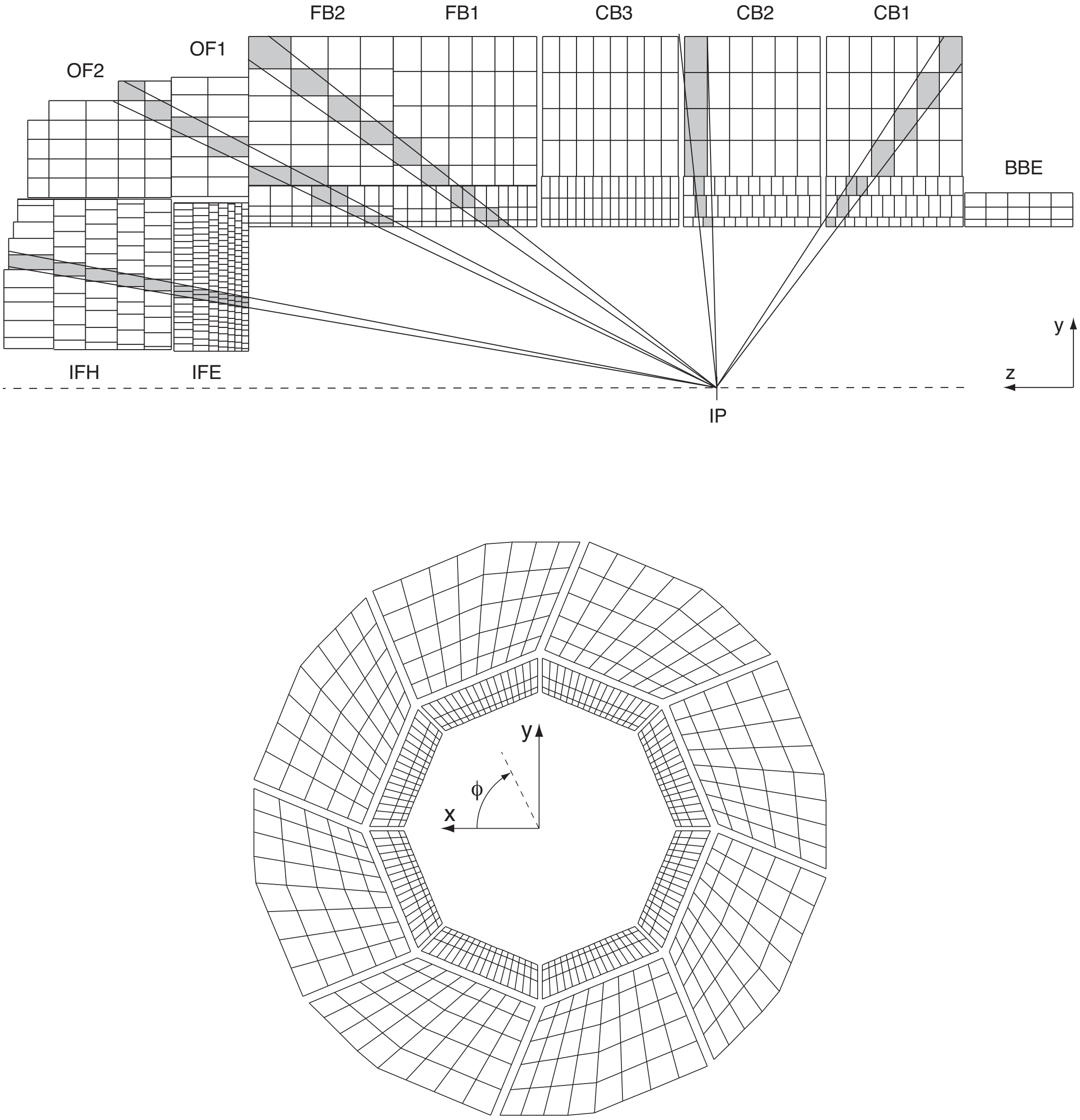


Figure 5.10: *Top:* Segmentation of the LAr calorimeter with a cut along the z-axis. Different incident angles are depicted by solid lines. The shaded areas show approximately the number of possible longitudinal samplings of a shower. *Bottom:* Segmentation of the LAr calorimeter in the radial direction. Shown is the FB1 module.

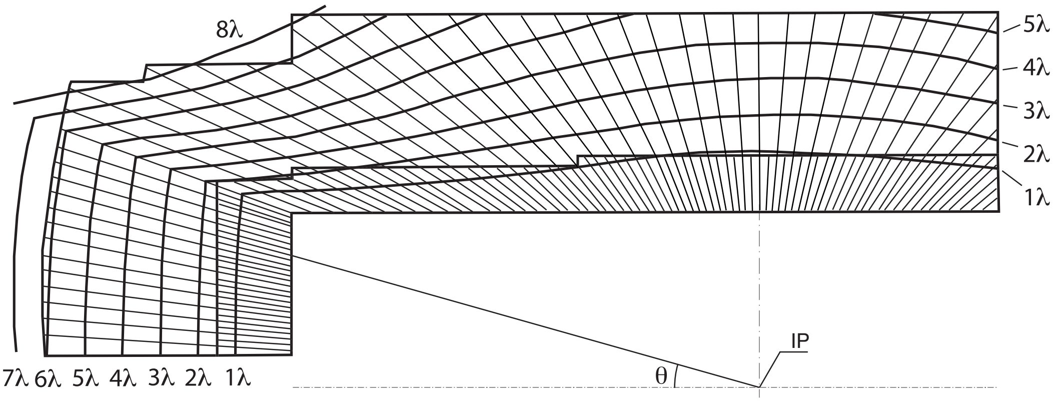


Figure 5.11: Lines of constant interaction lengths λ in the LAr calorimeter for different incident angles θ. The $1\,\lambda$ line corresponds to $20\,X_0$.

of these cracks corresponded to about $0.5\,R_{\mathrm{M}}$, and offline corrections were applied on an event-to-event basis to correct for this dead material. The electromagnetic stacks consisted of 3 (BBE, CBs, FB1), 4 (FB2) and 7 (IFE) longitudinal layers. In the radial direction the cell size was of the dimension of $1\,R_{\mathrm{M}}$, with exception of the backward wheels BBE, CB1 and CB2, where it was about $2.5\,R_{\mathrm{M}}$. The effective length of the electromagnetic calorimeter was 20–30 X_0, see figure 5.11.

Hadronic Calorimeter (HAC)

The hadronic stacks consisted of 16 mm stainless steel absorber plates and 2.4 mm liquid argon gaps. The readout boards were separated from the absorber plates so that variations in the thickness of the steel plates did not affect the size of the LAr gaps. The φ-cracks of the hadronic section were not pointing to the interaction point (see figure 5.10, bottom) with just minimal amount of dead material in them. The cell size was roughly twice as large as for the electromagnetic stacks. At the intersection between the electromagnetic and hadronic calorimeter no additional support structures were needed such that no dead material had to be introduced. The total effective length of the calorimeter (EMC and HAC) was 4.5–8 λ (figure 5.11) with a longitudinal separation of 7–10 layers. A summary of the characteristics of the LAr calorimeter is given in table 5.1.

Section	d_a (mm)	d_p (mm)	X_0 (cm)	R_{M} (cm)	E_c (MeV)	λ (cm)	f_{samp}
EMC	2.35	2.4	1.6	3.6	9.5	30.5	0.14
HAC	2×2.4	16+3	2.5	2.6	21.4	21	0.04

Table 5.1: Characteristic values for the electromagnetic (EMC) and hadronic (HAC) calorimeter. The variables d_a and d_p denote the thicknesses of the active and passive layers, respectively. The active and passive layers in the HAC have separate sections because of the independet readout boards.

5.4.2 Energy Measurement

Charge Collection and Electronic Chain

Free charges generated by particle showers in the liquid argon gaps were separated and collected by an applied high voltage across the gaps. The nominal field strength was 1 kV/mm which amounted to a total of approximately 2.3 kV. The gap width of 2.35 mm resulted in a drift time of 470 ns, which was the minimum integration time to accumulate the deposited charge[3]. The signals were collected by individual merging boards for each calorimeter module, where also a charge calibration from pulse generators with a known amplitude could be applied. The signals were then processed via a charge sensitive preamplifier and a shaping amplifier before they were digitised with analog-to-digital converters (ADCs). A detailed description of the electronic chain can be found in [H193a].

Noise Suppression

To estimate the contribution of electronic noise, the signal in each channel was measured in random events. The obtained noise distributions for channel i had mean energy E_n^i (on the em. scale, see below) and standard deviation σ_n^i. Typical values for the electronic noise (pedestal) from a single channel were about 20 MeV in the EMC and 40 MeV in the HAC. These pedestals were then subtracted from each channel. Further noise suppression was obtained by applying cuts on the measured energy of each channel. A signal in channel i was only kept if its measured energy E_0^i passed the conditions

$$|E_0^i| > 4\sigma_n^i \tag{5.38}$$

or

$$|E_0^i| > 2\sigma_n^i \quad \text{and} \quad |E_0^j| > 4\sigma_n^j\,, \tag{5.39}$$

where the cell of channel j is a neighbour of cell i. While σ_n is always positive, E_0 can be negative and thus provides compensation of the positive noise contributions.

Clustering

For efficient analyses of calorimeter data, neighbourig cells with a measured signal after the noise suppression were clustered together. The clustering was optimised to

- distinguish between signal and noise,
- minimise the number of clusters for single electromagnetic particles,
- separate signals from electromagnetic and hadronic subshowers originating from incident hadrons,
- optimise the two particle separation.

---

[3] Only half the deposited charge was collected, since the slow ions did not contribute to the signal.

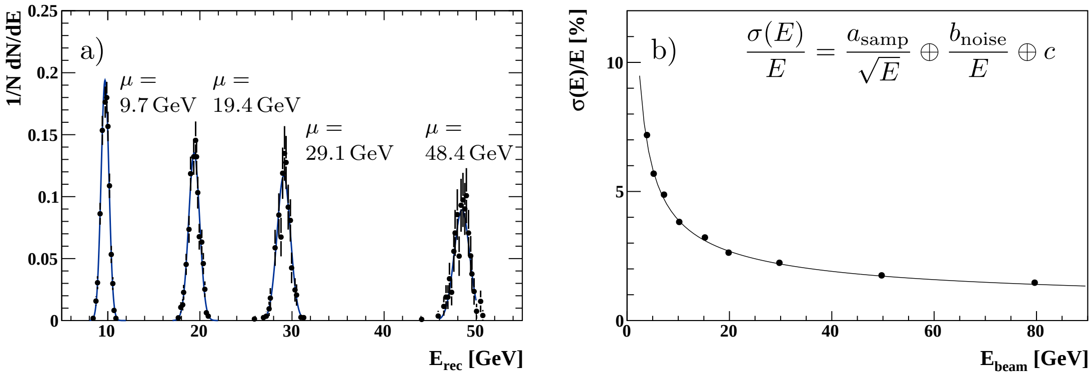


Figure 5.12: a) Reconstructed energies on the ideal electromagnetic scale for electron beam energies of 10, 20, 30 and 50 GeV in the BBE. The lines are Gaussian fits with mean values μ. b) Resolutions obtained from test beam measurements with electrons in the FBE [H194b]. The line is a fit by equation (5.31) with values $a_{\rm samp} = 11.2\%$, $b_{\rm noise} = 0.151\,\text{GeV}$ and $c = 0.6\%$.

The efficiency to reconstruct electrons between 2–100 GeV in a single cluster was between 95–98%. The mean cluster multiplicity for 100 GeV pions was found to be 4–9, depending on the calorimeter wheel. The separation of photons and pions generated with an angular distance of 5° was tested at various pion/photon energies. The separation probability was found to be between 50–90% [GW91].

Electromagnetic Scale

To reconstruct energy from measured charges for every calorimeter wheel w, constants of proportionality c_w need to be determined. They are defined through

$$E_{\rm em} = c_w \sum_{i}^{\rm channels} Q_i\,, \tag{5.40}$$

where Q_i denotes the measured charge in channel i in pC, $E_{\rm em}$ is the reconstructed energy in GeV and c_w has units of GeV/pC. The charge-to-energy conversion factors c_w are given for an ideal electromagnetic scale. This is defined such that all deposited energy inside the active volume is recovered. It implies corrections for the charge collection efficiency which depends on the applied HV, for the LAr purity, for cross-talk between channels, for the dead material inside the calorimeter and for detector imperfections. Thus, the ideal electromagnetic scale is not dependent on the noise suppression, leakage and the details of the experimental setup. The constants c_w have been obtained at test beam measurements with electrons at the CERN SPS [H194b]. Figure 5.12a shows results from these test beam measurements. Shown are distributions of reconstructed energy on the ideal electromagnetic scale for different test beam energies. The signals show Gaussian distributions which are shifted towards lower energies with respect to the beam energies. The shift is an effect of leakage and dead material in front of the calorimeter wheels. These losses are corrected for on the analysis level with calibration constants determined from *ep* data (section 8.1).

Using equations (5.30), (5.32) and d_a, d_p of the EMC (table 5.1) one obtains an expected resolution of $a_{\rm samp} = 11\%$. Results from the test beam measurements are

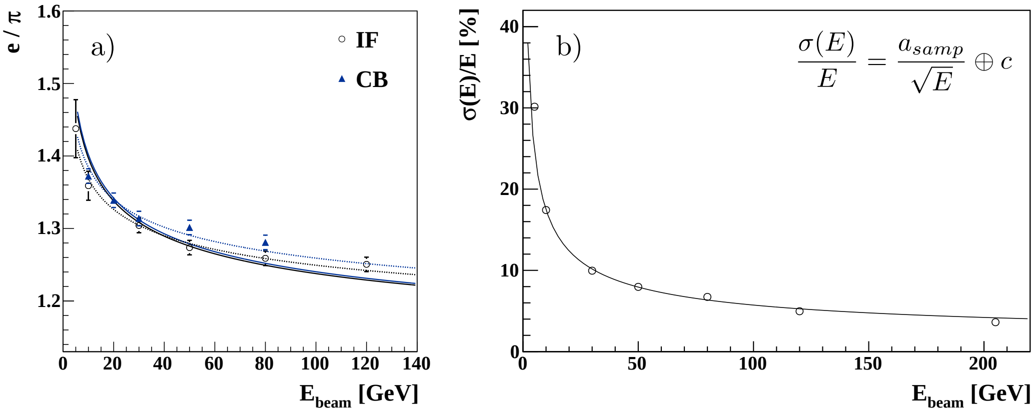


Figure 5.13: a) the e/π ratio of the H1 calorimeter in two different wheels as function of the pion energy. The lines are fits to equation (5.35) as described in the text. b) Obtained resolutions from test beam measurements with pions in the IF [H193b]. The line is a fit to equation 5.37. The obtained values are $a_{\text{samp}} = 55\%$ and $c = 1.6\%$.

shown in figure 5.12b together with a fit by equation (5.31). The obtained value $a_{\text{samp}} = 11.2\%$ is in good agreement with this expectation, and the small noise and constant terms of $b_{\text{noise}} = 0.151\,\text{GeV}$ and $c = 0.6\%$ show the excellent performance of the electromagnetic LAr calorimeter.

Hadronic Scale

Figure 5.13a shows the e/π ratio measured in different modules of the H1 calorimeter [H193b]. The lines are fits by equation (5.35). The solid lines have been obtained by fitting the data points from the IF with the parameters $k = 0.83$ and $E_0 = 1\,\text{GeV}$ (see equation (5.24)). The obtained values of $e/h \approx 1.7$ are similar for the IF and CB. The fit lies below the data points at high energies, which suggests a slightly higher value for k. Also the data points from the measurement in the CB exhibit a smaller slope towards higher energies. In a detailed study of the average electromagnetic fraction in hadronic showers, Gabriel *et al.* found a weak Z dependence of k and E_0. A simulation yielded values of $E_0 = 0.74$ and $k = 0.87$ in lead [GG+94]. These two parameters were used for the fits illustrated by the dashed lines. The resulting e/h ratio is 1.6 for the IF and 1.64 for the CB. Taking the uncertainties on the parameterisation of f_{em} into account the e/h value for the H1 LAr calorimeter lies in the range of 1.6–1.75 which compares well to obtained values of Fe/scintillator and Pb/scintillator calorimeters [AA+02].

To overcome the disadvantages of $e/h \neq 1$, namely a non-Gaussian lineshape and significant non-linearity (see section 5.3.2), the H1 collaboration developed a weighting scheme to compensate for losses due to the invisible energy component in hadronic showers [WK+94]. The applied software-weighting aims to find π^0 subshowers on the basis of their compactness and energy density. Hadronic showers are weighted with an exponential function depending on the measured energy density of cell i, given by $\rho_i^0 = E_{\text{em},i}/V_i$ with V_i denoting the cell's volume. The reconstructed energy of a hadronic shower is obtained via

$$E_{\mathrm{had}} = \sum_{i}^{\mathrm{channels}} E_{\mathrm{em},i} \left[C_1 \exp\left(-C_2\, \rho_i^0\right) + C_3 \right] \tag{5.41}$$

where the constants C_j have been obtained by an iterative method for jets, depending on their energy and polar angle. Below a jet energy of 7 GeV the iterative weighting function is replaced by an average e/π weighting factor, and in the region 7–10 GeV an interpolation between the weighting function and the e/π factor is used [H193b]. Figure 5.13b shows the obtained resolution from pion test beam measurements. The data points have been fitted by equation (5.37) without adding a noise term, since it has large ambiguities in a region dominated by the sampling term. The fit gives the value $a_{\mathrm{samp}} = 55\%$ with a small constant term of $c = 1.6\%$.

5.4.3 Simulation of Electromagnetic and Hadronic Showers

Simulated particles are tracked through the H1 detector using GEANT3 [BB+87] to simulate energy loss, multiple scatterings, inelastic interactions and so on. After the first inelastic interaction in the calorimeter, electromagnetic showers and the electromagnetic parts of hadronic showers are parameterised with GFLASH [GR+90, GP00]. For the hadronic part of the showers GEISHA [Fes85] as implemented in GEANT3 is used.

GFLASH generates particle shower shapes within a simplified geometry, where the active and passive materials of the calorimeter modules have been replaced by homogenous averaged media. In the first step GFLASH calculates the spatial distribution of the deposited energy. Only if no border crossing between the defined volumes occurs, is the shower parameterised. Otherwise the detailed shower simulation from GEANT is used. In the second step GFLASH computes the visible energy fraction in the active volume of the calorimeter. The calculation is based on the parameterisation (5.11) for the longitudinal shower development and an extension of equation (5.16) for the lateral shower development, where fluctuations on the mean and variances as well as their correlations are fully taken into account.

GEANT and GEISHA are used for tracking hadrons through the detector volume and simulating their interactions. The GEISHA program models hadronic interactions in successive steps. First the primary interaction with a quasi-free nucleon is simulated, followed by nuclear excitation, an intra-nuclear cascade and an evaporation step. Fission, capture reactions and coherent elastic scattering are also modeled through parameterised approaches.

The interface between the different simulation packages is handled by the program H1FAST [Kuh92] which is part of the detector simulation H1SIM. H1FAST has been extensively tested and it has been demonstrated many times that electromagnetic as well as hadronic showers are well simulated with this approach [GR+90,Pet92,H193b, H194a,GP00]. The description of electromagnetic showers has been further improved by tuning the free parameters of the parameterisations with deep-inelastic scattering data [EM+05a,EM+05b]. Recently H1 published a measurement of prompt photons in photoproduction, which relies heavily on well simulated showers to distinguish between photons, π^0 and hadronic background [H110a].

Chapter 6

Separation of Electromagnetic and Hadronic Showers

In the previous chapter it was shown that the electromagnetic component of hadronic showers exhibits an energy dependence. In non-compensating calorimeters with a different response to electromagnetic and hadronic showers ($e/h \neq 1$) this results in an energy dependence of the e/π ratio, which can lead to biases in hadron and jet measurements. Also, the resolution of the energy measurement of hadronic showers is worse than for electromagnetic ones. From equation (5.36) it can be seen that sampling fluctuations contribute only partially to the resolution, which is governed to a large extent by fluctuations in the electromagnetic and hadronic shower components, σ_e and σ_h. At very high energies the non-compensating nature $e/h \neq 1$ gives rise to an unwanted constant term in the resolution, proportional to $(1 - h/e)$.

The H1 software weighting technique aims to equalise the calorimeter response to electrons and hadrons by weighting hadronic showers depending on their energy density. Thus the weighting algorithm aims to correct for fluctuations in the electromagnetic content of hadronic showers on an event-by-event basis. It was shown that this leads to the expected energy resolution of $\sim 50\%/\sqrt{E}$ with a relatively small constant term of $\sim 2\%$. However, the uncertainty on the absolute energy scale for jets using this weighting method is about 5%. Therefore a jet calibration was developed which improved this uncertainty and corrected for the dead material in front of the calorimeter and the effect of the noise suppression algorithms applied [JZ+99]. The jet calibration works by weighting jets as a function of their energy in order to remove the observed residual non-linearity. The achieved jet energy measurement had an uncertainty of 2% [Sch02]. The development of an improved energy-flow algorithm and slightly modified noise and background reduction algorithms [PP+05] lead to a further improvement of the hadronic final state measurement in the track-dominated central region, where an uncertainty of 1.5% could be achieved at high transverse momenta [H110c]. In the forward region, where the calorimetric measurement dominates, the uncertainty could not be improved by this method.

The key to further improve the calorimeter performance is to estimate the electromagnetic content of jets on an event-by-event basis. Once this information is available it can be used in the calibration such that electromagnetic and hadronic deposits can be corrected accordingly. The event-by-event determination of the electromagnetic fraction f_{em} was shown to improve the energy measurement by the WA1 collaboration [AD+81, Wig02]. Recently the Dual-Readout Module (DREAM) collaboration showed that it is possible to eliminate deviations from the $E^{-1/2}$ scaling behaviour of the hadronic resolution by determining f_{em}, which was achieved by probing the different components of the hadronic shower [AC+05b].

In this chapter a method to separate electromagnetic from hadronic showers using statistical methods is introduced. The results obtained are used to preselect electromagnetic subshowers, which are in a first step excluded from the software weighting. This improves the reconstruction of the absolute energy scale as well as the resolution of the jet energy measurement. Further improvements are obtained by taking this information into account in the calibration of the hadronic final state measurement (chapter 8).

For the software weighting to work a fine longitudinal and lateral segmentation of the calorimeter is essential. The large number of readout channels not only provides information on the energy density of different parts of the shower, but also can be used to calculate characteristic shower shapes. Since electromagnetic and hadronic showers develop differently, these provide valuable information that can be used to distinguish between them. In an earlier work, Höppner and Wegener studied the use of neural networks for the energy reconstruction in the H1 calorimeter using shower estimators as input quantities [HW97, Höp97]. They trained neural networks in the FB region and tested the energy reconstruction with test beam data, simulated events and isolated tracks from 1995 data. The obtained resolutions were similar to those obtained with the standard H1 weighting method with improvements at low energies. However, the suitability of this method for jet data was not shown and it was not extended over the full calorimeter range. One major disadvantage of the method was that the absolute energy scale could only be obtained from simulated events, thus relying on the simulation of hadronic interactions. Deficiencies therein could lead to inaccurately simulated visible energy which in turn biases the energy reconstruction.

Instead of using neural networks to directly reconstruct the deposited energy, in this work they are used to estimate only the electromagnetic content of showers. Small deficiencies in the simulation of hadronic showers do not drastically change the simulated longitudinal and lateral shower shapes, thus enabling robust solutions which can be applied to data without biases. Subsequently this information can be used to improve the calibration. Through the over-constrained kinematics of NC events at HERA the hadronic final state can be calibrated for data and simulated events separately, without relying on the simulation of the correct absolute energy scale.

In figure 6.1 (left) the fractional energy deposited in the EMC is shown for simulated electrons and charged pions. Electrons (as well as photons) deposit essentially all of their energy in the EMC with only a negligible fraction in the HAC. In this energy range showers initiated by hadrons result on average in 4–5 calorimeter clusters with about 60% of the deposited energy contained in the EMC. In jets about one half of the energy is carried by photons from π^0 decays which are in close proximity to the hadronic jet constituents. Hence in a jet event the clusters from electromagnetic and hadronic showers will be close together and sometimes overlapping. Since the photon

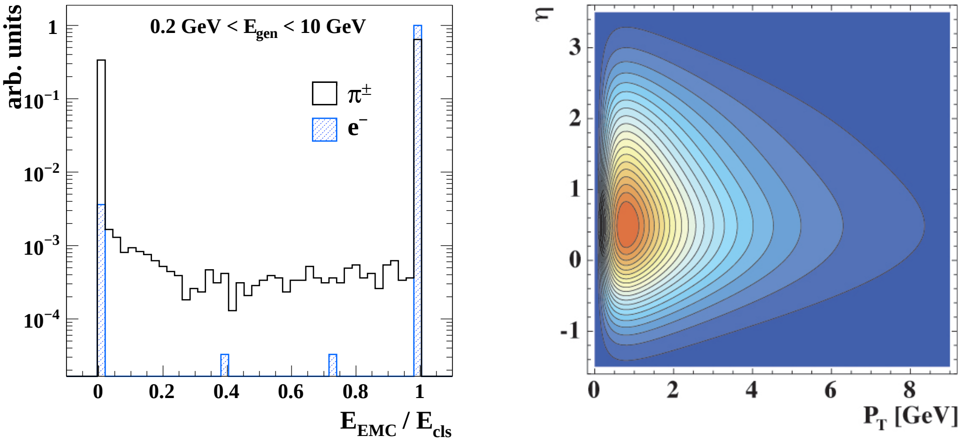


Figure 6.1: *Left:* Fractional energy deposited in the EMC by electrons and pions simulated with energies between 0.2–10 GeV. In this energy range pions deposit about 60% of their energy in the EMC. *Right:* P_T and η distribution of pions inside jets, obtained from simulated jets with $P_\mathrm{T}^\mathrm{jet} > 7\,\mathrm{GeV}$.

clusters should not be subject to the software weighting, a reliable preselection of the photon clusters will therefore improve the performance of the hadronic final state measurement.

No attempt to separate electromagnetic from hadronic showers in the HAC is made. Electromagnetically interacting particles deposit only an insignificant energy fraction in the HAC. Electromagnetic subshowers caused by hadronic showers are usually merged with the deposits from the hadronic shower into a single calorimeter cluster, due to the coarser granularity of the HAC compared to the EMC. Therefore, all clusters with more than 5% of the cluster energy measured in the HAC are subject to the software weighting, which takes the energy density of clusters and thus the electromagnetic fraction f_em into account.

The energy range for which the shower separation should be applicable can be deduced from the spectra of jet constituents. These spectra have been studied for various particles such as $\pi^\pm$, π^0, γ, $e^\pm$, K_L^0, p in jets with $P_\mathrm{T}^\mathrm{jet} > 7\,\mathrm{GeV}$. Figure 6.1 (right) shows a parameterisation of the η, P_T distribution of charged pions in jets. Whereas the majority of pions have momenta smaller than 2 GeV, the distribution has a long tail towards high momenta. At large pseudorapidities, jet constituents may carry momenta between 0.5–80 GeV. Hence the shower separation needs to be reliable over a large range in energy.

6.1 Shower Shape Estimators

Differences in the development of electromagnetic and hadronic showers can be used to distinguish between them. In sections 5.1 and 5.2 parameterisations of the electromagnetic and hadronic shower developments have been introduced. In particular, the radial and longitudinal shape of showers provide important information because of the different length scales in the shower development. Other useful variables are the compactness and energy density of a shower, as well as variables which take correlations between the longitudinal and radial shape of a shower into account.

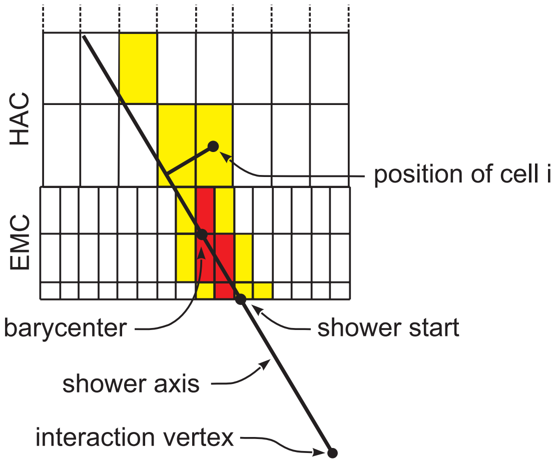


Figure 6.2: Definition of the coordinates used to determine cluster and cell positions. Cells with measurable signals are coloured and the four most energetic cells are depicted in red.

All shower shape variables described in the following are calculated for calorimeter clusters, i.e. they depend on the clustering algorithm described in reference [GW91].

Cluster coordinates

The shower axis $\boldsymbol{a}$ is determined by using the interaction vertex $\boldsymbol{v}$ and the cluster barycenter $\boldsymbol{b}$ (see figure 6.2) via $\boldsymbol{a} = \boldsymbol{b} - \boldsymbol{v}$. The barycenter is calculated as a weighted sum over all cells in one cluster, $\boldsymbol{b} = \sum_i w_i \boldsymbol{p}_i$. Here w_i is the weight of cell i and $\boldsymbol{p}_i$ its position in the H1 coordinate system. The weights are chosen to be the energy density on the ideal electromagnetic scale,

$$w_i = \rho_i = \frac{E_{\mathrm{em},i}}{V_i} \tag{6.1}$$

with V_i denoting the volume of cell i. The shower start $\boldsymbol{s}$ is defined to be the intersection of the shower axis with the front face of the calorimeter. Then the longitudinal distance of cell i from the shower start is given by

$$l_i = (\boldsymbol{p_i} - \boldsymbol{s}) \cdot \frac{\boldsymbol{a}}{|\boldsymbol{a}|}\,, \tag{6.2}$$

and its radial distance from the shower axis is

$$r_i = |\boldsymbol{p_i} \cdot \hat{n}| \tag{6.3}$$

with $\hat{n}$ denoting the unit vector perpendicular to $\boldsymbol{a}$.

Longitudinal and radial moments

The n^{th} moment of a distribution $f(x)$ is given by $\mu^n = \int x^n f(x)\mathrm{d}x$. The longitudinal and radial moments of the energy distributions of showers in a segmented calorimeter can be approximated by

$$\mu_L^n = \frac{1}{E}\int_0^\infty \frac{\mathrm{d}E}{\mathrm{d}l}\, l^n \,\mathrm{d}l \approx \frac{\sum_i w_i l_i^n}{\sum_i w_i} \tag{6.4}$$

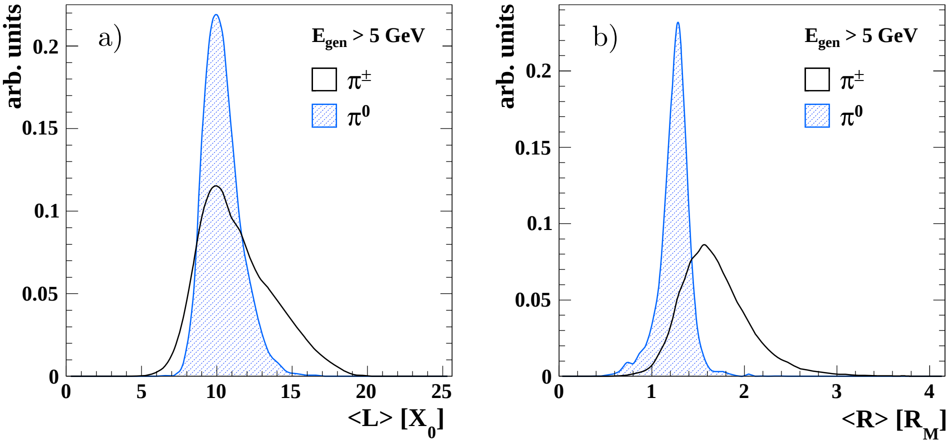


Figure 6.3: First moment of the longitudinal (a) and radial (b) energy distribution of calorimeter clusters in the EMC, originating from simulated charged and neutral pions with energies between 5 and 20 GeV incident on the CB3.

and

$$\mu_R^n = \frac{1}{E}\int_0^\infty \frac{\mathrm{d}E}{\mathrm{d}r}\, r^n \,\mathrm{d}r \approx \frac{\sum_i w_i r_i^n}{\sum_i w_i} \tag{6.5}$$

with $\frac{\mathrm{d}E}{\mathrm{d}l}$ and $\frac{\mathrm{d}E}{\mathrm{d}r}$ being the longitudinal and lateral deposited energy per unit length, respectively. The first moments μ_L^1 and μ_R^1 are the expectation (mean) values $\langle L\rangle$ and $\langle R\rangle$. They are shown in figure 6.3 for simulated charged and neutral pions. Clusters originating from π^0s and thus purely electromagnetic showers are more compact in the longitudinal and radial direction. The distributions are energy-dependent and at decreasing energies they become increasingly harder to distinguish.

Longitudinal and radial width

The standard deviations σ_L and σ_R are the widths of the longitudinal and radial distributions around their mean values,

$$\sigma_L = \sqrt{\mu_L^2 - (\mu_L^1)^2} \tag{6.6}$$

with a similar expression for σ_R. For particles with energies above 5 GeV σ_L is very similar for π^0 and $\pi^\pm$, while their radial width σ_R is very different, as indicated in figure 6.4. Electromagnetic showers are well contained within 1 R_M whereas hadronic showers have a larger lateral extent.

Cluster width

The n^th moment of the cluster's z-position is given by

$$\mu_z^n = \frac{\sum_i w_i z_i^n}{\sum_i w_i} \tag{6.7}$$

where z is measured in the H1 coordinate system and the standard deviation σ_z is given by

$$\sigma_z = \sqrt{\mu_z^2 - (\mu_z^1)^2}\,. \tag{6.8}$$

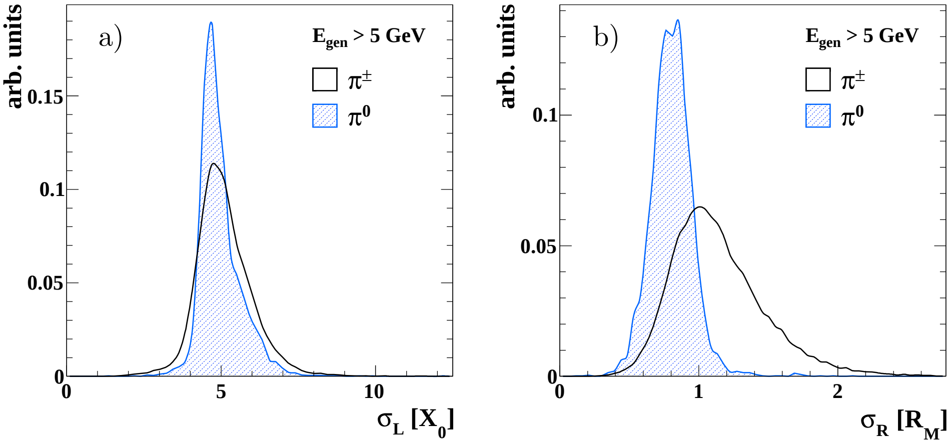


Figure 6.4: Standard deviation around the mean value of the longitudinal (a) and lateral (b) energy distribution of clusters in the EMC, originating from simulated charged and neutral pions with energies between 5 and 20 GeV incident in the CB3.

A vanishing value of σ_z implies that a cluster consists only of cells from one vertical layer. Such clusters are remnants from the clustering algorithm with usually very small energy. This estimator can thus be used to exclude clusters with unphysical shapes.

Covariance

The covariance between the longitudinal and radial energy distribution is defined as

$$c_{LR} = \mathrm{cov}(L, R) = \langle LR \rangle - \langle L \rangle \langle R \rangle. \tag{6.9}$$

It is a measure of the simultaneous longitudinal and radial variation of the shower. Positive values indicate that if the shower length is above the expectation value the radial width will be larger than the average width. Hadronic showers have the tendency to have larger tails in this distribution due to larger fluctuations in the shower development.

Kurtosis

The kurtosis is a measure of the ”peakedness” of a distribution. For the longitudinal energy distribution it is defined through

$$\kappa_L = \frac{\mu_L^4}{\sigma_L^4} - 3 \tag{6.10}$$

with an equivalent definition of the transversal kurtosis κ_R. A higher kurtosis means that large values of the variance are a result from infrequent extreme deviations opposed to a frequent small deviations.

Energy fraction in calorimeter layers

The fraction of the cluster's energy in calorimeter layer k is given by $f_k = E_k/E_{\mathrm{cls}}$. E_k is the energy in calorimeter layer k and E_{cls} is the total cluster energy, both

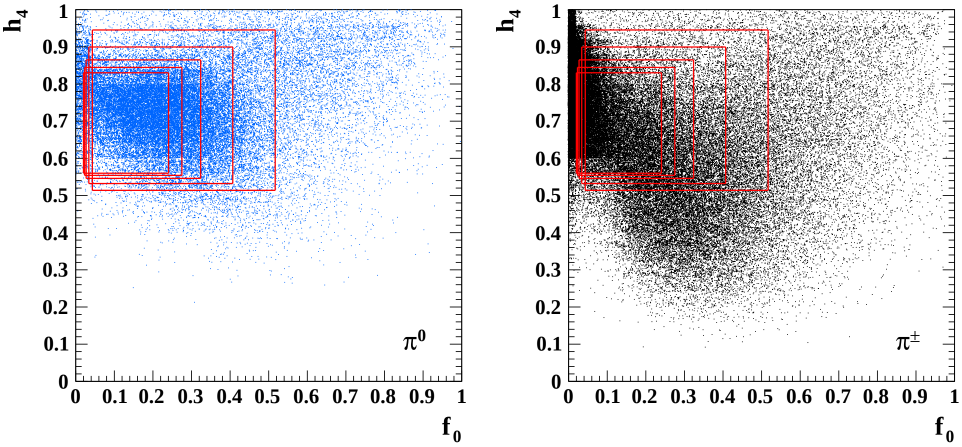


Figure 6.5: Energy fraction in the four most energetic cells h_4 versus energy fraction in the first calorimeter layer f_0 for clusters in the EMC. Shown are simulated π^0s and $\pi^\pm$ with energies between 0.5 and 30 GeV incident on the FB1. The red rectangles show the parameterisations, large rectangles depict small energies, becoming smaller with increasing energy.

taken on the ideal electromagnetic scale. f_k is a good estimator especially for low energetic clusters.

Energy fraction in hot core

The fraction of the cluster's energy in the N most energetic, neighbouring cells is called h_N and given by $h_N = E_{hN}/E_{\text{cls}}$. E_{hN} is determined by taking the cluster's most energetic cell as seed and summing up the energy of the $N-1$ most energetic neighbouring cells (see figure 6.2). Since electromagnetic deposits are more compact, h_N tends towards larger values for them. However, since the two photons from $\pi^0 \to \gamma\gamma$ have some opening angle between them, h_N is better for the discrimination between hadrons and single photons than between hadrons and π^0s.

Energy fraction outside a cylinder

The variable f_{R_M} denotes the energy fraction outside a cylinder with radius R_M. The cylinder axis is taken to be the shower axis $\boldsymbol{a}$, and the energy is measured on the ideal electromagnetic scale. Since 95% of the energy of electromagnetic showers is contained within a cylinder with radius R_M, f_{R_M} is small for electromagnetic clusters. The radial extent of hadronic showers is approximately λ and thus f_{R_M} tends towards larger values for them. For single particles f_{R_M} is a useful estimator, but it proved less viable in dense environments where showers from different particles overlap.

6.2 Standard Reconstruction

Two different approaches exist in H1's standard reconstruction to separate electromagnetic from hadronic showers. One approach is implemented in H1Rec. In this

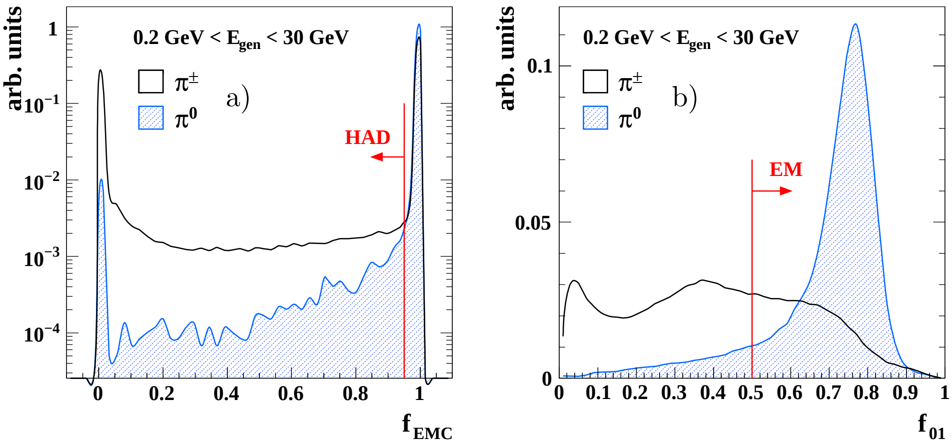


Figure 6.6: The two cuts used by Hadroo2 for the separation of electromagnetic and hadronic showers. The fractional energy in the EMC (a) and the fractional energy in the first two calorimeter layers (b).

approach clusters with more than 20% of their energy in the HAC are considered to originate from hadronic showers (compare figure 6.1). For clusters with a significant energy fraction in the EMC the energy dependence of the two estimators f_0 and h_4 is parameterised via

$$\begin{aligned} \max(0.01, 0.05 - 0.01 \ln E_{\mathrm{em}}) &< f_0 < \max(0.10, 0.60 - 0.12 \ln E_{\mathrm{em}}) && (6.11) \\ \min(0.60, 0.50 + 0.02 \ln E_{\mathrm{em}}) &< h_4 < \max(0.75, 0.98 - 0.05 \ln E_{\mathrm{em}}), && (6.12) \end{aligned}$$

where E_{em} is the cluster energy on the ideal electromagnetic scale [Loc92]. These parameterisations are shown in figure 6.5 together with results from simulated pions with energies between 0.5 and 30 GeV. The preselected regions contain a large portion of the π^0 initiated showers, which however cannot be selected without selecting a non-negligible fraction of hadronic showers. Wrongly preselected hadronic showers lead to an undesired systematic undercalibration such that the cuts were tightened by sacrificing some of the preselected electromagnetic energy [WK$^+$94].
With the development of the energy-flow algorithm Hadroo2 the selection of electromagnetic deposits was reviewed [PP$^+$05]. Two cuts have been introduced to decide between the electromagnetic and hadronic energy scale for clusters: if a cluster has more than 5% of its energy in the HAC, i.e.

$$f_{\mathrm{EMC}} < 0.95\,, \tag{6.13}$$

its energy is taken on the hadronic scale. All clusters in the EMC are separated depending on their energy fraction in the first two calorimeter layers. If the cluster has more than 50% of its energy in the first two calorimeter layers, i.e.

$$f_{01} > 0.5\,, \tag{6.14}$$

it is taken on the electromagnetic scale. This selection is shown in figure 6.6 for simulated charged and neutral pions. They were generated in an energy range between 0.5 and 35 GeV and evenly distributed over the whole calorimeter using H1Fast. The first requirement, $f_{\mathrm{EMC}} < 0.95$, results in a high efficiency for selecting clusters

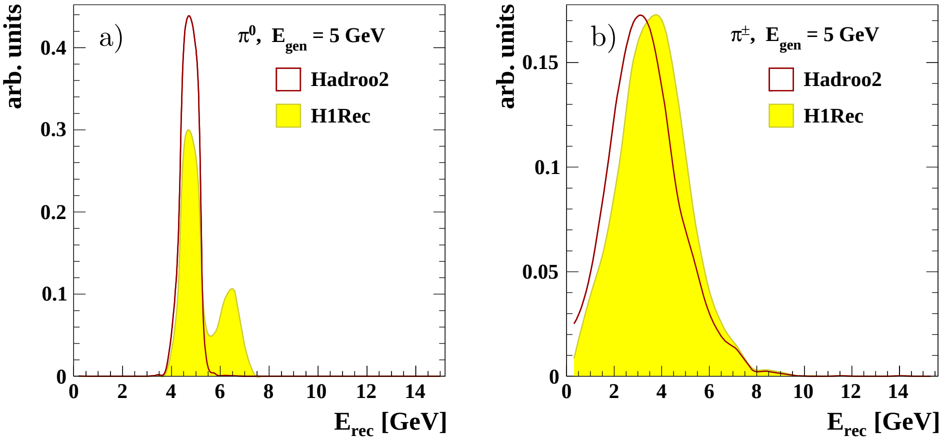


Figure 6.7: Reconstructed energies using the reconstruction algorithms Hadroo2 and H1REC for neutral (a) and charged (b) pions simulated with energies of 5 GeV.

originating from hadronic showers with only a very small fraction of wrongly classified clusters from electromagnetic showers. The EMC's effective length of 20–30 X_0 is enough to sufficiently contain even high energetic electromagnetic showers. The peak from π^0s at $f_{\text{EMC}} \approx 0$ results from photons which went into one of the ϕ-cracks and is suppressed by two orders of magnitude compared to photons fully contained in the EMC. The requirement $f_{01} > 0.5$ yields a good selection of electromagnetic showers, but more than 1/3 of the hadronic deposits are wrongly classified to be electromagnetic. The situation is worse for jets, where the jet energy is mainly coming from a large number of low energy hadrons, which have a higher probability to deposit their energy earlier in the calorimeter.

In both approaches selected electromagnetic clusters are excluded from the software weighting and get the energy on the ideal electromagnetic scale E_{em} assigned. For hadronic clusters the energy after the software weighting E_{had} is used. Reconstructed energies of simulated charged and neutral pions with energies of 5 GeV are compared for the two approaches in figure 6.7. For π^0s Hadroo2 yields good results with a Gaussian line shape and a mean reconstructed energy of 4.7 GeV. Using the selection of H1REC, about one third of the electromagnetic clusters are misidentified. The misidentified clusters are taken on the hadronic scale resulting in a significantly too high energy, peaking at ∼6.5 GeV (figure 6.7a). For charged pions the opposite picture emerges. Hadroo2 identifies about 60% of the clusters to be electromagnetic, compared to 10% from H1REC. This leads to smaller reconstructed energies in Hadroo2 which consequently leads to larger residual calibration factors. The best possible reconstructed energies from figure 6.7 are $\langle E_{\text{rec}}(\text{Hadroo2})\rangle \approx 4.7$ GeV and $\langle E_{\text{rec}}(\text{H1R}\textsc{ec})\rangle \approx 3.7$ GeV for π^0s and $\pi^\pm$, respectively. Despite H1REC's efficient selection of hadronic clusters the mean reconstructed energy is too small indicating that even after an efficient separation an additional calibration is necessary.

6.3 Requirements on a New Algorithm

The application of several cuts on the shower estimators introduced in section 6.1 and linear combinations of them have been studied. Only marginal improvements in the reconstruction of the hadronic energy with respect to the standard reconstructions could be achieved without spoiling the efficient selection of electromagnetic deposits by Hadroo2. This has several reasons:

- Shower shape estimators are approximated as discrete sums over cells. Due to the different geometries of the calorimeter wheels the estimators have different shapes for calorimeter wheels with different cell sizes.
- The amount of dead material traversed by incident particles before entering the calorimeter changes depending on the angle of incidence. This means that in some calorimeter regions the probability of a shower starting in front of the active volume is larger than in other regions. This alters the shower shapes depending on their position.
- The strong magnetic field from the superconducting solenoid in which the calorimeter was operated changes the shape of the estimators as well. The magnetic field lines are parallel to the z-axis. Hence, all showers caused by particles originating from the interaction point have a momentum component perpendicular to the field lines. The Lorentz force acting on positively and negatively charged particles, which are abundantly produced in any shower, is then directed away from the shower axis in opposite directions. This leads to a broadening of showers, resulting in smaller differences in the radial widths between electromagnetic and hadronic showers. Again, this effect depends on the shower position, since the momentum component perpendicular to z is proportional to $\sin\theta$.
- Most shower estimators are energy dependent. This means that they have to be studied additionally in different energy regions, if a reliable separation over the full energy range is the aim.

The new separation algorithm should be applicable to the whole hadronic final state. It should yield reliable results for isolated single particles as well as jets where overlapping incident electromagnetic and hadronic particles can spoil results obtained from single particle studies. Hence the new algorithm should be robust and result in an estimate of the amount of electromagnetic energy of a shower.

A solution to this complicated multi-dimensional problem is provided by the application of a neural network. Neural networks are designed to separate two classes of events using information from many input variables. An advantage of a selection based on a neural network over a cut-based selection lies in the networks' capability to extrapolate and to yield results in the case of missing input variables. Correlations between input variables are learned and taken into account. A well-trained neural network's fault-tolerance against noisy input data will be of special importance when applying obtained results from single particles to jets. In the following such a network will be outlined, followed by its application to the separation of electromagnetic and hadronic showers.

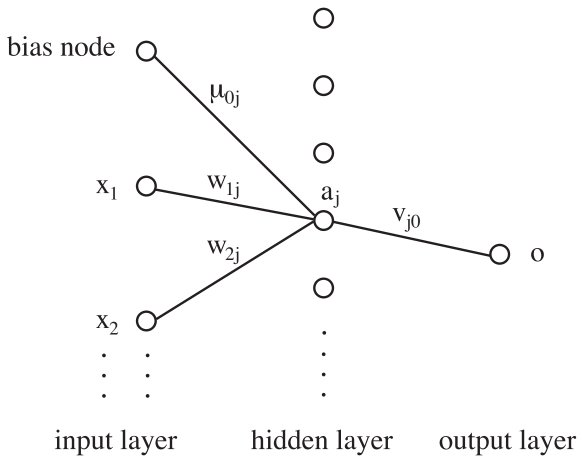


Figure 6.8: Architecture of a three layer neural network. Only the connections from node j in the hidden layer to the input and output layers are shown.

6.4 The Neurobayes Neural Network Package

The Neurobayes package is a tool for statistical data analyses based on Bayes' Theorem. It consists of a neural network combined with a complex preprocessing of the input variables. Bayesian regularisation techniques are used in the network training which also includes pruning techniques to eliminate the risk of overtraining [Fei04, FK06].

The Neural Network

A neural network consists of nodes (neurons) arranged in different layers. The neural network used in this work is a three layer feed forward back propagation neural network. The three layers consist of an input layer, a hidden layer and an output layer. The information flow in a feed forward neural network is strictly monotonic from the input layer to the output layer. Each input variable corresponds to one node in the input layer where an additional bias node with a constant value of 1 can improve the network performance. In the hidden layer the number of nodes can be chosen arbitrarily[1]. For binary classification problems the output layer consists of one node only which provides a continuous network output in the interval $[-1, 1]$. The architecture of a three layer neural network is depicted in figure 6.8. Let i be a node in the input layer and j a node in the hidden layer. The total number of nodes in the input layer is N and the hidden layer consists of M nodes. The input value of node i is denoted by x_i and the output of node j is a_j. The nodes of the input and hidden layer are connected with weights w_{ij}, such that for each node j a weighted linear combination of the inputs x_i is calculated,

$$a_j(\boldsymbol{x}) = \sum_{i=1}^{N} w_{ij} x_i + \mu_{0j} \, . \tag{6.15}$$

---

[1] The choice is not completely free, since on one hand enough freedom to learn the problem at hand needs to be allowed for, and on the other hand overlearning needs to be prevented. The achieved performance of a fully trained neural network should however be independent of reasonable variations of the number of nodes in the hidden layer.

The sum runs over all nodes in the input layer and μ_{0j} represents the threshold of node j, provided by its connection with the bias node. The output of each node in the hidden layer and the output node is computed via a non-linear activation function chosen to be a sigmoid-function

$$s_j = \mathrm{sigm}(a_j(\boldsymbol{x})) = \frac{2}{1+e^{-a_j(\boldsymbol{x})}} - 1 \tag{6.16}$$

which transforms the input interval $[-\infty, \infty]$ to $[-1, 1]$. The sigmoid function behaves approximately linearly around zero and saturates for small and large values of $a_j(\boldsymbol{x})$. The output o of this three layer feed forward neural network is given by

$$o = \mathrm{sigm}\left(\sum_{j=1}^{M} v_{j0}\ \mathrm{sigm}\left(\sum_{i=1}^{N} w_{ij}x_i + \mu_{0j}\right)\right), \tag{6.17}$$

where v_{j0} denotes the weight between node j in the hidden layer and the output node o.

Determination of Weights

The determination of the network's weights w_{ij} and v_{j0} is usually referred to as network training. This is achieved by feeding samples of known data to the network and minimising the deviation of the network output o from the desired result T. The minimisation is achieved by defining a cost function, which in this work is chosen to be an entropy

$$E = \sum_j \ln\left(\frac{1}{2}(1 + T_j\, o_j + \epsilon)\right), \tag{6.18}$$

where the sum runs over all patterns j in the training sample. T_j is the true value of pattern j and o_j is the corresponding network output. In order to avoid numerical problems for an untrained network a small regularisation constant ϵ is introduced which is set to 0 after a few iterations. At the beginning of the training the network weights are randomly distributed following a Gaussian distribution with mean 0 and width 1. The network weights are then adjusted iteratively by minimising the entropy function E by a modified backpropagation algorithm in combination with gradient descent, as described in [Fei04].

The training is performed in a quasi-online mode with weight adjustment every 300 events. For large samples with partly redundant information this ensures faster convergence compared to weight updates after the full sample has been processed. To remove biases from subsequent training patterns, if for example sorted input is used, the events are read in random order. Additionally, the input sample is separated randomly into training and test samples with a ratio of 1:1.

The Bayesian Approach

The Neurobayes neural network is trained with regularisation techniques based on Bayes' Theorem which provides a relation between a conditional probability and its inverse,

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}. \tag{6.19}$$

Here H denotes a hypothesis and D are measured data. Then $P(H|D)$ is the conditional probability of having H, given D, also called the posteriori probability. The inverse of it, $P(D|H)$ is the probability of observing D, assuming H. The probabilities $P(H)$ and $P(D)$ are the probabilities of having H and D without any further assumptions which are called a priori probabilities.
This theorem can be applied to the network training, where a measurement of the true quantity t yields the value x. Then the neural network can be used to estimate the conditional probability $P(t|x)$ for each event, where $P(t|x)$ is the probability of obtaining t if x has been measured. Hence the output of the neural network can be interpreted as a conditional probability, if the a priori probabilities at the training stage are realistic. In other words, the network output is the probability of finding a signal event, if the network is trained with the same signal to background ratio as found in the data.
Additionally, Bayesian statistics are used to determine the significance of the individual weights during the network training. Insignificant connections and network nodes are removed (*pruning*) to eliminate the risk of overtraining and ensure a minimal network topology to correctly reproduce the characteristics of the data.

Preprocessing

Prior to the training a sophisticated preprocessing of the input variables is performed. First a flattening of the input distribution is performed by a non-linear monotonous variable transformation such that the transformed variable lies between -1 and 1. This is advantageous since otherwise extreme outliers may saturate neurons and dominate the network output. In the next step the input variables are transformed into Gaussian distributions with mean 0 and width 1, such that they follow the same distributions as the random weights at the beginning of the network training. This provides an optimal initial condition for a fast convergence of the training.
It proved to be advantageous for the network training to decorrelate the input variables. Taking the $N-1$ input variables and the true value t (the training target), the $N \times N$ covariance matrix C is calculated via

$$C_{ij} = \frac{1}{n} \sum_{\text{events}} (x_i - \langle x_i \rangle) \cdot (x_j - \langle x_j \rangle) \,, \tag{6.20}$$

where n is the total number of events and x_i and x_j denote the input variables. The correlation ρ_{ij} between the input variables x_i and x_j can be expressed as

$$\rho_{ij} = \frac{C_{ij}}{\sqrt{V[x_i]}\sqrt{V[x_j]}} \tag{6.21}$$

with the variance of x denoted by $V[x]$. Neurobayes decorrelates all input variables by Jacobian rotations of the correlation matrix. These rotations are simultaneously applied to the correlation vector $\boldsymbol{\rho} = (\rho_{i,1}, \ldots, \rho_{N-1,1})$ of each variable's correlation to the target. The resulting vector $\tilde{\boldsymbol{\rho}}$ is used to calculate the total correlation of all variables to the target

$$\rho_T^2 = \sum_{i=1}^{N-1} \tilde{\rho}_i^2 \,. \tag{6.22}$$

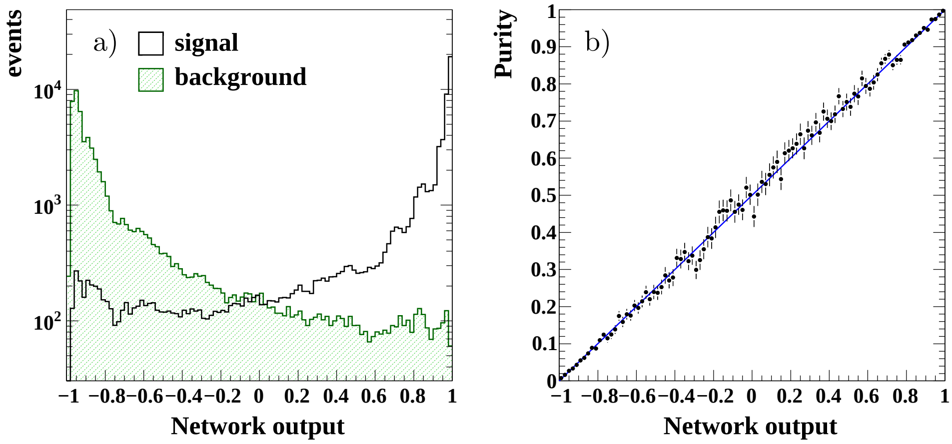


Figure 6.9: a) Output of a trained neural network for signal and background samples. Calorimeter clusters originating from charged pions are used for the signal sample and clusters from electromagnetic showers are used as background. b) Purity versus the network output which should lie on a straight line (equation (6.25)) for a well-trained network.

The loss of correlation when removing an input variable x_i is its significance $S(x_i)$,

$$S(x_i)^2 = n\left(\rho_T^2(x_1,\ldots,x_{N-1}) - \rho_T^2(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_{N-1})\right). \tag{6.23}$$

It serves as a measure of the additional information obtained by adding input variable x_i to the training. During the training process the significances of all variables are monitored and variables with significances smaller than $3\,\sigma$ are removed from the input to improve the signal-to-noise ratio.
The vector $\tilde{\boldsymbol{\rho}}$ contains the full linear correlation to the target, with all other correlations being zero. Hence it is the optimal correlation that can be achieved with a linear method. Improvements are possible by exploiting the non-linear nature of the neural network with the preprocessed variables as input.

Training Results

The purity P is defined as fraction of signal events in a given sample, $P = s/(s+b)$ where s and b denote the number of signal and background events, respectively. The function [Fei04]

$$\chi^2 = P(o-1)^2 + (1-P)(o+1)^2 \tag{6.24}$$

is a measure of the achieved separation between signal and background, assuming that the network output o corresponds to -1 for background and to $+1$ for signal events. For a well trained network χ^2 is minimal, $\mathrm{d}\chi^2/\mathrm{d}o = 0$, which leads to

$$P = \frac{o+1}{2}. \tag{6.25}$$

Consequently the network output scaled to the interval $[0,1]$ corresponds to the probability of observing a signal event. The signal and background contributions to the network output of a well trained network are shown in figure 6.9a. Clusters from simulated charged pions are used for the signal sample and clusters originating from photons from π^0 decays are used for the background sample. The quality of

cut	rel. occurrence		decision
	$\pi^{\pm}$	π^0	
$f_0 = 1$	3%	2%	EM
$f_1 = 1$	2%	1%	EM
$f_2 = 1$	3.5%	¡0.1%	HAD
FB1, $\sigma_z < 0.1\,\mathrm{cm}$	3.5%	2.5%	EM
FB2, $\sigma_z < 0.1\,\mathrm{cm}$	7%	1%	HAD

Table 6.1: Cuts 3 and 4 applied for the cluster preselection and the decision made for clusters not considered in the separation by the neural network. The relative occurrence denotes the fraction of clusters removed by the cuts for simulated single pions in an energy range between 0.5 and 30 GeV.

the network training can be assessed by calculating the purity for each output bin which is shown in figure 6.9b. For a well trained network the points should lie on the diagonal given by equation (6.25).

6.5 Cluster Selection

Prior to the network training clusters with no or very little information in the cluster estimators are identified. They contribute only to additional noise in the shower separation. In the following the selection of clusters for which the separation with the neural network is defined is summarised.

1. $f_{\mathrm{EMC}} > 0.95$
2. $E_{\mathrm{em}} > 0.2\,\mathrm{GeV}$
3. $f_k \neq 1$ for $k = 0, 1, 2$
4. $\sigma_z > 0.1\,\mathrm{cm}$
5. CB1 and CB2: $\langle R \rangle < 7\,\mathrm{cm}$ for clusters with $E_{\mathrm{em}} > 3\,\mathrm{GeV}$

The first cut is identical with the one used in Hadroo2, shown in figure 6.6a. It is used to preselect clusters contained in the EMC. All deposits in the HAC are assumed to originate from hadronic showers. Their electromagnetic content is taken into account by the software weighting. The second cut removes clusters close to the noise level of the calorimeter. Clusters with an energy smaller than 0.2 GeV are taken on the electromagnetic scale. Cuts number 3 and 4 are employed to discard clusters consisting only of cells in a single horizontal or vertical calorimeter layer. These clusters are usually remnants from the clustering and contain very little energy. The decision made for these clusters are based on single particle MC simulations and are summarised in table 6.1. In two cases the relative occurrence is larger for clusters originating from hadrons and still the electromagnetic scale for them is used. This choice is made to reduce non-gaussian tails for electromagnetic particles which can not be corrected for afterwards. The slightly smaller resulting energy for hadronic clusters can be corrected for with the subsequent calibration. Cut number 5 is only applied to clusters in the CB1 and CB2. Due to the coarser granularity in these two calorimeter wheels deposits from many particles are merged to single clusters in dense environments. These clusters have large transverse extensions which cannot be reproduced by single particle simulations. Since the network cannot be trained for them they are taken on the electromagnetic scale. This is a valid approximation since in the CB region the track finding efficiency is high and most of the energy from charged hadrons is accounted for by tracks. Most of the remaining energy originates from photons with only a small fraction coming from neutral hadrons. Again, the calibration can correct for the small remaining amount of invisible energy.

6.6 Neural Network Training

For the network training simulated single particles with the full detector simulation and reconstruction are used. In order to take the effect of calorimeter noise into account the events are overlaid with electronic noise recorded in periods without beam. The simulated particles are generated uniform in θ and ϕ covering the full calorimeter acceptance. A mixture of neutral pions and single photons is used for the generation of electromagnetic showers and negatively and positively charged pions are used for hadronic showers. The ratio of electromagnetic to hadronic showers is 1:1, approximately corresponding to their abundance in the data. For validating the network training also electrons, neutral and charged kaons, neutrons, protons and anti-protons are simulated.

6.6.1 Geometrical Division

Cluster shapes differ considerably among different calorimeter wheels. The reason for this is a varying amount of dead material in front of them, their different geometries[2], a varying angle of incidence and the strong magnetic field. Attempts to train the neural network with the cluster position and angle of incidence as additional input variables were not successful, since the full correlation to the cluster shapes could not be learned by the 3 layer neural network. Instead, the calorimeter is organised in different regions corresponding to the different calorimeter wheels and individual neural networks are trained for each wheel. After this division the significance of the cluster position drops below the $3\,\sigma$ limit. Furthermore, since the particles' angle of incidence varies only slightly going from the outer to the inner edge of a single wheel it becomes insignificant as well. Only the cluster's proximity to a ϕ or z-crack is significant due to its effect on the cluster shape. This is taken into account by introducing two boolean variables which are set to 'true' if the cluster consists of cells adjacent to a crack.

---

[2] The absorber plates are oriented paralell or perpendicular to the beam axis, depending on the calorimeter wheel. Different wheels also have different cell sizes.

6.6.2 Choice of Input Variables

From the large number of possible input variables introduced in section 6.1 only some can be used for the network training. The reason is twofold. Firstly, only variables well described by the MC simulation should be used. The network training relies on the detector simulation and shortcomings in the simulation can result in biases when the results are applied to data. Secondly, the input variables should not differ greatly between single particle and jet data. Since the trained networks will be used in the reconstruction of the hadronic final state, they have to yield reliable results for isolated particles as well as jet clusters. The possible input variables were carefully studied for all calorimeter wheels and only the ones fulfilling these requirements are used for the network training.

In order to allow the neural network to learn the correlation between the cluster's energy and the shower shapes, the energy is chosen as an input variable. Care has to be taken, however, that the network does not base its decision on the energy, but instead learns the correlation to the other input variables. This is achieved by identical energy distributions for electromagnetic and hadronic clusters in the training. To monitor the effect of the energy as network input, the purity as function of efficiency calculated for the cluster energy only can be used. In all network trainings this distribution is flat at a constant value of $\sim 50\%$ showing that there is no direct information in the cluster energy. The correlation to other variables is usually higher than 30%.

6.6.3 Energy Dependence

The shapes of hadronic and electromagnetic showers have an intrinsic energy dependence as shown in chapter 5. The clustering algorithm applied in H1 tends to split hadronic clusters with higher energies into more clusters, which also influences the shape of the resulting clusters. The measured energy spectrum of clusters in jets is shown in figure 6.10a. It resembles the distribution of simulated jet particles (cf. figure 6.1, right). To have the same apriori probabilities in the training sample as in the data, the training samples are weighted to the physical energy spectra of jet particles.

When training one network for the full energy range the result is statistically dominated by clusters with energies smaller than 5 GeV which make up more than 90% of the sample (see figure 6.10a). The extrapolation of the network training to higher energies becomes uncertain. Above 20 GeV with effectively no clusters available for the training, the purity drops rapidly. However, a wrong classification of high energy clusters has a large effect on the hadronic final state measurement. To overcome this and to ensure a good separation over the full energy range the training is split up into low and high energies. The training for low energy clusters is performed with clusters in the interval 0.5–5 GeV whereas the high energy network is trained with clusters between 4–30 GeV. In order to have a smooth transition between the output of the low and high energy networks o_L and o_H, the results are combined via a transition function $f(E)$,

$$o = f(E)\, o_H + (1 - f(E))\, o_L\,. \tag{6.26}$$

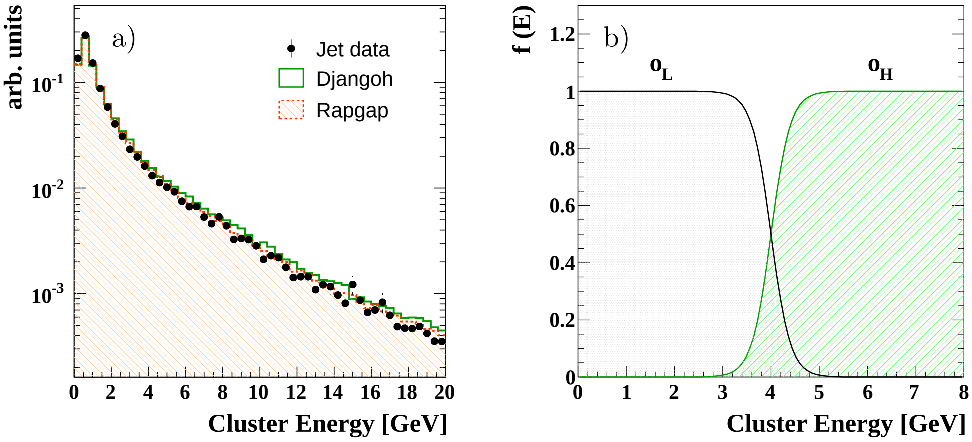


Figure 6.10: a) Measured energy distribution of clusters in jets, compared to DJANGOH and RAPGAP after the full detector simulation. b) Multiplicative factors for the low and high energy network outputs o_L and o_H as function of the cluster energy.

The function $f(E)$ is a sigmoid function chosen to be a simple logistic function,

$$f(E) = \frac{1}{1 + \exp(-(E - E_{\mathrm{tr}})/\sigma_E)} \tag{6.27}$$

with the parameters $E_{\mathrm{tr}} = 4\,\mathrm{GeV}$ and $\sigma_E = 0.2\,\mathrm{GeV}$. The combination of the network outputs is illustrated in figure 6.10b. Detailed tests for clusters with energies in the transition region between 3–5 GeV were performed to ensure stable results over the full energy range. In the forward region the combination is extended since the cluster energy distribution extends up to 100 GeV there. Hence in the IF and FB2 three networks are trained and their results combined via

$$o = f_2(E)\, o_H + f_1(E)(1 - f_2(E))\, o_M + (1 - f_1(E))\, o_L \tag{6.28}$$

with the above parameters for f_1 and $E_{\mathrm{tr}} = 30\,\mathrm{GeV}$ and $\sigma_E = 1\,\mathrm{GeV}$ for f_2. In this case the medium energy network is trained in the range 5–30 GeV and the high energy network uses clusters in the range 30–100 GeV for the training.

6.6.4 Training

Due to the above described separation into calorimeter wheels and energy ranges a total of 16 neural networks is trained. For each of the 16 networks the input variables are carefully chosen and the training results are verified by independent samples.

Small deficits in the simulation of the details of the shower development are only visible in the total number of cells in a cluster. In this distribution the data tend to larger mean values than what is obtained from the simulation. However, in the calculation of all used shower variables the cells are weighted with their energy density, which results in an excellent description of mean values and their correlations over the full energy range in the whole calorimeter as already shown by Grindhammer and Peters [GP00].

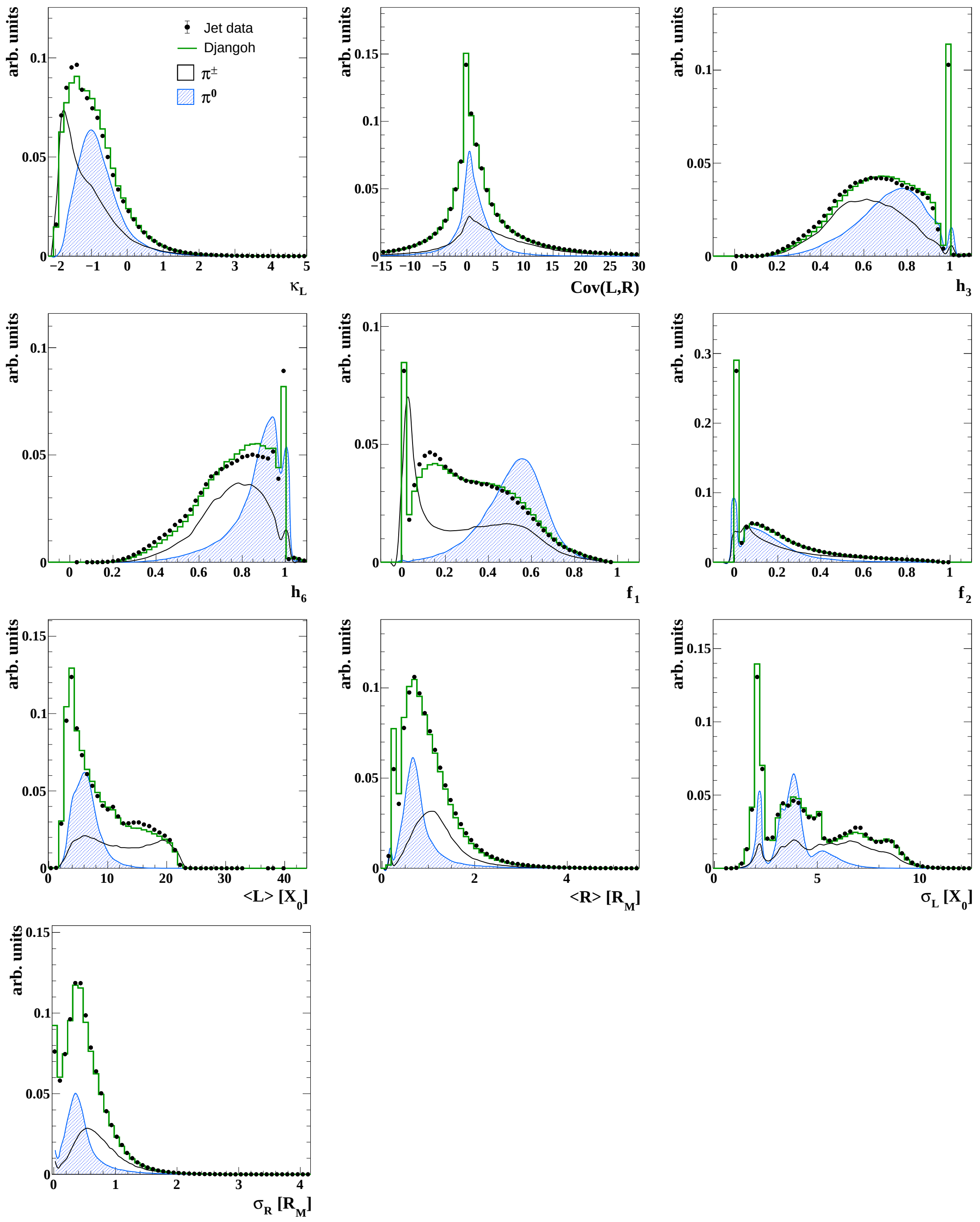


Figure 6.11: Variables used for the network training for energies smaller than 5 GeV in the IF. Shown are measured distributions from a jet sample together with simulated jet events from DJANGOH and simulated electromagnetic and hadronic single particles.

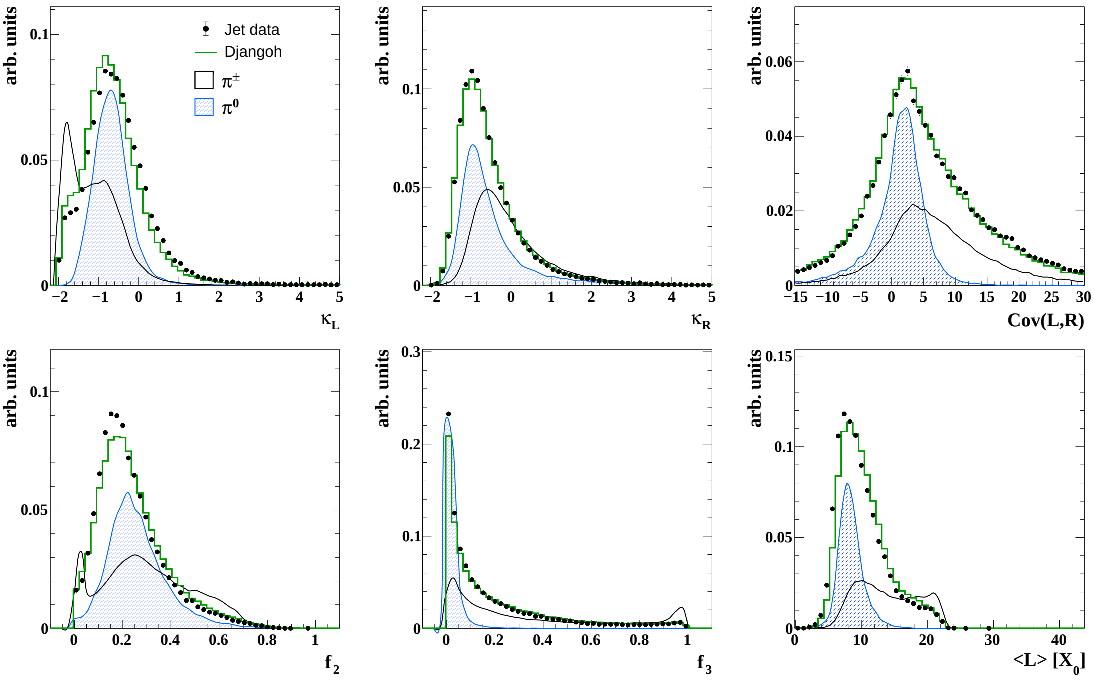


Figure 6.12: Variables used for the network training for energies between 4 and 30 GeV in the IF. Data and simulated clusters as in figure 6.11.

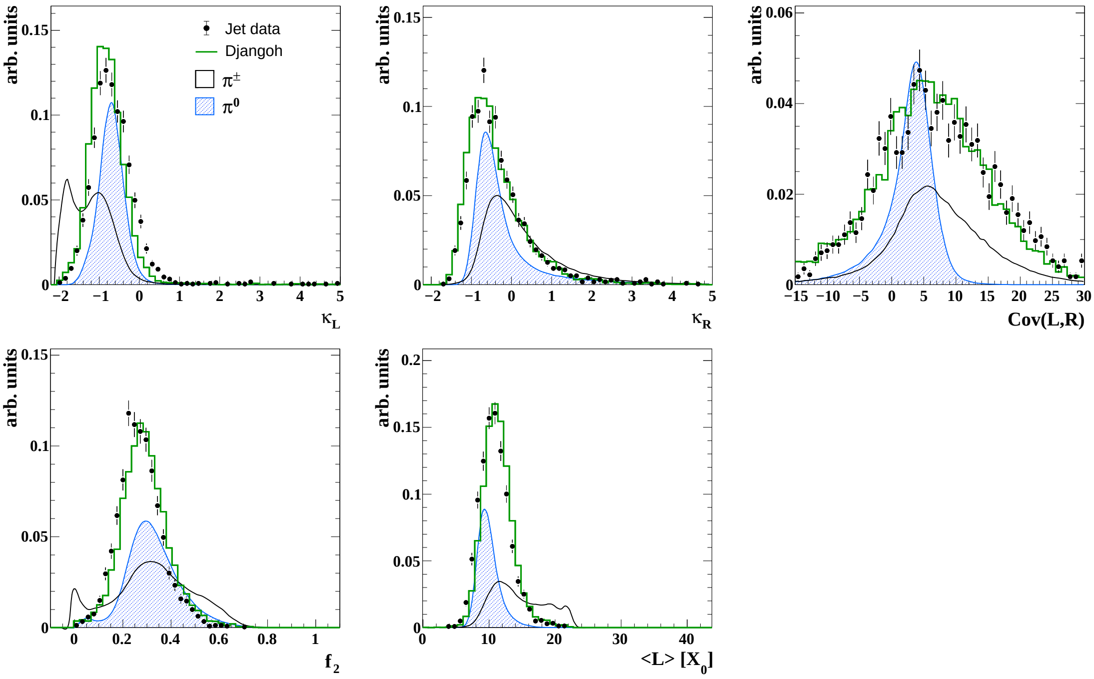


Figure 6.13: Variables used for the network training for energies above 30 GeV in the IF. Data and simulation as in figure 6.11.

As an example the input variables used for the three networks in the IF are shown in figures 6.11–6.13. Information on the longitudinal shower development proves an important input to the neural networks in all energy regimes. These can be the fractional energy in the N$^{\text{th}}$ calorimeter layer f_N or the mean shower length $\langle L \rangle$. These distributions are fairly unaffected by going from single particles to dense environments. The energy dependence of $\langle L \rangle$ can be seen when going from a mean energy of $\sim$1.5 GeV (fig. 6.11) to an energy of $\sim$15 GeV (fig. 6.12) and to $\sim$60 GeV (fig. 6.13). The mean value for electromagnetic showers shifts from $5\,X_0$ at low energies to $10\,X_0$ at the highest energies, the tails towards higher values are mostly due to hadronic showers. The distribution falls off rapidly when approaching $25\,X_0$ which is the effective length of the EMC.

Lateral variables like the mean cluster radius $\langle R \rangle$ and the standard deviation σ_R are only useful at low energies. At higher cluster energies overlapping incident particles shift these distributions to larger values than what can be found in single particle simulations. The energy fraction in the hottest N cells, h_N can be quite different for single photons and π^0s. However, it is still a useful variable at low energies if the network is trained with a mixture of single photons and π^0s. At high energies it suffers the same problem as lateral distributions. In this case the mean value is shifted to lower values owing to the fact that more deposits at some distance from the shower core are merged into a single cluster.

Higher moments of the energy distribution like the longitudinal and radial kurtoses κ_L and κ_R are relatively robust against changes due to multiple incident particles. They prove especially useful at high energies, where the number of input variables is generally smaller.

Also the covariance between the longitudinal and lateral shower extent, $\mathrm{cov}(L, R)$, is very well described over the full energy range. It becomes broader at higher energies and is a good estimator in all regions. At large covariances, i.e. when the radial and longitudinal dimensions of the shower are both large, the distribution is dominated by hadronic deposits.

Clearly some of the input variables are correlated. The correlations can be computed from equation (6.21) and are taken into account in the training. An example of the full correlation matrix ρ is shown in figure 6.14. It is obtained from the network training in the IF for the energy range 4–30 GeV, corresponding to the input variables shown in figure 6.12. Here z_c denotes the boolean variable indicating if a cluster is close to a z-crack[3] or not. One of the requirements in the training is that z_c and the cluster energy E are not directly correlated to the target value, their correlation to the target being 0. The energy fraction in the third layer f_3 shows the largest correlation to the target, followed by the mean length $\langle L \rangle$. However, they are strongly correlated and contain mostly redundant information. By adding variables related to the radial shower width like the covariance, here denoted by c_{LR}, and the radial kurtosis κ_R additional information is made available. The role of the crack information can be seen on the variables f_2 and f_3. z_c is correlated to f_2 and anti-correlated to f_3. This can be understood by looking at the detector geometry in figure 5.10. The z-crack between the IF and FB2 is located at the front face of the IF. Hence, a larger energy fraction in deeper layers reduces the probability of

---

[3] There are no ϕ-cracks in the IF hence the corresponding ϕ_c did not enter the network training.

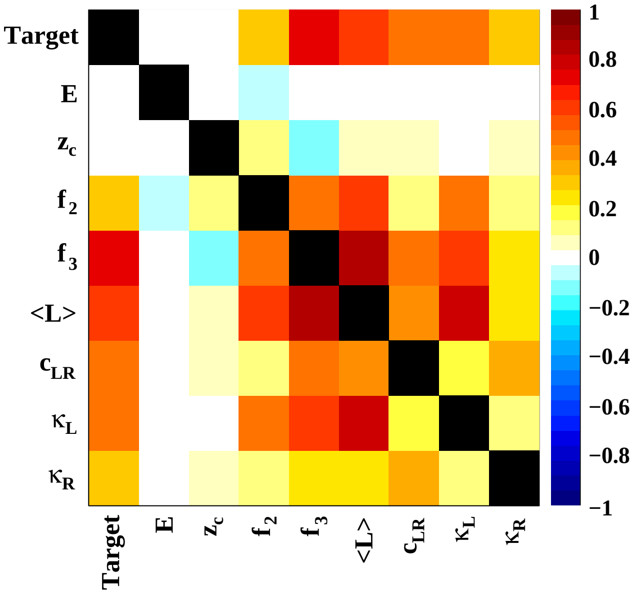


Figure 6.14: Correlation Matrix ρ for the IF network in the energy range 4–30 GeV. Detailed explanation in the text.

the cluster being close to the crack region.
The results from the network training are summarised in figure 6.15. The significance S calculated from equation (6.23) is shown in figure 6.15a. As expected from the correlation matrix, f_3 has the largest significance. The variable with the second largest direct correlation ot the target, $\langle L \rangle$, has a comparable small significance due to its large correlation to f_3. The second most important input variable is c_{LR} followed by the crack-information z_c and the longitudinal and transversal kurtoses. Even though z_c has no direct correlation to the target its significance is comparable to the kurtoses because of its anti-correlation to f_3. The cluster energy E as input variable is rather insignificant in this case, with its significance of $\sim 4\,\sigma$ being close to the threshold.
In figure 6.15b the network output is shown for the training sample. Electromagnetic showers can be found at -1 whereas hadronic showers peak at $+1$. In the shown example the exact value of -1 is not reached because of a diagonalisation procedure which forces the signal purity versus the network output to be distributed along the diagonal after the preprocessing. The hadronic contribution on the left side is expected since high energy hadronic showers are often clustered into several calorimeter clusters. If one of these clusters fully consists of a prominent electromagnetic subshower it is identified to be electromagnetic. The contribution of electromagnetic showers on the hadronic side is one order of magnitude smaller in comparison to the right side. In figure 6.15c the purity as function of the network output is shown. It lies reasonably well on the line defined by equation (6.25).

6.7 Results with the Shower Separation

After the successful training of all 16 neural networks their applicability on isolated single particles as well as on jets has been studied carefully. The results are reviewed

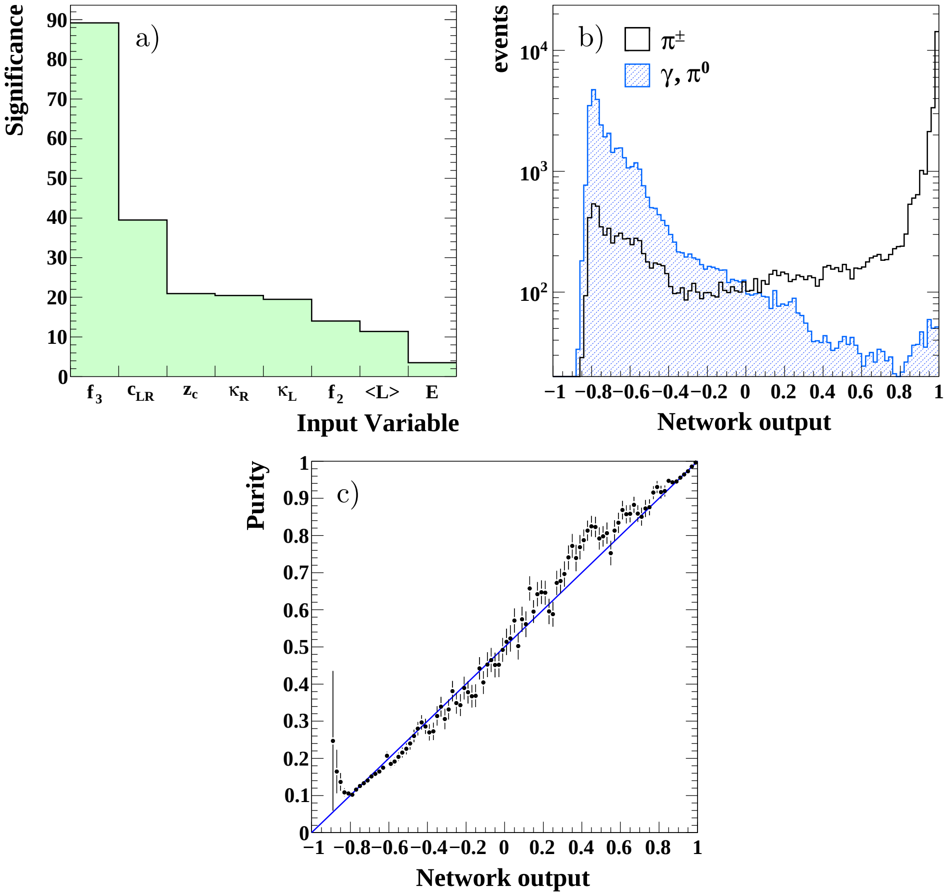


Figure 6.15: Significance S of the input variables for the IF network for energies between 4–30 GeV (a). Output of the trained network (b) and purity versus the network output (c).

in this section.

The output of the neural network is used for the energy reconstruction. If a cluster has a probability of more than 50% to be electromagnetic, i.e. $o < 0$ (see figure 6.15b,c) its energy is taken to be the ideal electromagnetic energy E_{em} (5.40). Else the hadronic energy from the software weighting E_{had} defined in equation (5.41) is used.

It is important to note that the neural networks introduced here do not estimate the intrinsic electromagnetic content of hadronic showers. They rather yield a probability of a cluster originating from an electromagnetic or hadronic particle. Fluctuations in f_{em} happen on much smaller scales than the length scales of calorimeter clusters. Whereas calorimeter clusters can consist of tens of cells, electromagnetic subshowers may be fully contained in a single cell. This means that the software weighting, which takes into account the energy density of single cells, is still crucial for the reconstruction of the hadronic energy.

6.7.1 Single Particles

In addition to the training samples, additional single particle simulations of photons, neutral and charged pions are used to validate the trained networks. In order to verify their general validity also electrons, neutrons, neutral and charged kaons, protons and anti-protons were simulated. The particles were generated over the full calorimeter acceptance with uniform distributions in θ and ϕ in an energy region between 0.5 to 35 GeV. For the forward region, i.e. $\theta < 30^\circ$ particles with energies up to 100 GeV were simulated.

With the cut on $o < 0$ the probability for clusters from single photons and neutral pions to be recognised as electromagnetic is above 90%. This includes electrons and positrons from photon conversions in the inactive detector material. For simulated electrons the probability to recognise them to be electromagnetic drops to about 80% which is not crucial since the contribution of electrons and positrons to the hadronic final state is negligible. This is because the scattered electron and electrons from charm and beauty decays are identified first with a dedicated finder which will be described in section 7.2. The remaining electrons in the hadronic final state originate from fragmentation processes in jets and are very rare. They contribute on average less than 0.1% to the jet's energy. With the same cut clusters from hadronic showers are wrongly identified as electromagnetic only in 10–15% of all cases. This result changes very little for different types of hadronic particles (mesons and baryons). The transition regions at 4 GeV and in the forward region also at 30 GeV are carefully studied. The results neither deteriorate nor are any discontinuities in the network output visible.

The effect of cutting on the neural network output on the reconstructed energies can be studied with the relative difference between the generated and reconstructed energy,

$$\Delta_E = \frac{E_{\text{gen}} - E_{\text{rec}}}{E_{\text{gen}}} \,. \tag{6.29}$$

This quantity is shown for various types of particles in figure 6.16. The results show the success of the developed approach for the separation of electromagnetic and hadronic showers. The two approaches so far available at H1 are efficient for selecting either clusters from electromagnetic showers or from hadronic showers, but they do not succeed to separate them sufficiently. Whereas the selection implemented in Hadroo2 is very efficient for the selection of electromagnetic showers, too many clusters from hadronic showers are misidentified to be electromagnetic. The selection of H1REC correctly identifies most of the clusters from hadronic showers, but it fails to identify clusters from electromagnetic showers.

In the left column of figure 6.16 the distributions of Δ_E are shown for electromagnetic particles where a logarithmic scale is chosen to visualise the tails of the distributions. The distributions obtained with Hadroo2 have small tails owing to the efficient selection of electromagnetic clusters. The distributions of Δ_E obtained with the cuts of H1REC exhibit large tails towards small values of Δ_E. This is comparable to what was shown in figure 6.7 and is a result of the weighting algorithm applied to purely electromagnetic showers. The results using the neural network follow closely the results obtained with Hadroo2, showing an equally high efficiency for the selection of electromagnetic clusters. For electrons this efficiency is slightly smaller

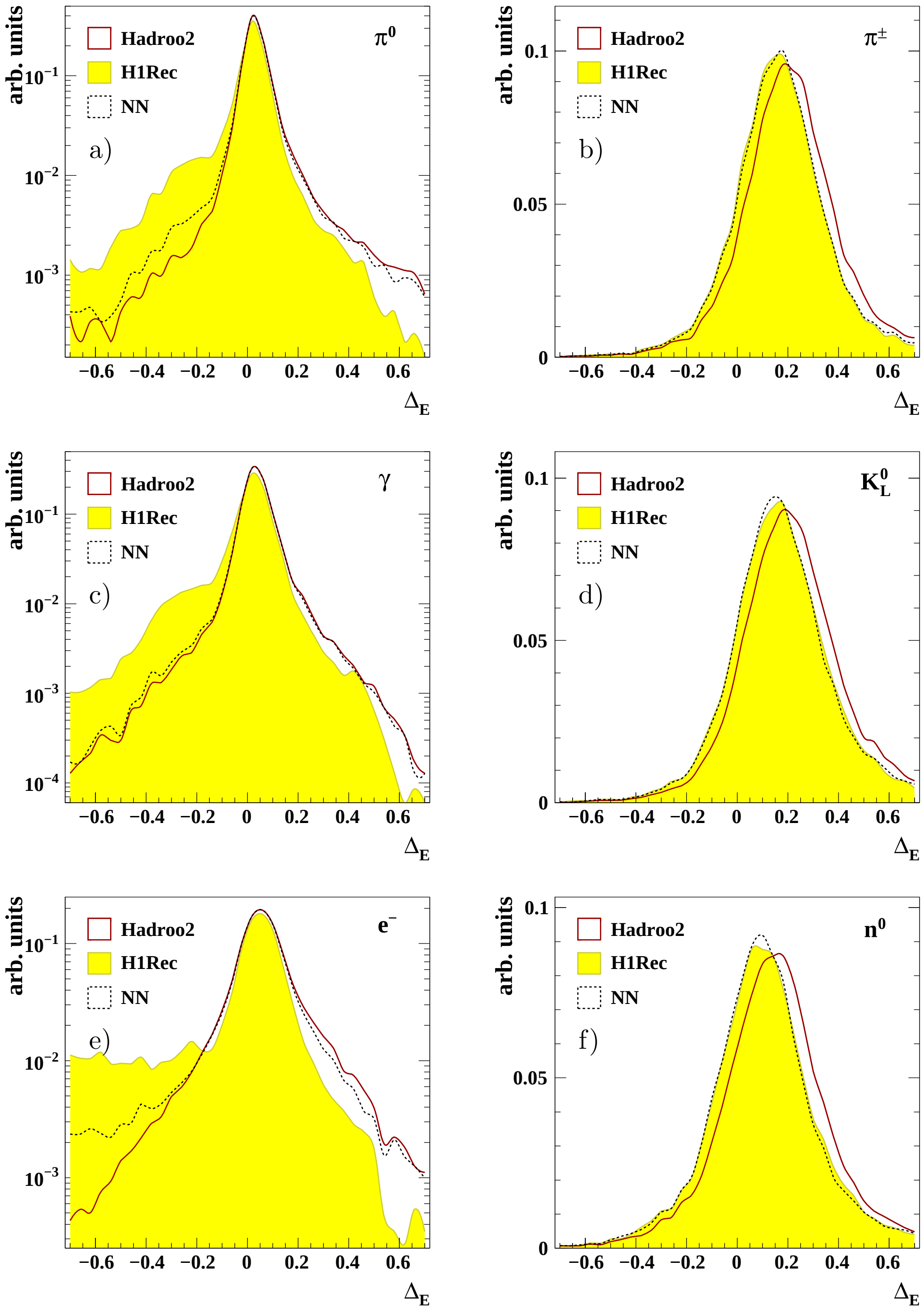


Figure 6.16: Difference between generated and reconstructed energies for single particles in the energy range between 0.5 and 30 GeV in the FB1. Results with the algorithms implemented in H1REC and Hadroo2 are compared to results using the neural network. In the left column results from particles producing electromagnetic showers are shown: neutral pions (a), photons (c) and electrons (e). On the right results from hadronic particles are are shown: charged pions (b), neutral kaons (d) and neutrons (f).

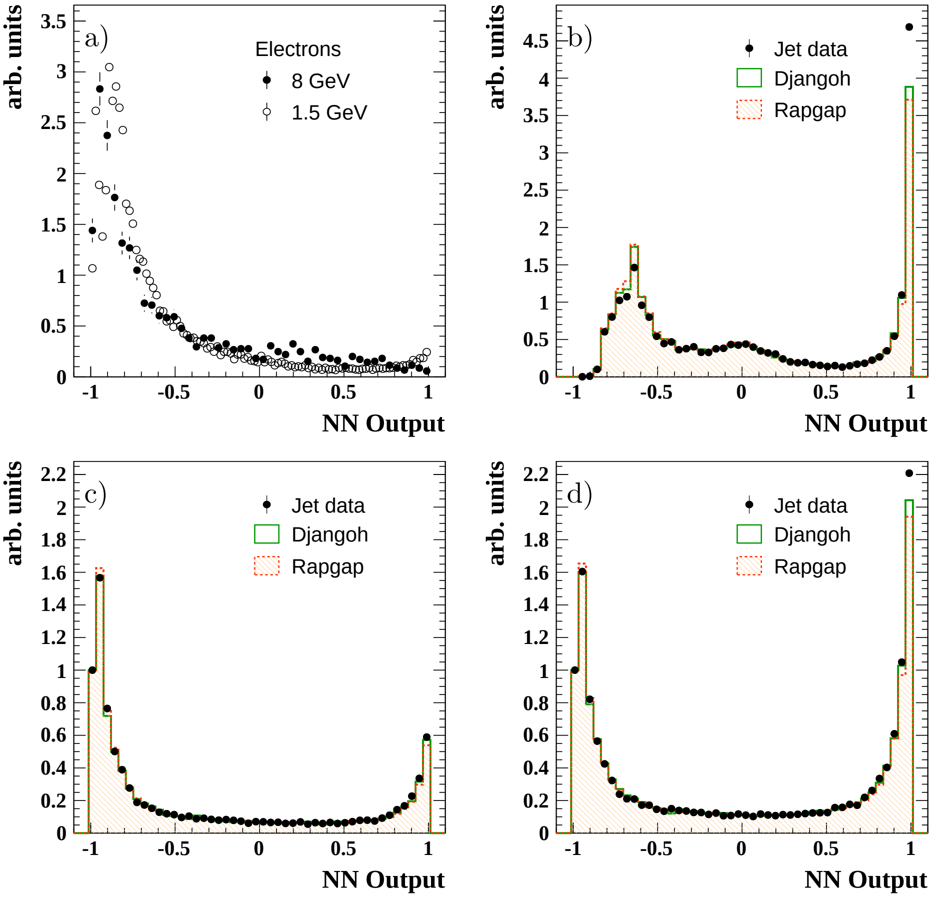


Figure 6.17: Output of the neural network for identified electrons selected with two different finders in the CB2 (a). Neural network output for clusters in jets in the IF (b) and CB2 (c) after the full reconstruction. Also shown is the neural network output of jet clusters in the CB2 without using track information (d).

than for single photons and π^0s which is apparent in the slightly larger tail in figure 6.16e. However, this tail is nearly one order of magnitude smaller than in the H1Rec case.

For hadrons (right column of figure 6.16) the selection of H1Rec results in larger reconstructed energies compared to the selection of Hadroo2. This is due to too many clusters taken on the electromagnetic scale in the Hadroo2 case. The neural network's efficiency of identifying clusters to orginate from hadronic showers is as high as H1Rec's, resulting in larger reconstructed energies compared to Hadroo2. At the same time the widths of the distributions are not altered, showing also the success of the software weighting.

The fact that the distributions of Δ_E are not centred around 0, particularly for hadrons, means that there is the need for an additional calibration. In the course of this work a method to calibrate the hadronic final state has been developed, utilising the information from the neural network. It will be introduced in section 8.2. Here it should be remarked that the larger reconstructed energies with the new shower separation reduce the magnitude of the calibration constants with respect to former calibration methods.

To test the neural network on single particles in data, it was applied to identified electrons. The neural network output is shown in figure 6.17a for two electron energies. Electrons with energies larger than 8 GeV were found by the standard electron finder [Bru98] as implemented in the H1 Object Oriented Analysis Software (H1OO). Electrons with energies around 1.5 GeV are obtained from $J/\Psi \rightarrow e^+e^-$ decays. They are selected with the help of a neural network using tracking and calorimeter information [Sau09]. In order to reduce background the invariant mass of the e^+e^- pair was required to be $2.9 < M_{e^+e^-} < 3.2\,\mathrm{GeV}$. The neural network behaves as expected on data, with the distributions peaking close to -1. The tails towards positive values of o are due to background in the electron samples which can never be fully excluded. An efficiency of about 80% for identifying electron showers is obtained, which is in good agreement with the results from simulated electrons.

6.7.2 Jets

The jet sample used to verify the neural network output is identical to the calibration sample introduced in section 8.2. Jets with a transverse momentum of 7 GeV or higher are selected within the acceptance of the calorimeter $7^\circ < \theta_{\mathrm{jet}} < 155^\circ$.

In figures 6.17b-d the neural network output is shown for clusters belonging to jets in data and MC simulations. The data are very well described by the simulations, demonstrating the applicability of the neural network for data. A well-modelled sharing of the electromagnetic and hadronic energy in jets can also be observed. The neural network output in the IF (figure 6.17b) shows an enhancement in the region of $o \approx 0$. This is due to jet clusters originating from more than one incident particle. Deposits from overlapping electromagnetic and hadronic showers are clustered together and cannot be separated. In figure 6.17c the output for clusters in the CB2 is shown. The suppressed peak on the hadronic side (+1) is due to the energy flow algorithm used to reconstruct the hadronic final state. Calorimeter deposits which can be assigned to tracks are removed from the hadronic final state. These tracks originate from charged hadrons, resulting in a decrease of hadronic clusters. In figure 6.17d the neural network output in the CB2 is shown with the track finding switched off. The peak on the hadronic side is recovered, also showing the good selection efficiency of the neural network.

For a simulated jet the generated electromagnetic fraction is given by

$$f_{\mathrm{em,jet}}^{\mathrm{gen}} = \frac{1}{E_{\mathrm{jet}}^{\mathrm{gen}}} \left(\sum_{\mathrm{photons}} E_\gamma + \sum_{\mathrm{electrons}} E_{\mathrm{e}} \right) , \tag{6.30}$$

where the sums run over photons, electrons and positrons contained in the jet. The reconstructed electromagnetic fraction can be calculated from

$$f_{\mathrm{em,jet}}^{\mathrm{rec}} = \frac{1}{E_{\mathrm{jet}}^{\mathrm{rec}}} \left(\sum_{\mathrm{clusters}} P_{\mathrm{em}}^i\, E_{\mathrm{cls}}^i + \sum_{\mathrm{em}} E_{\mathrm{em}}^j \right) , \tag{6.31}$$

where the first sum runs over all jet clusters and the second one runs over all particles found by the electron finder. P_{em}^i is the the probability obtained from the neural network of cluster i to originate from an electromagnetically showering particle,

$$P_{\mathrm{em}} = \frac{1-o}{2} . \tag{6.32}$$

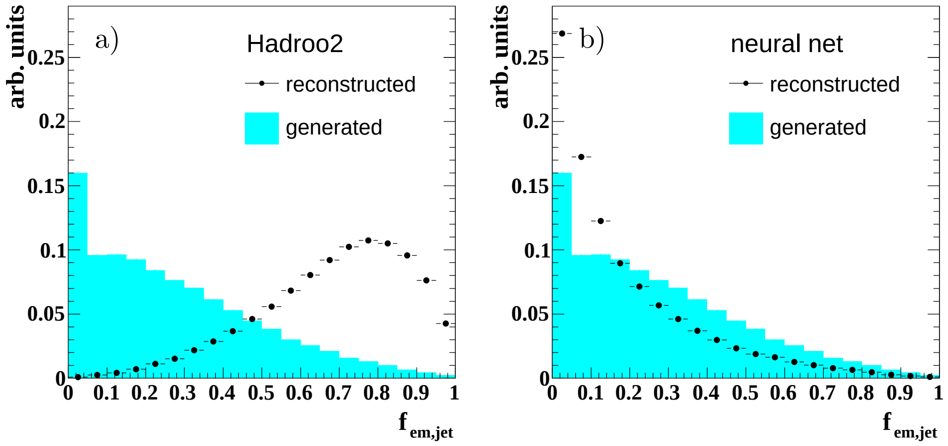


Figure 6.18: Generated and reconstructed electromagnetic fractions of jets simulated inside the calorimeter acceptance. The standard reconstruction Hadroo2 (a) is compared to results obtained with the neural networks (b).

The jet's generated electromagnetic fraction $f^{\rm gen}_{\rm em,jet}$ (6.30) represents the jet's fractional energy carried by photons and electrons after fragmentation and hadronisation processes. The fractional energy carried by electromagnetic particles after the full shower development in the calorimeter is different and cannot be accessed. No attempt to take the expected electromagnetic fraction of the hadronic shower into account is made. Since in the calculation of the reconstructed electromagnetic fraction $f^{\rm rec}_{\rm em,jet}$ (6.31) the electromagnetic fraction of hadronic clusters does not contribute, the same simplification is made on generated and reconstructed level. This means that on both levels the electromagnetic fraction will be underestimated by approximately the same amount.

In figure 6.18 the generated and reconstructed electromagnetic fraction for jets within $7^\circ < \theta_{\rm jet} < 155^\circ$ is shown for the standard reconstruction and using the neural networks. To calculate $f^{\rm rec}_{\rm em,jet}$ for the standard reconstruction $P_{\rm em}$ is taken to be 1 for electromagnetic clusters and 0 for hadronic ones. As can be observed, Hadroo2 assigns too much energy on the electromagnetic scale, shifting the peak of the $f^{\rm rec}_{\rm em,jet}$ distribution to ~ 0.8 (figure 6.18a). With the neural network the correspondence between the generated and reconstructed electromagnetic fraction is much better with a slight tendency to overestimate the hadronic component (figure 6.18b).

The correlation between $f^{\rm rec}_{\rm em,jet}$ and $f^{\rm gen}_{\rm em,jet}$ is shown in figure 6.19 for different calorimeter regions with the standard reconstruction (left column) and the neural networks (right column). For Hadroo2 the electromagnetic component is overestimated everywhere, whereas when utilising the neural networks the distributions are close to the diagonal. A slight tendency towards a too large hadronic component can be noted, which is more pronounced in the IF and FBs than in the CBs.

With the new algorithm a reliable estimate of the jet's electromagnetic component is possible, which is of major importance for the calibration described in section 8.2.

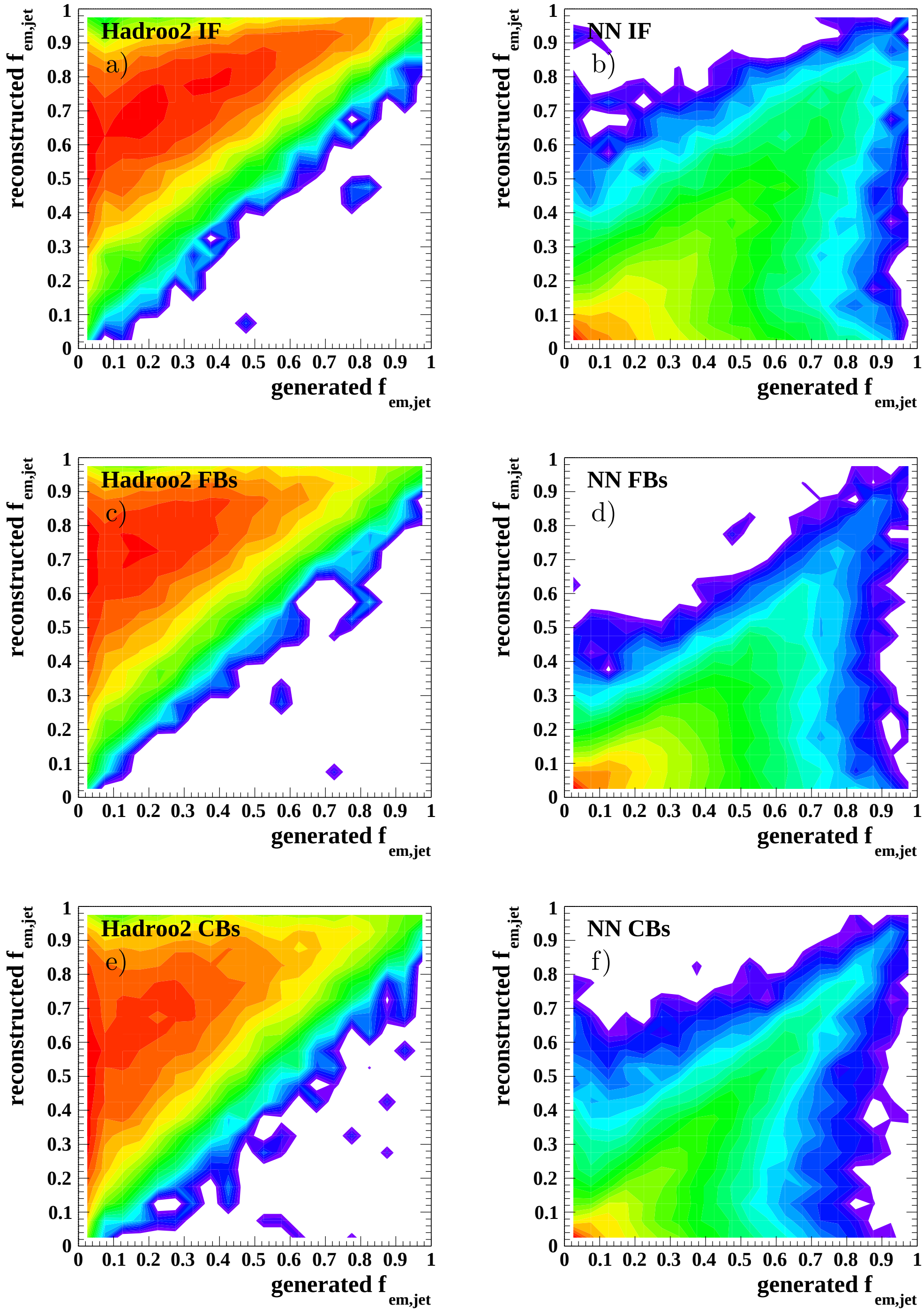


Figure 6.19: Generated and reconstructed electromagnetic fractions of jets in different calorimeter regions. Shown are results with the standard reconstruction Hadroo2 and the reconstruction using the neural network for the IF (a and b), the FBs (c and d) and the CBs (e and f).

Chapter 7

Reconstruction

In this chapter an overview of the reconstruction algorithms used in H1 is given. These include track and vertex finding algorithms, sub-detector alignment strategies, the treatment of the scattered electron and the reconstruction of the hadronic final state. In the last years an effort has been made by the H1 collaboration to improve on all aspects of the reconstruction. Improvements in the track and vertex finding and the alignment of the tracking detectors have been implemented in the reconstruction software H1REC, and the resulting output is known as DST7. In the object-oriented analysis software H1OO the alignment of the LAr calorimeter with respect to the CTD, the calibration of the electron and the reconstruction of the hadronic final state have been improved. The hadronic final state measurement gained in precision through the developments discussed in the previous chapter and a newly developed calibration technique which will be introduced in the following. The corresponding software version used in this work is 4.0.11.

The emphasis in this chapter will be on achieved precisions and improvements with respect to reconstruction and software versions used in previous analyses.

7.1 Track and Vertex Finding

The idealised parameterisation of a charged particle's trajectory in a homogenous magnetic field with $\boldsymbol{B} = (0, 0, B_z)$ is a helix. The parameterisation has 5 free parameters: the curvature κ which is the signed inverse radius; the distance d_{ca}, the angle ϕ_0 and the z-value z_0 at the point of closest approach to the z-axis; the slope $\tan\lambda = \cot\theta$. The helix is a circle in the $r\phi$-plane, which is given by

$$\frac{1}{2}\kappa(x_i^2 + y_i^2 + d_{\mathrm{ca}}^2) - (1 + \kappa d_{\mathrm{ca}})(x_i \sin\phi_0 - y_i \cos\phi_0) + d_{\mathrm{ca}} = 0\,, \tag{7.1}$$

and a straight line in the sz-plane

$$z_i = z_0 + (\tan\lambda)s_i \tag{7.2}$$

where s_i is the circle's arc-length at point (x_i, y_i). The track's curvature is related to the particle's transverse momentum via

$$p_{\mathrm{T}} = \frac{0.3 \cdot B_z}{\kappa}[\mathrm{GeV/c}] \tag{7.3}$$

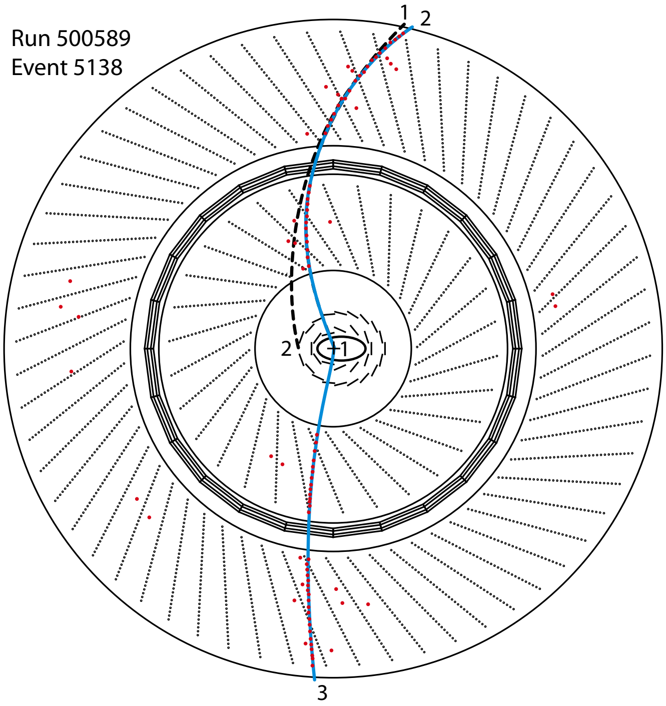


Figure 7.1: Example of a broken-helix track used in the reconstruction version DST7. Shown is the $r\phi$ view of the CTD. Hits in the CJC1 anc CJC2 are illustrated as red circles, tracks obtained with the broken-helix fit are shown as solid blue lines. Whereas the high momentum track 3 stays unchanged, the low momentum track "1+2" gets a significant correction compared to the single helix fits 1 and 2, illustrated as dashed lines. [Pit09].

which is given in units of GeV/c if B_z is given in Tesla and κ in m^{-1}. The reconstruction versions used in previous related works [Nik07, Gou08, H110c] assumed a single helix for each track. However, various effects can result in a deviation from the ideal helix trajectory. The most prominent ones are multiple scattering, magnetic field inhomogenities and continuous energy loss along the particle's trajectory. These effects can be taken into account by a track-fit algorithm based on broken lines [Blo06a]. In H1 a similar approach is used which approximates the particle's trajectory with two helixes and allows for a discontinuity at the connection point.

The track reconstruction in H1 starts with combining hits in the CJC to obtain seeds for the track finding. Approximate track parameters are then fitted with the helix parameterisation given in equations (7.1) and (7.2) by successively adding hits which are consistent with the track assumption. After the track reconstruction with the CJC additional information from the COZ, CIP and CST is added, and the track fit is repeated. Finally, the track parameters obtained are used as starting values in a broken-helix fit to acquire corrections to the track parameters due to multiple scattering and other perturbations. In this fit a discontinuity of the scattering angle and a track offset between the CJC1 and CJC2 is allowed due to the amount of dead material between them[1]. An example of a broken helix fit is shown in figure 7.1.

A bias in the measurement of the polar angle θ of tracks in previous reconstruction

---

[1] The COP and COZ Rohacell structures with their Aluminum/Kapton coatings are treated as thick scatters, also see figure 4.5

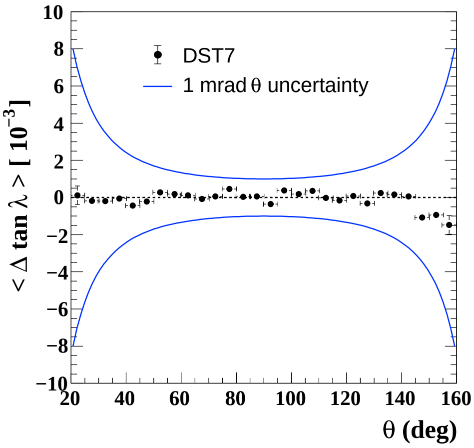


Figure 7.2: Comparison of the slope parameter tanλ measured with the CJC and CST after the application of the z-correction to CJC hits. The effect of a systematic uncertainty of 1 mrad on the track's θ measurement is shown by the blue lines [Pit09].

versions could be traced back to a bias of the z measurement of CJC hits. The reason for it were additional hits on the same wire from earlier bunch crossings which were not taken into account in the pulse subtraction. A correction procedure has been developed which takes previous and next pulses on the same wire into account [Pit09]. With the correction applied, the track's slope $\tan\lambda$ measured with the CJC agrees within 0.1% with the measurement from the CST (see figure 7.2). An uncertainty of 1 mrad on the measurement of the track's polar angle θ is assigned, which covers the uncertainty coming from the measurement of $\tan\lambda$ in the central region.

Before using the CJC tracks for the vertex determination, an algorithm to identify nuclear interactions in the detector material and in-flight decays is employed. It uses pairwise track intersects as seeds. Tracks identified to originate from a nuclear interaction are excluded from the averaging procedure employed to determine the position of the primary vertex. Since nuclear interactions can have large charged particle multiplicities, they have been reported to spoil the vertex reconstruction in rare cases [Nik07]. This effect is considerably reduced with the nuclear interaction finder.

In principle it is possible to deduce the primary event vertex from the z_0 parameters of the reconstructed tracks. However, a more precise measurement can be obtained by using the event timing t_0 and information from the silicon track detectors CST, BST and FST. The primary event vertex is first reconstructed in the $r\phi$ plane where it can vary only within a few 100 µm from the xy run vertex due to the small radial beam dimensions. Then from all CJC and silicon tracks with acceptable d_{ca} and comparable timing a weighted average is built to obtain the z-coordinate of the primary vertex [Blo04].

Finally, all tracks originating from the the primary vertex are refitted using the vertex position as a constraint. A similar procedure is employed for all tracks originating from secondary vertices from particle decays, photon conversions and nuclear interactions. A similar strategy as for CJC tracks is used for FTD tracks. In the region where the acceptance of the CJC and FTD overlaps a combined track fit to

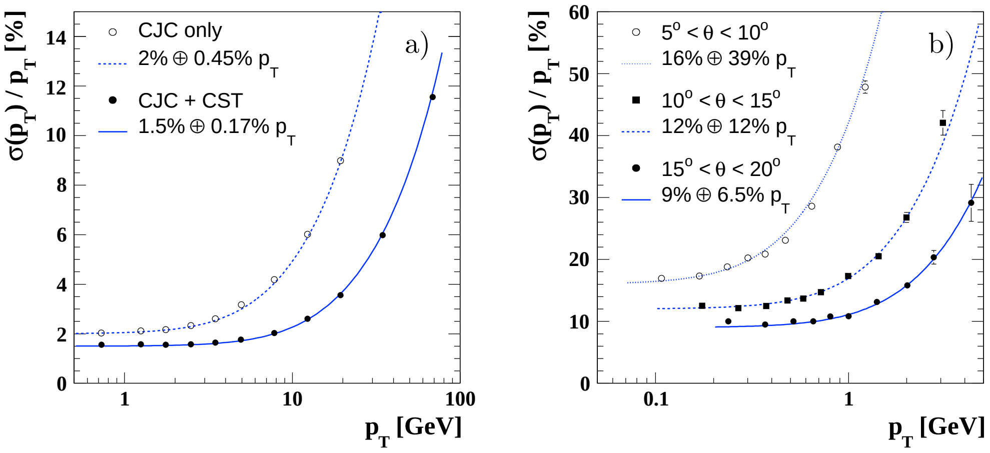


Figure 7.3: a) Momentum resolution of CJC tracks measured with cosmic muons with and without using the CST. b) Momentum resolution of vertex fitted FTD tracks depending on the polar angle, obtained from simulated data (both from [Pit09]).

hits from both detectors is tried. If the combined fit fails, standalone FTD tracks are fitted to the primary vertex. In this case a matching procedure is employed to avoid double-counting.

The track and vertex finding described above combines information from various sub-detectors. Essential for a precise track reconstruction is an accurate knowledge of their respective alignments. Alignment parameters for the CJC, silicon track detectors and FTD are obtained by using cosmic and beam-halo muons and tracks from *ep* interactions. Since not only the positions of the track detectors but also their components like CJC cells, anode wires and silicon sensor pads can be misaligned, a large number of shifts and rotations has to be determined. In order to achieve this and take correlations among the parameters into account they are fitted in a global procedure [Blo06b]. The alignment together with improvements of the electric and magnetic field description results in an improvement of the track parameter resolutions by an overall factor of two [Kle07].

The momentum resolution for CJC tracks can be measured with cosmic muons. In figure 7.3a such a measurement is shown using CJC hits alone and including hits from the CST. The constraints from CST hits have a similar impact on the resolution of CJC tracks as the vertex constraint in *ep* data. The achieved resolution is $\sigma(p_\mathrm{T})/p_\mathrm{T} = 1.5\% \oplus 0.17\% \cdot p_\mathrm{T}/\mathrm{GeV}$, limited by multiple scattering at low momenta and by small track curvatures at high momenta. Additionally, the resolution of the angle ϕ_0 at the interaction point σ_{ϕ_0} improves significantly by adding CST hits, from 1 mrad to 0.25 mrad at high transverse momenta. The distance of closest approach to the beam in $r\phi$ can be measured with an accuracy of $\sigma_\mathrm{dca} = 0.22\,\mathrm{mm} \oplus 0.52\,\mathrm{mm} \cdot \mathrm{GeV}/p_\mathrm{T}$. The track finding efficiency for CJC tracks is 99% for muons and about 94% for hadrons at high p_T, which is limited by nuclear interactions. Below transverse momenta of 0.5 GeV the efficiency drops steeply for kaons and protons, for pions it stays above 80% down to $p_\mathrm{T} \sim 0.1$ GeV.

The momentum resolution for FTD tracks is shown in figure 7.3b. It depends on the track length and consequently on the polar angle θ of the track. Below a track momentum of 3 GeV it is limited by multiple scattering. The best resolution

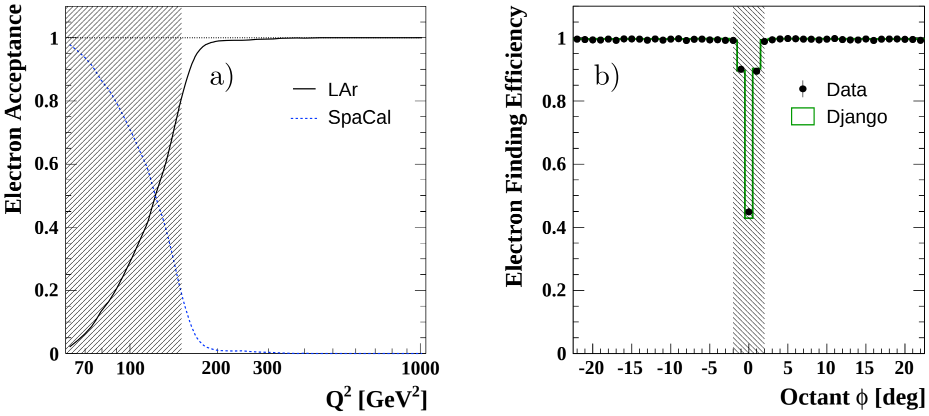


Figure 7.4: a) Geometrical acceptance of the SpaCal and LAr calorimeter for the scattered electron as a function of virtuality. The shaded region represents the excluded phase space in this analysis, the inelasticity is required to be $0.2 < y < 0.7$. b) Electron finding efficiency as function of ϕ, folded into one octant. The shaded region corresponds to the ϕ-cracks and is excluded.

is $\sigma(p_\mathrm{T})/p_\mathrm{T} = 9\% \oplus 6.5\% \cdot p_\mathrm{T}/\mathrm{GeV}$ in a region between 15 and 20 degrees. Even though the track finding efficiency is between 50–60% in single supermodules, tracks included in the hadronic final state have to pass a number of quality requirements which include a minimal track length and a minimal number of hits. After their application the track finding efficiency is only about 10% such that FTD tracks do not contribute largely to the hadronic final state.

7.2 Electron Identification

Depending on the polar angle θ_e of the scattered electron it impacts either at the SpaCal or the LAr calorimeter. In figure 7.4a the acceptance of both calorimeters as function of the virtuality Q^2 is shown. The LAr calorimeter acceptance rises from close to 0 at $Q^2 = 60\,\mathrm{GeV}^2$ to 1 at $Q^2 = 200\,\mathrm{GeV}^2$. The acceptance does not exhibit a sharp rise but is smeared out because of the distribution of the vertex z-position. The region excluded in this analysis is shown as a shaded area. In the considered phase space with $0.2 < y < 0.7$ and $Q^2 > 150\,\mathrm{GeV}^2$, the geometrical acceptance of the LAr calorimeter is always greater than 0.8. Hence it is sufficient to use only the LAr calorimeter for the reconstruction of the scattered electron.

For the electron identification the QESCAT algorithm is used [Bru98]. It distinguishes electrons from hadrons based on cluster estimators. The suppression of background from photons (mainly from π^0 decays) is achieved with the requirement of a track pointing from the primary vertex to the cluster. In a first step all clusters in the EMC with an energy above 2 GeV and a transverse momentum larger than 1 GeV are used as seeds. The axis of a cone is defined through the primary vertex and the cluster's centre of gravity. All cells within an opening angle of 7.5°, starting at a distance one metre from the calorimeter surface, are merged to the original seeds (see figure 7.5). Then requirements on the following estimators (cf. 6.1) are applied:

- The energy fraction in the EMC, f_EMC, should be larger than $0.94{+}0.05\cos 2\theta_\mathrm{e}$,

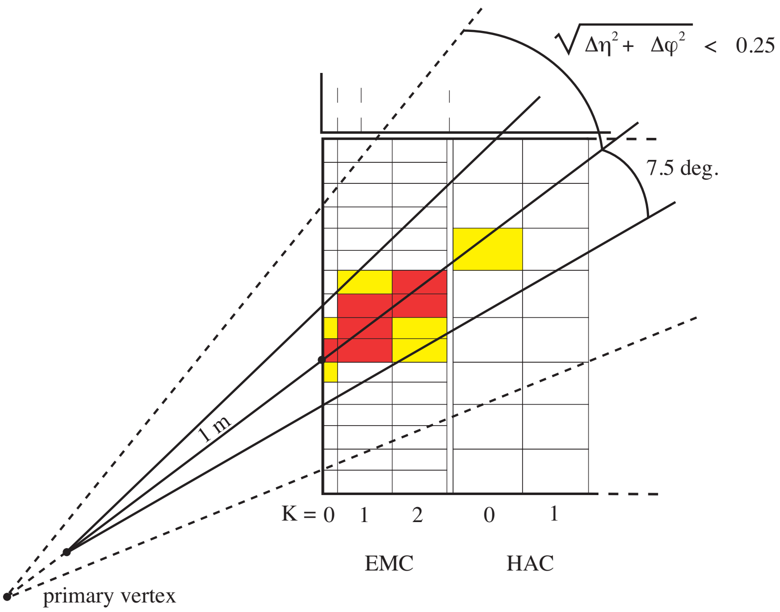


Figure 7.5: Schematic view of the two cones used by the electron finder. All cells within a cone with an opening angle of 7.5°, starting at a distance of one meter from the calorimeter surface, are merged to the original seed (solid lines). The cone used for the isolation criterium is illustrated with dashed lines.

where θ_e denotes the polar angle of the electron candidate.

- The energy fraction in the hottest N cells h_N, where N varies between 4 to 12 depending on the calorimeter wheel, should be larger than 0.5 to 0.8, depending on θ_e.
- The lateral standard deviation σ_R has to be smaller than a limit which depends strongly on the calorimeter wheel.
- The number of cells contributing to the cluster has to be greater than three.
- The electron candidate has to be isolated in order to suppress electrons originating from jet fragmentation processes. The energy fraction of the electron candidate in a cone with radius 0.25 in (η, ϕ) has to be larger than 0.98 (0.95, if 0.3 GeV in the cone originate from the HAC).

All electron candidates fulfilling the above criteria are matched with tracks to distinguish between clusters originating from electrons or photons. The distance of closest approach between a track and the cluster is calculated by extrapolating the track with the helix parametrisation from equations (7.1) and (7.2) to the calorimeter surface. The track with the smallest value of d_{ca} is stored. The selection of scattered electron candidates with the help of associated tracks is discussed in section 9.4.
With the above requirements on electron clusters the rejection efficiency for hadrons is above 99%. In the ϕ-crack regions and the z-crack between the CB2 and CB3 the electron finding efficiency is only about 40% (see figure 7.4b). These regions are excluded from the measurement.
After the application of these cuts the electron finding efficiency is established with an independent track-based electron finder and shown in figure 7.6 [Shu10]. It is

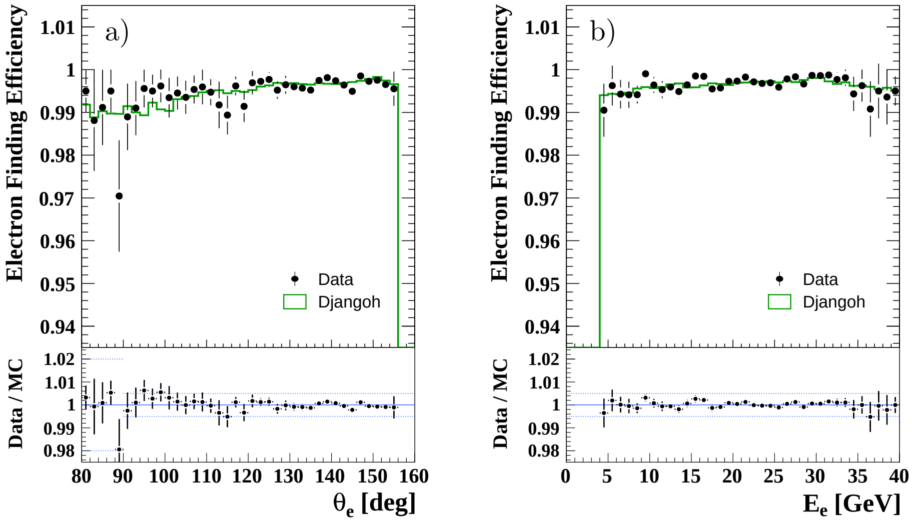


Figure 7.6: Electron finding efficiency as function of the electron's polar angle θ_e (a) and its energy E_e (b). The uncertainty of 0.5% is illustrated by the dotted lines. In the region $z_{imp} > 0$, corresponding approximately to $\theta_e < 90°$, the assigned uncertainty is 2%.

found to be 99.5%, largely independent of the electron energy and polar angle. It is well described by the simulation and no correction for it is applied. The uncertainty on the electron finding efficiency is 0.5% in the central region ($z_{imp} < 0$) and 2% in the forward region ($z_{imp} > 0$) [Nik07].

7.3 Calorimeter Alignment

At the HERA upgrade and during shutdown periods the CTD was removed from the H1 detector for maintenance. Due to mechanical imprecisions at the re-insertion of the CTD, the CTD may have been shifted with respect to its nominal position. The global alignment of the track detectors (section 7.1) is used to determine the shift of the CTD with respect to the origin of the H1 coordinate system. In order to correct for a possible misalignment of the CTD with respect to the LAr calorimeter, the relative position of the calorimeter is measured with tracks associated with scattered electrons in a NC DIS sample. The tracks are extrapolated to the surface of the calorimeter taking the magnetic field into account. The impact positions calculated from these extrapolated tracks are then compared to the ones obtained from the electron clusters. In this sample electron tracks are required to have COZ and CST hits in order to yield a precise measurement of the z-coordinate in addition to a superior resolution in ϕ with respect to clusters.
To correct for the misalignment of the calorimeter, its position is allowed to be shifted along the x-, y- and z-directions with simultaneous rotations around these axes. In the z-direction the calorimeter wheels BBE, CB1,2 and 3 and FBs are allowed to shift relative to each other. Prior to the determination of the alignment parameters a correction for the calorimeter shrinkage due to the cooling down to its working temperature of 72 K is applied. The resulting 10 parameters describing the rotations and translations are obtained in a global minimisation procedure [Tra10]. The obtained parameters have been tested by applying the corrections to the cluster

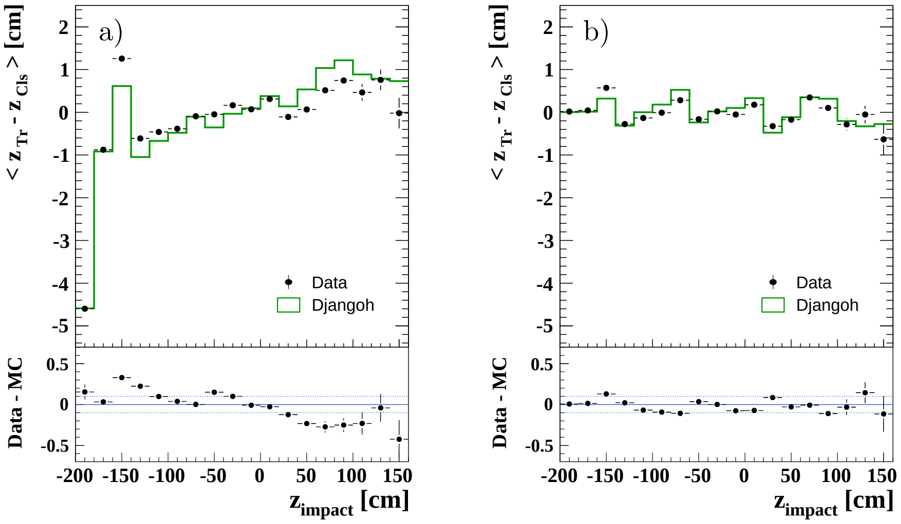


Figure 7.7: Difference between the scattered electron's z-position measured with the track and cluster before (a) and after the calorimeter alignment (b). At the bottom the difference between data and simulated events is shown. The dotted lines illustrate a deviation of 0.1 cm between data and MC.

positions and re-measuring the calorimeter's alignment. The resulting parameters are compatible with no further shifts and rotations. Additionally, the procedure has been tested on simulated data. In the detector simulation no shift of the calorimeter with respect to the CTD is applied. The measured alignment parameters are compatible with zero. Also a simulated arbitrary shift of the calorimeter could be reproduced by the alignment procedure [Tra10].
In figure 7.7 the difference between the impact z-position measured with electron clusters and tracks is shown before (a) and after the calorimeter alignment (b). Before the alignment the difference amounts up to 4.5 cm in the BBE. After the alignment the cluster and track measurements agree within 0.6 cm. Small fluctuations are well described by the simulation. The agreement between data and MC improves considerably and is within 0.1 cm, which translates into a difference of less than 1 mrad for the θ measurement.

7.4 Hadronic Reconstruction

For the reconstruction of the hadronic final state an energy flow algorithm is employed [Pee03, PP$^+$05]. The algorithm known as Hadroo2 (Hadronic Reconstruction in H1OO 2) combines information from tracks with calorimetric energy deposits. The momenta of charged particles traversing the track detectors can be either measured via their track's curvature or through the energy deposited in the calorimeter. Depending on the momentum of the incident particle the track or calorimetric measurement has a better resolution. The estimated resolution is used to choose between the track or the calorimetric measurement and the redundant measurement is removed from the Hadronic Final State (HFS) to avoid double counting. A refined treatment is necessary to prevent clusters originating from neutral particles being associated to tracks in dense environments. The output of the algorithm are HFS objects which are subsequently calibrated (section 8.2) and used for the jet finding and reconstruction of the kinematics. Before the algorithm itself is described, the selection of tracks and clusters used as input is reviewed.

7.4.1 Tracks

Since four-vectors of particle momenta can be assigned only to vertex-fitted tracks, solely they enter the hadronic final state. In the case of a track matching with the primary and a secondary vertex, the primary vertex is prioritised. Three possible types of tracks are available: CJC tracks (central tracks) which may be combined with CST and COZ hits with a geometrical acceptance of $20° \leq \theta \leq 160°$, tracks from a combination of hits in the CJC and FTD (combined tracks) with an angular coverage of $6° \leq \theta \leq 40°$, standalone FTD tracks (forward tracks) in the range $6° \leq \theta \leq 25°$. The selection criteria for them are based on studies by Lee West [Wes00]. Central and combined tracks are required to have a minimum transverse momentum of 0.12 GeV with additional requirements on the track starting radius and vertex-matching. Forward tracks have to have a momentum of at least 0.5 GeV with a minimum number of hits in subsequent radial and planar chambers. A detailed list of all cuts applied is given elsewhere [PP+05]. For tracks in the overlap region of the CJC and FTD, which fulfil more than one selection criterion, the order of preference is central, combined and forward.

7.4.2 Clusters

Calorimeter clusters are treated in two steps before they are compared to selected tracks. In the first step several noise suppression algorithms are applied to remove electronic and background noise. In the second step the cluster four-vectors determining the kinematics are obtained.

Noise Suppression

First, all clusters consisting of only one cell and clusters with an energy of less than 0.2 GeV on the electromagnetic scale are removed. Then, four different noise suppression algorithms are applied [PP+05]. The FSCLUS and NEWSUP algorithms remove isolated low-energetic clusters originating from electronic noise. They compare energy deposits inside a sphere of radius 40 cm around low-energetic clusters (20 cm in the IF) with a threshold energy. Clusters with no or only little energy in their vicinity are flagged as noise and suppressed. The HALOID algorithm removes energy deposits from halo muons overlaying *ep* events by selecting clusters in narrow cylinders parallel to the z-axis. Clusters originating from cosmic muons and coherent noise are suppressed by the HNOISE algorithm. It suppresses energy deposits in the HAC with neither a signal in the EMC nor a track pointing into the cluster direction.

The performance of the noise suppression algorithms has been thoroughly studied [CZ99,Pee03,PP+05]. They significantly improve the calorimetric measurement with very small inefficiencies of 0.2–1% in not recognising low energy signal clusters.

Cluster Kinematics

The barycenter $\boldsymbol{b}$ of each cluster is calculated as sum over all contributing cells, $\boldsymbol{b} = \sum_i w_i \boldsymbol{p}_i$, where the weight w_i of cell i is chosen to be $\sqrt{E_{em,i}}$. The position $\boldsymbol{p}_i$ of each cell is shifted with respect to its nominal value through the alignment procedure described in section 7.3. For the reconstruction of the kinematics of the incident particle a four-vector with energy E is assigned to the calorimetric deposit. The energy is chosen to be either E_{em} (5.40) or E_{had} (5.41), depending on the cluster's probability to originate from an electromagnetic particle, P_{em} (6.32). The azimuthal and polar angle of the four-vector are calculated from the shower axis $\boldsymbol{a} = \boldsymbol{b} - \boldsymbol{v}$, where $\boldsymbol{v}$ denotes the position of the primary vertex (cf. figure 6.2). A small correction to them is applied due to the beam tilt, which is an inclination of the beams with respect to the H1 coordinate system. This tilt is determined in each run and is smaller than 1 mrad.

The use of the electromagnetic probability P_{em} from chapter 6 to choose between E_{em} and E_{had} leads to a considerable improvement of the absolute energy scale and the resolution with respect to the standard cut of $f_{01} > 0.5$, which is shown in section 8.2.

7.4.3 The Hadroo2 Energy Flow Algorithm

The selected tracks and noise-suppressed and corrected clusters are the input to the energy flow algorithm. For a decision which measurement is preferred the error from the track fit is compared to the expected resolution of the calorimetric measurement. Each track is assumed to originate from a pion and its energy is set to

$$E^2_{\text{track}} = P^2_{\text{track}} + m^2_\pi = p^2_{\text{T}} / \sin^2\theta + m^2_\pi \,. \tag{7.4}$$

Neglecting correlations between p_{T} and θ, the error on the track's energy measurement is

$$\frac{\sigma(E)_{\text{track}}}{E_{\text{track}}} = \frac{1}{E_{\text{track}}} \sqrt{\frac{p^2_{\text{T}}}{\sin^4\theta} \cos^2\theta\, \sigma^2_\theta + \frac{\sigma^2_{p_{\text{T}}}}{\sin^2\theta}} \,, \tag{7.5}$$

where σ_θ and $\sigma_{p_{\text{T}}}$ are the errors obtained from the track fit. In figure 7.8a the errors on the energy measurement from central tracks calculated with equation (7.5) are compared to the hadronic resolution of the LAr calorimeter from section 5.4.2. The errors on the tracks are well described by the simulation. The crossover point where the calorimetric measurement is to be preferred is at ∼25 GeV. Hence, the track is used for the creation of the HFS object if

$$\left(\frac{\sigma(E)}{E}\right)_{\text{track}} < \left(\frac{\sigma(E)}{E}\right)_{\text{LAr}} \,. \tag{7.6}$$

For the resolution of the LAr calorimeter an approximate expression is used

$$\left(\frac{\sigma(E)}{E}\right)_{\text{LAr}} = \frac{\sigma(E_{\text{track}})_{LAr}}{E_{\text{track}}} = \frac{0.5}{\sqrt{E_{\text{track}}}} \,. \tag{7.7}$$

The track's energy is used for the evaluation of the expected calorimeter resolution because the number of induced clusters by the incident particle is a priori not known.

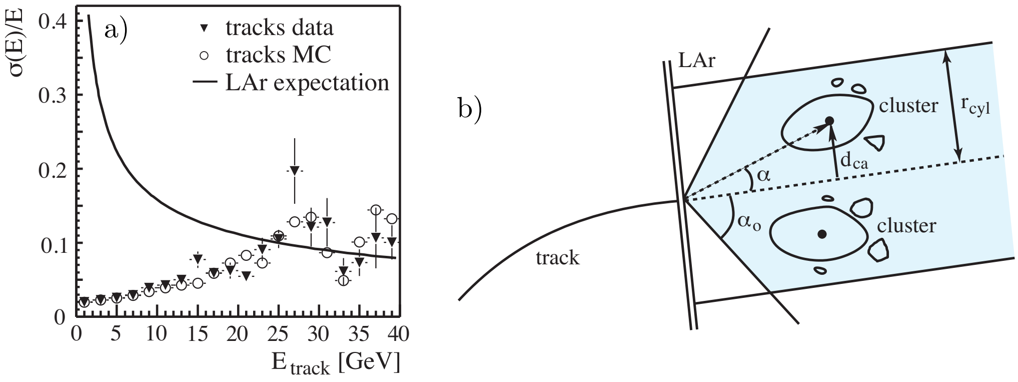


Figure 7.8: a) Measured and simulated track energy resolutions compared to the expected resolution of the LAr calorimeter [PP$^+$05]. b) Definition of the distance of closest approach d_{ca} between clusters and tracks pointing to the calorimeter. The cylinder radius r_{cyl} together with the opening angle α_0 of a cone define the volume (coloured area) within which clusters are collected before their energy is compared to the track's energy [Per06].

If the energy resolution from all clusters behind a track is compared with the track's energy resolution the result is biased towards the calorimetric measurement, since clusters from proximate neutral particles contribute to the cluster energy. To remove this bias the track energy is preferred to estimate the calorimetric resolution. Only after this comparison the calorimeter clusters are collected as described below. To optimise the performance of the algorithm, tracks are sorted according to their transverse momenta. The algorithm then starts to match calorimeter clusters with the well-measured low p_{T} tracks first.

Cluster Collection

To estimate the deposited energy originating from a track it is extrapolated to the calorimeter surface by a helix trajectory. Inside the calorimeter its trajectory is continued with a straight line as illustrated in figure 7.8b. All clusters inside an overlapping volume of a cone with opening angle $\alpha_0 = 67.5°$ and two cylinders with radii $r^{\mathrm{e}}_{\mathrm{cyl}}$ and $r^{\mathrm{h}}_{\mathrm{cyl}}$ for the EMC and HAC are then collected.
95% of a hadronic shower's energy are contained in a cylinder of radius λ (see section 5.2.3), which is approximately 30 cm for the EMC. However, naively choosing $r^{\mathrm{e}}_{\mathrm{cyl}} = \lambda_{\mathrm{EMC}}$ results in too much cluster energy collected. This is because the algorithm described here uses the barycenters of clusters as distance measure such that cells further away than $r^{\mathrm{e}}_{\mathrm{cyl}}$ can contribute to the energy as long as they belong to a cluster with its barycenter inside the cylinder. Choosing not too large cylinder radii is especially important for the EMC, since on average about half of a jet's energy is carried by photons from π^0 decays, and their electromagnetic showers are fully contained in the EMC.
Suitable cylinder radii can be estimated from simulated single particles, where the calorimeter deposits behind a track can be unambiguously associated to the incident particle. In figure 7.9 the deposited energy as a function of the distance of closest approach d_{ca} of the cluster's barycenter from the extrapolated straight line is shown for the different EMC wheels. Different charged hadrons behave very similar

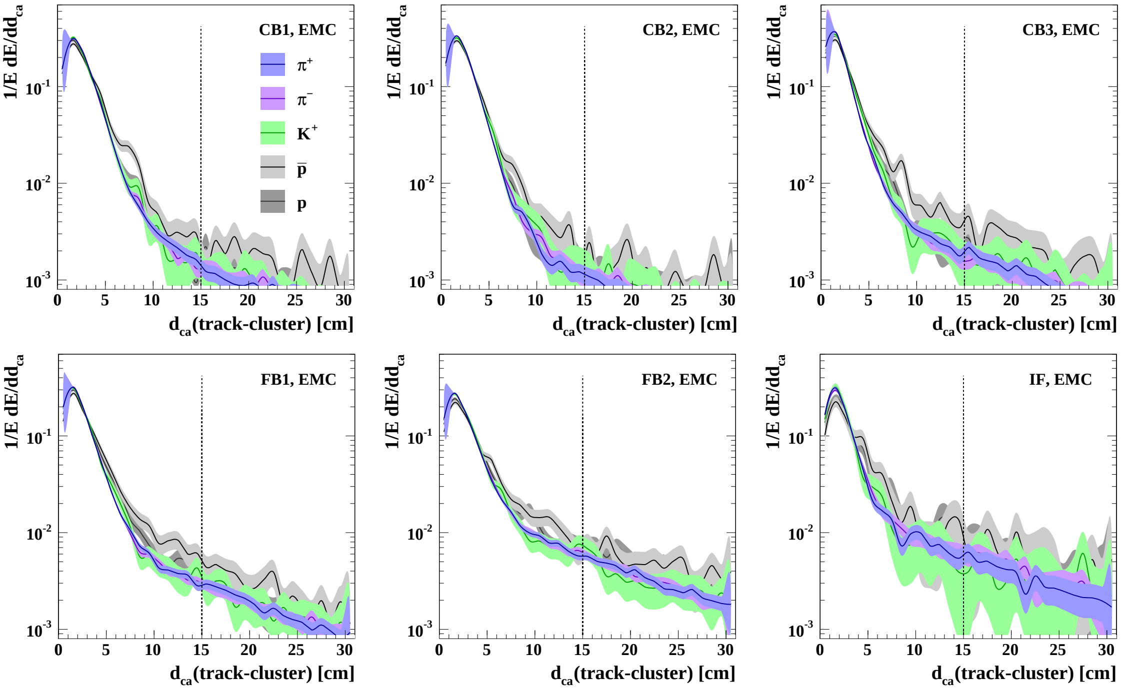


Figure 7.9: Deposited energy as function of the distance of closest approach between the barycenters of EMC clusters and the extrapolated track of incident charged pions, kaons and protons. The chosen cylinder radius $r^{\mathrm{e}}_{\mathrm{cyl}} = 15\,\mathrm{cm}$ is shown as dashed lines. The particles were simulated in an energy range between 0.5 and 30 GeV.

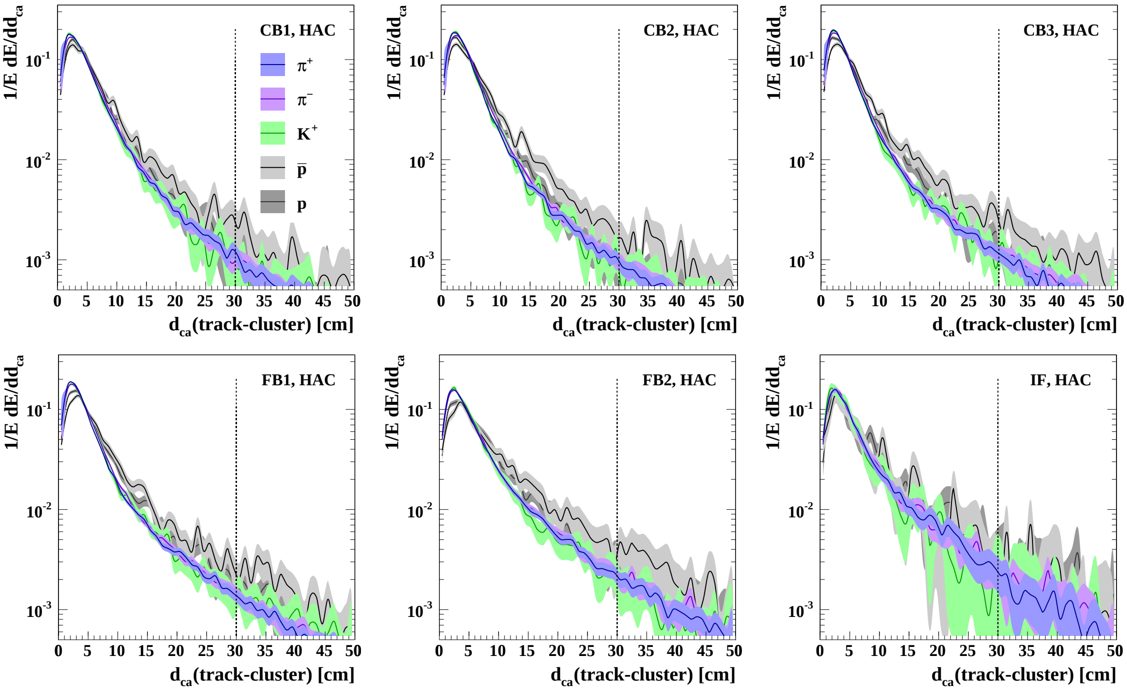


Figure 7.10: Same as in figure 7.9 but for HAC clusters. The cylinder radius for the HAC is chosen to be $r^{\mathrm{h}}_{\mathrm{cyl}} = 30\,\mathrm{cm}$ and is shown as dashed line.

owing to their similar shower development. The energy dependence is found to be negligible. Choosing $r_{\mathrm{cyl}}^{\mathrm{e}}$ to be 15 cm, between 97–99% of the shower energy are contained inside the cylinder. Only in the FB2 the tail is slightly larger with about 94% shower containment in the cylinder. Figure 7.10 shows the same distributions for the HAC. A cylinder radius $r_{\mathrm{cyl}}^{\mathrm{h}}$ of 30 cm leads to a 95% shower containment, with the exception of the FB2 and IF, where it is again about 92%. The twice as large cylinder radius can be attributed to the about twice as large cell sizes in the HAC.

The total cluster energy inside the collection volume E_{clus} is taken on the final scale, i.e. clusters with $P_{\mathrm{em}} > 0$ are taken on the electromagnetic scale E_{em} (5.40) and for all others the hadronic scale E_{had} (5.41) is used.

Track Measurement Preferred

If equation (7.6) is fulfilled, the track is used to create an HFS object. The energy deposited in the calorimeter due to the charged particle which generated the track has to be set to zero to avoid double counting of the track energy. However, just subtracting E_{track} from E_{clus} would bias the measurement towards a too large total energy because of upwards fluctuations from the calorimetric measurement. Instead, the cluster energy E_{clus} is set to zero if

$$E_{\mathrm{clus}} < E_{\mathrm{track}} \left(1 + 1.96 \sqrt{ \left(\frac{\sigma(E)}{E} \right)^2_{\mathrm{track}} + \left(\frac{\sigma(E)}{E} \right)^2_{\mathrm{LAr}} } \right) , \tag{7.8}$$

which takes into account the fluctuations of both measurements within a 95% confidence interval. If equation (7.8) is not fulfilled, the cluster's excess energy is assumed to originate from a neutral particle or another track, and the energy E_{track} is subtracted from E_{clus}. The remaining clusters are kept and used in the subsequent steps of the algorithm.

Calorimetric Measurement Preferred

In the case of equation (7.6) being false the cluster energy is compared to the track energy,

$$E_{\mathrm{clus}} - 1.96\sigma_{\mathrm{clus}} < E_{\mathrm{track}} < E_{\mathrm{clus}} + 1.96\sigma_{\mathrm{clus}} \tag{7.9}$$

with $\sigma_{\mathrm{clus}} = 0.5/\sqrt{E_{\mathrm{clus}}}$. If this relation is true the measurements are considered to be compatible and the track is removed. The clusters are then used for the creation of an HFS object.

If the track's energy is smaller than $E_{\mathrm{clus}} - 1.96\sigma_{\mathrm{clus}}$, it is assumed that E_{clus} originates from several particles, only one of which resulted in the track pointing to the calorimetric deposit. The track is then used for the creation of an HFS object and calorimetric energy is subtracted as above.

On the other hand, if $E_{\mathrm{track}} > E_{\mathrm{clus}} + 1.96\sigma_{\mathrm{clus}}$ the track measurement is usually not reliable due to a high p_{T} track. In this case the track is removed. Once all tracks have been treated, the remaining clusters are used for the creation of HFS objects.

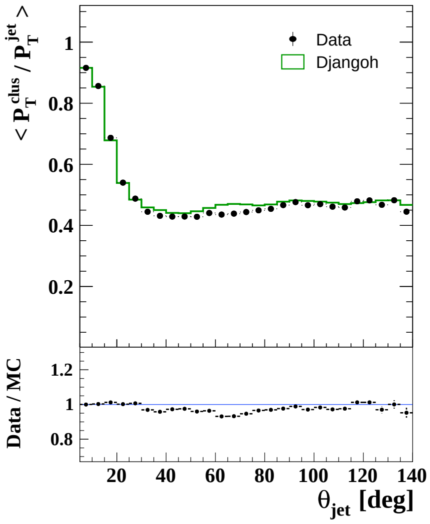


Figure 7.11: Average fraction of the jet transverse momentum from clusters as function of the jet's polar angle.

Performance

After the application of the energy flow algorithm the hadronic final state is a mixture of tracks and clusters. These HFS objects are used for the reconstruction of the event kinematics and the jet finding. In figure 7.11 the average fractional jet transverse momentum from clusters is shown for jet data in DIS and simulated jet events. For $\theta_{\text{jet}} < 10^\circ$ the fractional jet P_{T} from clusters is larger than 90% due to the low efficiency of the FTD. In the region of $10^\circ < \theta_{\text{jet}} < 25^\circ$ combined and central tracks start to contribute thus lowering the fraction of clusters. Inside the full acceptance of the CJC the fractional P_{T} from clusters is about 45–50%. The data are well described by the simulation over the full angular coverage of the detector. The calibration of the HFS in the context of the energy flow algorithm Hadroo2 is described in section 8.2.

7.5 Definition of Kinematic Variables

In deep-inelastic scattering the event kinematics are constrained by only two of the three variables x, y and Q^2 at a fixed centre-of-mass energy $\sqrt{s}$. Because of energy and momentum conservation any one of them can be measured either through the scattered electron or the hadronic final state, if the electron and proton beam energies E_{e}^0 and E_{p}^0 are known. This results in various possibilities of reconstructing the event kinematics, each with its advantages and disadvantages depending on the values of Q^2 and y. The kinematic reconstruction methods and variables used in this work are reviewed in the following.

Scattered Electron

The azimuthal angle of the scattered electron ϕ_{e} is obtained from the associated track, if it is a vertex fitted track and thus benefits from the good $r\phi$-resolution of

the CJC. If a non-vertex fitted track is associated to the electron ϕ_e is obtained from the cluster. The polar angle θ_e is always calculated from the cluster position, since electron tracks without COZ or CST hits do not have a very precise $\tan\lambda$ and consequently θ measurement. After the alignment procedure (section 7.3) the agreement between measuring the angles ϕ_e and θ_e with tracks and the cluster position is within 1 mrad, and no additional systematic uncertainty is introduced.
For the energy of the scattered electron E_e the calorimetric energy measurement is used, with a calibration uncertainty of 0.5% in most of the kinematic coverage (see section 8.1). The excellent electromagnetic resolution of the LAr calorimeter yields resolutions of 4–6% over the full energy range considered in this analysis [Nik07]. The fourvector of the scattered electron is given by

$$\begin{pmatrix} E_\mathrm{e} \\ p_{x,\mathrm{e}} \\ p_{y,\mathrm{e}} \\ p_{z,\mathrm{e}} \end{pmatrix} = \begin{pmatrix} E_\mathrm{e} \\ E_\mathrm{e}\sin\theta_\mathrm{e}\cos\phi_\mathrm{e} \\ E_\mathrm{e}\sin\theta_\mathrm{e}\sin\phi_\mathrm{e} \\ E_\mathrm{e}\cos\theta_\mathrm{e} \end{pmatrix} . \tag{7.10}$$

Electron-Sigma Method

The electron-Sigma (eΣ) method is a combination of two reconstruction methods introduced to reduce the effects of QED radiation and losses, due to the limited detector acceptance, on the kinematic reconstruction [BB95]. It is the most precise reconstruction method in the kinematic range of this analysis [Gou08] and is used to reconstruct the DIS kinematic variables x, y and Q^2 throughout this work.
The reconstruction of Q^2 is identical with the electron-method, using the energy of the scattered electron and its polar angle,

$$Q^2 \equiv Q^2_\mathrm{e} = 4\,E^0_\mathrm{e}\,E_\mathrm{e}\,\cos^2\left(\frac{\theta_\mathrm{e}}{2}\right) . \tag{7.11}$$

The longitudinal momentum balance of the hadronic final state is defined as

$$\Sigma = \sum_\mathrm{h} E_\mathrm{h} - p_{z,\mathrm{h}} , \tag{7.12}$$

where the sum runs over all HFS objects[2]. Energy and momentum conservation imply that

$$E - p_z = \Sigma + E_\mathrm{e} - p_{z,\mathrm{e}} = 2E^0_\mathrm{e} , \tag{7.13}$$

which is only approximately fulfilled for an experimental apparatus with limited resolution and acceptance. However, $(E - p_z)/2$ can be used as an alternative expression for the electron beam energy E^0_e. Thus it is possible to correct for a reduction of E^0_e prior to the hard interaction due to initial state radiation escaping through the beam pipe in the negative z-direction. Utilising this, the inelasticity y and the virtuality Q^2 can be written as

$$y_\Sigma = \frac{\Sigma}{E - p_z}, \quad Q^2_\Sigma = \frac{E^2_\mathrm{e}\sin^2\theta_\mathrm{e}}{1 - y_\Sigma} . \tag{7.14}$$

---

[2] In this analysis all particles except for the scattered electron are considered to contribute to the HFS.

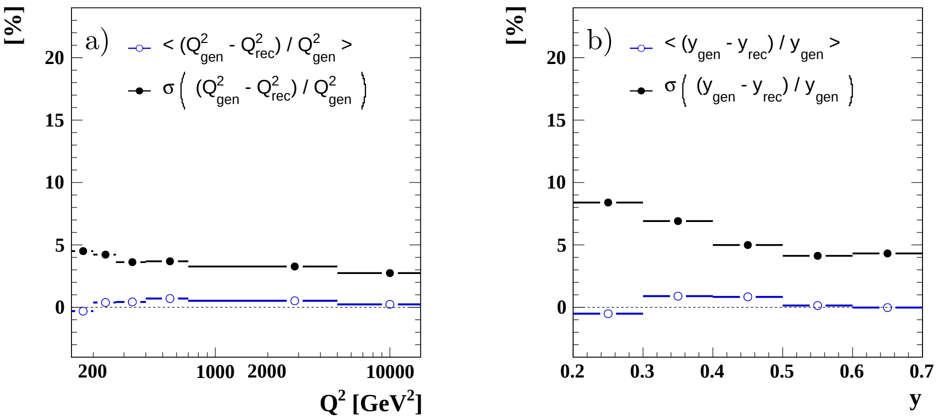


Figure 7.12: Systematic biases and resolutions of the reconstruction of Q^2 (a) and y (b) with the $e\Sigma$-method obtained with the Djangoh MC. Shown are the mean values and variances of fitted Gaussian distributions. Identical results are obtained with the Rapgap MC.

The Bjorken scaling variable x is then obtained through the relation $Q^2 = xys$,

$$x \equiv x_\Sigma = \frac{Q^2_\Sigma}{s\, y_\Sigma}\,. \tag{7.15}$$

Since the centre-of-mass energy $\sqrt{s}$ is calculated using the nominal beam energies, x_Σ retains some sensitivity to initial state radiation. Finally, the inelasticity y is calculated from a mixture of both reconstruction methods

$$y \equiv y_{e\Sigma} = \frac{Q^2_{\mathrm{e}}}{s\, x_\Sigma} \tag{7.16}$$

and thus benefits from the excellent resolution of the electron method for the reconstruction of Q^2 and the relatively low sensitivity to initial state radiation for the reconstruction of x [BB99]. In figure 7.12 the achieved resolutions with this reconstruction method are shown together with the introduced systematic biases. The generated Q^2 and y are defined by taking radiated photons in close proximity to the scattered electron into account (see section 10.2). The mean differences between the reconstructed and generated values of Q^2 and y are smaller than 1%. The introduced biases from the eΣ-method are thus negligible. The obtained resolutions are 4.5–2.5% for the reconstruction of Q^2 and 8–4% for the reconstruction of y.

Double-Angle Method

The double-angle method uses the angles of the scattered electron θ_{e} and the inclusive hadronic angle γ_{h} for the reconstruction of kinematic variables. It is, in a first approximation, insensitive to the absolute energy scales of the electron and HFS measurements and is used for the determination of calibration constants and the evaluation of energy scale uncertainties. Since angles can be measured very precisely, the double-angle method yields good resolutions. It is, however, very sensitive to initial state QED radiation and losses due to the limited detector acceptance. To obtain a sample without any systematic biases special selection criteria have to be applied. Hence, this method is not used for the reconstruction of x, y and Q^2 for the measurement of the jet cross sections, but only for the reconstruction of variables relevant for the calibration.

The angle of the hadronic system is defined via

$$\tan\frac{\gamma_\mathrm{h}}{2} = \frac{\Sigma}{P_\mathrm{T}^\mathrm{h}}\,, \tag{7.17}$$

where the transverse momentum of the hadronic final state P_T^h is given by

$$P_\mathrm{T}^\mathrm{h} = \sqrt{\left(\sum_\mathrm{h} P_{x,\mathrm{h}}\right)^2 + \left(\sum_\mathrm{h} P_{y,\mathrm{h}}\right)^2}\,. \tag{7.18}$$

The sums extend over all HFS objects. In equation (7.17) the nominator and denominator are both calibrated with the same method such that calibration effects mostly cancel. In the simple QPM the angle γ_h corresponds to the polar angle of the struck quark. The transverse momentum of the hadronic final state can thus be expressed in terms of the angles θ_e and γ_h

$$P_\mathrm{T}^\mathrm{da} = \frac{2E_\mathrm{e}^0}{\tan\frac{\gamma_\mathrm{h}}{2} + \tan\frac{\theta_\mathrm{e}}{2}}\,. \tag{7.19}$$

Writing the electron energy in terms of γ_h and θ_e one obtains

$$E^\mathrm{da} = \frac{2E_\mathrm{e}^0 \sin\gamma_\mathrm{h}}{\sin\gamma_\mathrm{h} + \sin\theta_\mathrm{e} - \sin(\gamma_\mathrm{h} + \theta_\mathrm{e})}\,. \tag{7.20}$$

The resolution of the double-angle measurement and biases due to unmeasured particles are covered in section 8.2.

Chapter 8

Calibration

This chapter covers the calibration of the LAr calorimeter for the scattered electron and the hadronic final state. Whereas the energy of the scattered electron is measured purely with the calorimeter, the hadronic final state is reconstructed with the energy flow algorithm combining tracks with the measurement of calorimetric energy deposits, as described in the previous chapter.

The calibration algorithms developed make use of the over-constrained kinematics in NC DIS scattering to calculate the expected energy of the scattered electron or the HFS. Calibration functions are derived for the data and the simulation by comparing the expected with the measured energy. The calibrations obtained are therefore independent of the simulation of the absolute energy scale. Prior to the application of the calibrations presented here, the energy scale is known within 2–3% for the electron energy and 4–5% for the HFS.

8.1 Electron Calibration

The electromagnetic energy scale of the LAr calorimeter has been obtained from test beam data [H194b]. During H1's operation the stability of the calorimeter over time has been carefully monitored and charge calibrations of the read-out system have been performed at least once per day. However, changes of the LAr purity, the pedestal subtraction and the charge collection efficiency can lead to small deviations from the electromagnetic scale over time. Furthermore, modifications to the inner part of the detector introduced additional dead material that reduces the energy of the scattered electron before it enters the active part of the calorimeter. Since these partly time-dependent effects are only to a certain degree modelled in the simulation, the deviations between data and MC simulations can be larger than 2%. The electron calibration aims to correct for these effects and to obtain the absolute energy scale in the data and simulation independently.

The procedure to obtain calibration constants for the electron energy has been developed elsewhere [Hei99, Nik07] and has been improved recently [Nik10]. In a special sample, where the double-angle method yields unbiased reconstructed energies, the energy of the scattered electron E_{e} is compared to the energy obtained with the double-angle method E^{da}. Deviations of the ratio $E_{\mathrm{e}}/E^{\mathrm{da}}$ from one are then used to derive the calibration constants.

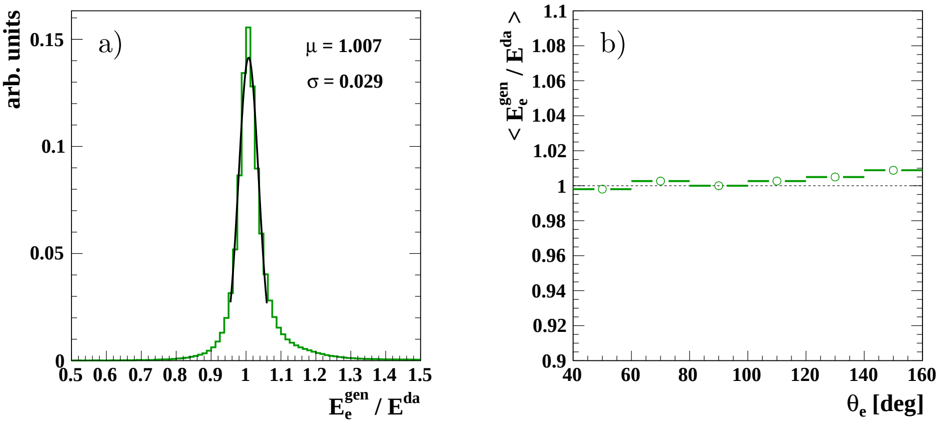


Figure 8.1: a) The ratio of the generated energy of the scattered electron to the double-angle energy E^{da} after the application of the electron calibration selection. The peak is fitted by a Gaussian with mean μ and variance σ. b) Mean values of $E_{\mathrm{e}}/E^{\mathrm{da}}$ as a function of the polar angle of the scattered electron.

The trajectory of the scattered electron can be reconstructed very precisely using the primary vertex, the electron's track and impact positions on the calorimeter. Because of the high statistics of the HERA-2 data set, the calibration factors can be determined for a finely segmented grid in z_{impact} and ϕ_{e}. Here z_{impact} denotes the z-coordinate of the electron's impact position on the calorimeter and ϕ_{e} is the azimuthal angle of the scattered electron. Local changes of the amount of material in front of the calorimeter can thus be taken into account. The calibration is performed in three steps. First, calibration factors are obtained for each octant of the BBE and CBs. Because of decreasing statistics for smaller values of θ_{e}, only one calibration factor is obtained for the FBs. In the next step 200 calibration factors depending on z_{impact} are determined. In the backward region a segmentation of 1 cm is chosen, becoming larger with increasing z_{impact}. In the last step 128 calibration factors depending on ϕ_{e} are obtained for each of the BBE, CB1 and CB2 calorimeter wheels. The calibration factors are determined for eight different HERA-2 run periods to account for a possible time dependence of the calorimeter response.

To test the electron calibration the following selection criteria are applied additionally to the DIS selection discussed in section 9.6:

- $E_{\mathrm{e}} > 14\,\mathrm{GeV}$
- $y_{\Sigma} < 0.3$ for $z_{\mathrm{impact}} \leq 20\,\mathrm{cm}$
- $y_{\Sigma} < 0.5$ for $20 < z_{\mathrm{impact}} \leq 100\,\mathrm{cm}$
- $\gamma_{\mathrm{h}} > 10^{\circ}$

The first cut ensures well reconstructed electrons and a clean sample with neglibile contamination from non-*ep* background. The other three cuts ensure a good measurement of the hadronic final state and consequently an unbiased double-angle measurement. The effect of this selection on the measurement of E^{da} is shown in figure 8.1a using MC simulation. E^{da} shows good correspondence with the generated

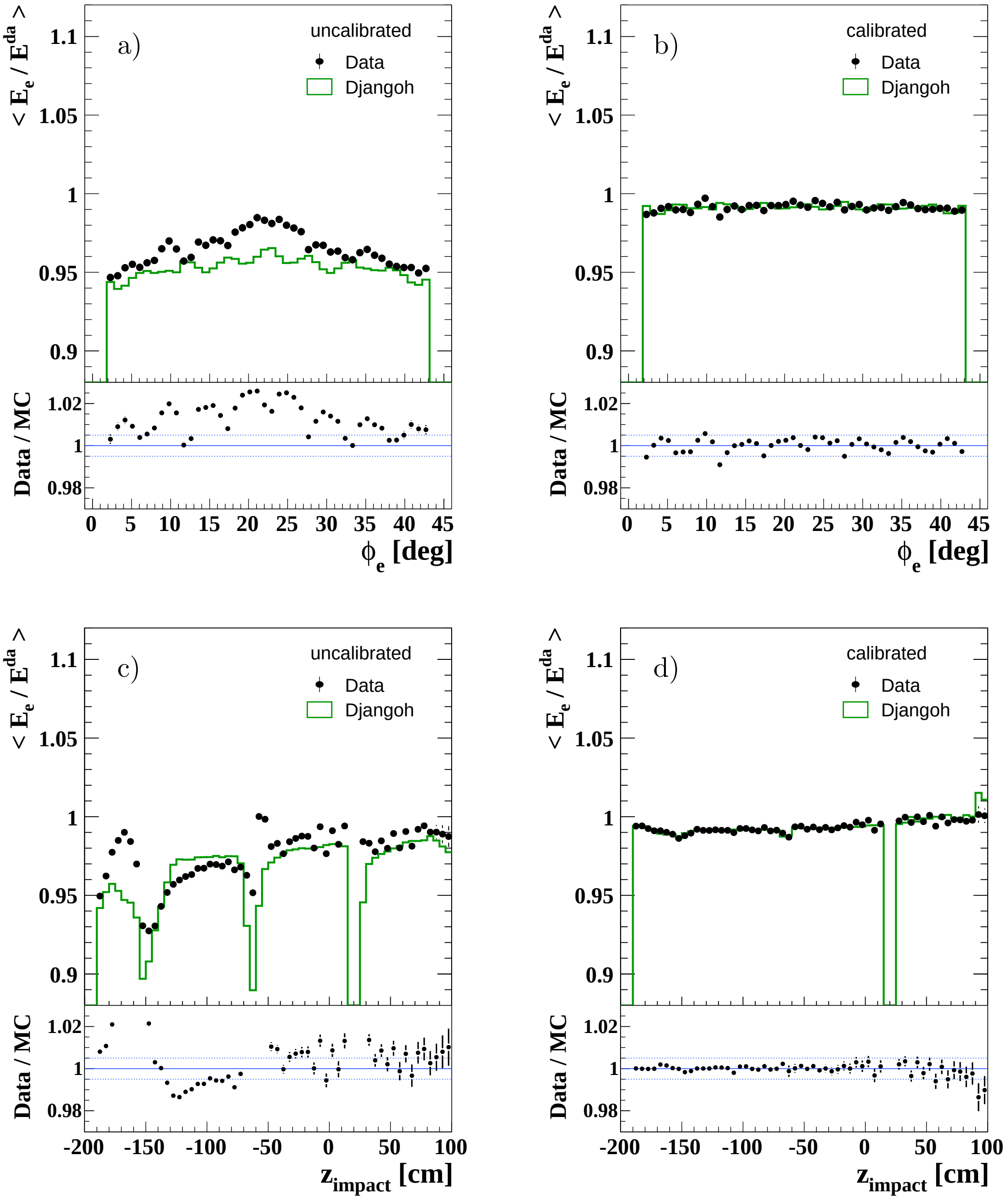


Figure 8.2: Mean values of the ratio of the electron energy E_e to the energy obtained with the double-angle method E^da as function of ϕ_e folded into one octant (a), (b) and depending on z_impact (c), (d). Data from the full HERA-2 data set are compared to simulated events with the ratio shown at the bottom, the dotted lines denote a 0.5% difference. Before applying the calibration the reconstruction of E_e depends strongly on z_impact and ϕ_e, the difference between data and simulation is 2–3% (a), (c). After the calibration E_e agrees well with E^da, independent of the azimuthal angle or the impact position, and the ratio is described by the simulation within 0.5% (b), (d).

energy of the scattered electron, almost independent of the electron scattering angle (figure 8.1b). The resolution of the double-angle measurement is 2–3%, slightly varying with θ_e.

Comparisons of the electron energy measurement with E^{da} are shown in figure 8.2 for data and simulated events before and after the calibration. The uncalibrated electron energy is too small and depends strongly on the impact position. Visible structures in the E_e/E^{da} ratio are only approximately described by the simulation. The largest deviations are observed in the z-crack regions and in the BBE ($z < -155\,\text{cm}$). Electron clusters within $\pm 5\,\text{cm}$ of the z-crack between the CB2 and CB3 are excluded from this analysis, as well as clusters within $\pm 2^\circ$ of ϕ-cracks. After the application of the calibration E_e agrees well with E^{da} independent of the impact position. The difference between data and simulated events is within 0.5% for $z_{\text{impact}} < 100\,\text{cm}$ which is assigned as systematic uncertainty on the measurement of E_e. For the region $z_{\text{impact}} > 100\,\text{cm}$ an uncertainty of 1% is assigned [Mai07].

8.2 Calibration of the Hadronic Final State

The calibration of the HFS is more involved than the electron calibration. The HFS is a composition of tracks and calorimeter clusters due to the application of the energy flow algorithm Hadroo2 (section 7.4.3). Tracks entering the HFS are well measured and do not have to be corrected. Calorimeter clusters on the other hand can originate from hadronic or electromagnetic showers or mixtures of them and their correction factors will be very different. Even after the application of the software weighting an energy dependence is observed for low energy clusters which has to be taken into account. However, the same principle as for the electron calibration can be applied. Due to the over-constrained kinematics at HERA it is possible to define a reference measurement independent of the hadronic energy scale. The transverse momentum of the HFS is then compared to the reference measurement and calibration functions are defined.

8.2.1 Calibration Sample

The double-angle method (see section 7.5) is to a good approximation independent of the absolute energy scale. Hence it is possible to use P_T^{da} as a reference, if biases due to losses and photon radiation from the incoming and scattered electron due to QED processes can be eliminated. In a first step a sample with a well measured scattered electron in the LAr calorimeter and a jet that balances its transverse momentum is defined via the following selection criteria:

- $Q^2 > 100\,\text{GeV}^2$,
- $P_T^e > 10\,\text{GeV}$,
- background suppression and electron quality cuts,
- only one jet.

The first three cuts ensure a well measured scattered electron and a background free sample, where the third item stands for the same topological background finders and requirements on the primary vertex and the scattered electron as used for the selection of NC DIS events (sections 9.3 and 9.6).

The last cut in the list is used to select events with only one high-P_T jet, which balances the transverse momentum of the electron to a good approximation. For this purpose the longitudinally invariant k_T-algorithm with $R_0 = 1$ and a minimum jet P_T of 2.5 GeV is used. The small minimum transverse momentum used at the jet finding stage ensures that in the selected events only a small fraction of the transverse momentum of the HFS originates from objects outside the jet. This is because if there is a second jet with $P_\mathrm{T} > 2.5\,\mathrm{GeV}$ found in the event, the event is rejected.

In the next step a selection is employed to remove biases of the P_T^da measurement originating from losses due to the limited detector acceptance. No cuts on variables depending on the HFS measurement are applied in order not to bias the subsequent calibration. The following cuts are used:

- no significant signal in the SpaCal,
  $P_\mathrm{T}^\mathrm{spacal}/P_\mathrm{T}^\mathrm{da} < 0.01$ and $E^\mathrm{SpaCal}/E^\mathrm{da} < 0.01$,
- no significant signal in the Tail Catcher,
  $P_\mathrm{T}^\mathrm{TC}/P_\mathrm{T}^\mathrm{da} < 0.01$ and $E^\mathrm{TC}/E^\mathrm{da} < 0.01$,
- $P_\mathrm{T}^\mathrm{e}/P_\mathrm{T}^\mathrm{da} > 0.88$,
- $P_\mathrm{T}^\mathrm{e}(\mathrm{cluster})/P_\mathrm{T}^\mathrm{e}(\mathrm{track}) < 1.25$,
- $|\theta_\mathrm{jet} - \gamma_\mathrm{h}|/\theta_\mathrm{jet} < 1.5$.

This list is in the following termed "calibration sample selection" or CSS. The first cut ensures that no hadronic energy was lost in the backward direction because of a punch-through in the SpaCal with its small effective length of two nuclear interaction lengths. The second cut ensures that no energy escaped out of the LAr calorimeter. The third and fourth cuts considerably reduce the effects of initial and final state QED radiation on the double-angle measurement. The last cut ensures a well contained hadronic final state when the jet is measured in the forward direction at small θ_jet. The effect of these cuts on the measurement of P_T^da is shown in figure 8.3. The generated P_T, which is the P_T of the generated HFS and consequently of the generated scattered electron in simulated events, is compared to the measured P_T^da. This is done after the electron and one-jet selection and after additionally applying the CSS cuts from above. Without the CSS cuts, the double-angle measurement is biased towards too large transverse momenta with a strong P_T dependence. After the application of all cuts, P_T^da agrees within 1% with the generated P_T in simulated events. A small remaining bias in the central region ($80^\circ < \theta_\mathrm{jet} < 135^\circ$) is being accounted for by a correction function. For this purpose a smooth parameterisation depending on P_T^da and θ_jet is constructed and fitted to the relative difference over the full measurement range. The obtained correction function is displayed as black line.

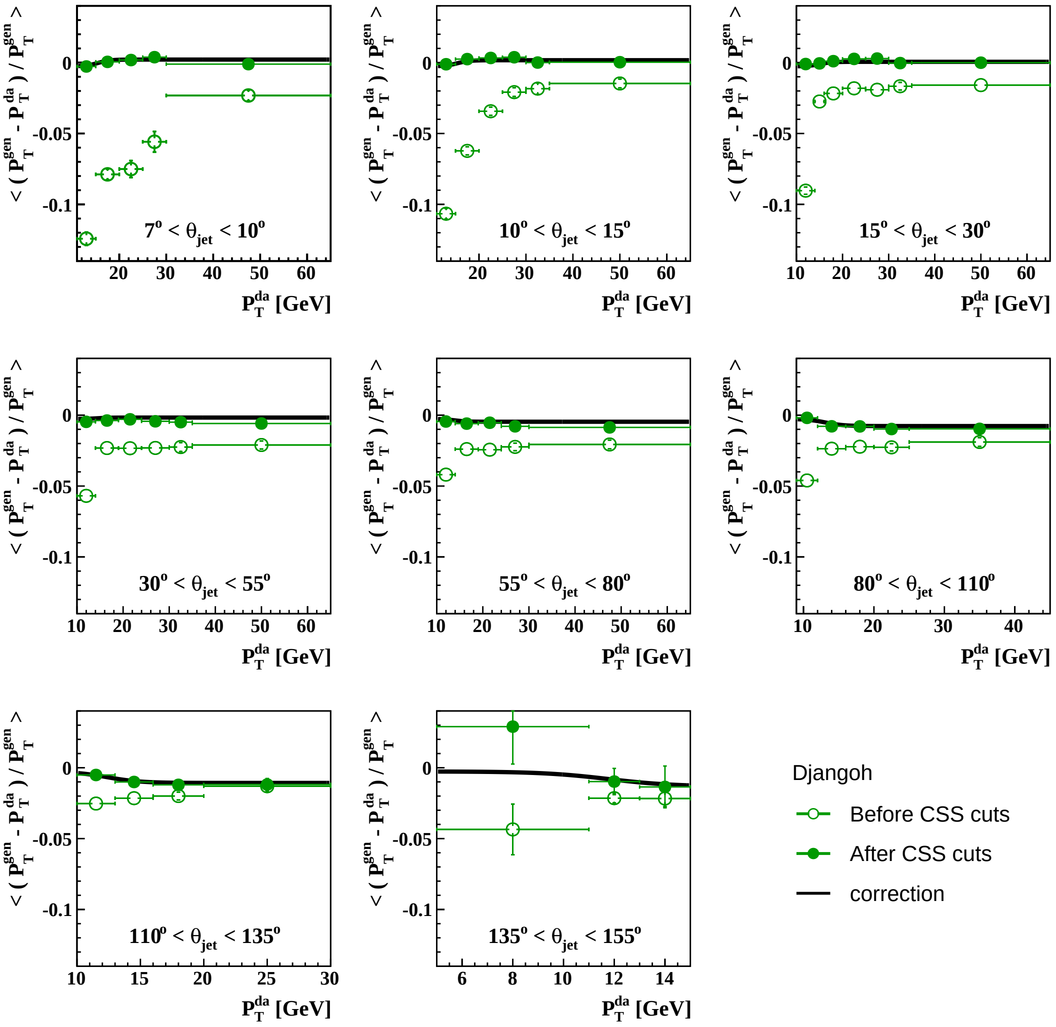


Figure 8.3: Relative difference for simulated DIS events between the generated P_T and P_T^{da} as function of P_T^{da} in various regions of θ_{jet} for the calibration sample. Shown are the P_T^{da} measurement before (open circles) and after (filled circles) the application of selection criteria which ensure a well-contained HFS (CSS). The correction applied to the P_T^{da} measurement is shown as black line.

After the application of this correction the double-angle measurement is tested for the calibration sample. The agreement between the generated P_T and P_T^{da} is within 0.3% as shown in figure 8.4. Only in the direction of the BBE ($135° < \theta_{jet} < 155°$) a larger deviation than 0.3% is observed at small values of P_T^{da}. However, the statistical uncertainty in this region is large and a further study of this effect is left for a dedicated analysis at low transverse momenta. The influence of the calibration on the double-angle measurement is also shown in figure 8.4. After the full calibration of the HFS P_T^{da} still agrees with the generated P_T within 0.3% validating the assumption of P_T^{da} being independent of the absolute energy scale. In this calibration sample the resolution of the P_T^{da} measurement is approximately 2% for $\theta_{jet} \lesssim 80°$ and 3–4% in the region of the CB1 and CB2, being by a factor of three to four smaller than the expected resolution of the HFS measurement. In the

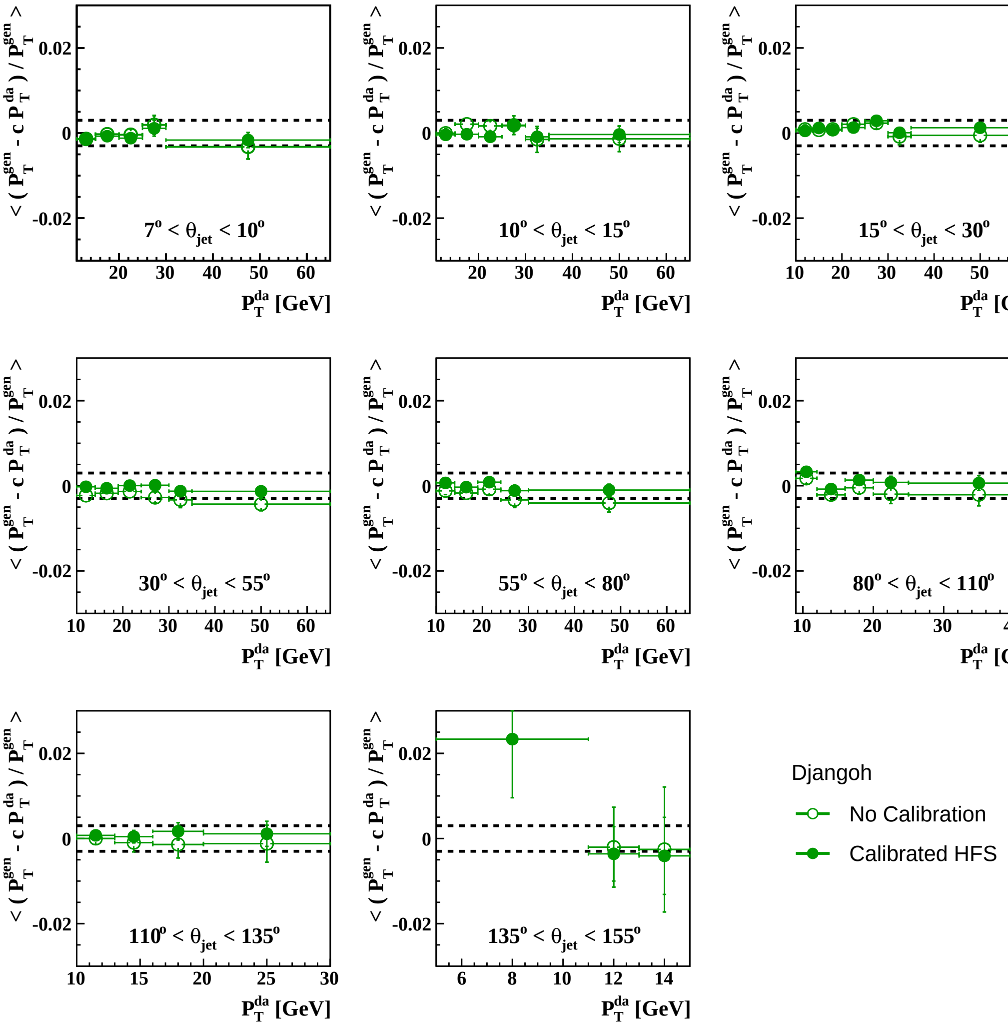


Figure 8.4: Relative difference for simulated DIS events between the generated P_{T} and corrected $P_{\mathrm{T}}^{\mathrm{da}}$ as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in various regions of θ_{jet} for the calibration sample. The dashed lines illustrate a deviation of 0.3%. Shown is the $P_{\mathrm{T}}^{\mathrm{da}}$ measurement before (open circles) and after (filled circles) the HFS has been calibrated.

following $P_{\mathrm{T}}^{\mathrm{da}}$ stands always for the double-angle measurement multiplied with the derived correction function (figure 8.3).

In figure 8.5 control distributions of the calibration sample are shown for one period of the HERA-2 data and simulated DIS events. In general, all relevant variables are well described by the simulation. The distribution of $P_{\mathrm{T}}^{\mathrm{da}}$ (8.5a) peaks at approximately 14 GeV which is due to the requirement of the scattered electron being measured in the LAr calorimeter and the cuts on Q^2 and $P_{\mathrm{T}}^{\mathrm{e}}$. The distribution has a significant tail towards large values of $P_{\mathrm{T}}^{\mathrm{da}}$ providing a reference for jets with large P_{T}. The distribution of the polar angle of jets θ_{jet} is shown in figure 8.5b. It peaks at small values of θ_{jet} and is relatively flat in the central region. In the backward region ($\theta_{\mathrm{jet}} > 135°$) a comparable small number of jets is observed, resulting in limited statistics for the calibration of the BBE. Large statistics in the forward direction allow for a finer binning in the IF and FB calorimeter wheels. Also shown is the elec-

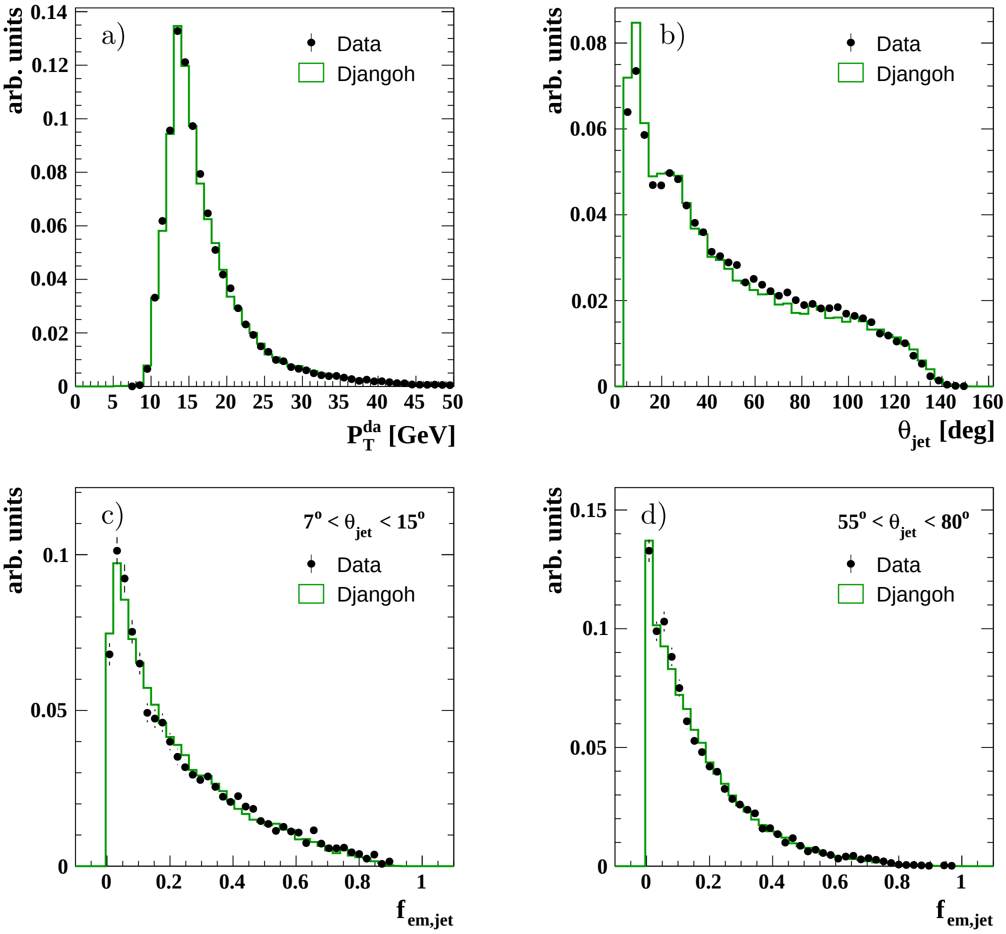


Figure 8.5: Distributions of $P_{\mathrm{T}}^{\mathrm{da}}$ (a) and θ_{jet} (b) for the calibration sample. The jets' electromagnetic energy fraction $f_{\mathrm{em,jet}}$ is shown for the forward (c) and central (d) region of the detector. All distributions are normalised to their area.

tromagnetic fraction of jets $f_{\mathrm{em,jet}}$ (6.31) for two different regions of θ_{jet} (figures 8.5c and d). $f_{\mathrm{em,jet}}$ is important for the subsequently developed calibration, which treats electromagnetic showers differently from hadronic ones. In the forward direction the electromagnetic fraction is calculated to a very large extent from calorimeter clusters only, whereas in the central region tracks contribute to the calculation. In both cases $f_{\mathrm{em,jet}}$ is very well modelled by the simulation.

8.2.2 Starting Scale

The absolute scale of the energy measurement is given by the mean value of the distribution of the P_{T}–balance, defined as $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{da}}$, in the calibration sample. To estimate the mean value and width of the P_{T}–balance distribution it is fitted by a Student's t-distribution given by

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \tag{8.1}$$

with

$$t = \frac{x-\mu}{\sigma_s}. \tag{8.2}$$

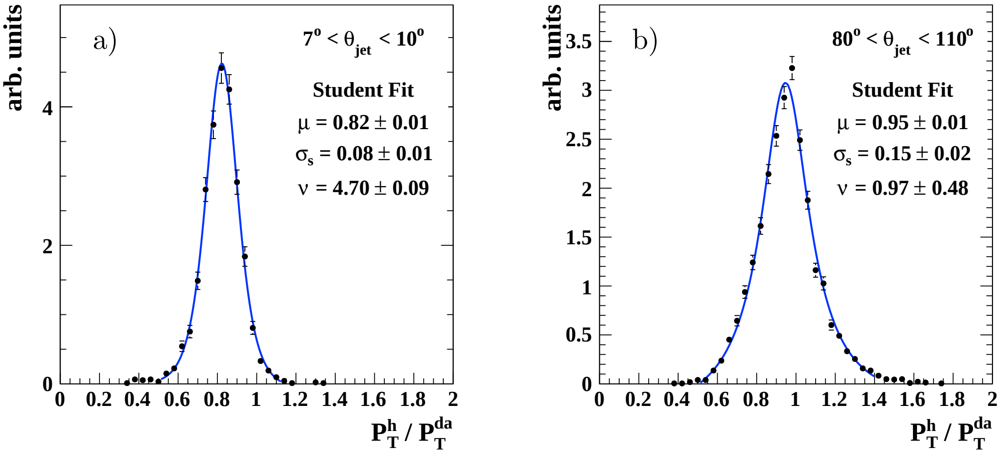


Figure 8.6: Fits of the parameterisation (8.1) to distributions of $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{da}}$ in the forward (a) and central (b) regions.

Γ stands for the Gamma-function, the parameter μ corresponds to the mean value of the distribution, and ν is referred to as number of degrees of freedom. The distribution resembles a Breit-Wigner distribution for small values of ν and approaches a Gaussian distribution as ν becomes large. In this case the parameter σ_s resembles the variance of the Gaussian distribution. The advantage of using equation (8.1) for fits of the P_{T}–balance distribution instead of a Gaussian lies in the varying shape of the P_{T}–balance distribution depending on the jet's polar angle. In the forward region, where the energy measurement is dominated by the calorimetric measurement, the P_{T}–balance distribution is Gaussian shaped. In the central region, where the contribution from tracks is large, a Breit-Wigner distribution describes the shape of the P_{T}–balance better. The Student's t-distribution allows a smooth transition between these two cases without an a-priori assumption on the shape of the P_{T}–balance. It allows also a fit over the full range instead of restricting the fit only to the peak region. Examples of fits to P_{T}–balance distributions are given in figure 8.6 for the forward (a) and central (b) regions. In the forward region the value of ν is large with an approximately Gaussian distribution whereas in the central region ν takes usually values around one. Since both, ν and σ_s, control the width of the distribution the variance is defined via the full width at half maximum (FWHM) as

$$\sigma = \frac{\mathrm{FWHM}}{2\sqrt{2\ln 2}} \tag{8.3}$$

which resembles the variance of a Gaussian in the limit of large ν. The error on σ is determined by independently varying ν and σ_s until a change of χ^2 by one is obtained. The error on σ is consequently defined as the change of the FWHM by this variation.

In figure 8.7 the mean values (a) and relative resolutions σ/μ (b) of the P_{T}–balance for the full HERA-2 dataset are shown as function of θ_{jet}. Results obtained with the standard implementation of Hadroo2 are illustrated as open circles. In this implementation a cut on the energy fraction in the first two calorimeter layers is used to decide whether a cluster is taken on the electromagnetic or hadronic energy scale (section 6.2). Additionally it has larger cylinder radii $r_{\mathrm{cyl}}^{\mathrm{e}}$ and $r_{\mathrm{cyl}}^{\mathrm{h}}$ for the

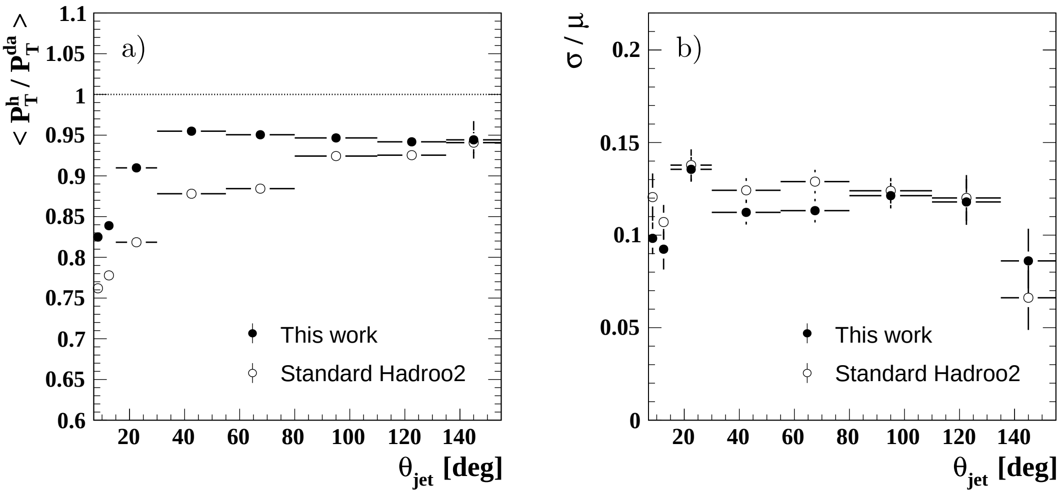


Figure 8.7: Mean values (a) and relative resolutions (b) of P_T-balance distributions in bins of θ_{jet}, obtained with fits to equation (8.1). Shown are data from the full HERA-2 dataset. Results obtained with the standard hadronic reconstruction are compared to results using neural networks for the choice of the electromagnetic or hadronic energy scale and optimised values of the cylinder radii $r_{\text{cyl}}^{\text{e}}$ and $r_{\text{cyl}}^{\text{h}}$ (see section 7.4.3).

collection of clusters behind tracks (see section 7.4.3). The utilisation of the electromagnetic probability P_{em} from equation (6.32) for the separation of electromagnetic and hadronic showers (filled black circles) significantly improves the absolute energy scale by 8% in the forward region. Simultaneously the resolution improves by 10–15% in the IF ($\theta_{\text{jet}} < 15°$). In the region of the FB2 and OF, corresponding to $15° < \theta_{\text{jet}} < 30°$, the absolute energy scale improves by 10% with an unaltered resolution. In the central region the visible improvement of the absolute energy scale results from the combination of the shower separation and optimised cylinder radii. In the region of the FB1 and CB3 ($30° < \theta_{\text{jet}} < 80°$) the improvement is larger due to a higher granularity compared to the CB1 and CB2. In the FB1 and CB3 also the resolution improves. For $\theta_{\text{jet}} > 80°$, corresponding to the CB1 and CB2, the absolute energy scale improves only slightly by about 2% with no change of the relative resolution.
The mean values of the P_T–balance distributions are the starting point for the calibration. On average the deviation from the absolute energy scale is 10–15% in the forward region and 5% in the central region. Prior to the calibration, besides the dependence on the jet's polar angle shown, also a dependence on the incident jet energy and the jet's electromagnetic fraction is observed. These dependencies have to be taken into account to achieve a flat detector response in these variables.

8.2.3 Calibration Procedure

The four-vector of the hadronic final state h is given by

$$h = \sum_{i \in \text{jet}} (\alpha^i \, c_{\text{jet}}^i + t_{\text{jet}}^i) + \sum_{j \in \text{out}} (\beta^j \, c_{\text{out}}^j + t_{\text{out}}^j) \,, \tag{8.4}$$

where c denotes a HFS object with its four-vector obtained from clusters and t an object with a track four-vector. The sums run over all HFS objects in the event whereby a distinction between objects inside and outside jets is made, with cluster calibration factors α and β. Tracks can be taken as well measured and no calibration is applied to them. During the first calibration step the calibration factors α and β are assumed to be equal

$$\alpha = \beta = P_{\text{em}}\, g_e(\eta_{\text{cls}}, E_{\text{cls}}; \vec{v}_e) + (1 - P_{\text{em}})\, g_h(\eta_{\text{cls}}, E_{\text{cls}}; \vec{v}_h)\,. \tag{8.5}$$

The calibration functions for electromagnetic and hadronic clusters g_e and g_h are smooth parameterisations covering the full detector acceptance. They depend on the cluster's pseudorapidity η_{cls} and energy E_{cls} and are given by

$$g(\eta_{\text{cls}}, E_{\text{cls}}; \vec{v})^{-1} = N(\eta_{\text{cls}}, v_{0\text{–}4})\, S(\eta_{\text{cls}}, E_{\text{cls}}, v_{5\text{–}10})\,. \tag{8.6}$$

The functions g factorise into a normalisation part N and a shape S, where both are chosen to be sigmoid functions (6.16) with polynomials in η_{cls} as arguments. These functions are n times continuously differentiable and in contrast to plain polynomials they do not diverge outside the fit range. Smoothness of the definitions (8.5) and (8.6) ensures that no discontinuities are present, for example from the binning chosen to obtain the free parameters.

The 22 free parameters $\vec{v}_e$ and $\vec{v}_h$ of the calibration functions g_e and g_h are obtained in a global χ^2 minimisation procedure with χ^2 being defined as

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{1}{\sigma_i^2} \left(\left\langle \frac{P_{\text{T}}^{\text{h}}}{P_{\text{T}}^{\text{da}}} \right\rangle_i - 1 \right)^2 . \tag{8.7}$$

It is a measure of the deviation of P_{T}^{h} from P_{T}^{da} where the former stands for the calibrated transverse momentum of the HFS,

$$P_{\text{T}}^{\text{h}} = \sqrt{h_x^2 + h_y^2}\,. \tag{8.8}$$

The sum in equation (8.7) extends over bins defined double-differentially in $(\eta_{\text{jet}}, P_{\text{T}}^{\text{da}})$ and $(\eta_{\text{jet}}, f_{\text{em,jet}})$. During the minimisation the mean value of the P_{T}–balance is obtained for each bin i through a summation over all events inside the bin, N_{ev}^i, and is given by

$$\left\langle \frac{P_{\text{T}}^{\text{h}}}{P_{\text{T}}^{\text{da}}} \right\rangle_i = \frac{1}{N_{\text{ev}}^i} \sum_{j=1}^{N_{\text{ev}}^i} \left(\frac{P_{\text{T}}^{\text{h}}}{P_{\text{T}}^{\text{da}}} \right)_j . \tag{8.9}$$

The error σ_i is the standard deviation of the distribution of $P_{\text{T}}^{\text{h}}/P_{\text{T}}^{\text{da}}$ in bin i. χ^2 is minimised with a least-squares technique using second order derivatives. The used algorithm is similar to the Minuit algorithm [JR75], except that the calculation of the first derivatives is performed analytically. The exact covariance matrix is also calculated using the first derivatives, resulting in a more efficient algorithm than the original one. This is necessary since the calculation of χ^2 uses all events in the calibration sample in each step of the minimisation.

In the second step of the calibration procedure a small correction factor is applied to clusters in jets to remove a residual energy dependence on the jet's incident energy,

$$\alpha = \beta \cdot f_{\text{j}}(\eta_{\text{jet}}, P_{\text{T}}^{\text{jet}}, \vec{v}_{\text{j}}) \tag{8.10}$$

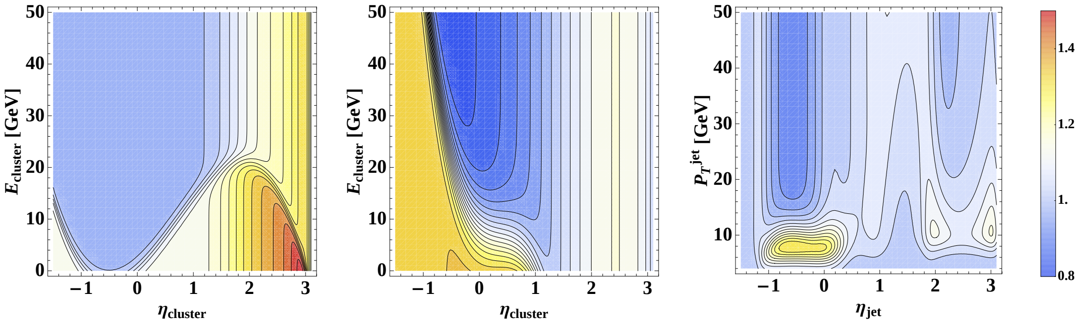


Figure 8.8: The calibration functions g_e, g_h and f_{j} (from left to right) for data.

which depends on the jet's pseudorapidity η_{jet} and transverse momentum $P_{\text{T}}^{\text{jet}}$. In this step of the calibration process the previously obtained values for $\vec{v}_e$ and $\vec{v}_h$ are kept fixed and only the parameters $\vec{v}_{\text{j}}$ are determined. The parameterisation of f_{j} is chosen to be a superposition of three sigmoidal functions to allow enough flexibility to correct for differences between the different regions of the H1 detector. It was found to be sufficient to use one sigmoidal function for the calibration of jets in the central, track-dominated region, in the transition region between the CJC and FTD in front of the calorimeter and in the forward (IF) region. A smooth interpolation between them is employed such that no discontinuities are present. In the angular regions where the fit is not constrained by data, i.e. $\eta_{\text{jet}} < -1.5$, $\eta_{\text{jet}} > 3.0$ the function f_{j} is constructed to smoothly approach unity. In the region of low jet transverse momenta, $P_{\text{T}}^{\text{jet}} < 7\,\text{GeV}$ a functional form motivated by the calibration of jets at low $P_{\text{T}}^{\text{jet}}$ is used [OS09].

The obtained calibration functions are shown in figure 8.8 for data. The electromagnetic cluster function g_e is flat with a value of one in the CBs. In regions with more dead material in front of the calorimeter, i.e. the BBE and in the forward direction, it takes values larger than one. The largest value of about 1.5 is observed for low energy clusters in the very forward region, $\eta \sim 3$, where particles traverse a long stretch of the beam pipe and the support structure of the FTD (see figure 8.9). The hadronic function g_h exhibits a fall-off from values of ~ 1.3 at low energies to 0.9 at high energies in the central region. This energy dependence is expected since low energy hadronic clusters are so far only calibrated with an average e/π ratio [WK+94]. In the forward region, $\eta > 1$, no energy dependence is observed with values of g_h between 1.1 and 1.2. g_h exhibits a strong rise towards large energies and small values of η and saturates at values of ~ 1.3 in the BBE region. In this region of pseudorapidity considerable amounts of energy are lost due to the missing hadronic calorimeter in the backward direction. The correction function for jet clusters, f_{j}, takes values between 1.25 an 0.95. The largest P_{T}-dependence is observed in the backward region with $\eta < 0$ for jets with low transverse momenta. In the forward region the P_{T}-dependence is reduced with correction factors close to unity.

After each step of the minimisation procedure the results are verified using the mean values of the P_{T}–balance distributions obtained from fits to equation (8.1). In figure 8.10 the obtained mean values of the P_{T}–balance distributions (parameters μ from the fits) are shown prior to the calibration, after the first step and after the total calibration in $(\eta_{\text{jet}}, P_{\text{T}}^{\text{da}})$ bins. The uncalibrated P_{T}–balance shows a strong

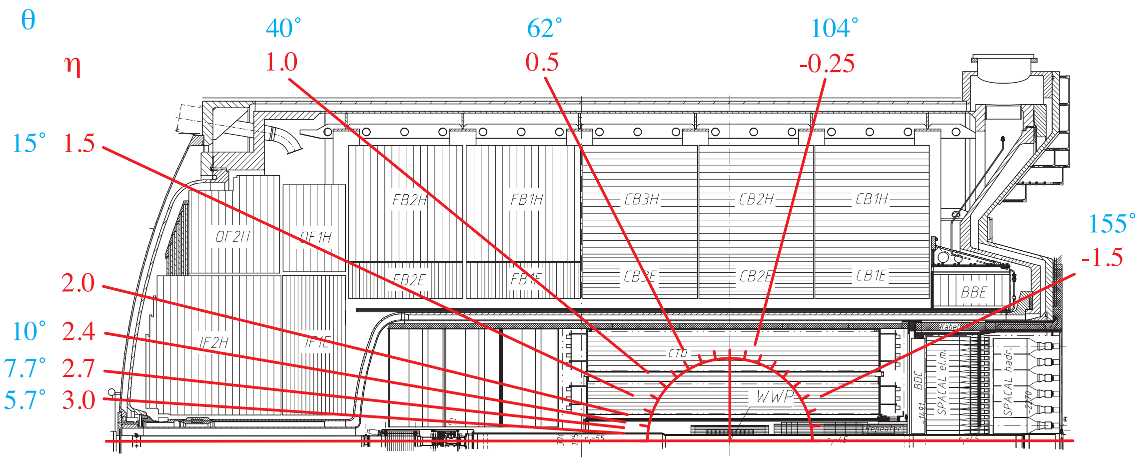


Figure 8.9: Pseudorapidity regions in the H1 detector.

dependence on η_{jet} and a weak dependence on $P_{\mathrm{T}}^{\mathrm{da}}$ which is more pronounced in the forward region. At larger values of $P_{\mathrm{T}}^{\mathrm{da}}$ the uncalibrated P_{T}–balance is closer to unity owing to the on average larger electromagnetic content of high energy hadronic showers. The first calibration step, i.e. the application of g_e and g_h, mostly removes the dependence on η_{jet}, also achieving a flat P_{T}–balance as function of $f_{\mathrm{em,jet}}$. After this step the P_{T}–balance is within 5% around unity with a residual dependence on $P_{\mathrm{T}}^{\mathrm{da}}$. Only in the very forward region $3.0 < \eta_{\mathrm{jet}} < 2.7$, corresponding to θ_{jet} of 5–7°, the P_{T}–balance is about 8% low after the first step. After the application of the full calibration, including f_j, the P_{T}–balance is flat in $P_{\mathrm{T}}^{\mathrm{da}}$, η_{jet} and $f_{\mathrm{em,jet}}$.

8.2.4 Results

The free parameters of the calibration functions are obtained for four different running conditions of the Hera-2 phase. The time dependence is found to be small but non-negligible if a precision of 1% should be achieved. Djangoh is used to determine the calibration constants for the simulation.

Subsequently the results are verified for each of the Hera-2 run periods. The data are compared to Djangoh simulations with a different set of PDFs and modified hadronisation and parton shower settings and to simulated events from Rapgap. Three different samples are used to test the calibrations. All periods show the same quality of agreement and thus allow to summarise the results for the full Hera-2 dataset as shown in the following.

In figure 8.11 the mean P_{T}–balance is shown as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in bins of η_{jet} with the selection criteria of section 8.2.1. The mean values are obtained with fits to Student's t-distributions (8.1). The P_{T}–balance is sufficiently flat and close to unity. In the central region at low values of $P_{\mathrm{T}}^{\mathrm{da}}$ deviations from unity of the order of 2% can be observed. These result from migrations into the sample with respect to the sample used for the determination of the calibration functions. Furthermore, during the calibration procedure the arithmetic mean values of the P_{T}–balance distributions were required to be unity, whereas now the distributions are fitted with Student's t-distributions. The different method of obtaining the mean values can result in small differences of the order of 1-2%. These effects are well modelled. The double-ratio,

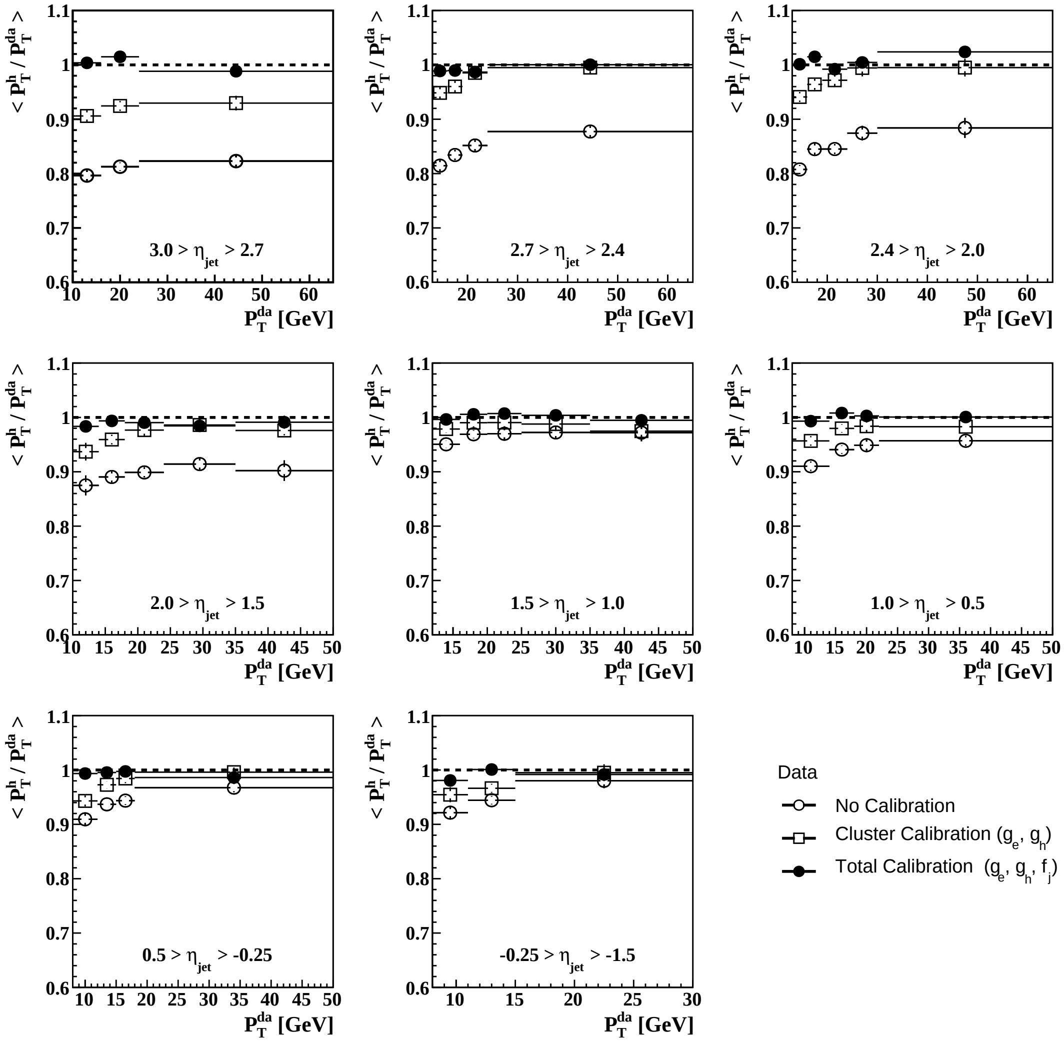


Figure 8.10: Mean values of the P_{T}-balance distributions as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in bins of η_{jet}. Shown are results obtained prior to the calibration (open circles), with the cluster calibration (open squares) and with the full calibration (filled circles). The mean values were obtained with fits using the Student's t-distribution (8.1).

defined as

$$\mathrm{dr} = \left\langle \frac{P_{\mathrm{T}}^{\mathrm{h}}}{P_{\mathrm{T}}^{\mathrm{da}}} \right\rangle_{\mathrm{Data}} \Bigg/ \left\langle \frac{P_{\mathrm{T}}^{\mathrm{h}}}{P_{\mathrm{T}}^{\mathrm{da}}} \right\rangle_{\mathrm{MC}} \tag{8.11}$$

is shown in figure 8.12. Both models describe the data within 1% over the full detector acceptance. In figure 8.13 the mean P_{T}–balance is shown as function of the electromagnetic energy fraction of jets $f_{\mathrm{em,jet}}$ in bins of η_{jet}. The P_{T}–balance is flat around unity in the calorimeter dominated forward region with $\eta_{\mathrm{jet}} > 2$ where a flat response in $f_{\mathrm{em,jet}}$ corresponds to a correctly reconstructed jet energy independent of the jet's particle content. This shows the merit of the separation of electromagnetic and hadronic showers (chapter 6) together with a calibration that takes the probability of a calorimeter cluster to originate from electromagnetic showers into account. In the full acceptance of the CJC the P_{T}–balance is between 0.96 and 0.98 for small values of $f_{\mathrm{em,jet}}$. In this region a large hadronic energy fraction

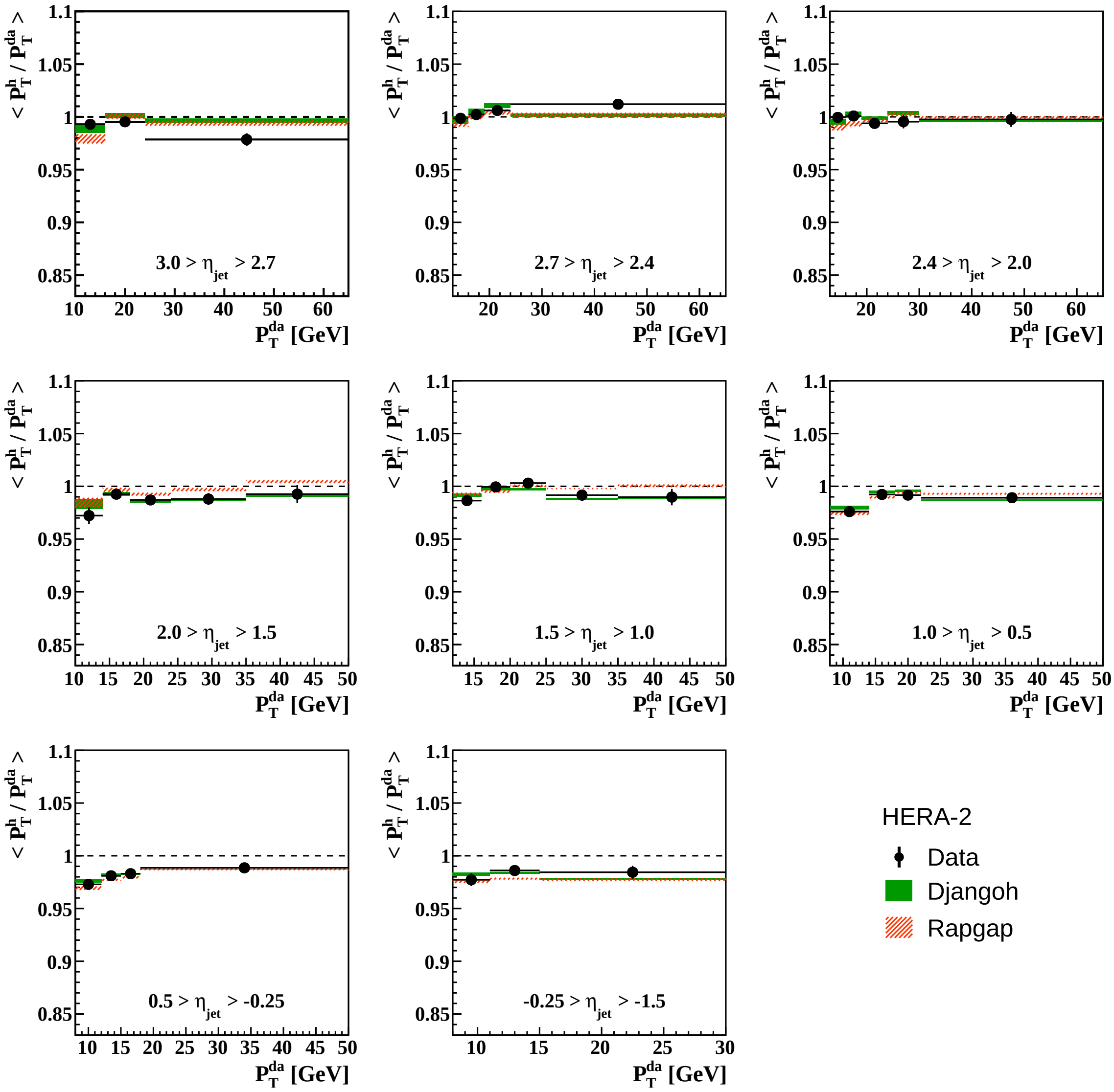


Figure 8.11: Mean values of the P_T-balance distributions as function of P_T^da in bins of η_jet using the full HERA-2 dataset. Shown are data, DJANGOH and RAPGAP after the full calibration with the cuts from section 8.2.1. The mean values are obtained from fits to Student's t-distributions (8.1).

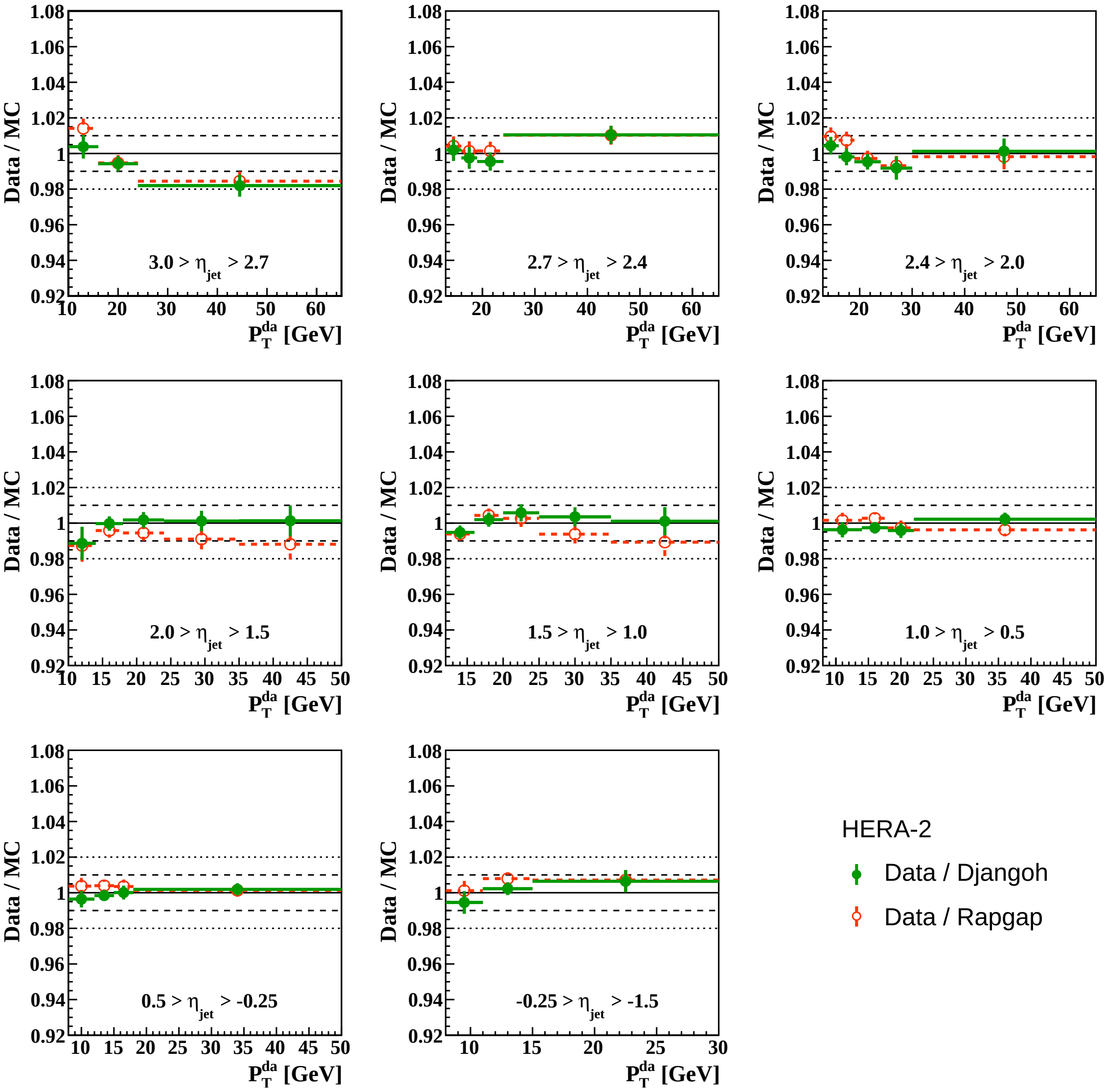


Figure 8.12: Double-ratio of the P_{T}-balance as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in η_{jet} bins corresponding to figure 8.11. The dashed (dotted) line represents a 1% (2%) deviation.

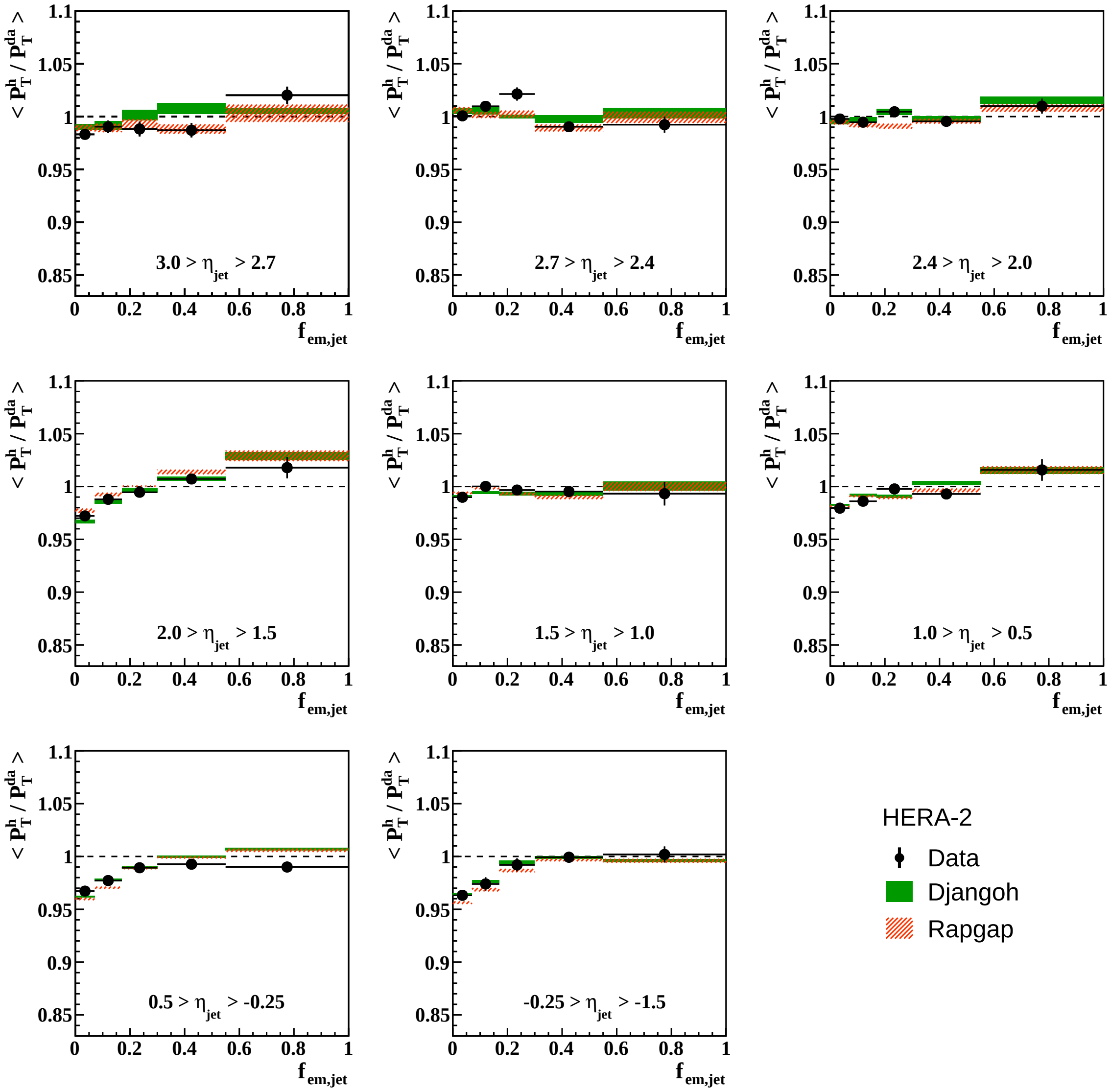


Figure 8.13: Mean values of the P_T-balance distributions as function of $f_{\mathrm{em,jet}}$ in bins of η_{jet} for the full HERA-2 dataset. Other details are given in the caption of figure 8.11.

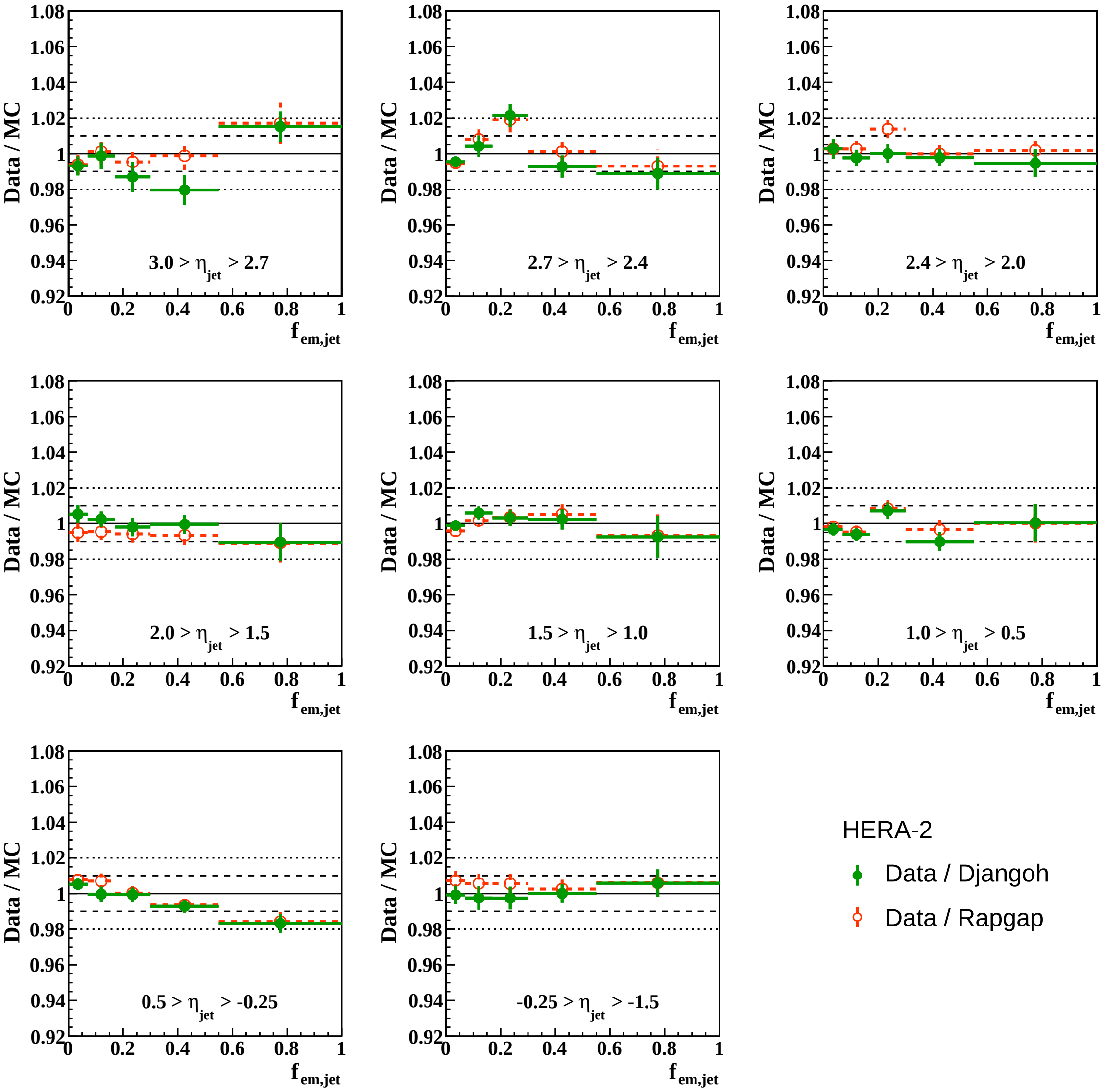


Figure 8.14: Double-ratio of the P_{T}-balance as function of $f_{\mathrm{em,jet}}$ in η_{jet} bins corresponding to figure 8.13. The dashed (dotted) line represents a 1% (2%) deviation.

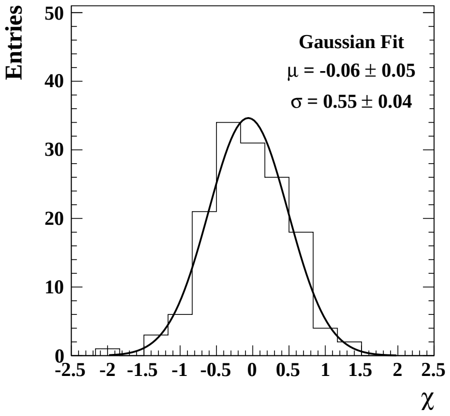


Figure 8.15: Pull distribution (8.12) obtained for mean values of P_T-balance distributions in bins of P_T^da, $f_\mathrm{em,jet}$ and η_jet. The assumption of $\epsilon_\mathrm{h} = 1\%$ was used.

is equivalent to a large track fraction and only a small contribution from hadronic clusters to the jet's reconstructed energy. This means that the energy flow algorithm removed too many clusters from the list of HFS objects and only unphysical large values of the calibration functions could bring the P_T–balance to one in this region. At large values of $f_\mathrm{em,jet}$ the P_T–balance approaches one. This behaviour is well described by the simulation. In figure 8.14 the double-ratio is shown as function of $f_\mathrm{em,jet}$ in bins of η_jet. The data is well described by both models with deviations not larger than 1%.

A way to quantify the differences between data and simulation uses the value χ, also called pull, which is defined as

$$\chi = \frac{\mu_\mathrm{Data} - \mu_\mathrm{MC}}{\epsilon_\mathrm{Data} \oplus \epsilon_\mathrm{MC} \oplus \epsilon_\mathrm{h}} \,. \tag{8.12}$$

The values μ_Data and μ_MC in equation (8.12) are the mean values obtained from fits to the P_T–balance distributions and ϵ_Data and ϵ_MC are their respective errors. The value ϵ_h is the uncertainty assigned to the HFS energy measurement. The distribution of pulls is illustrated in figure 8.15 for all data points from figures 8.11 and 8.13 for the assumption of $\epsilon_\mathrm{h} = 1\%$. The distribution is centred around zero owing to the good agreement between data and simulation. The width of the Gaussian fit is 0.55 suggesting that an error of 1% is already a conservative estimate. Setting the error from the fits to the P_T–balance distributions ϵ_Data and ϵ_MC to zero, the width of the pull distribution becomes 0.65, staying well below one.

The calibration is also tested on an independent DIS sample by requiring two jets instead of only one together with the calibration sample selection (see section 8.2.1). The cut on the distance between the jet's polar angle and the hadronic angle is released, which results in a sample with more hadronic activity outside the jets. The transverse momenta of jets lie between 6 and 50 GeV. The mean values of the P_T–balance together with the double-ratios as function of the hadronic angle γ_h are shown in figure 8.16a. For $\gamma_\mathrm{h} < 120°$ the P_T–balance is flat around unity, but when the hadronic system is measured in the backward direction the P_T–balance drops to values of about 0.93. The reason for this are losses in the backward direction due to the missing hadronic calorimeter behind the BBE. This behaviour is well

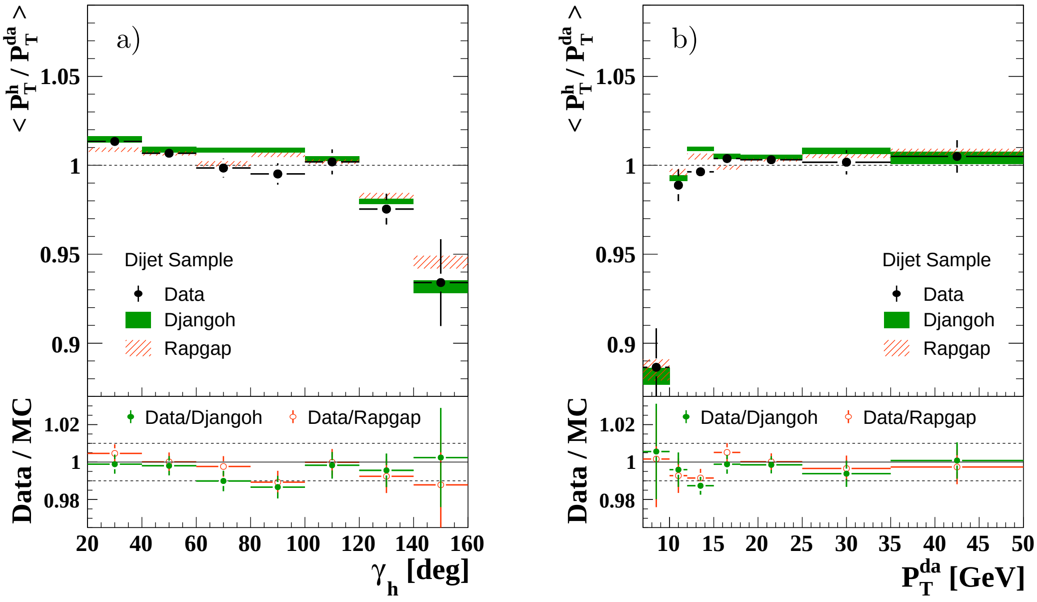


Figure 8.16: Mean values of the P_T-balance as function of γ_h (a) and P_T^da (b) for a dijet selection using the full HERA-2 dataset. Shown are data, DJANGOH and RAPGAP with the respective double-ratios after the full calibration. The mean values are obtained from fits to a Student's t-distribution (8.1). The dashed lines in the ratio plot illustrate a 1% deviation.

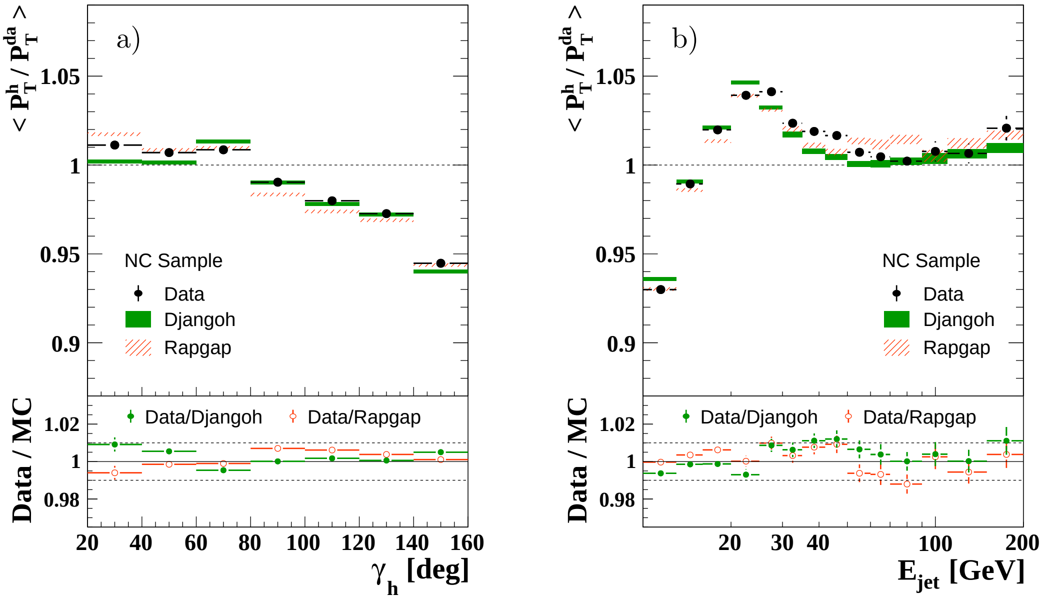


Figure 8.17: Mean values of the P_T-balance distributions as function of γ_h (a) and E_jet (b) for a neutral current selection using the full HERA-2 dataset. Other details are given in the caption of figure 8.16.

simulated and the double-ratio is within 1% for both MC models. In figure 8.16b the P_{T}–balance is shown as function of $P_{\mathrm{T}}^{\mathrm{da}}$. Values of $P_{\mathrm{T}}^{\mathrm{da}}$ smaller than 10 GeV correspond to large values of γ_{h} with mean values of the P_{T}–balance around 0.9. For $P_{\mathrm{T}}^{\mathrm{da}}$ ¿10 GeV the P_{T}–balance is flat around unity and described by the simulation within 1%.

The calibration scheme developed is furthermore tested on an inclusive neutral current sample with $Q^2 > 150\,\mathrm{GeV}^2$ (see section 9.6 for the selection) and one jet with $P_{\mathrm{T}}^{\mathrm{jet}} > 6\,\mathrm{GeV}$ in the final state. No explicit cuts to suppress QED radiation or losses are applied, which results in a very different distribution of the jets' polar angle compared to the calibration sample (also see figure 8.5). In the inclusive NC sample less activity in the forward region and more jets in the central and backward direction make it a good testing ground for the calibration in the central region of the detector. The mean values of the P_{T}–balance as function of γ_{h} are shown in figure 8.17a together with the double-ratios. In this sample reduced values of the P_{T}–balance are already visible for $\gamma_{\mathrm{h}} \sim 90^\circ$. The distribution is well described by both models within 1% uncertainty with very high statistics even in the backward direction $140^\circ < \gamma_{\mathrm{h}} < 160^\circ$. In figure 8.17b the P_{T}–balance as function of the reconstructed jet energy is shown. The shape of the P_{T}–balance is the result from a bias towards downwards fluctuations for small values of E_{jet} and upwards fluctuations for large values of E_{jet}. Also biases in the double-angle measurement towards too small values of $P_{\mathrm{T}}^{\mathrm{da}}$ (cf. figure 8.3) influence the shape of this distribution, being responsible for values larger than one for jet energies around 20 to 40 GeV. The observed peak in this energy range is absent when the calibration sample selection is applied, which removes the biases of the double-angle measurement towards too small values. All observed features are well modelled and the data are described within 1% over an energy range between 10 to 200 GeV.

8.2.5 Resolutions

The resolution of the jet measurement can be obtained either directly from the data with the help of a well defined reference measurement or indirectly from the simulation.

The first method makes use of the double-angle measurement with the appropriate cuts defined above (section 8.2.1). Only events with one reconstructed jet are considered such that the P_{T}–balance can be measured as function of the polar angle of the jet. The transverse momentum of the generated jet agrees within 5% with the total generated transverse momentum and thus with $P_{\mathrm{T}}^{\mathrm{da}}$ in this sample. Consequently the resolution of the jet transverse momentum measurement, $\sigma(P_{\mathrm{T}})/P_{\mathrm{T}}$, can be derived from measuring the width of the P_{T}–balance distribution $\sigma\left(P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{da}}\right)$, which is shown in figure 8.18 as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in different bins of η_{jet} for data and simulation. The data are well described by the simulation over the full detector acceptance and over the full range of transverse momenta. The data points are not corrected for the resolution of the $P_{\mathrm{T}}^{\mathrm{da}}$ measurement, $\sigma(P_{\mathrm{T}}^{\mathrm{da}})$, which is three to four times smaller than the resolution of the HFS measurement. Thus, the data points are approximately equal to $\sigma(P_{\mathrm{T}})/P_{\mathrm{T}}$. The resolution shows a tendency to improve towards larger transverse momenta which is in good agreement with the expectations. The worst resolution of about 14% is observed for low transverse momenta in

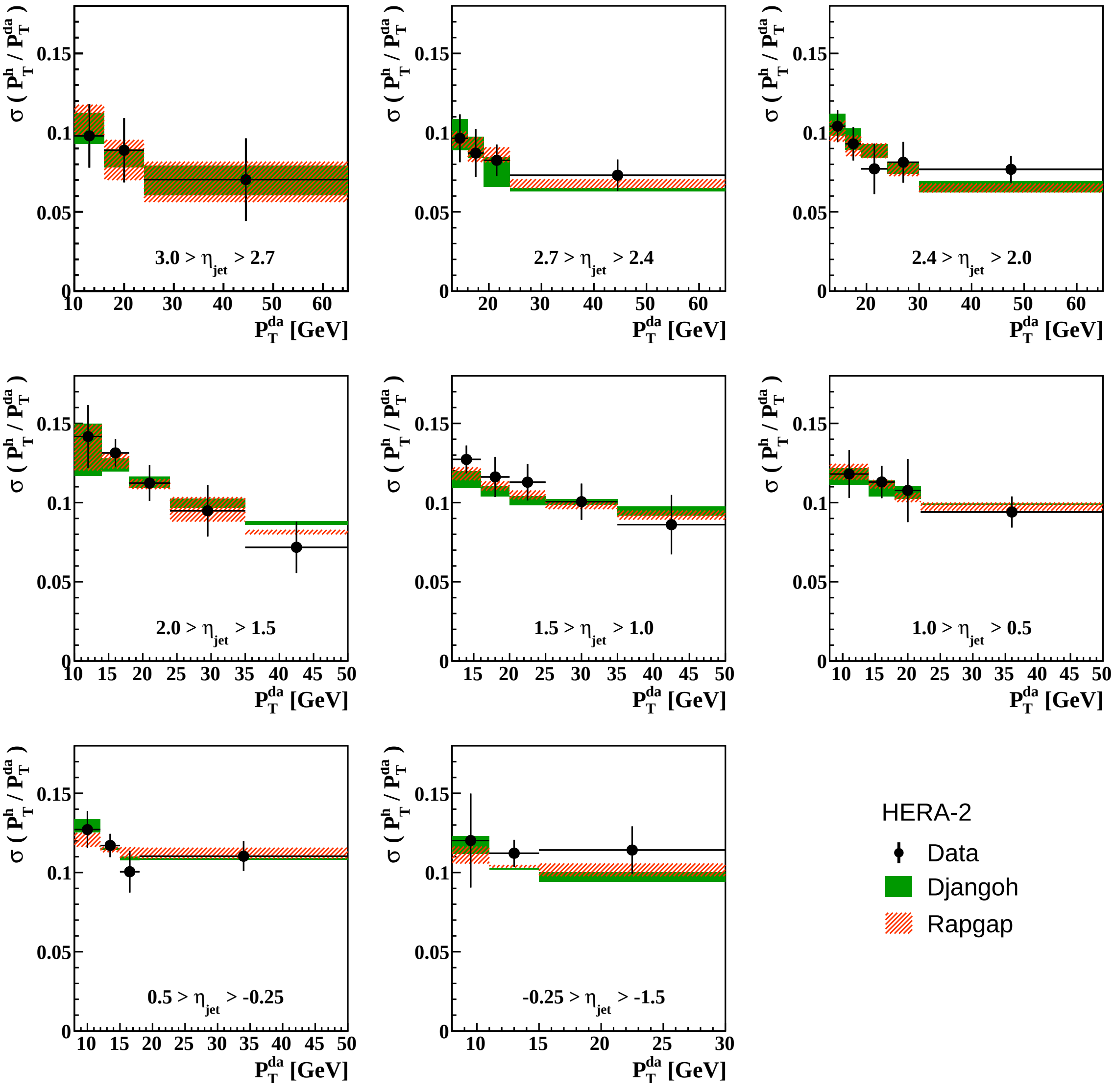


Figure 8.18: Variance of the P_{T}-balance as function of $P_{\mathrm{T}}^{\mathrm{da}}$ in various η_{jet} bins. Shown are results obtained with the full HERA-2 dataset for data and simulation.

the region $2 < \eta_{\mathrm{jet}} < 1.5$, which corresponds to the FB2/OF. Jets in this region of pseudorapidity traverse a relatively large amount of dead material before entering the active medium of the calorimeter. This leads to worse calorimetric resolutions, especially at small energies, as well as worse resolutions for the separation of electromagnetic and hadronic showers. Also a low track-finding efficiency due to a small overlap with the CJC (see figure 8.9) leads to worse resolutions for jets with low transverse momenta in this angular region. Jets with high P_{T} measured in the OF have a resolution of 8%, which is comparable with the resolution achieved in the IF ($\eta_{\mathrm{jet}} > 2$). In the central region the resolutions are relatively flat as function of the transverse momentum compared to the calorimetric expectation which is a consequence of the used tracks and a high track finding efficiency.

To subtract the resolution of the $P_{\mathrm{T}}^{\mathrm{da}}$ measurement from the measured ratio $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{da}}$, the resolution of the double-angle measurement is obtained from the simulation. It is 1.4–2% if the jet is measured in the forward region and between 3% and 4% if

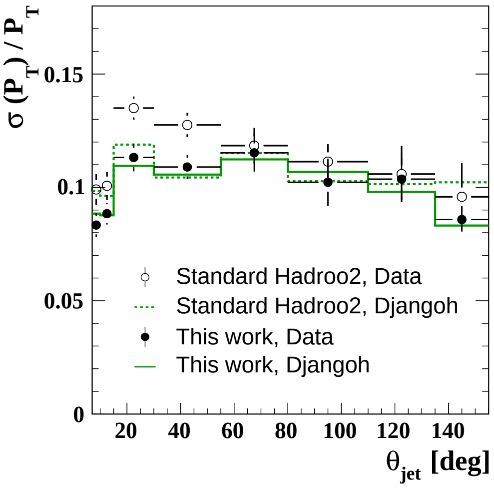


Figure 8.19: Resolution of the jet transverse momentum measurement as function of θ_{jet}. The resolution is obtained from the variance of the P_{T}-balance, where the resolution of the double-angle measurement is subtracted assuming Gaussian error propagation. Shown are results obtained with the standard reconstruction (open circles, dashed lines) and with the improved reconstruction together with the new calibration (filled circles, solid lines) for data and simulation.

the jet has a polar angle larger than 80°. It is subtracted from the P_{T}–balance measurement assuming Gaussian error propagation,

$$\sigma(P_{\text{T}})/P_{\text{T}} = \sqrt{\left[\sigma\left(P_{\text{T}}^{\text{h}}/P_{\text{T}}^{\text{da}}\right)\right]^2 - \left[\sigma\left(P_{\text{T}}^{\text{da}}/P_{\text{T}}^{\text{gen}}\right)\right]^2}\,. \tag{8.13}$$

Equation (8.13) is to a good approximation fulfilled, since $P_{\text{T}}^{\text{h}}/P_{\text{T}}^{\text{da}}$ and $P_{\text{T}}^{\text{da}}/P_{\text{T}}^{\text{gen}}$ are very close to unity. The obtained resolutions are shown in figure 8.19 as function of θ_{jet} where the standard reconstruction Hadroo2 [PP+05] is compared with the reconstruction developed in the course of this work. For $\theta_{\text{jet}} < 55°$ improvements between 10–15% are achieved for the data.

In the simulation the largest improvement is obtained for jets incident in the IF with $\theta_{\text{jet}} < 15°$, in the OF region the improvement is of the order of 5%. In the central region of the detector with $55° < \theta_{\text{jet}} < 135°$ the resolution is dominated by tracks and stays unchanged. The achieved resolutions are about 9% in the IF and between 10 and 11% elsewhere. Additionally to an overall improved HFS measurement the resolutions are better described by the simulation than with the standard reconstruction. The resolution is described within the uncertainty over the full detector acceptance.

The second method to derive the resolution of the jet measurement is an indirect method. Since all quantities related to the HFS measurement like the absolute value and width of the P_{T}–balance are very well modelled the resolutions can be obtained from the simulation. This brings the advantage of the exact knowledge of the incident energy. Reconstructed jets are matched with jets on hadron level in the η-ϕ plane, and only jets with a match are considered in this study. The relative difference between the reconstructed and generated jet P_{T} is given by

$$\Delta_{P_{\text{T}}}^{\text{jet}} = \frac{P_{\text{T,gen}}^{\text{jet}} - P_{\text{T,rec}}^{\text{jet}}}{P_{\text{T,gen}}^{\text{jet}}} \tag{8.14}$$

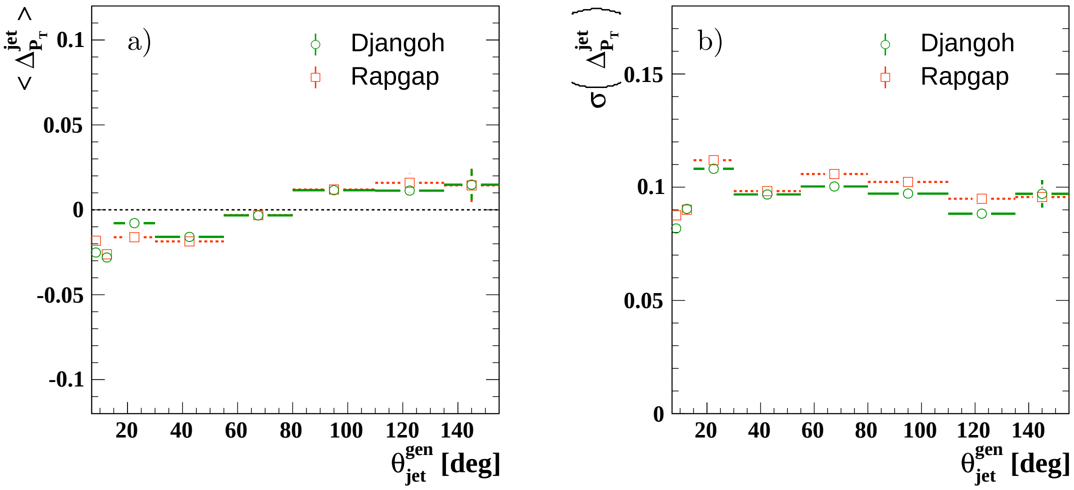


Figure 8.20: Mean values (a) and widths (b) of the resolution of the jet transverse momentum $\Delta_{P_\mathrm{T}}^{\mathrm{jet}}$ as function of the polar angle of the generated jet.

in analogy to the definition of Δ_E (6.29). The mean values of $\Delta_{P_\mathrm{T}}^{\mathrm{jet}}$ are shown in figure 8.20a. Overall good agreement between the reconstructed and generated transverse momenta is observed. In the forward region with $\theta_{\mathrm{jet}} < 55°$, the reconstructed P_T is on average 2–3% too large, whereas in the backward region it is about 1% too small. These differences are not necessarily only due to the calibration. Particles attributed to the jet on hadron level but not on reconstructed level or vice-versa may have an effect on $\Delta_{P_\mathrm{T}}^{\mathrm{jet}}$. The observed shifts are small and are taken into account when jets are corrected for detector effects. In figure 8.20b the resolution of the jet transverse momentum measurement as defined in equation (8.3) is shown. The resolution is about 9% in the IF and 11% in the region of the OF, which is in excellent agreement with the results obtained for data (cf. figure 8.19). In the central region of the detector the obtained resolutions from both MC models agree with the results from the double-angle method within one absolute percent.

In order to obtain resolutions directly comparable to the calorimeter resolutions acquired from test beam measurements, generated jets are selected in a small interval of 2% around a central energy E_{gen}. In each of these energy intervals $[0.98E_{\mathrm{gen}}, 1.02E_{\mathrm{gen}}]$ the distribution of the reconstructed jet energy is then fitted to equation (8.1) for all matching jets. An example of this is given in figure 8.21 for jets incident on the CB3. From these fits the widths of the distributions are obtained from the FWHM (8.3) and corrected for the effect of the smeared generated energy by $\sigma(E) = \sigma - 0.017\,E_{\mathrm{gen}}$. This correction is calculated from the width of a superposition of Gaussian distributions with shifted means of ±2%. The acquired resolutions for the different calorimeter wheels are compared with the calorimeter resolution from test beams with charged pions (see section 5.4.2) in figure 8.22.

In the region where tracks contribute largely to the HFS, $30° < \theta_{\mathrm{jet}} < 140°$, corresponding to the BBE, CBs and FB (cf. figure 7.11), the achieved resolution is improved with respect to the calorimetric expectation. Jets in this angular interval are measured with a precision of approximately 10% on average over the full energy range. In the CB1 and CB2 the resolution improves from about 11% at $\sim$8 GeV to about 8% at 11–13 GeV and flattens out around 10% at the highest kinematically

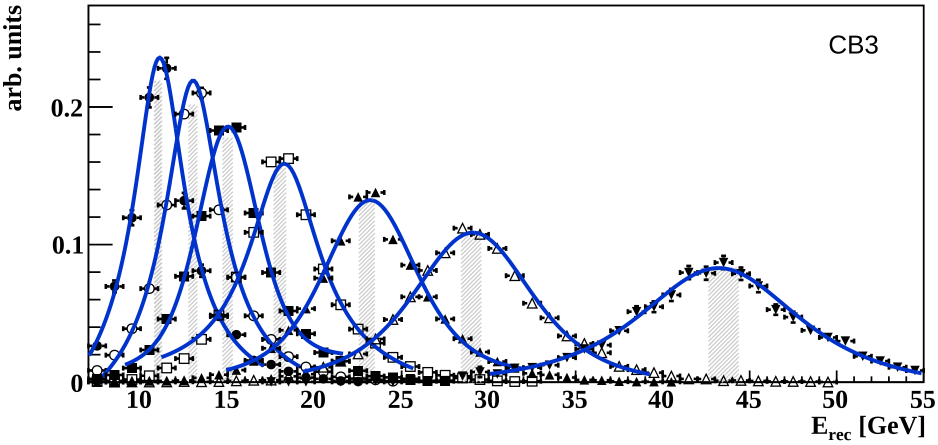


Figure 8.21: Distributions of reconstructed jet energies for simulated jets incident on the CB3. The generated energies are shown as dashed grey areas. Fits to equation (8.1) are shown with blue lines.

possible energies. The best resolution of about 8% corresponds to an energy region with high track finding efficiency, good tracking resolution and good calorimetric resolution. At low energies the tracking provides a precise measurement, but the calorimetric resolution is worse which reflects in a a slightly worse resolution for jets. At the highest energies the resolution is dominated by the calorimetric measurement, however, low energy hadrons which are abundantly produced even in high energy jets are accounted for by tracks. At these energies the resolutions for jets become comparable to the calorimetric resolution for charged pions over the full coverage of the CJC. In the CB3 and FB the resolution is relatively flat at a value of 10% over the full energy range. In the forward region, $\theta_{\mathrm{jet}} < 30^\circ$, corresponding to the OF and IF, the contribution from tracks to the jet's transverse momentum becomes small, and the measurement can be considered as purely calorimetric. The obtained resolutions are $72\%/\sqrt{E} \oplus 3.5\%$ in the OF and $83\%/\sqrt{E}$ in the IF, which is worse than the resolution from test beam measurements and single hadrons (see below). This can be explained by the large number of low energy hadrons in jets, which result in an overall smaller electromagnetic fraction of showers compared to the one induced by single high energy hadrons. Also fluctuations from the jet-defining algorithm can contribute to larger resolutions for jets [Wig00, p254].

In the OF a constant term of $\sim 3.5\%$ is observed, which can be attributed to the amount of dead material in front of this calorimeter wheel. In the IF no constant term is needed to describe the resolution, which shows that compensation is achieved over the full energy range with the H1 software-weighting in combination with the separation of electromagnetic and hadronic showers and the calibration method developed.

It is interesting to compare for the IF jet energy resolutions to the resolutions obtained from simulated neutral and charged single pions (see figure 8.23). Electromagnetic showers are measured with a comparable resolution as obtained from electron test beams, with a slightly larger noise term of $\sim$0.6 GeV. For charged pions, events with more than 1% of the total measured energy in the tail catcher were discarded to obtain a calorimetric measurement not biased by energy leakage. The resolution

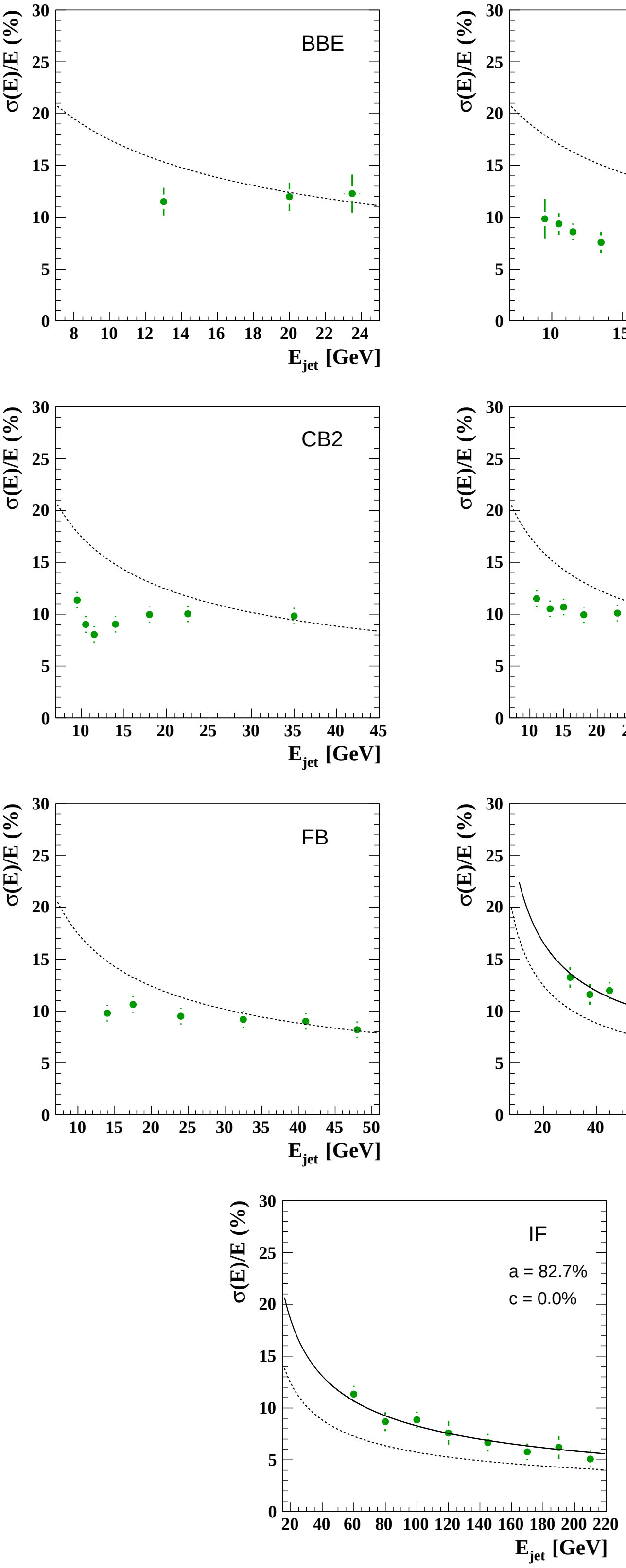


Figure 8.22: Resolutions for jets incident on different calorimeter wheels obtained with the Djangoh MC. The dashed line represents the resolution determined from pion test beams with $a_{samp} = 55\%$ and $c = 1.6\%$ (see figure 5.13b). In the calorimeter regions where tracks play a minor role a fit to equation (5.37) was performed, the obtained values are shown in the corresponding figures.

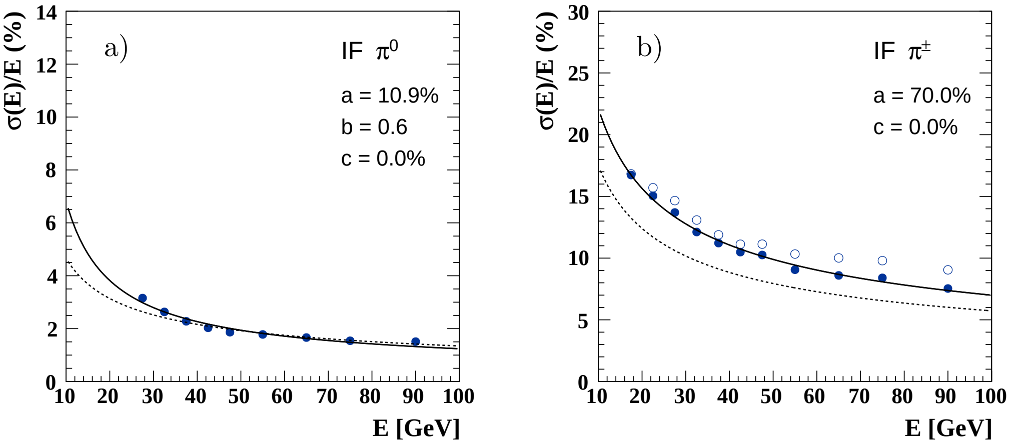


Figure 8.23: Resolutions for simulated neutral (a) and charged (b) single pions incident on the IF. The dashed lines represent the resolutions determined from test beams with $a_{\mathrm{samp}} = 13.3\%$, $b_{\mathrm{noise}} = 0.2\,\mathrm{GeV}$ and $c = 0$ for electromagnetic showers [H194b] and $a_{\mathrm{samp}} = 55\%$ and $c = 1.6\%$ for hadronic ones (see figure 5.13b) in the IF. In the case of hadronic showers the obtained resolution for events with no signal in the tail catcher (filled circles) and without any further selection (open circles) is shown. The results from fits to equations (5.31) and (5.37) for electromagnetic and hadronic showers, respectively, are given in the figures.

obtained for hadronic showers is $70\%/\sqrt{E}$ which is slightly worse than the resolution from test beams. However, no constant term is observed showing the applicability of the developed calibration method to single particles. Releasing the condition on the tail catcher and thus allowing shower leakage the resolution becomes worse (open circles in figure 8.23b). The deterioration of the resolution due to leakage can be approximated by

$$\left(\frac{\sigma(E)}{E}\right)_f = \left(\frac{\sigma(E)}{E}\right)_{f=0} (1+4f)\,, \tag{8.15}$$

if a shower fraction f leaked out of the calorimeter [Mav04]. Using this equation and the difference in resolutions with and without a measurable signal in the tail catcher, the leakage is about 1% for 20 GeV pions, 3% for 55 GeV pions and 5% for 90 GeV pions. These values are in good agreement with the calculated average longitudinal energy leakage for the ATLAS TileCal with a comparable effecitve length (7.9λ at $|\eta| = 0.35$) [AA$^+$10b]. Applying the condition on the tail catcher for the jet sample the obtained resolutions do not change measureably.

Chapter 9

Data Selection

The measurement of jet cross sections in neutral current (NC) deep-inelastic scattering (DIS) presented in this work is based on the full HERA-2 dataset. Selection criteria are applied on the recorded dataset in order to reduce the sample, which contains a significant amount of background from *ep* and non-*ep* interactions, to the data sample used to extract the cross sections. The inclusive NC sample is defined through the scattered electron, which is also used as the main requirement for the trigger. The jet samples are subsamples of the NC sample and therefore benefit from the high trigger efficiency and excellent signal-to-background ratio.

In the following the applied selection criteria for the inclusive NC sample and the jet samples are discussed. Trigger efficiencies and the efficiency of the electron selection through its associated track are determined. Cuts to reduce non-*ep* background are applied, and the remaining background from *ep* scattering is estimated with dedicated MC simulations. This is followed by the reconstruction of the boost to the Breit frame and the jet selection. The chapter concludes with control distributions comparing the data with simulated events for the various jet samples.

9.1 Run Selection and Luminosity Measurement

During the data taking events were recorded in time intervals called runs. A run is defined as a period with nominally stable experimental conditions, with a maximum duration of two hours. Based on criteria like the overall detector performance, beam, background and readout conditions, a run is classified as “good”, “medium” or “poor”. In this analysis a run has to be of good or medium quality with a minimal integrated luminosity of $0.2\,\mathrm{nb}^{-1}$ in order to be selected. In addition, the following subdetectors are required to be fully operational and included in the readout: the LAr calorimeter and calorimeter trigger, the CJC1 and 2, the CIP, the ToF and the luminosity system. The status information from the various subdetectors was stored on the database every ten seconds during data taking. A full run is rejected, if any of the mentioned subsystems was not operational for a large fraction of time during this run. Single events are rejected, if any of the subdetectors was off when it was recorded.

MC events are simulated for each run separately, taking the detector status into account. The run selection used for the data is applied to simulated events, such that the same experimental conditions are found in the simulation as in the data.

Period	Lepton	Run Range	f_{QEDC}	L in pb^{-1}
03/04	e^+	357160 – 392213	1.005	52.3
05	e^-	399629 – 436893	0.997	101.9
06	e^-	444312 – 466997	1.015	57.7
06/07	e^+	468530 – 500611	1.045	139.7
HERA-2		357160 – 500611	1.020	351.6

Table 9.1: Selected run periods together with the total integrated luminosities. The average luminosity correction factor from the QEDC analysis is denoted by f_{QEDC}.

The total integrated luminosity is calculated by summing over the integrated luminosities of all selected runs. At the time of writing, the offline analysis of the data from the luminosity system is not completed (cf. section 4.2.4). Taking into account the uncertainty on the pile-up and background subtraction, the correction for the beryllium shielding and the acceptance correction, the uncertainty of the luminosity measurement from the luminosity system is 3–4.5%, depending on the run period. A way to reduce the systematic uncertainty is to use the rate of QEDC events, where the scattered electron and the radiated photon are both measured in the central part of the H1 detector. This measurement is dominated by different sources of systematic uncertainties than the measurement from the luminosity system. The most prominent systematic uncertainties are the subtraction of background from Deeply Virtual Compton Scattering (DVCS) processes, the vertex reconstruction from CIP and SpaCal data only, and the reconstruction of the four-vector of the radiated photon from the SpaCal cluster [SS09]. A systematic uncertainty of 2.5% of the luminosity measurement is obtained by combining the two independent measurements. A summary of the selected run ranges together with the corresponding integrated luminosities and the applied average correction factors is given in table 9.1.

The standard model predicts a difference in the NC cross section for leptons with different helicity states. To measure this effect spin rotators were installed at HERA-2 to provide longitudinally polarised lepton beams. The measured polarisation asymmetry increases with increasing Q^2 and is negative in e^-p and positive in e^+p scattering [Sou10]. The asymmetry arises from electroweak effects and is not included in present NLO QCD calculations. However, a measurement of the polarisation of the full HERA-2 dataset with the polarimeters TPOL and LPOL [LS02] yields an average polarisation of $-2.1 \pm 3\%$, which is consistent with no polarisation. The effect of the polarisation asymmetry can thus be neglected in this analysis.

9.2 Trigger Efficiency

The trigger used in this analysis is the calorimeter subtrigger S67. In addition to information from the LAr calorimeter, it uses information from the CIP and the ToF for the rejection of background. It is a combination of three level one trigger conditions,

$$\mathrm{S67} = t_{\mathrm{LAr}} \ \wedge \ t_{\mathrm{T0}} \ \wedge \ t_{\mathrm{veto}}\,, \tag{9.1}$$

where $\wedge$ represents the logical conjunction. The calorimeter condition t_{LAr}, the timing requirement t_{T0} and the veto condition t_{veto} are described in the following as well as the determination of their efficiencies.

9.2.1 Calorimeter Trigger Element

The t_{LAr} condition is fulfilled if a large amount of energy is deposited in the LAr calorimeter. For this purpose the 45,000 calorimeter cells are grouped into 256 big towers. The signals measured in each big tower are summed up and if the deposited energy is above 4.8 GeV in the BBE and CBs or above 6.2 GeV to 25.6 GeV in the FBs and IF, the t_{LAr} condition is fulfilled. The big towers can be triggered by the scattered electron or the HFS. These conditions are independent and thus the sample triggered by the electron can be used to monitor the HFS-triggered sample and vice versa. Let n_{eh} be the number of events that were triggered by both, the electron and HFS, n_{e} the number of events triggered by the electron only and n_{h} the number of events triggered by the HFS only. Then the efficiency of the trigger to be fired by the electron $\epsilon(\mathrm{e})$ is given by

$$\epsilon(\mathrm{e}) = \frac{n_{\mathrm{eh}}}{n_{\mathrm{h}} + n_{\mathrm{eh}}} \tag{9.2}$$

with a similar expression for $\epsilon(\mathrm{h})$, which is the efficiency of the trigger to be fired by the HFS. The total efficiency of the t_{LAr} condition, $\epsilon(\mathrm{LAr})$, can then be written as

$$\epsilon(\mathrm{LAr}) = \epsilon(\mathrm{e}) + \epsilon(\mathrm{h}) - \epsilon(\mathrm{e})\epsilon(\mathrm{h})\,, \tag{9.3}$$

where $\epsilon(\mathrm{e})$ and $\epsilon(\mathrm{h})$ denote the efficiency of the LAr calorimeter trigger to be fired exclusively by the electron or HFS, respectively.

The total efficiency $\epsilon(\mathrm{LAr})$ is close to 100%. However, in some localised regions of the calorimeter lower efficiencies can be observed. Typically, the reasons are switched off calorimeter cells because of high noise levels or malfunctioning hardware. A detailed investigation of the trigger efficiency as function of the impact position of the electron on the calorimeter reveals regions with low trigger efficiencies. The removal of these inefficient calorimeter regions constitutes the fiducial volume cut, where regions in the ϕ- and z-coordinate of the electron impact position are excluded from the analysis. This fiducial volume cut is time dependent and is determined for each HERA-2 period separately [Hab09].

Figure 9.1a shows a summary of the LAr calorimeter trigger efficiency $\epsilon(\mathrm{LAr})$ for the full HERA-2 dataset. The efficiency is shown as function of the energy of the scattered electron E_{e} before and after the fiducial volume cut. Prior to the cut the efficiency has a strong dependence on E_{e}, rising from 95% at $E_{\mathrm{e}} \approx 11$ GeV to 100%

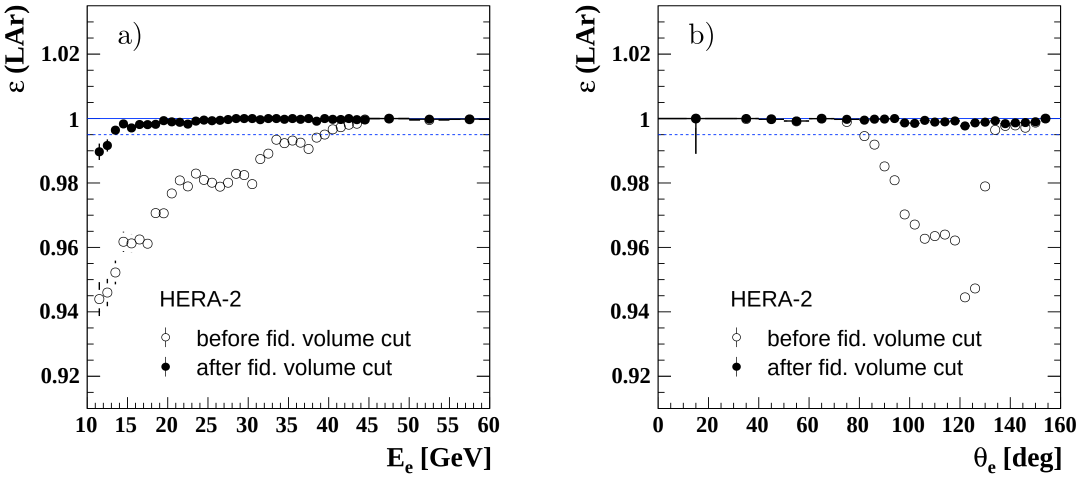


Figure 9.1: The efficiency of the calorimeter trigger element, $\epsilon(\mathrm{LAr})$ as function of the energy (a) and polar angle (b) of the scattered electron. Shown are the measured efficiencies for the full HERA-2 dataset before and after the fiducial volume cut is applied. The dotted line illustrates an efficiency of 99.5%.

at $E_\mathrm{e} \approx 45\,\mathrm{GeV}$. The fiducial volume cut leads to a sufficiently flat efficiency above $E_\mathrm{e} \approx 13\,\mathrm{GeV}$, with values larger than 99.5%. Only at energies between 11–13 GeV an energy dependence of the trigger efficiency is still observed. In this energy region simulated events are weighted to correct for this effect. To avoid regions with trigger efficiencies smaller than 99%, E_e is required to be larger than 11 GeV.

The trigger efficiency as function of the polar angle of the scattered electron θ_e is shown in figure 9.1b. After the fiducial volume cut it is above 99.5% over the full calorimeter acceptance.

The efficiency of the LAr trigger element has also been investigated for each HERA-2 period separately. All periods show identical behaviour after the application of the fiducial volume cut. The efficiencies agree with the observation for the combined HERA-2 dataset, with efficiencies always above 99.5%. An uncertainty of 0.5% is attributed to the efficiency of the t_LAr trigger condition.

9.2.2 Timing Condition

The timing condition t_T0 combines the event timing T_0 from the LAr calorimeter and the CIP. An event is accepted if either the CIP or the LAr calorimeter trigger provide a T_0 signal. The CIP also provides timing information from an earlier bunch crossing which is used to reject events with a wrong calorimeter timing: if the T_0 signal from the LAr calorimeter coincides with the T_0 from the CIP from an earlier bunch crossing, the event is rejected.

The efficiency of the timing condition $\epsilon(\mathrm{T0})$ can be monitored similarly to the efficiency of the calorimeter trigger element $\epsilon(\mathrm{LAr})$. The timing information from the CIP is used to validate the timing information from the LAr calorimeter and vice versa. The efficiency $\epsilon(\mathrm{T0})$ can then be calculated similarly to equation (9.3). The efficiency $\epsilon(\mathrm{T0})$ is illustrated in figure 9.2 as function of the energy (a) and polar angle (b) of the scattered electron for the NC DIS sample of the full HERA-2

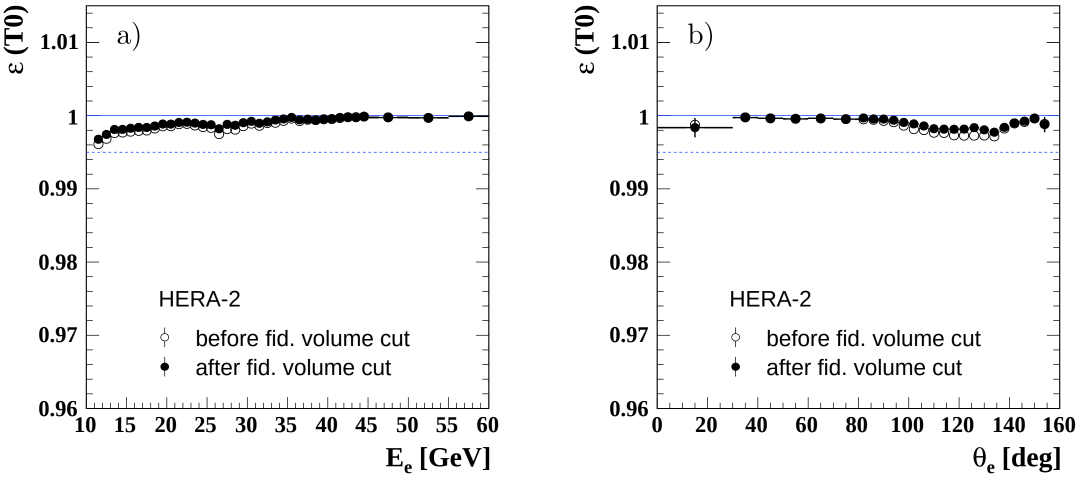


Figure 9.2: Efficiency of the event timing condition t_{T0} for the full HERA-2 dataset. The efficiency is shown as function of the energy (a) and polar angle (b) of the scattered electron. The dotted lines illustrate an efficiency of 99.5%.

dataset. It is always above 99.5%, already before the fiducial volume cut, which improves it slightly. The individual HERA-2 periods show an identical behaviour, with the efficiency ϵ(T0) always above 99.5%. No correction for the timing condition is applied and an uncertainty of 0.5% is assigned to it.

9.2.3 Veto Condition

The veto condition t_{veto} of the subtrigger S67 was used to reject background at the triggering stage. It consists of information from three different subdetectors: the ToF, the CIP and the TC. The ToF trigger bits were delivered from the various plastic scintillation detectors described in section 4.2.3. The CIP provided a fast estimate of the number of tracks in the central region of the detector which could be used to reject background from beam-gas interactions. In 2006 an additional condition was implemented to reject high background rates from beam halo muons which caused the calorimeter trigger element t_{LAr} to fire. It comprised of hits in the backward TC in conjunction with no tracks measured by the CIP. The three different veto conditions arising from the different subdetectors are monitored with the monitor trigger S57, which has no veto condition, except for runs with run numbers smaller than 382137. For these runs a special sample composed of QED Compton events can be used to determine the veto efficiency ϵ(Veto) [Hab09]. The efficiencies of the trigger condition t_{veto} are summarised in table 9.2. They are applied as weights on MC events to simulate the fraction of rejected ep events. The veto efficiency for the e^+p 03/04 period is consistent with the analysis of QED Compton events.

In figure 9.3 the efficiency ϵ(Veto) is shown as function of the run number for the full HERA-2 dataset. The efficiency is stable over time and no significant outliers are observed. A systematic uncertainty of 1% is assigned to the determination of ϵ(Veto).

	Efficiency [%]			
Period	ToF	CIP	Muon	ϵ(Veto)
03/04	99.2 ± 0.2	99.6 ± 0.1	N/A	98.8 ± 0.2
05	99.6 ± 0.1	99.6 ± 0.1	N/A	99.2 ± 0.1
06	99.8 ± 0.1	99.5 ± 0.1	N/A	99.3 ± 0.1
06/07	99.9 ± 0.0	99.6 ± 0.1	100 ± 0.0	99.5 ± 0.1
HERA-2	99.7	99.6	100	99.3

Table 9.2: Summary of the veto efficiencies for the four HERA-2 periods. The statistical uncertainties given for the four subperiods are negligible for the full HERA-2 dataset.

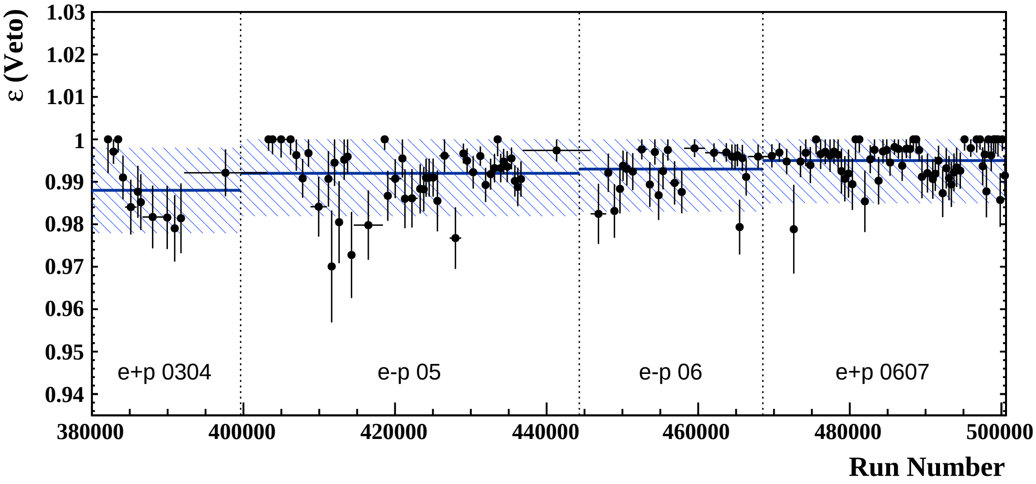


Figure 9.3: Efficiency of the veto condition of the subtrigger S67 as function of the run number. The mean values, which are applied as a correction to the MC simulation (cf. table 9.2), are shown as solid lines for the four different HERA-2 periods separately. The assigned uncertainties are illustrated as dashed areas.

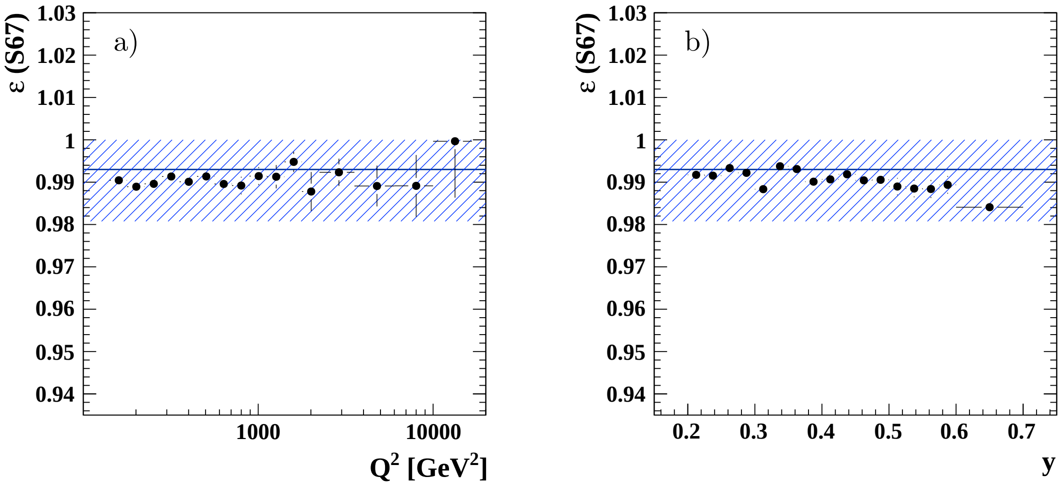


Figure 9.4: Efficiency of the subtrigger S67 as function of Q^2 (a) and y (b) for the full HERA-2 dataset. The MC prediction, after the correction for the veto efficiency has been applied, is illustrated as solid line. The hatched area corresponds to the total systematic uncertainty assigned to the efficiency of the S67.

9.2.4 Efficiency of the S67

Having determined the efficiencies of the three individual components t_{LAr}, t_{T0} and t_{veto} of the subtrigger S67, the total efficiency of the trigger is given by

$$\epsilon(\mathrm{S67}) = \epsilon(\mathrm{LAr})\,\epsilon(\mathrm{T0})\,\epsilon(\mathrm{Veto}). \tag{9.4}$$

This efficiency is shown as function of Q^2 and y for the NC DIS sample of the full HERA-2 dataset in figure 9.4. The solid line shows the MC prediction after the correction for the veto efficiency has been applied. The hatched area corresponds to the systematic uncertainties on $\epsilon(\mathrm{LAr})$, $\epsilon(\mathrm{T0})$ and $\epsilon(\mathrm{Veto})$ added in quadrature. The trigger efficiency $\epsilon(\mathrm{S67})$ is flat at $\sim 99\%$ over the full kinematic range of this analysis, being well within the total systematic uncertainty of 1.2%.

9.3 Rejection of non-*ep* Background

After the veto requirements on the trigger, a small number of events originating from beam-gas collisions, halo muons and cosmic muons is still present in the selected sample. These events are identified with dedicated background finders [Vel02]. These finders are based on topological and timing criteria. The ones used in this analysis are summarised in table 9.3. Events are only rejected if they are classified as background by two finders at least, or if additionally to one of the finders longitudinal or transverse momentum conservation is not fulfilled. The following criteria are applied:

- Ibg0 and Ibg1 or two out of Ibg5, Ibg6 and Ibg7
- Ibg5 or Ibg6 together with $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}} < 0.5$
- Ibg7 and $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}} < 0.1$

Finder	Name	Background	Description
Ibg0	HaLAr	beam-halo μ	Longitudinal energy deposit in the LAr calorimeter
Ibg1	HaMuLAr	beam-halo μ	Longitudinal energy deposit in the LAr calorimeter with energy deposit in the backward TC
Ibg5	CosMuMu	cosmic μ	Two opposite matching muon tracks
Ibg6	CosMuLAr	cosmic μ	At least one muon with more than 90% of its energy deposited in the LAr calorimeter
Ibg7	CosTaLAr	cosmic μ	Two opposite clusters in the TC with matching deposits in the LAr calorimeter
Ibgam0	BeamGas	beam-gas	Ten or more non-vertex fitted tracks in the backward region

Table 9.3: Non-ep background finders used in this analysis. The H1 internal numbers and names are given in the columns one and two.

- Ibgam0 with $E - p_z > 57\,\mathrm{GeV}$ and $N_{\mathrm{part}} > 50$

The first three criteria are efficient for the rejection of cosmic and halo muons [Nik07], where the fraction of rejected events in the inclusive NC sample is 0.2%. The last criterion is used to reject events from beam-gas collisions, which are characterised by a large number of non-vertex fitted tracks together with a large particle multiplicity and a longitudinal momentum imbalance. The cut was optimised in a previous jet analysis on HERA-2 data to efficiently reject beam-gas events [Gou08]. In the inclusive NC sample the fraction of rejected events by this cut is 0.4%. After the application of all background cuts the inclusive NC sample is effectively free of non-ep background.

9.4 Electron Track Requirement

After the identification of an electron candidate through cluster estimators (see section 7.2), a track is associated with it. This track can either be a vertex-fitted track, termed DTRA track, or a non-vertex fitted track which is referred to as DTNV track. The requirement of a track pointing to the identified calorimeter cluster of the scattered electron reduces the background from neutral particles misidentified as the scattered electron, which is mainly due to photons from π^0 decays. Additionally, the track's excellent resolution for the reconstruction of the azimuthal angle of the scattered electron is of importance for the reconstruction of the boost to the Breit frame.

Figure 9.5a shows the distance of closest approach between the cluster of the electron candidate and the associated DTRA track, $d_{\mathrm{ca}}(\mathrm{cluster}, \mathrm{track})$. The distribution peaks around 1 cm, falls off steeply until it reaches values of ~ 5 cm and exhibits a long tail with a relatively gentle slope. The background sample consists of misidentified scattered electrons from simulated low Q^2 and photoproduction events and is normalised to the predicted cross sections from the MC event generators. The distribution of d_{ca} for wrongly identified scattered electrons does not peak at small values of d_{ca}, but rises slowly and flattens out at $d_{\mathrm{ca}} \sim 2$ cm. Therefore, in the peak region with $d_{\mathrm{ca}} < 5$ cm, the data are predominantly genuine NC DIS data. For increasing values of d_{ca} the signal-to-background ratio decreases.

In order to reduce the contamination from wrongly identified scattered electrons, the d_{ca} between the DTRA track and the electron cluster is required to be smaller than 8 cm, if the scattered electron is within the acceptance of the CJC ($\theta_{\mathrm{e}} > 30°$). If no DTRA but a DTNV track is found within 8 cm, the DTNV track is used for the electron validation. In this case the DTNV track is used to determine the position of the primary vertex. The requirement of $d_{\mathrm{ca}} < 8$ cm reduces the amount of wrongly identified scattered electrons from 1.2% to 0.3%.

In simulated NC DIS events the peak of the distribution of $d_{\mathrm{ca}}(\mathrm{cluster}, \mathrm{track})$ lies at smaller values of d_{ca} than in the data. The tail of the distribution has the same slope, but lies below the data by a factor of 1.5. Therefore, the requirement for the electron validation leads to different efficiencies in the data and simulation.

The efficiency of the electron validation is closely related to the efficiency of the vertex finding, since vertex-fitted tracks are required to originate from the same primary vertex. The efficiency of the requirement of having found a primary vertex together with a track associated with an electron cluster is studied with a dedicated NC sample. This sample is obtained with hard cuts on the longitudinal and transverse momentum balance and the application of additional background finders. No constraints on the vertex or on the electron track are made [Nik07, Tra10]. The efficiency of the vertex-track-cluster link requirement, $\epsilon(\mathrm{vtc})$, is calculated as the number of events with a primary vertex and a scattered electron found, divided by the total number of events in this sample. In figure 9.5b the efficiency obtained is shown as a function of the polar angle of the scattered electron. The average efficiency is about 95% in the data. In simulated events the efficiency is between 0.5–2% higher, depending on θ_{e}. The ratio between data and MC is used to derive a correction for $\epsilon(\mathrm{vtc})$, which is determined for the four different HERA-2 periods separately.

The efficiency $\epsilon(\mathrm{vtc})$ is shown for the full HERA-2 dataset in figure 9.6, after the application of the correction to the efficiency. The data are described within 1% by the simulation, independent of the impact position of the scattered electron or the polar angle of the hadronic final state. A systematic uncertainty of 1% due to the vertex and electron-track requirement is assigned.

9.5 Interaction Vertex

A precise knowledge of the primary interaction vertex is important for the calculation of the kinematic variables and the reconstruction of the particle four-vectors in an event. Due to the small transverse size of the beams of $\sigma_x \times \sigma_y = 112 \times 30\,\mu\mathrm{m}$, only small variations of the x- and y-coordinates of the primary vertex are possible

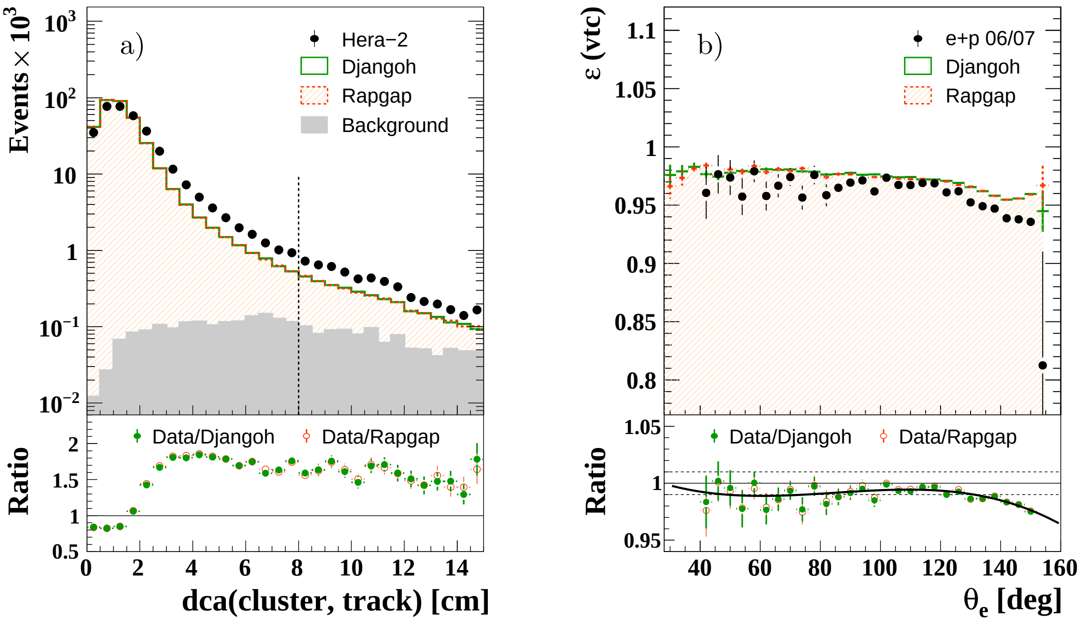


Figure 9.5: a) Distance of closest approach between the cluster of the scattered electron candidate and the associated vertex-fitted track. The requirement $d_{\mathrm{ca}} < 8\,\mathrm{cm}$ is illustrated as dashed line. b) Efficiency of the requirement of a link between the primary vertex, the electron track and the electron cluster as function of the polar angle of the scattered electron for the $e^{+}p$ 06/07 period. The ratio of data to MC is used to derive a correction for the simulation, shown as solid line.

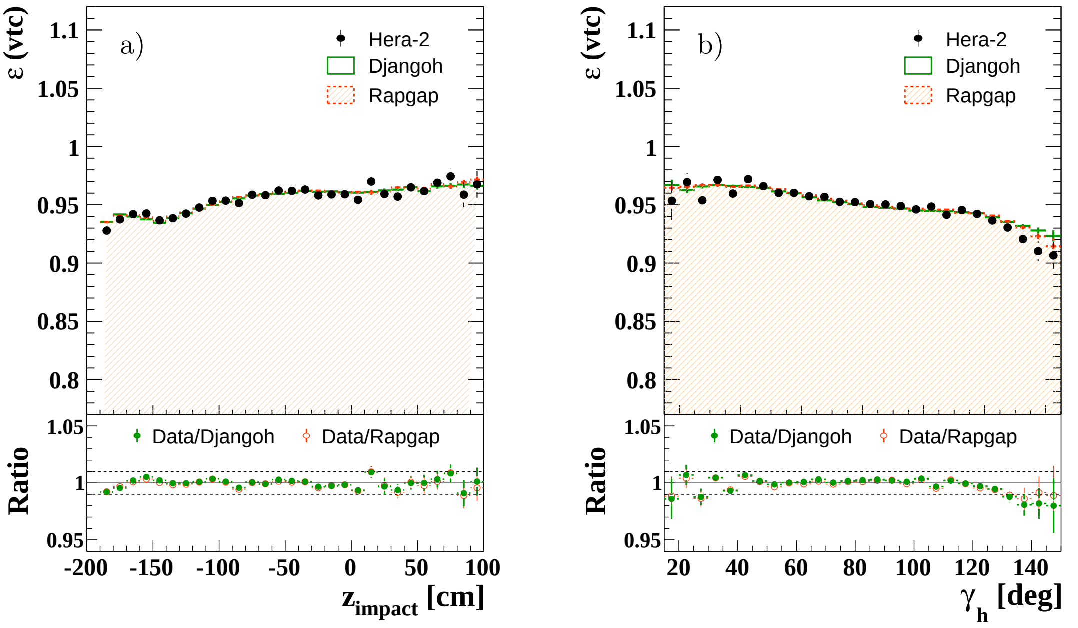


Figure 9.6: Efficiency of the requirement of a link between the primary vertex, the electron track and the electron cluster after the application of the requirement $d_{\mathrm{ca}} < 8\,\mathrm{cm}$. The efficiency correction for $\epsilon(\mathrm{vtc})$ has been applied to the simulation. The efficiency is shown as function of the z-coordinate of the impact position of the scattered electron (a) and the angle of the hadronic final state (b) for the high Q^2 NC DIS sample obtained from the full HERA-2 dataset.

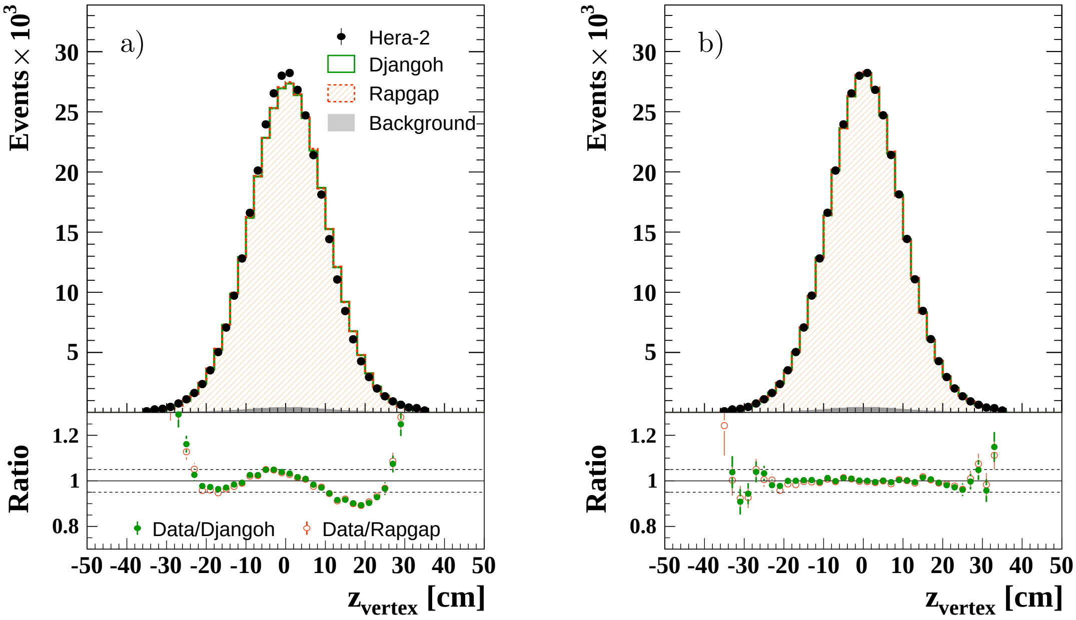


Figure 9.7: Distribution of the measured z-position of the primary vertex. NC DIS data from the full HERA-2 dataset are compared to simulated events prior (a) and after (b) the z-vertex weights have been applied.

during one fill of HERA. The distribution of the z-coordinate of the primary vertex has a larger width because of the larger longitudinal beam sizes ($\sigma_z^{\mathrm{p}} \approx 13\,\mathrm{cm}$ and $\sigma_z^{\mathrm{e}} \approx 2\,\mathrm{cm}$, for the proton and electron beam, respectively). The z-position of the primary vertex is required to be within approximately three standard deviations around the nominal interaction point,

$$-35\,\mathrm{cm} < z_{\mathrm{vertex}} < 35\,\mathrm{cm}, \tag{9.5}$$

to ensure events to be well contained in the detector acceptance. This requirement on the vertex position is also imposed on the luminosity calculation.

Since for each fill the distributions of the x-, y- and z-positions of the primary vertex are approximately Gaussian, the integrated distributions over a full dataset are superpositions of a large number of Gaussian distributions. In order to simulate this effect, the mean and standard deviation of the position of the primary vertex are measured for each run. The simulated position of the primary vertex is then distributed run-dependently according to the measured values.

The distribution of the z-position of the primary vertex is shown in figure 9.7a. A small residual shift between the data and simulation can be observed. This shift is corrected for by a weight applied to simulated events [Shu10]. The resulting distribution of z_{vertex} is shown in figure 9.7b. After the reweighting excellent agreement between data and the simulated z-position of the primary vertex is obtained.

In this sample vertices from electron DTNV tracks are included (see section 9.4). When the electron is validated by a DTNV track only, the HFS particles are likely to have undergone one or more nuclear interactions in the beam pipe. The hadronic vertex may then be displaced from the primary interaction vertex, and the electron track cannot be fitted to the hadronic vertex. However, the electron DTNV track points to the correct primary vertex, which is used for the reconstruction of kinematic

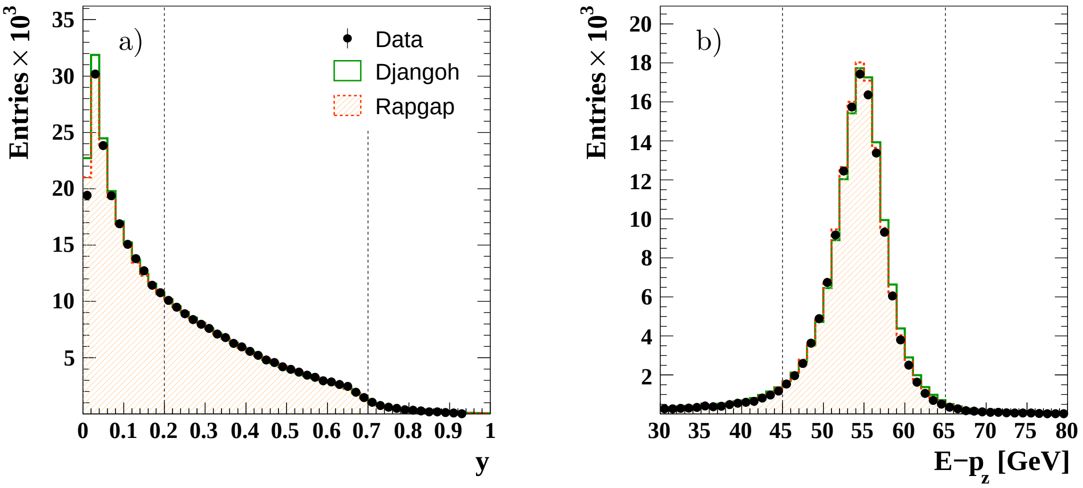


Figure 9.8: The inelasticity y (a) and the longitudinal momentum balance $E-p_z$ (b) for the e^+p 06/07 period. The applied cuts are illustrated as dashed lines.

variables. Tracks originating from HFS particles have been fitted to the correct secondary vertex and can be used without a correction. The application of a nuclear interaction finder at the vertex fitting reduces the amount of events with a DTNV vertex considerably with respect to a previous analysis [Gou08]. The fraction of such events is only about 0.3% and is well modelled by the simulation.

9.6 Neutral Current Sample

The inclusive NC sample is defined by the kinematic phase space

$$150 < Q^2 < 15000\,\mathrm{GeV}^2, \tag{9.6}$$

$$0.2 < \; y \; < 0.7\,. \tag{9.7}$$

The restriction to high photon virtualities Q^2 results in a good geometrical acceptance of the LAr calorimeter for the detection of the scattered electron. The upper cut on Q^2 is not necessary from an experimental point of view, but improves the convergence of pQCD NLO calculations. Also at very high Q^2 the electroweak corrections become sizeable with only a small gain in statistics. The cut on the inelasticity y is illustrated in figure 9.8a. The lower cut restricts the polar angle of the HFS towards the central region, which is important for the reconstruction of the boost to the Breit frame, discussed in section 9.7. The distribution starts to fall off steeper than kinematically imposed at $y \sim 0.65$, which is due to the cut on the electron energy, $E_\mathrm{e} > 11\,\mathrm{GeV}$. The upper cut on y thus imposes a kinematical constraint on the energy of the scattered electron, which is experimentally favourable. Also, the upper cut on y prevents the hadronic angle γ_h to become too large and limits the size of corrections for higher order QED effects.

Additionally to the phase space requirement, a cut on the longitudinal momentum balance is imposed. $E-p_z$ is required to be within ±10 GeV around its expectation value,

$$45 < E - p_z < 65\,\mathrm{GeV}. \tag{9.8}$$

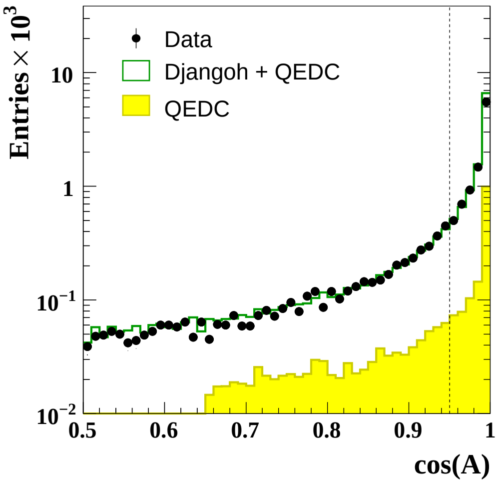


Figure 9.9: Cosine of the acoplanarity A, the applied cut is illustrated as dashed line.

Background	Amount
QED Compton	1.0%
low Q^2 DIS	0.1%
Photoproduction	0.1%
Lepton Pair Production	<0.01%
CC DIS	<0.01%
Total	1.2%

Table 9.4: Estimated *ep* background in the inclusive NC sample, given as fraction of the total number of events.

This cut rejects events with photons radiated collinear to the electron beam, escaping the detector. For a perfectly measured event it corresponds to a rejection of radiative photons with energies above 5 GeV. This cut was shown to improve the boost to the Breit frame [Wob00]. Furthermore, photoproduction events with a misidentified scattered electron, where the true scattered electron escapes undetected in the backward direction, have small values of $E - p_z$ and are suppressed by this requirement.

9.6.1 *ep* Background

After application of the requirements on the scattered electron, the inelasticity y and $E - p_z$, the background from wrongly identified scattered electrons is small. The dominant source of background are elastic and inelastic QEDC events ("wide-angle bremsstrahlung"), with the transverse momentum of the scattered electron and the radiated photon being significantly larger than zero. Since processes of this type are not included in pQCD NLO calculations, the MC simulation is used to correct for them. QEDC processes are included in Djangoh and Rapgap, but an improved simulation is available by the Compton generator. Hence, QEDC events are removed from the signal MCs and the Compton generator is used instead.

Elastic QEDC events can be identified through the acoplanarity, which is defined as [H104]

$$A = |\pi - \Delta\phi| , \tag{9.9}$$

where $\Delta\phi$ denotes the difference in the azimuthal angle between the scattered electron and the radiated photon. In elastic QEDC events the scattered electron and photon have a back-to-back topology with small acoplanarities [Kel04]. In figure 9.9 the cosine of the acoplanarity is shown for events with at least two particles found by the electron finder. Both particles have to have an energy of at least 4 GeV and their energy sum has to be larger than 18 GeV. The distribution peaks at $\cos(A) \sim 1$ for QEDC events and a cut on $\cos(A) < 0.95$ reduces the contribution from QEDC

Phase space	$150 < Q^2 < 15000\,\mathrm{GeV}^2$ $0.2 < y < 0.7$
Primary Vertex	$-35 < z_{\mathrm{vertex}} < 35\,\mathrm{cm}$
Trigger	subtrigger S67
Electron	LAr electron with $z_{\mathrm{imp}} > -190\,\mathrm{cm}$ $E_{\mathrm{e}} > 11\,\mathrm{GeV}$ reject z-crack at $15 < z_{\mathrm{imp}} < 25\,\mathrm{cm}$ reject ϕ-cracks at $2° < \phi_{\mathrm{e}}(\mathrm{octant}) < 43°$ fiducial volume cut $\theta_{\mathrm{e}} < 30°$ or $d_{\mathrm{ca}}(\mathrm{track, cluster}) < 8\,\mathrm{cm}$
Background rejection	$45 < E - p_z < 65\,\mathrm{GeV}$ cosmic and beam halo muons, beam-gas (section 9.3) elastic QED Compton cut (section 9.6.1)
Detector	good and medium quality runs subdetectors: LAr calorimeter, calorimeter trigger, CJC1, CJC2, CIP, ToF, luminosity system

Table 9.5: Summary of all requirements which are applied to obtain the inclusive NC DIS sample.

events by 50% in the NC sample. The remaining events are mostly inelastic QEDC events which cannot be easily distinguished from NC DIS events.
The total amount of *ep* background is estimated from MC simulations. After all applied cuts it amounts to 1.2% in the inclusive NC sample. A summary of the different background contributions is given in table 9.4.

9.6.2 Summary of Requirements and Control Distributions

A summary of all requirements which are applied to obtain the inclusive NC sample is given in table 9.5. In this list all cuts, except for the phase space definition, are technical cuts and are corrected for to obtain the cross section measurements.
A quantity that can be used to monitor the stability of the detector performance over time is the event yield Y, which is defined as the number of detected events per amount of recorded luminosity. A constant event yield over time thus indicates stable trigger, background and detector conditions as well as a reliable luminosity measurement. The acceptance cut of the LAr calorimeter, $z_{\mathrm{imp}} > -190\,\mathrm{cm}$, can bias the event yield due to different average vertex positions during run periods. To remove this bias, the polar angle of the scattered electron θ_{e} is required to be

smaller than 145° for the measurement of Y [Hab09].

The event yield of inclusive NC events per pb^{-1} for the four different HERA-2 run periods is shown in figure 9.10 as function of the run number. The average event yield $\langle Y \rangle$ for each run period is illustrated as solid line, with an uncertainty of 2.5% from the luminosity measurement. There is good agreement between the values of $\langle Y \rangle$ measured for the four HERA-2 periods. For single runs no significant deviations from $\langle Y \rangle$ are observed. The size of the step in the event yield in the e^+p 06/07 period at run number ~ 477000, which has been reported previously [Hab09], is reduced due to the luminosity correction factors from the QEDC analysis. The pull variable as defined in equation (8.12) can be used to test the consistency of the dataset with the assumption of the uncertainty from the luminosity measurement. In the case of the event yield, the pull variable χ is defined as

$$\chi = \frac{Y - \langle Y \rangle}{\sigma_Y \oplus \epsilon_{\mathrm{lumi}}}, \tag{9.10}$$

where σ_Y is the statistical uncertainty of the event yield Y and ϵ_{lumi} is the uncertainty from the luminosity measurement, chosen to be 2.5%. The distribution of pull values for the full HERA-2 dataset is fitted by a Gaussian and shown in figure 9.11. By construction the distribution is centred around zero. No significant outliers are observed and the variance σ is equal to one, expressing that the data are consistent with an uncertainty ϵ_{lumi} of 2.5%.

In figures 9.12-9.13 control distributions of the inclusive NC sample are shown. The background consists of MC simulations of the five different processes summarised in table 9.4. It amounts to 1.2% of the total number of events and is added to the signal MCs. Besides the efficiency corrections discussed above, weights are applied to the signal MCs to improve the description of control distributions for the jet samples, as described below in section 9.7.4. The ratio of data to simulation is shown at the bottom of each plot. The level of agreement between data and both signal MCs is excellent for the distributions of the kinematics and the scattered electron in the inclusive NC sample (figure 9.12).

The distribution of the transverse momentum of the HFS is well described by both, RAPGAP and DJANGOH. The models differ in the description of the longitudinal momentum component of the HFS, p_z^{h}. The peaks of the distributions from both MCs are shifted to larger values than what is observed in the data. The distribution of p_z^{h} falls off more steeply in RAPGAP than in DJANGOH, such that at large values of p_z^{h} RAPGAP is below and DJANGOH above the data. This behaviour is also observed in the distribution of the energy of the HFS E_h, such that the effect on Σ, which is the longitudinal momentum balance of the HFS, $\Sigma = E_h - p_z^{\mathrm{h}}$, largely cancels. A residual shift between the data and the MCs has already been reported previously [Gou08], and is corrected for by weighting the distribution of Σ (all weights are applied on the generated distributions, see section 9.7.4). After the weights have been applied, Σ is very well described by both models, as shown in figure 9.13d. The transverse momentum balance $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}}$ is shown in figure 9.13c. Here the transverse momentum of the HFS is balanced by the transverse momentum of the electron, which is less susceptible to losses and effects from QED radiation than $P_{\mathrm{T}}^{\mathrm{da}}$. The distribution of $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}}$ is well centred around one and well described by the MC models. DJANGOH shows a larger tail at large values of $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}}$ than the data, while RAPGAP describes

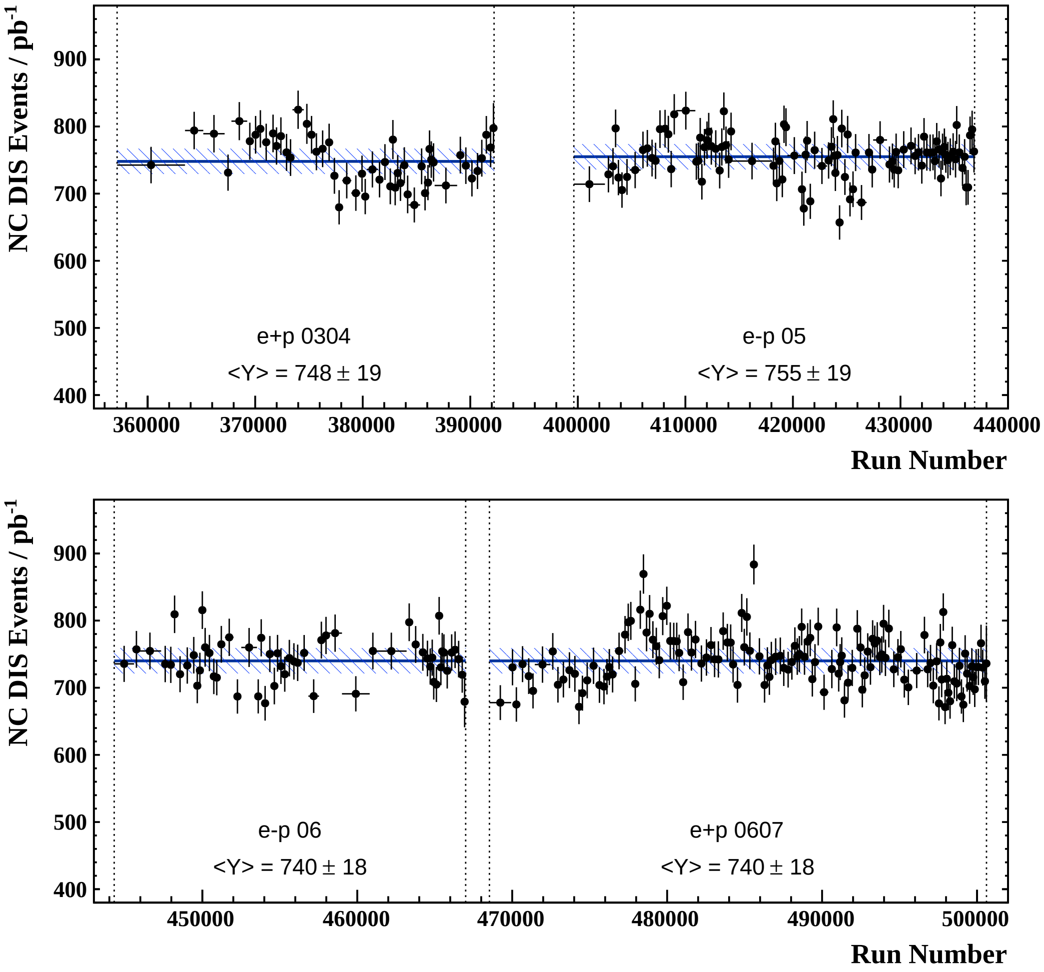


Figure 9.10: Event yield of inclusive NC events for the four HERA-2 periods. The average event yield $\langle Y\rangle$ is shown as solid line together with a band illustrating the 2.5% uncertainty from the luminosity measurement.

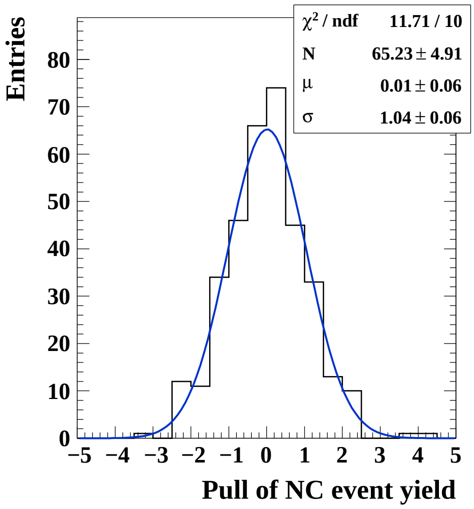


Figure 9.11: Pull distribution of the event yield of inclusive NC events, including all HERA-2 data points from figure 9.10. The distribution is fitted by a Gaussian with mean μ and variance σ.

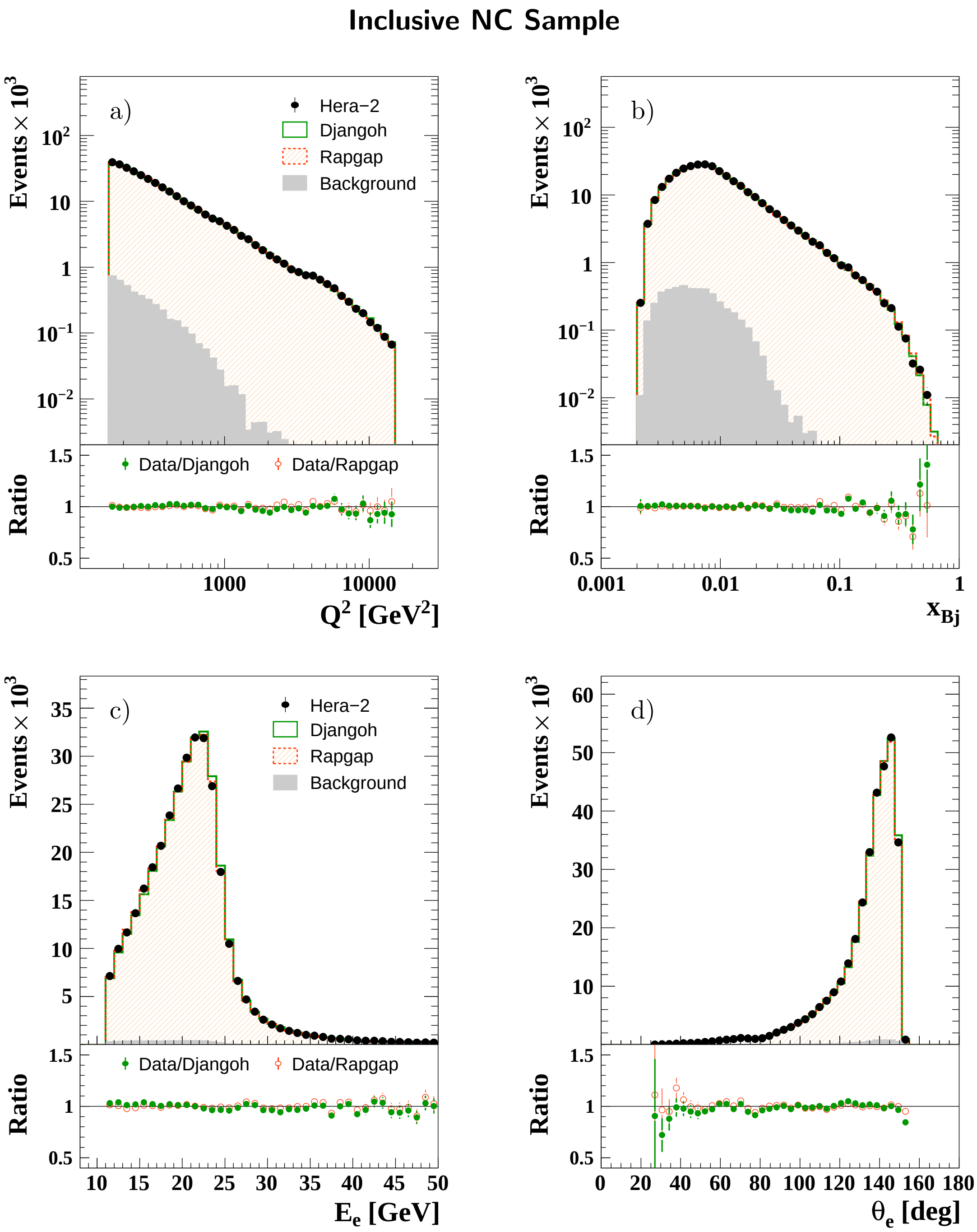


Figure 9.12: Distribution of the virtuality Q^2 (a) and Bjorken x (b) for the inclusive NC sample for the full HERA-2 dataset. Below the distributions of the energy (c) and polar angle (d) of the scattered electron are shown. The signal MCs were weighted to improve the description of jet data. The background contribution is added to the signal MCs.

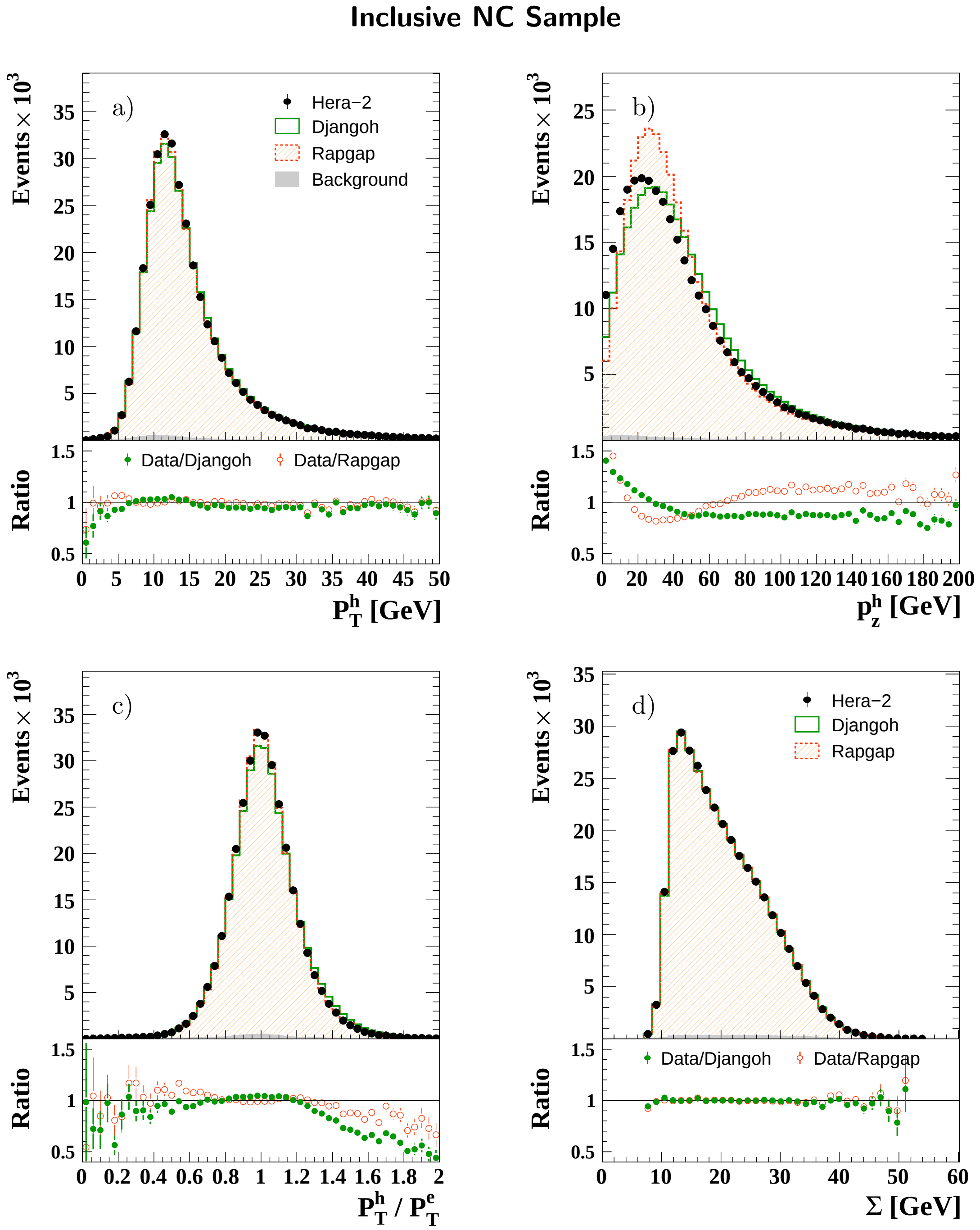


Figure 9.13: Distribution of the transverse (a) and longitudinal (b) momentum component of the HFS for the inclusive NC sample for the full HERA-2 dataset. The distributions of the transverse (c) and longitudinal (d) momentum balance are shown below. The signal MCs were weighted to improve the description of jet data. The background contribution is added to the signal MCs.

the data well over the full range of the distribution. The larger tail in DJANGOH is due to the on average larger energy of the HFS, which leads to a worse resolution in the measurement of E_h and consequently to a broader distribution of $P_{\mathrm{T}}^{\mathrm{h}}/P_{\mathrm{T}}^{\mathrm{e}}$. Overall, both models describe distributions of observables related to the HFS well.

9.7 Jet Selection

The inclusive NC sample is used as the basis of the jet measurement. This has the advantage of an essentially background-free sample with high trigger efficiencies. Furthermore, since the jet sample is a subsample of the inclusive NC sample, it can be used to normalise jet observables, which improves the experimental and theoretical precision. Jets are found and clustered in the Breit frame of reference, where the scattered electron is of importance for the reconstruction of the Lorentz transformation to this frame.

9.7.1 Reconstruction of the Boost to the Breit Frame

The transformation to the Breit frame, as defined in equation (3.4), can be reconstructed with the knowledge of the kinematic variables y and Q^2, as well as the direction of the exchanged virtual boson. QED effects can influence the reconstruction of the Lorentz transformation, if a radiated photon escaped detection through the beam pipe. It was shown in an earlier analysis [Wob00] that the reconstruction of the boost vector with the $e\Sigma$-method is less effected by QED radiation than with the electron method. The cut on $E-p_z$ helps to further improve the reconstruction of the boost vector. The energy and polar angle of the scattered electron can be expressed through the kinematic variables Q^2 and y, which are reconstructed using the $e\Sigma$-method,

$$E_{\mathrm{e}}^{\mathrm{boost}} = \frac{Q^2}{4E_{\mathrm{e}}^0} + E_{\mathrm{e}}^0(1-y)\,, \tag{9.11}$$

$$\theta_{\mathrm{e}}^{\mathrm{boost}} = \arccos\left(\frac{1-a}{1+a}\right) \quad \text{with} \quad a = \left(\frac{2E_{\mathrm{e}}^0}{Q}\right)^2 (1-y)\,, \tag{9.12}$$

and the azimuthal angle of the scattered electron in the laboratory frame ϕ_{e}, is taken from the electron track. The value of ϕ_{e} is used to rotate the Breit frame such that the x-axis is parallel to p_x of the scattered electron in the Breit frame, such that $\phi_{\mathrm{e}}^{\mathrm{Breit}} = 0$. The boost vector

$$b = 2xP + q \tag{9.13}$$

is reconstructed through the four-vector of the exchanged virtual photon, $q = p_e - p_{e'}$, where P, p_e and $p_{e'}$ are the four-vectors of the beam proton, the beam electron and the scattered electron, respectively.

The velocity of the Breit frame with respect to the laboratory rest frame is given by $\beta = p_b/E_b$, where p_b and E_b are the momentum and energy of the boost vector. The value of β approaches one if y becomes large. In this case the Lorentz factor $\gamma = \frac{1}{\sqrt{1-\beta}}$ becomes large, which is experimentally unfavourable, since the detector in the Breit frame becomes distorted. To avoid this region, it is possible to impose

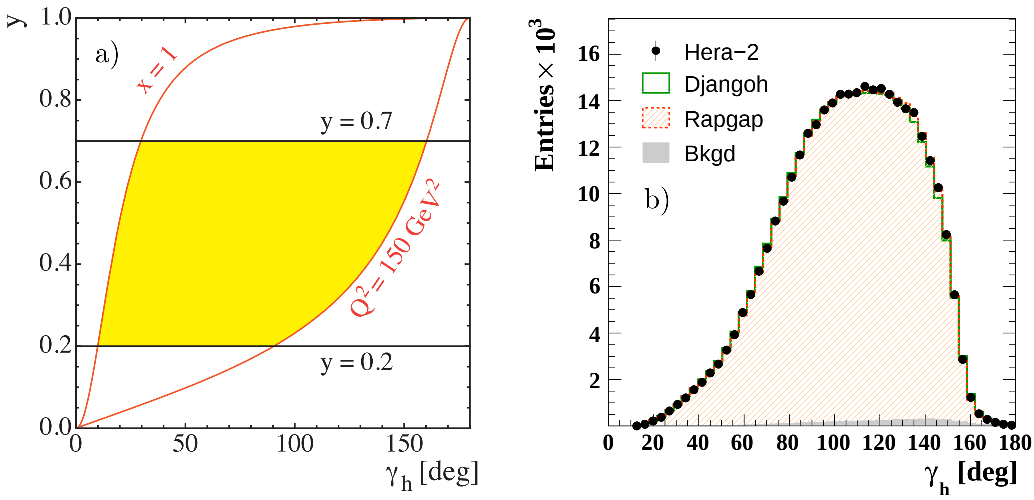


Figure 9.14: a) Allowed kinematic region in the γ_{h}–y plane. The cuts on Q^2 and y and the kinematic limit $x = 1$ are shown as solid lines. b) Distribution of the hadronic angle γ_{h} for the HERA-2 dataset.

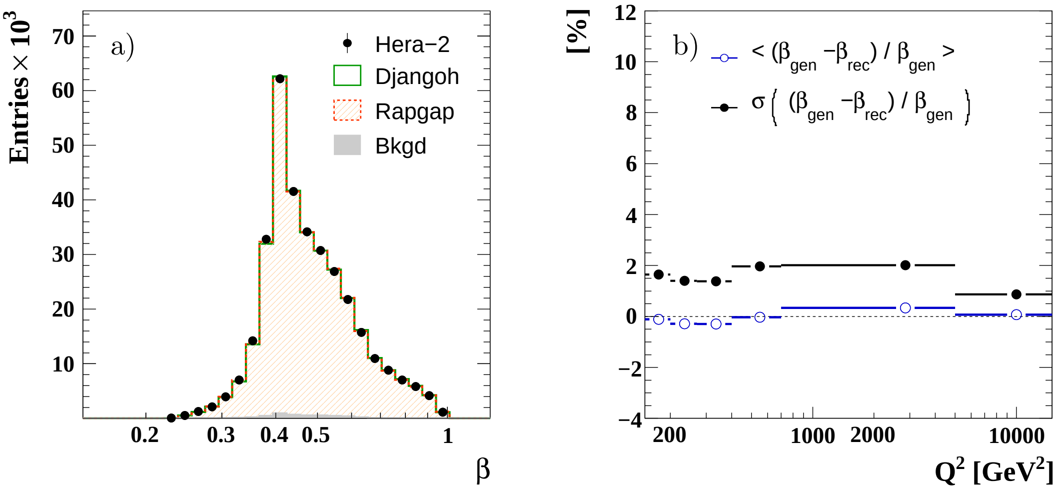


Figure 9.15: Distribution of the reconstructed velocity β of the Breit frame. Systematic bias and resolution of the reconstruction of β of the Breit frame as function of Q^2, obtained from the DJANGOH MC.

a cut on γ_{h} [ZEUS10b]. However, the restriction of this analysis to the inelasticity $0.2 < y < 0.7$ imposes a constraint on the boost vector, which can also be seen from the kinematically allowed region of the hadronic angle, as shown in figure 9.14a. Small values of γ_{h} are thus kinematically forbidden without a cut on the HFS. The distribution of γ_{h} in the inclusive NC sample is shown in figure 9.14b. No events with $\gamma_{\mathrm{h}} < 10°$ are observed, with the distribution only gently rising between $10° < \gamma_{\mathrm{h}} < 50°$.

The distribution of the reconstructed values of β is shown in figure 9.15a. It peaks at values of 0.4 and falls off steeply towards smaller values. The slope of the distribution towards larger values of β is less steep owing to the less rapidly falling distribution of γ_{h} towards small values. The agreement between data and simulation for the reconstruction of the Lorentz transformation to the Breit frame is excellent. In figure

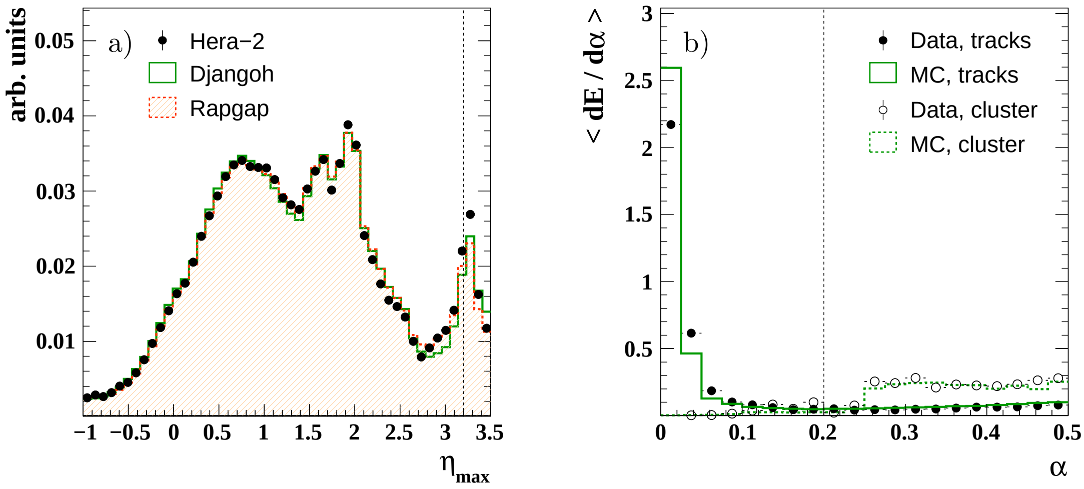


Figure 9.16: a) Maximum value of η in the laboratory rest frame of HFS objects in inclusive jets. b) Average energy between two concentric cones with radii α and $\alpha + \Delta\alpha$ around the electron. Energy measured from tracks is shown with filled circles and solid lines, energy from clusters is shown with open circles and dashed lines. The two cuts applied on the HFS are illustrated with dashed lines in a) and b).

9.15b the resolution of the reconstructed β is shown as function of Q^2. No systematic shift in the reconstructed values of β is observed. The precision of the reconstruction of β is relatively flat over the full range in Q^2 with achieved resolutions of $\sigma \approx 2\%$. In the highest Q^2 bin the resolution is even better due a better reconstruction of Q^2 and y.

9.7.2 Hadronic Final State

All HFS objects are defined and calibrated in the laboratory rest frame, as described in chapters 7 and 8. They are then boosted to the Breit frame of reference, where they are the input to the jet finding algorithm.

When deposits in the LAr calorimeter are merged into clusters, neighbouring cells with a measured signal are combined. For this purpose cells adjacent to the beam hole in the forward direction are considered as neighbours. When cells of two opposite sides of the beam hole are merged into one cluster, clusters with unphysically large values of the pseudorapidity η can result. In figure 9.16a the pseudorapidity in the laboratory rest frame of HFS objects belonging to inclusive jets is shown. Only the pseudorapidity of the HFS object with the largest value of η, compared to all other jet constituents, enters the distribution. The structure between $1.5 < \eta_{\max} < 2.2$ corresponds to the region between the FB2 and the IF and is due to the detector geometry. The peak at large values of $\eta_{\max}$ is due to clusters stretching over opposite sides of the beam hole. To avoid these clusters, only HFS objects with $\eta < 3.2$ are accepted and boosted to the Breit frame (also see figure 8.9).

When the scattered electron traverses the detector material it loses energy due to the radiation of bremsstrahlung photons. The cross section of radiating a photon with energy E_γ is proportional to $1/E_\gamma$, resulting in a small but non-vanishing probability for the radiation of photons with energies above hundreds of MeV. These high energy photons may in turn undergo pair production resulting in an e^+e^- pair, travelling collinear with the original scattered electron. The energy of the bremsstrahlung photons and the electron-positron pairs is merged to the electron cluster, if they deposit their energy within a cone of radius 7.5° in the calorimeter (section 7.2). However, the tracks from the e^+e^- pairs may still appear in the list of HFS objects. In 3% of the inclusive NC events more than one track is found in proximity to the scattered electron, in the simulation this is the case for 2.2% of the events. In figure 9.16b the average energy flow from tracks and clusters around the electron is shown for these events. The tracks originate from pair-conversions, and their energy is accounted for by the energy measurement of the electron cluster. To avoid double counting, tracks are removed from the HFS, if they are within a distance of 0.2 in η, ϕ to the scattered electron.

9.7.3 Jet Definition

Jets are found in the Breit frame with the inclusive, longitudinally invariant k_T-algorithm, as implemented in the FastJet package [CS06]. The distance parameter R is set to one and the P_T-recombination scheme is used to obtain the jet four-vectors (see section 3.4). Unless otherwise indicated, all variables related to jets are measured in the Breit frame of reference, for example $P_\mathrm{T} \equiv P_\mathrm{T,Breit}$, $\eta \equiv \eta_\mathrm{Breit}$.
In order to ensure jets to be well within the acceptance of the LAr calorimeter, the jet four-vector is boosted from the Breit frame back to the laboratory rest frame, where the pseudo-rapidity of each jet is required to be

$$2.5 < \eta_\mathrm{lab} < -1.0\,. \tag{9.14}$$

This cut is part of the phase space definition and is applied to each jet of the inclusive, dijet and trijet samples. The effect of this cut is illustrated in figure 9.17a.

Two technical cuts, which are not part of the phase space definition, are applied on reconstructed level to ensure well measured jets. The MC simulations are used to correct for the effects of these cuts at the cross section measurement. All jets consisting of only one HFS object are removed from the sample. These jets consist of the radiated QED photon and are usually in close proximity to the scattered electron. After the removal of these jets from the sample, the number of jets close to the scattered electron is negligible, see figure 9.17b.
For all reconstructed jets a minimal transverse momentum in the laboratory rest frame is required,

$$P_\mathrm{T,lab} > 2.5\,\mathrm{GeV}. \tag{9.15}$$

The purpose of this cut is to remove very soft jets which cannot be measured accurately. This cut is illustrated in figure 9.18 for the full Hera-2 sample, split into two Q^2 ranges. The cut removes more events at high Q^2, where the jet transverse momenta in the Breit and laboratory frame can be very different.

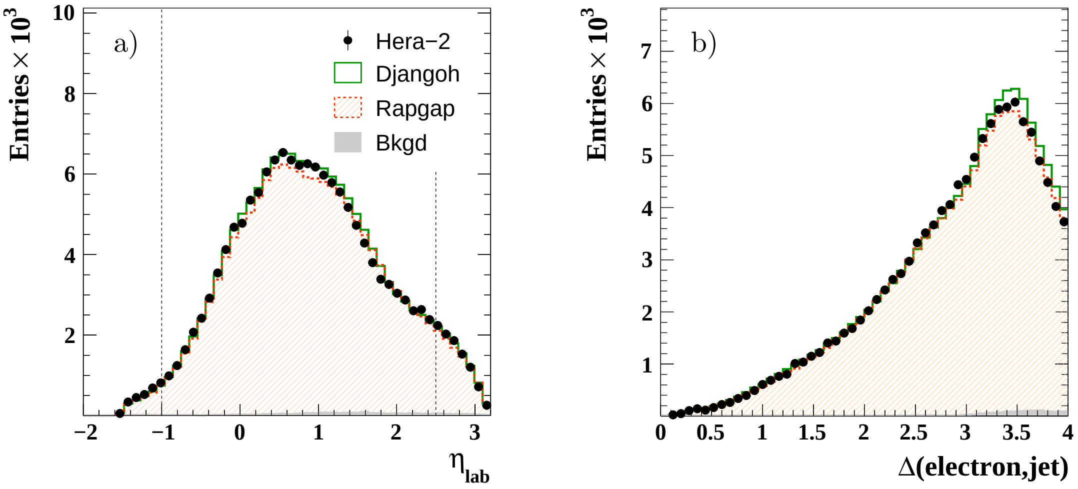


Figure 9.17: a) Pseudo-rapidity of inclusive jets after they were boosted from the Breit frame to the laboratory rest frame. b) Distance between the scattered electron and jets, defined as $\Delta = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, after the removal of jets consisting of only one HFS object. The signal MCs were weighted to improve the description of jet data.

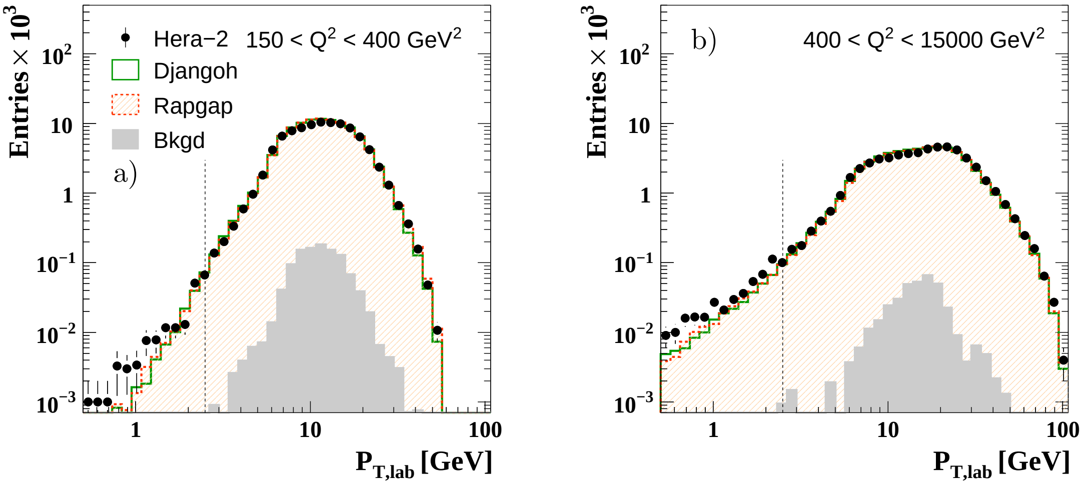


Figure 9.18: Distribution of the transverse momentum in the laboratory rest frame of inclusive jets for low (a) and high (b) virtualities Q^2. The applied cut is illustrated with a dashed line. The signal MCs were weighted to improve the description of jet data.

9.7.4 MC Weighting

Simulated events from MC event generators, which are subject to the detector simulation and reconstruction, are used to correct for detector effects for the extraction of jet cross sections. They are also used to estimate the size of the systematic errors due to measurement uncertainties. In order to have a reliable estimate of migrations between the measurement bins as well as migrations in and out of the samples used for the extraction of the cross sections, it is necessary that the simulation describes the data well.

The detector response and reconstruction efficiencies have been carefully studied, such that remaining discrepancies between the data and MC models are attributed

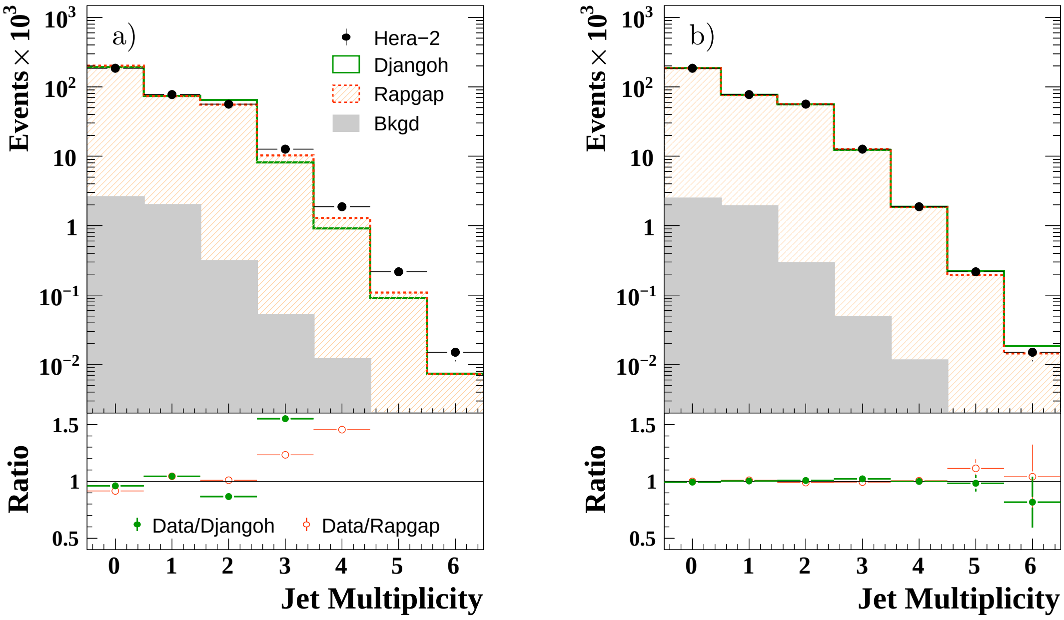


Figure 9.19: Jet multiplicity of jets with $P_{\mathrm{T}} > 5\,\mathrm{GeV}$, before (a) and after (b) the weighting is applied to the DJANGOH and RAPGAP MCs.

to deficiencies in the simulation of the underlying physics, such as missing higher orders in the perturbative series. In order to improve the agreement between data and reconstructed observables, weights are applied to adjust the generated distributions. All weights are obtained from the ratio of data to the reconstructed MC distributions. They are then applied on generated level, such that they are independent of the detector simulation.

In general, observables in the NC sample are well described by the MC simulations. Only the distribution of the longitudinal momentum balance of the hadronic final state, $\Sigma = E_h - p_z^{\mathrm{h}}$, is shifted between the data and the MCs. This is however an important quantity which enters in the calculation of x (7.15) and consequently the boost to the Breit frame. To improve the description of Σ, Σ_{gen} is weighted with a linear polynomial of the form $w = a + b \cdot \Sigma_{\mathrm{gen}}$ [Gou08], where a and b are adjusted for RAPGAP and DJANGOH separately. The resulting distributions of Σ and x have been shown for the NC sample in figures 9.12 and 9.13. Control distributions of the jet samples are shown below.

An observable of great importance in this analysis is the jet multiplicity. Only if it is well described by the MC simulation, will the contribution from one, two and more jet events to the inclusive jet sample be comparable in data and simulations. This is not only important for the description of jet observables, but also for kinematic observables like the inelasticity, which tends towards larger values for multijet events. Additionally, the description of the absolute normalisation of the jet samples is improved once the jet multiplicity is well described. Since higher orders in the perturbative series are only approximately included in the MCs through parton showers in the leading logarithmic or the colour dipole approximation, the distribution of the jet multiplicity is not expected to be well described for high jet multiplicities. A set of weights is derived, where each is a function of $\log(Q^2)$, for the different numbers of observed jets. The effect of the weighting of the jet multiplicity is shown in figure

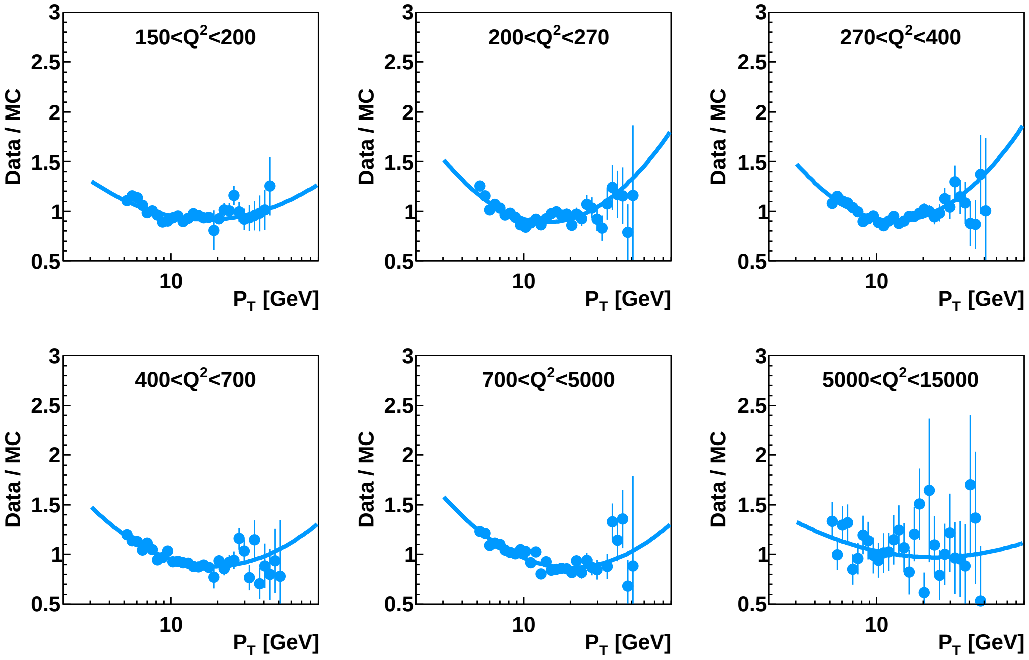


Figure 9.20: Ratio of the transverse momentum distribution for data to Djangoh, used to obtain the weight functions, which are shown as solid lines. They are two-dimensional functions of Q^2 and P_T. For their illustration they are evaluated at the mean value of Q^2 in each bin.

9.19. The unweighted jet multiplicity is described within 10% for up to two jets by both MCs. At larger jet multiplicities the difference between data and the simulation becomes larger. Rapgap describes the data slightly better, it undershoots the data by about 30% and 50% for three and four jet events, respectively. After the application of the MC weighting, the jet multiplicity is very well described by both MC simulations.

Another shortcoming of the used MC event generators is their inability to reproduce the P_T spectrum of jets at large P_T. Also at large rapidities, where higher orders become more important, they underestimate the number of jets. Since migrations in both variables, P_T and η, are of importance in this analysis, the generated spectra of P_T and η are weighted. For this purpose, the pseudorapidity of the most forward jet (the jet with the largest value of η) in events with at least two jets is chosen. For the purpose of the weighting procedure a larger region in P_T and η_lab than the analysis phase space is used, such that the description of migrations into the sample is improved as well. The larger phase space in P_T and η_lab results in an increase of the fraction of dijet events with respect to one-jet events in the inclusive jet sample. The forward jet is used for the weighting because of its good correspondence between reconstructed and generated level.

The weighting is performed in two steps. In the first step, the pseudorapidity distributions of the forward jet in different Q^2 regions are used to determine a weight function in η_fwd and $\log(Q^2)$. In the next step, the P_T spectra of the forward jet are used to obtain a weight function in P_T and $\log(Q^2)$. The weight functions are polynomials of second order such that they have no discontinuities because of the binning chosen for their determination. As an example, the obtained function used

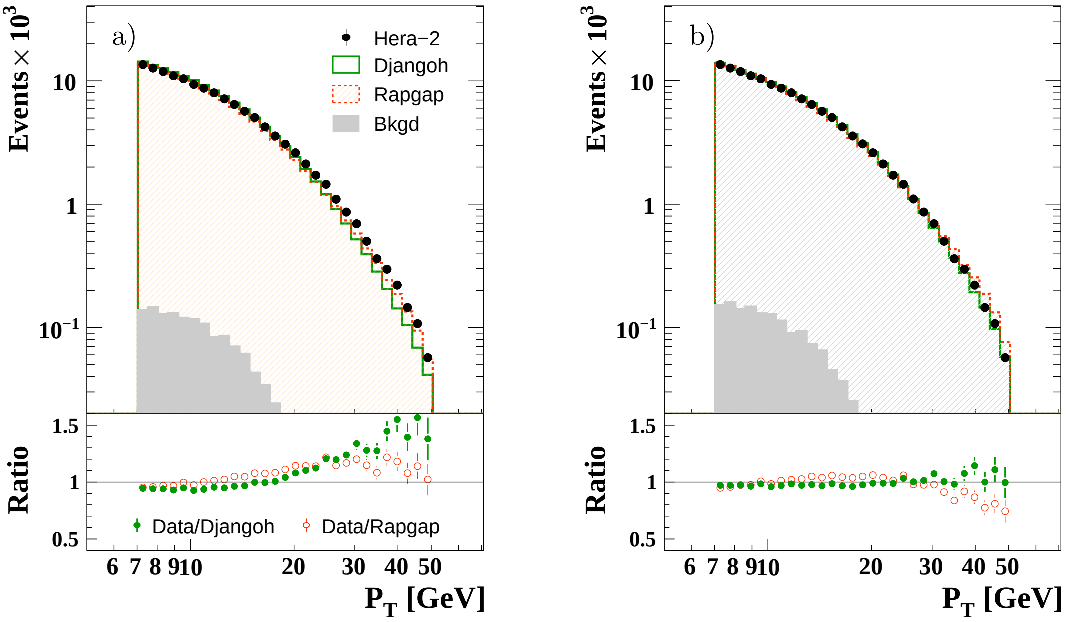


Figure 9.21: Transverse momentum of inclusive jets before (a) and after (b) the weighting is applied to the DJANGOH and RAPGAP MCs.

for the weighting in P_{T} for the DJANGOH MC is illustrated in figure 9.20 for different regions in Q^2. The function is well behaved over the full P_{T} and Q^2 range and does not oscillate outside the fit range. The obtained function for the RAPGAP MC shows a similar behaviour with a comparable absolute magnitude. In the P_{T} and η range of this analysis the applied weights take values between 0.7 and 1.3.
The transverse momentum spectrum of inclusive jets before and after the weighting is shown in figure 9.21. The unweighted simulated spectra are too soft, RAPGAP and DJANGOH undershoot the data by about 20% and 50% at large P_{T}, respectively. After the weights are applied, the data are described within 10% over most of the P_{T} range. Deviations larger than 10% can only be seen for RAPGAP at $P_{\mathrm{T}} > 40\,\mathrm{GeV}$, where RAPGAP lies about 20% above the data.

9.8 Inclusive Jet Sample

In the inclusive jet sample each jet contributes to the measurement. This is different from the other measurements presented in this thesis, where each event of the respective sample contributes once per observable. For example, an event containing three jets within the defined phase space will be counted three times for the determination of the inclusive jet cross section. Additionally to the requirement on the pseudo-rapidity (9.14), each jet is required to have

$$7 < P_{\mathrm{T}} < 50\,\mathrm{GeV}\,. \tag{9.16}$$

The P_{T} spectrum of all jets in the HERA-2 dataset is shown in figure 9.22a, where the above cut is indicated by dashed lines. The MC P_{T} spectrum has been weighted, as described in the previous section, to improve the description of the data. Excellent agreement is observed over the full accessible range, with deviation between data

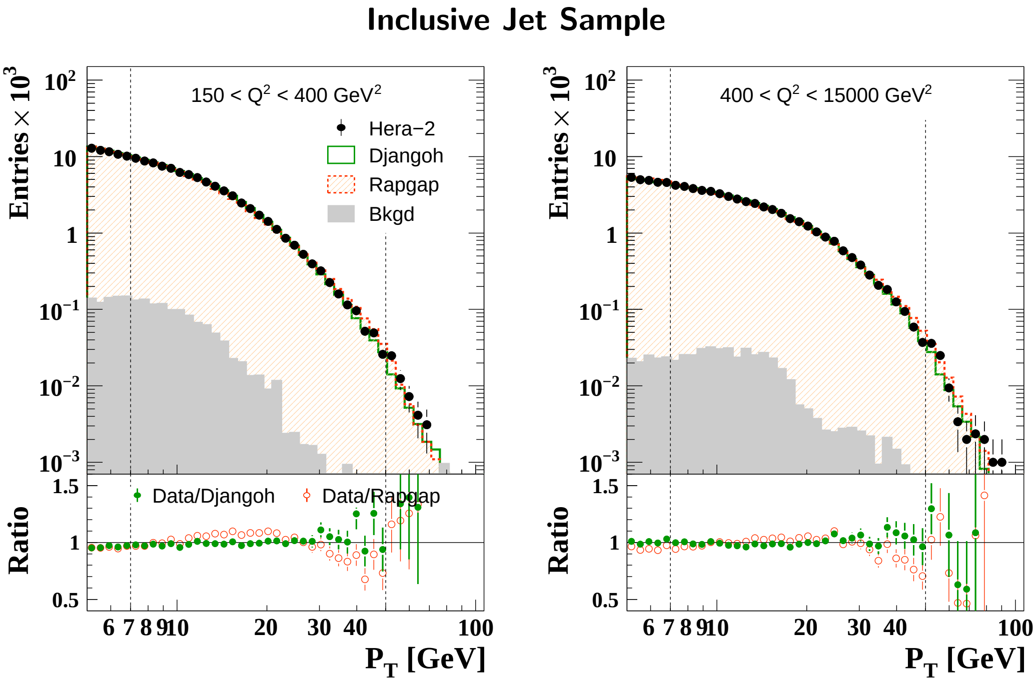


Figure 9.22: Distribution of the transverse momentum in the Breit frame of inclusive jets for the full HERA-2 dataset, shown for two different Q^2 ranges. The applied cuts are illustrated by dashed lines.

and the simulation never larger than 15%. The inclusive jet sample consists of 133 621 jets in 90 328 events. 39% of the jets are found in events with only one jet, 49% originate from events with two jets and 12% are found in events with three and more jets. Because of the weighting of the jet multiplicities, these fractions are well described by both MC models, differing by 1% at most.

The event yield Y, similarly defined as for NC events, is shown in figure 9.23 for events with at least one jet with $P_T > 7\,\mathrm{GeV}$. It is flat for the full HERA-2 dataset. The mean values determined for the four different run periods agree with each other within the uncertainty from the luminosity measurement. The assumption of the 2.5% uncertainty due to the luminosity measurement is tested similar to the NC sample, with the pull variable defined in equation (9.10). The distribution of pulls is fitted by a Gaussian and shown in figure 9.24. No statistically significant outliers are observed. The variance of the Gaussian σ is one to a good approximation, showing the consistency of the data with an uncertainty from the luminosity measurement of 2.5%. An individual test for each of the four different run periods leads to the same conclusion.

The total background from ep scattering processes in the inclusive jet sample is estimated to be 1.1%, which is slightly less than what is found in the inclusive NC sample. The main source of background are inelastic QEDC processes, where the jet balances the radiated photon, amounting to 0.6% of the events in the inclusive jet sample. Other sources of background are low Q^2 and photoproduction events, originating from a wrongly identified electron, amounting to 0.3% and 0.2% of the events in the inclusive jet sample, respectively. The background contribution from CC and lepton-pair production is negligible.

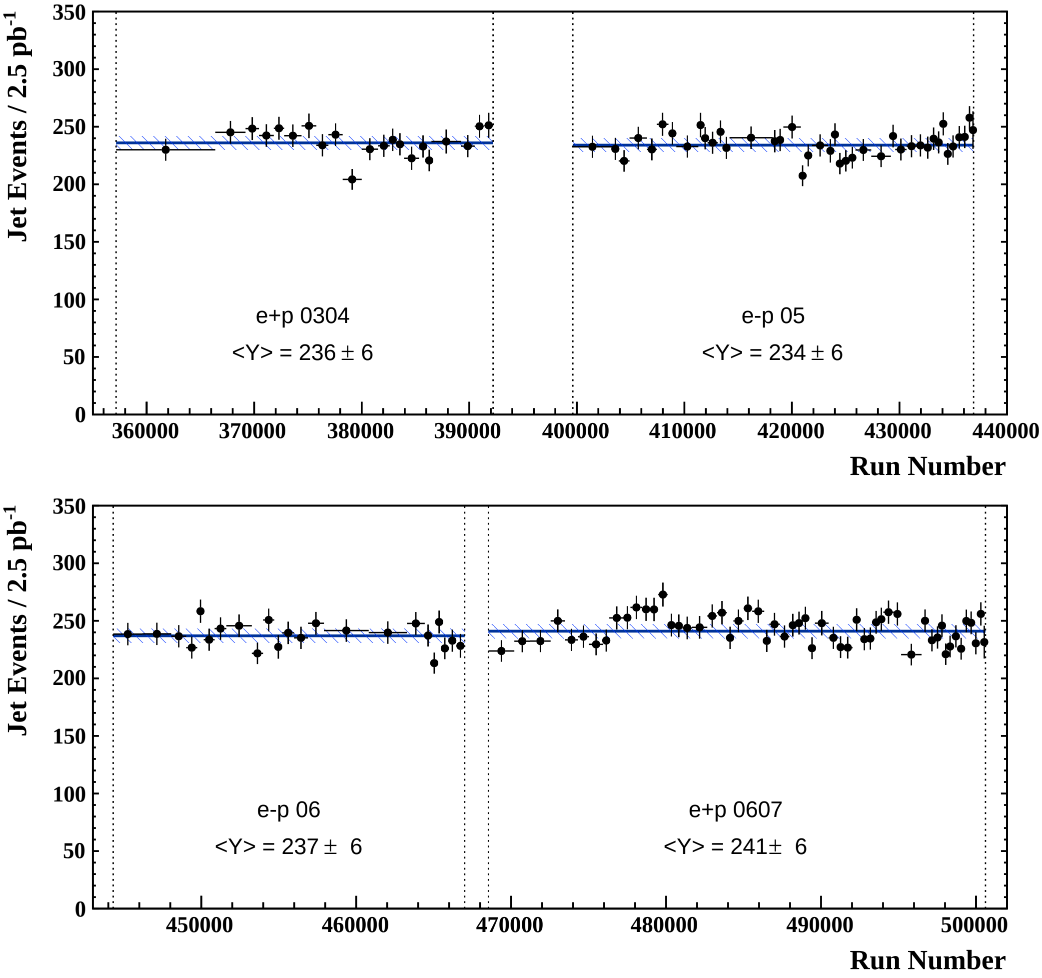


Figure 9.23: Event yield Y of events with at least one jet with $P_{\mathrm{T}} > 7\,\mathrm{GeV}$ for the four HERA-2 periods. The average yield $\langle Y \rangle$ is shown as solid line together with a 2.5% uncertainty from the luminosity measurement.

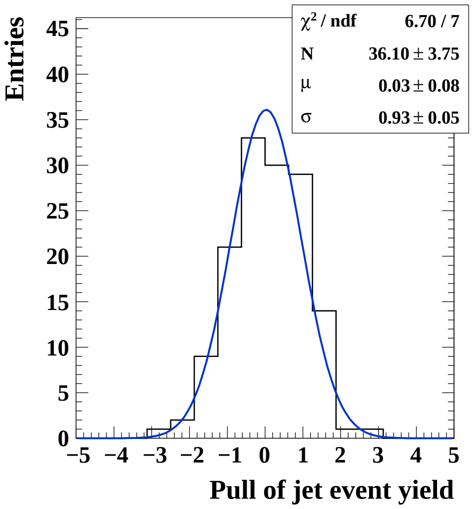


Figure 9.24: Pull distribution of the event yield Y of events with at least one jet with $P_{\mathrm{T}} > 7\,\mathrm{GeV}$, including all data points from figure 9.23. The distribution is fitted by a Gaussian with mean μ and variance σ.

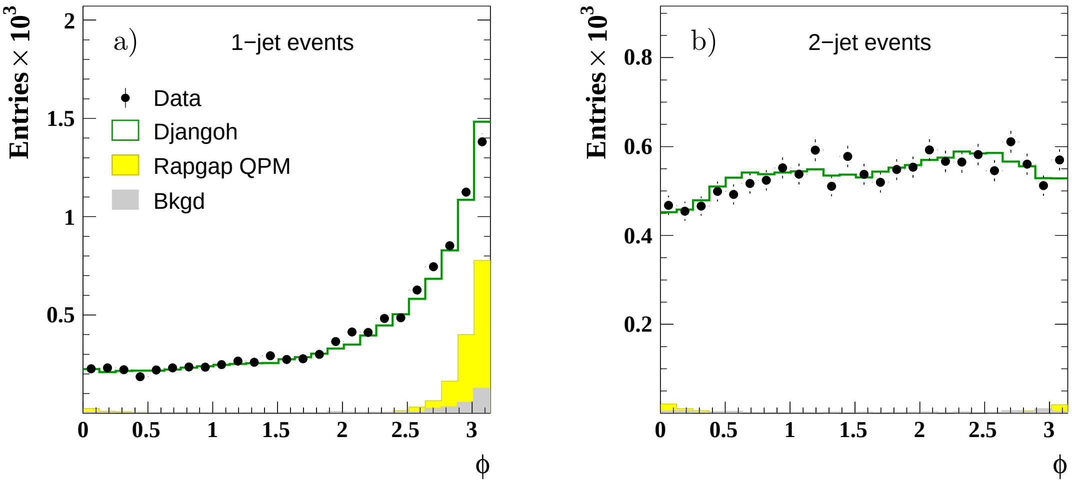


Figure 9.25: Distribution of the azimuthal angle of inclusive jets in the Breit frame for 1-jet (a) and 2-jet events (b). The Rapgap MC consists of QPM evens only. In the Breit frame $\phi = 0$ corresponds to the azimuthal angle of the scattered electron.

A different source of background, which is however included in the signal MCs, originates from QPM processes. If in a QPM process a radiated photon is present, the jet originating from the scattered quark may acquire significant P_T in the Breit frame. This happens when the photon escapes detection through the beam pipe which leads to a deterioration of the boost vector. These types of processes are studied with the Rapgap MC, selecting only QPM-type events. The estimated fraction of jets from QPM-type events is 5% of the total number of jets in the inclusive jet sample. The distribution of the azimuthal angle ϕ of the jet in the Breit frame is shown in figure 9.25 for one- and dijet events. Jets originating from QPM processes contribute only to one-jet events, with the jet being back-to-back with the scattered electron. In dijet events the contribution is negligible, indicating that QPM processes with a photon radiated at large angles, and thus balancing the jet in transverse momentum, are largely suppressed due to the small cross section. The azimuthal distribution of jets in one-jet events is well described by Djangoh, but Rapgap differs by about 15% from the data at large values of ϕ. Hence, it is refrained from cutting on ϕ to reject QPM-type events. Further studies of the topology of these types of events did not reveal an observable which can be used to safely reject them. However, QPM-type events are only part of the jet sample in MC simulations including higher order QED effects. The cross sections are corrected for higher order QED effects by utilising MC simulations without QED radiation. No jets from QPM processes are observed to migrate into the inclusive jet sample in these simulations. The measured cross sections are thus corrected for the contamination from QPM events.

Distributions of the kinematics of inclusive jet events are shown in figure 9.26. Both MC simulations describe the data well over the full range of the observables. Given the phase space restrictions on Q^2 and y, the Bjorken scaling variable x lies within a range of approximately $2 \cdot 10^{-3} < x < 0.3$, with a pronounced maximum at 0.01. The peak of the distribution of $E - p_z$ lies at 54.9 GeV in the data, which is close to the expectation value of $2E_\mathrm{e}^0 = 55.2\,\mathrm{GeV}$. It is well described by Rapgap, in

Inclusive Jet Sample

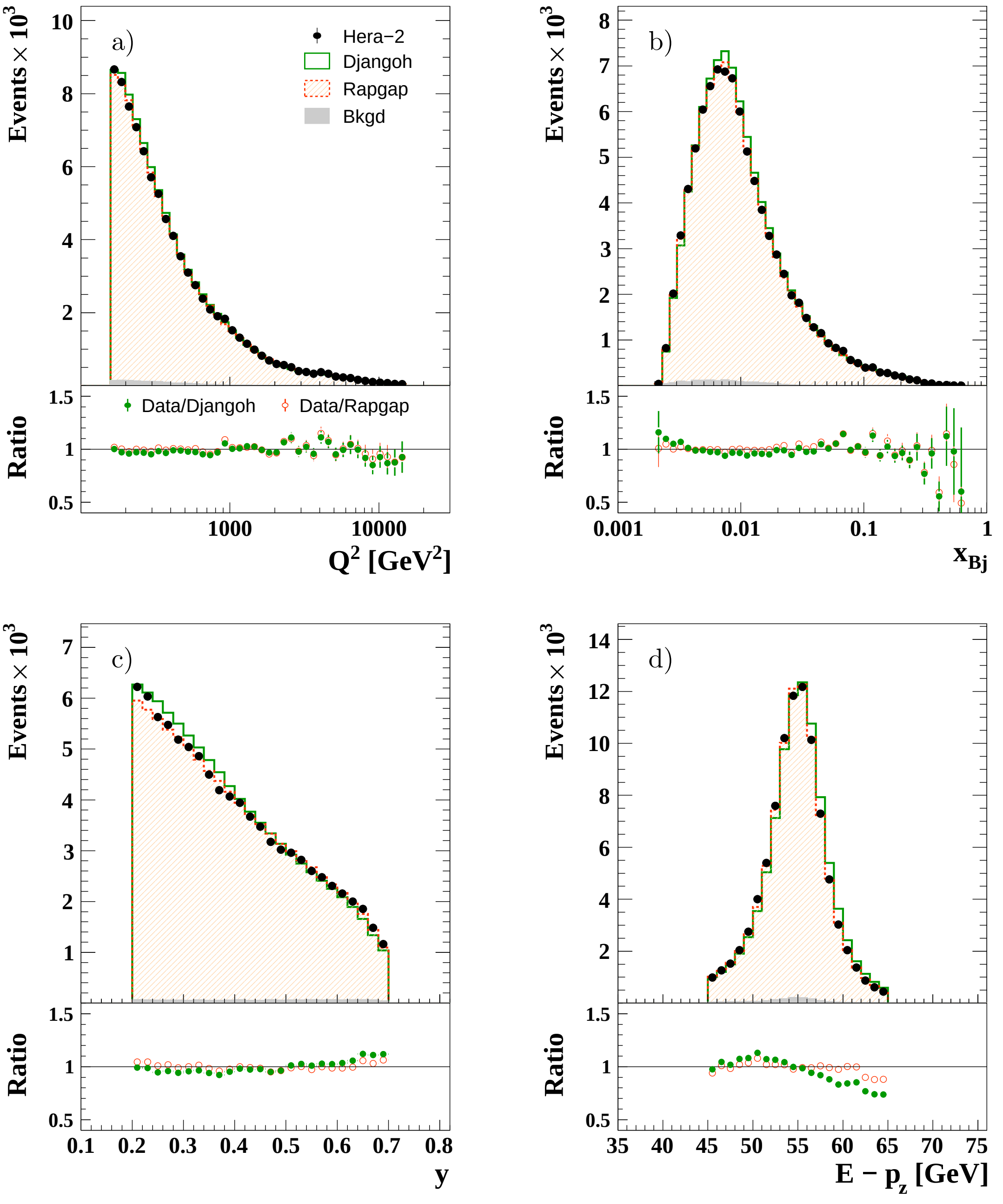


Figure 9.26: Distribution of the virtuality Q^2 (a), Bjorken x (b), the inelasticity y (c) and the longitudinal momentum balance $E - p_z$ (d) for the inclusive jet sample for the full HERA-2 dataset. The background contribution is added to the signal MCs.

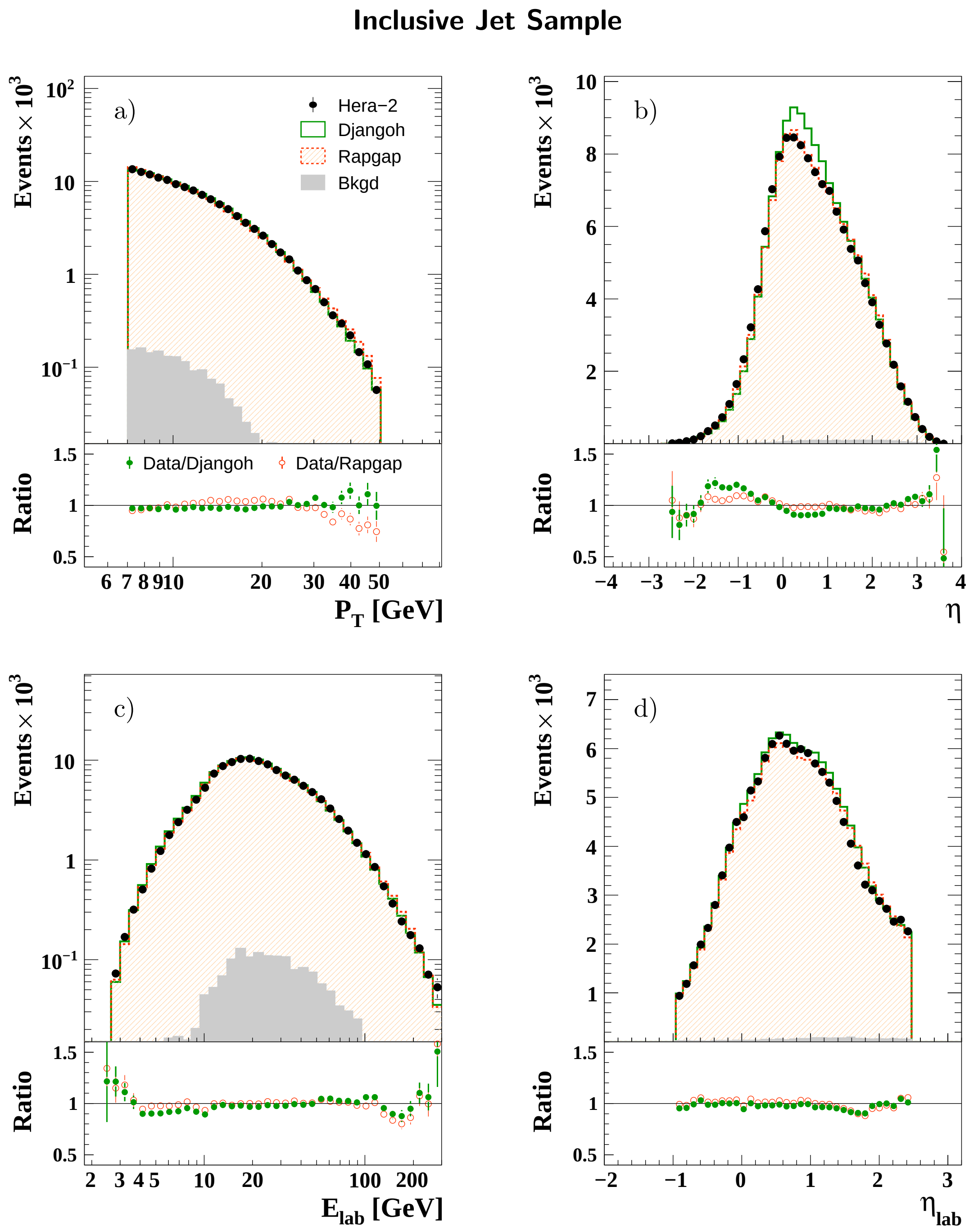


Figure 9.27: Distributions of jet observables for the inclusive jet sample for the full HERA-2 dataset. The transverse momentum P_{T} (a) and pseudorapidity η (b) are measured in the Breit frame and the jet energy E_{lab} (c) and pseudorapidity η_{lab} (d) are measured in the laboratory rest frame. The background contribution is added to the signal MCs.

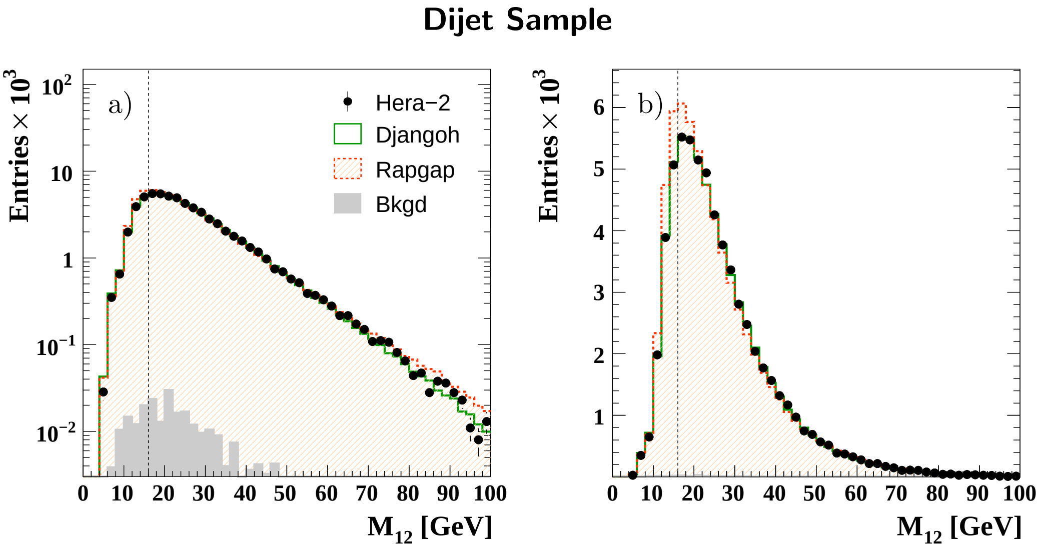


Figure 9.28: Distribution of the invariant mass of the two leading jets in the dijet sample, shown on a logarithmic (a) and linear (b) scale. The applied cut is illustrated by a dashed line.

DJANGOH the peak is shifted by 0.2 GeV towards larger values. This difference is taken into account by the systematic uncertainty due to the model dependence and the electron and HFS measurements.

Figure 9.27 shows distributions of jet observables of the inclusive jet sample for the full HERA-2 dataset. The data are described within 10% by the simulations in most regions, with the largest deviations never exceeding 20%. The P_T spectrum is a steeply falling distribution with an exponential behaviour. It is well described by $e^{-\lambda P_\mathrm{T}}$, with $\lambda \approx 0.12$. The distribution of the jet pseudorapidity in the Breit frame changes its shape with varying Q^2. At low values of Q^2 jets are approximately restricted to the target region ($\eta > 0$), whereas for large values of Q^2 the distributions shifts towards the current region ($\eta < 0$), which is an indication of multiple QCD radiation due to an increased phase space. The energy of jets measured in the laboratory rest frame extends over a large range, from 3 GeV up to 300 GeV, with an average value of 20 GeV. The pseudorapidity distribution in the laboratory frame shows a maximum at $\eta_\mathrm{lab} \sim 0.5$ and falls off steeply towards smaller values. Towards large values of η_lab the slope of the distribution is less steep, with a change in the slope around $\eta_\mathrm{lab} \sim 1.8$. This change is mainly due to multi-jet events, where more jets are observed at large values of η_lab.

9.9 Dijet Sample

The dijet sample consists of events with at least two jets, where the transverse momentum of every jet is required to be

$$5 < P_\mathrm{T} < 50\,\mathrm{GeV}\,. \tag{9.17}$$

Additionally, the invariant mass of the two leading jets[1] is required to be

$$M_{12} > 16\,\mathrm{GeV}\,, \tag{9.18}$$

which ensures convergence of fixed-order pQCD calculations. The distribution of M_{12} is shown in figure 9.28 together with the imposed cut. The logarithmic scale (a) shows the good agreement between data and the simulations up to large values of M_{12}. The linear scale reveals that the distribution of M_{12} tends towards smaller values in RAPGAP than in DJANGOH. DJANGOH describes the data well, whereas RAPGAP lies above the data for small values of M_{12}. The resulting differences in the migrations are taken into account by the model uncertainty. The requirement on M_{12} causes P_{T} of the leading jet to be greater than 7 GeV in 95% of all dijet events, which makes the dijet sample almost a subsample of the inclusive jet sample. In the HERA-2 dataset a total of 52 028 dijet events is found.

The estimated background from ep processes is less than 0.4% and is not shown in the control distributions. The background originates from wrongly identified scattered electrons, with 0.2% originating from each, photoproduction and low Q^2 DIS processes. The contamination from QEDC processes and QPM-type events is negligible. The variable ξ, which is in LO dijet production the proton's momentum fraction carried by the emerging parton, can be written as

$$\xi = x\left(1 + \frac{M_{12}^2}{Q^2}\right) = x + \frac{M_{12}^2}{ys}\,. \tag{9.19}$$

Using the hadron method [JB79] for the reconstruction of y leads to a partial cancellation of mismeasurements in M_{12} due to the jet energy scale uncertainty [Wob00].

In figure 9.29 distributions of the kinematics of dijet events are shown. Overall good agreement between the data and both simulations is observed. The distributions of Q^2 and ξ are well described by both models over the full accessible range. The mean value of ξ is about 0.03, with a covered range in the parton momentum fraction of $8 \cdot 10^{-3} < \xi < 0.3$. It is interesting to note that the inelasticity y is on average larger than in the inclusive jet sample. It is well described by both models, where at low values of y the data lie between RAPGAP and DJANGOH. Only at large values of y deviations larger than 5% are observed, with both MC simulations being below the data by about 10–15%. The same shift of the mean value of $E - p_z$ in DJANGOH with respect to the data as in the inclusive jet sample is observed.

In figure 9.30 the distributions of the transverse momenta in the Breit frame $P_{\mathrm{T},1}$ and $P_{\mathrm{T},2}$, and pseudorapidities in the laboratory frame, $\eta_{\mathrm{lab},1}$ and $\eta_{\mathrm{lab},2}$, of the two leading jets are shown. The mean value of $P_{\mathrm{T},1}$ is around 12 GeV, being 4 GeV larger than the mean value of $P_{\mathrm{T},2}$. The distribution of $P_{\mathrm{T},1}$ falls off steeply towards smaller values, which is a consequence of the requirement $M_{12} > 16\,\mathrm{GeV}$. The P_{T} spectra of both jets are well described by the simulations. At large values of $P_{\mathrm{T},2}$ discrepancies of about 50% between DJANGOH and RAPGAP can be observed, with the data lying between the two MCs. The distributions of $\eta_{\mathrm{lab},1}$ and $\eta_{\mathrm{lab},2}$ are reasonably well described, with typical deviations between the data and simulations of the order of 15%. In DJANGOH the distribution of $\eta_{\mathrm{lab},2}$ is shifted with respect to the data towards smaller values of $\eta_{\mathrm{lab},2}$.

---

[1] A leading jet is defined as the jet with the highest transverse momentum in the event.

Dijet Sample

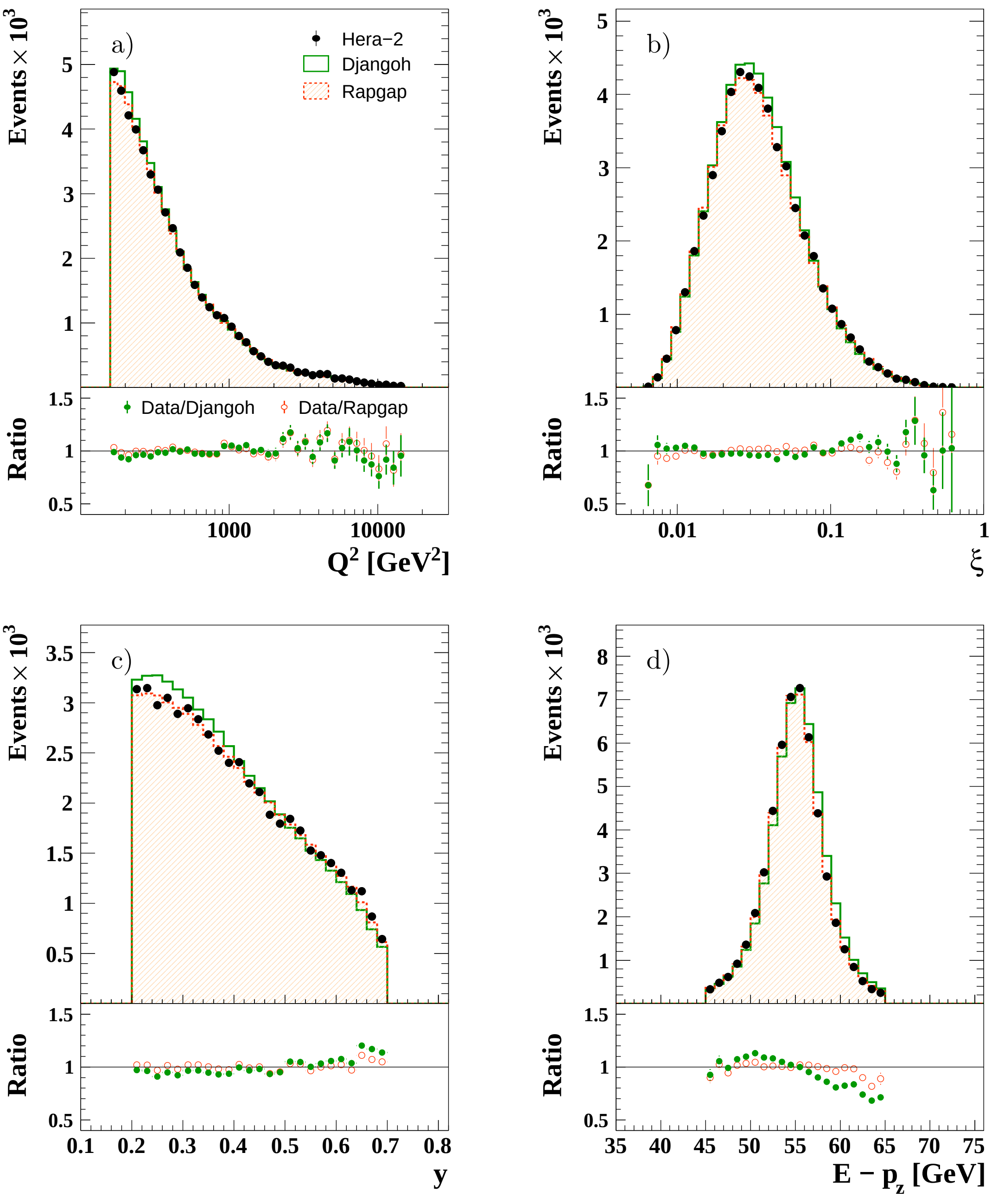


Figure 9.29: Distribution of the virtuality Q^2 (a), ξ (b), the inelasticity y (c) and the longitudinal momentum balance $E - p_z$ (d) for the dijet jet sample for the full HERA-2 dataset. Not shown is the contribution from background, which is less than 0.4%.

Dijet Sample

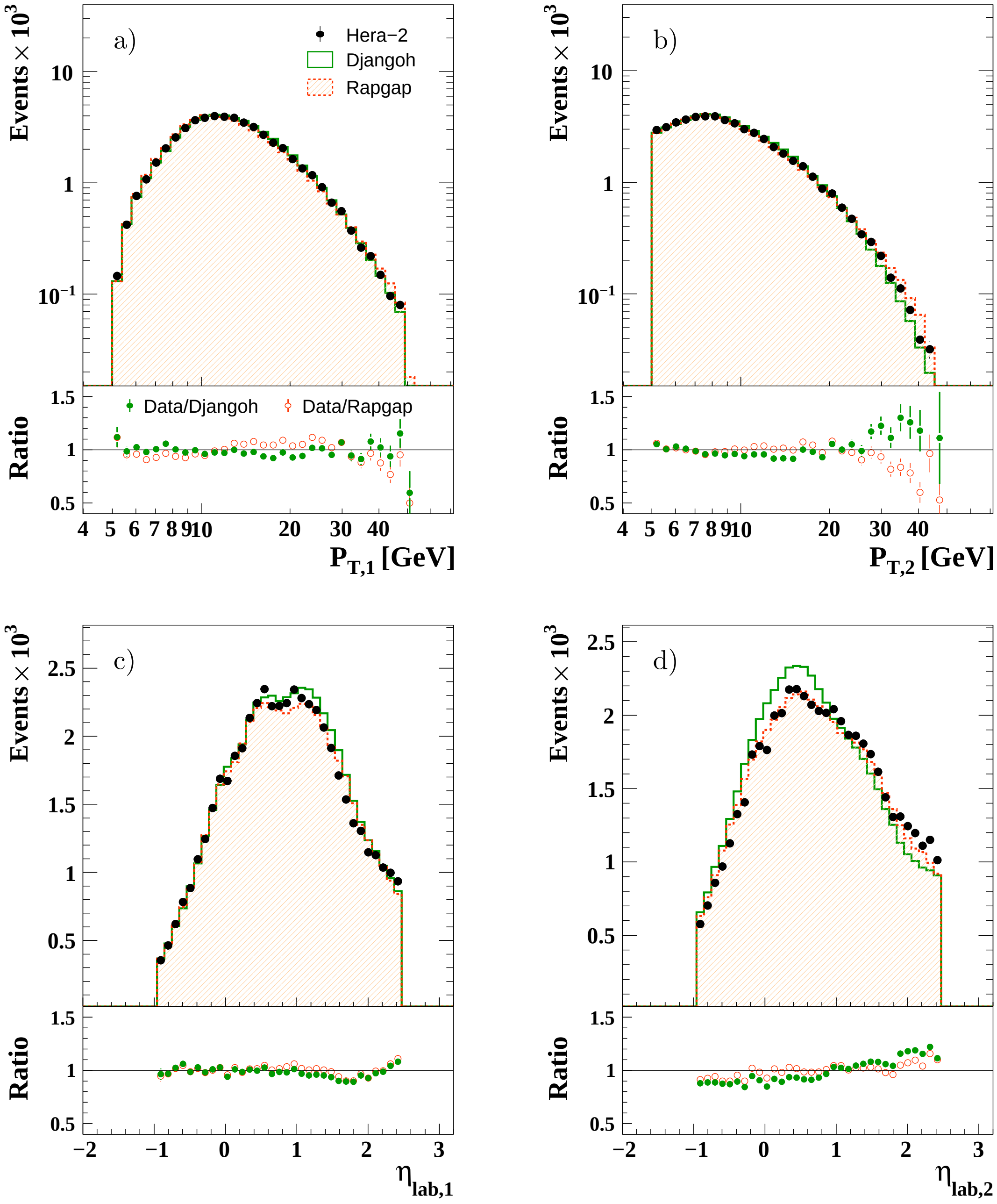


Figure 9.30: Distribution of the transverse momentum in the Breit frame of the leading (a) and subleading (b) jet. Also shown are the pseudorapidity distributions of the two jets in the laboratory frame, (c) and (d).

Dijet Sample

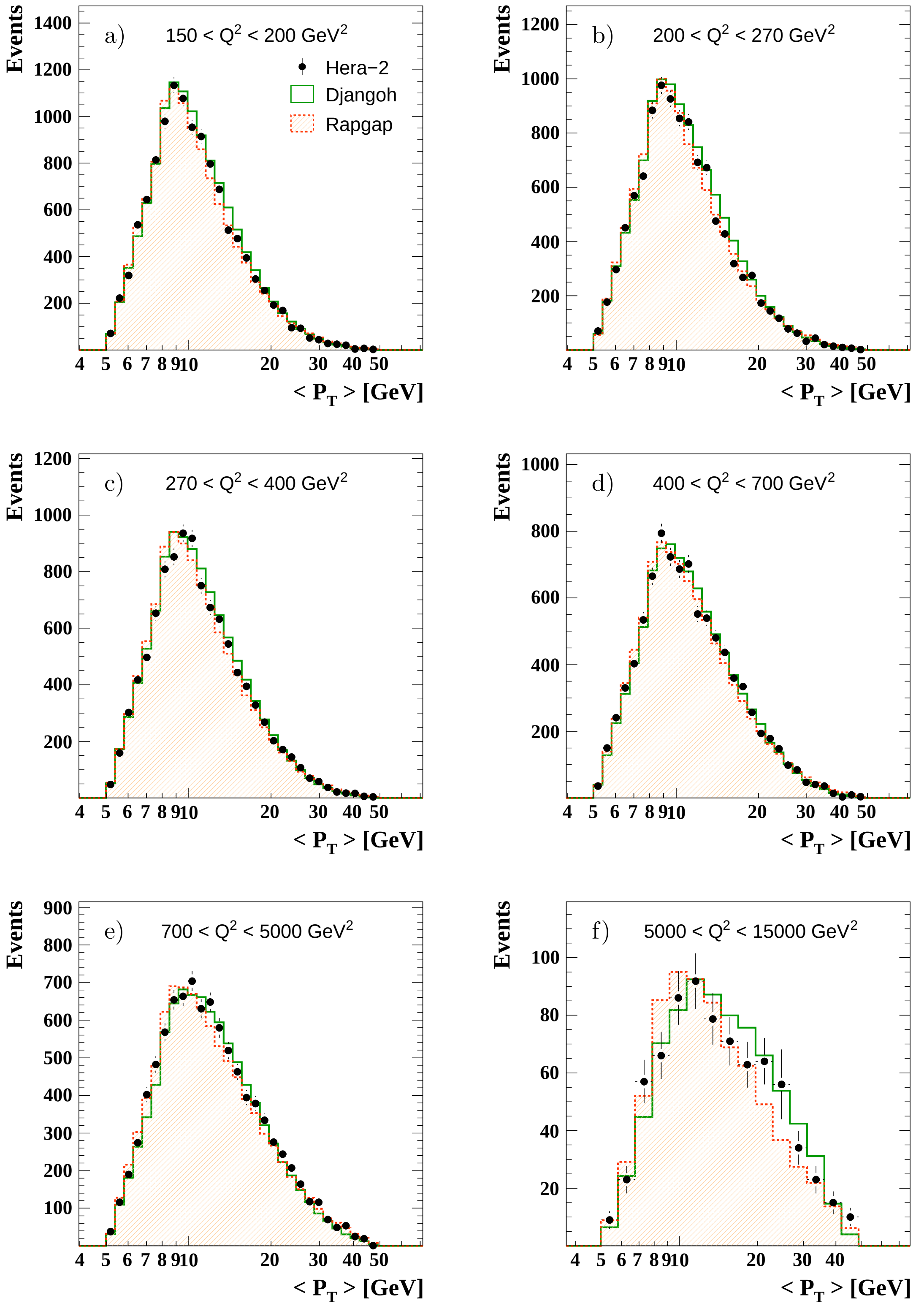


Figure 9.31: Mean transverse momentum in the Breit frame of the two leading jets, defined as $\langle P_{\rm T} \rangle = \frac{1}{2}(P_{\rm T,1} + P_{\rm T,2})$, shown for different Q^2 ranges.

In figure 9.31 the mean transverse momentum of the two jets, defined as

$$\langle P_{\mathrm{T}} \rangle = \frac{1}{2}(P_{\mathrm{T},1} + P_{\mathrm{T},2}) \tag{9.20}$$

is shown for different ranges of Q^2. The data are well described by the simulations over the full kinematically allowed range of this analysis.

9.10 Trijet Sample

If an event contains at least three jets, with the same requirements on P_{T} and M_{12} as in the dijet sample, it is part of the trijet sample. The requirement on M_{12} is neither necessary from the experimental nor from the theoretical point of view. However, it ensures that the trijet sample is a subset of the dijet sample, which is advantageous if these two samples are used in a common fit [H110c]. The requirement $M_{12} > 16\,\mathrm{GeV}$ removes 11% of the trijet events. In 97% of all trijet events the leading jet has a transverse momentum greater than 7 GeV, also making the trijet sample almost a subsample of the inclusive jet sample. In the HERA-2 dataset a total of 10 351 trijet events is found. The total amount of background from *ep* processes is 0.4%, shared to equal amounts between photoproduction and low Q^2 DIS processes, similar to the dijet sample. The contribution from QEDC scattering to the background is negligible in the trijet sample. No jets originating from QPM-type processes are found in the sample.

In LO three jet production the proton's momentum fraction carried by the emerging parton can be calculated similarly to ξ (9.19), by replacing M_{12} by the invariant mass of the three leading jets, M_{123}. In the trijet sample ξ is reconstructed using the relation

$$\xi = x + \frac{M_{123}^2}{ys}\,, \tag{9.21}$$

where y is reconstructed using the hadron method, similar to the dijet sample. Distributions of kinematic variables of the trijet sample are shown in figure 9.32. Both MC simulations describe the data very well. The mean value of ξ is with 0.06 larger than in the dijet sample, the covered parton momentum fraction is approximately $0.01 < \xi < 0.5$. The inelasticity is shifted to larger values with respect to the inclusive jet and dijet sample, peaking at $y \sim 0.35$. The shift in the longitudinal momentum balance between DJANGOH and the data is smaller than in the other samples.

Distributions of jet observables are shown in figure 9.33, for the three leading jets in the trijet sample. The mean value of P_{T} shifts from approximately 15 GeV for the leading jet to 6 GeV for the third jet. Simultaneously, the distribution of the pseudorapidity in the laboratory frame is shifted towards larger values, going from the leading to the third jet. Therefore, the extended η_{lab} range of this analysis with respect to a previous jet measurement from H1 [H110c], leads to a large improvement of statistics in the trijet sample, with about 40% more observed trijet events. The distributions of the transverse momenta of all individual jets are well described by both MC simulations. In η_{lab} the models show larger differences. The distribution of $\eta_{\mathrm{lab},2}$ in DJANGOH is shifted towards smaller values compared to RAPGAP and

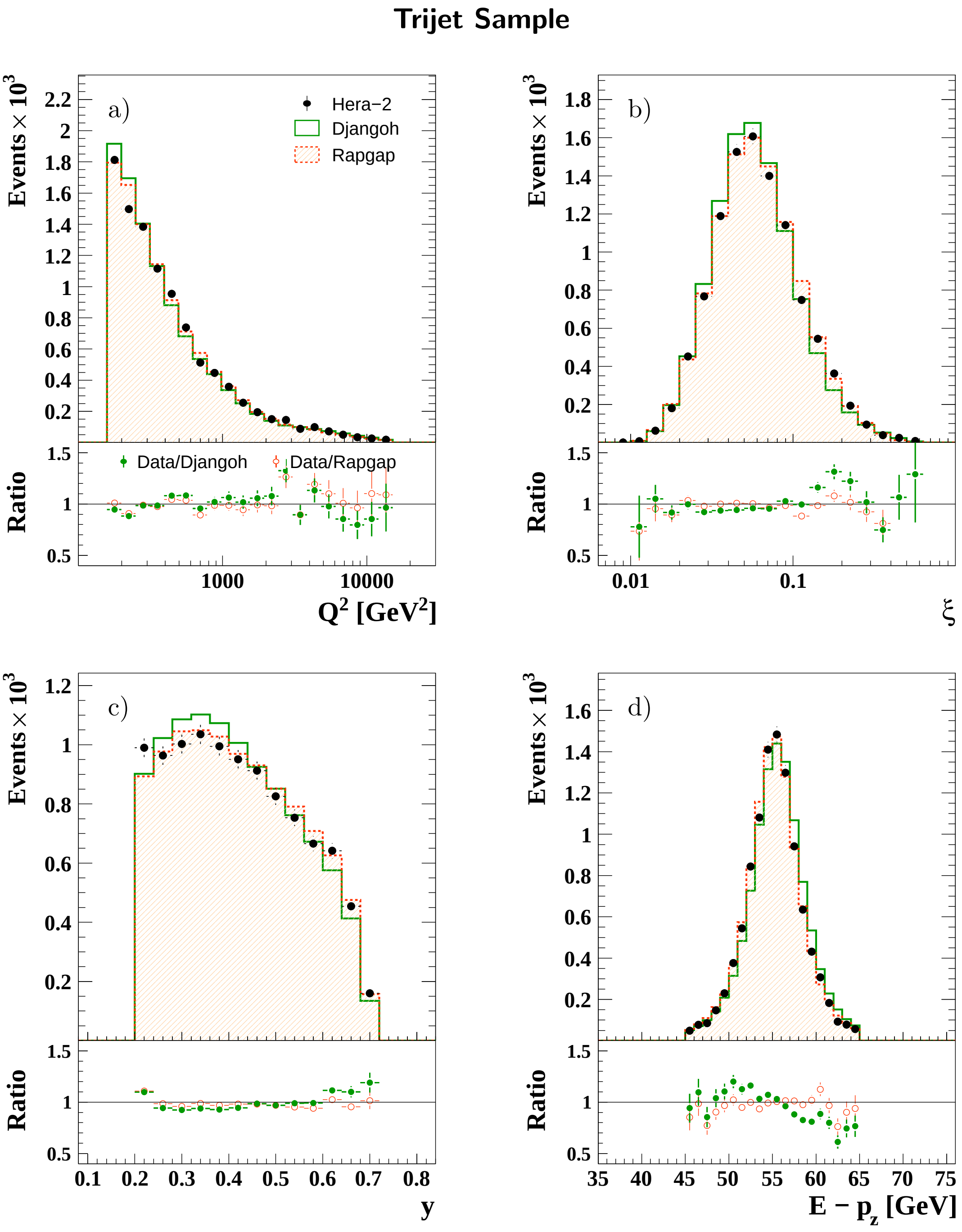
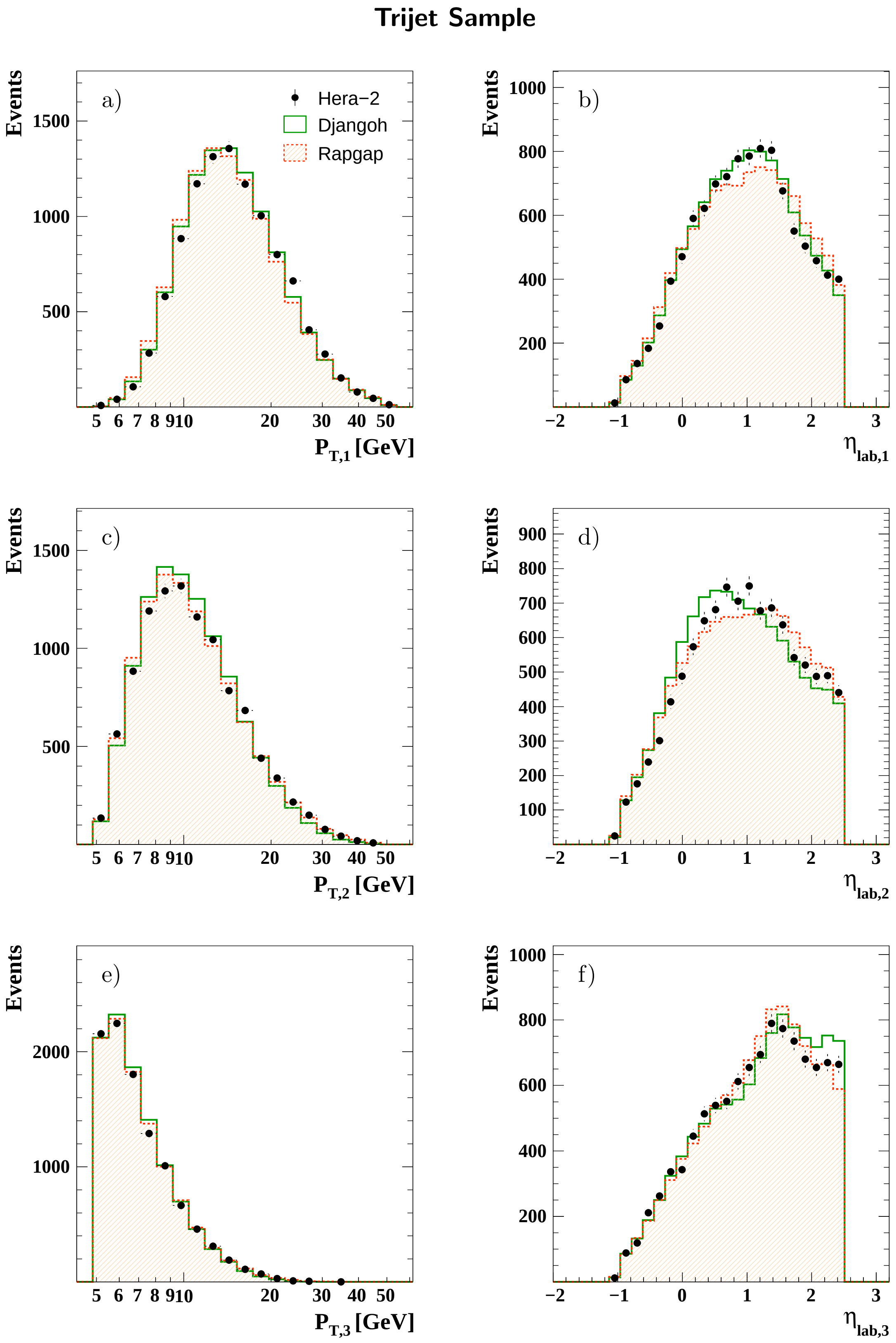


Figure 9.32: Distribution of the virtuality Q^2 (a), ξ (b), the inelasticity y (c) and the longitudinal momentum balance $E - p_z$ (d) for the trijet jet sample for the full HERA-2 dataset. The background contribution is approximately 0.4% and not shown.



Trijet Sample

Figure 9.33: Distributions of the transverse momentum in the Breit frame and the pseudorapidity in the laboratory frame for the three leading jets in the trijet sample. The background contribution is approximately 0.4% and not shown.

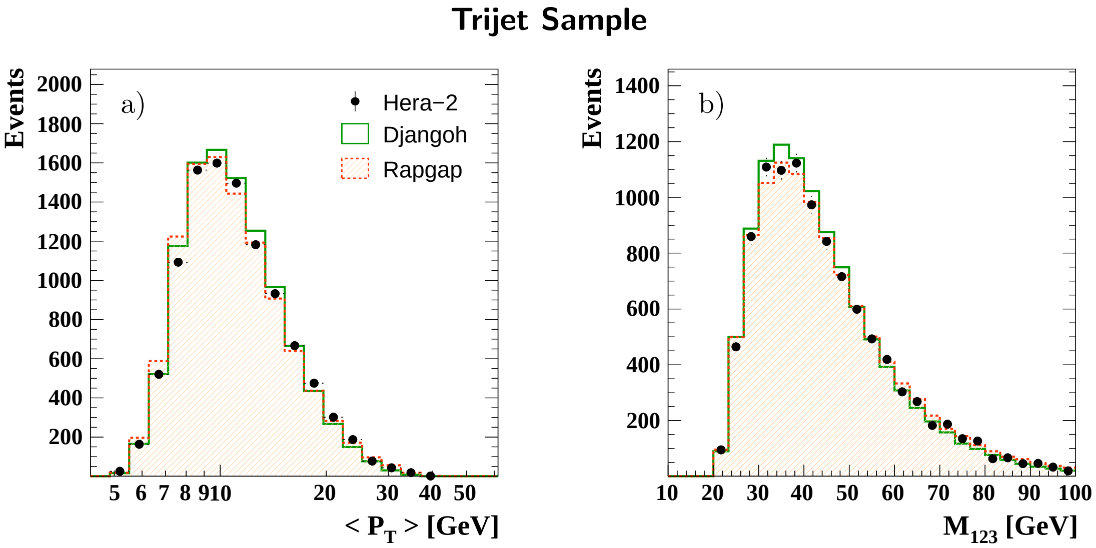


Figure 9.34: Distributions of the mean transverse momentum in the Breit frame (a) and the invariant mass (b) of the three leading jets for trijet events. The background contribution of 0.4% is not shown.

the data. At large values of the pseudorapidity of the third jet, $\eta_{\mathrm{lab},3}$, Djangoh lies above Rapgap, with the data lying between the predictions. These differences are taken into account by the systematic uncertainty due to the model dependence of the detector correction.

The mean transverse momentum in trijet events is defined as

$$\langle P_{\mathrm{T}} \rangle = \frac{1}{3}(P_{\mathrm{T},1} + P_{\mathrm{T},2} + P_{\mathrm{T},3})\,. \tag{9.22}$$

The distribution of $\langle P_{\mathrm{T}} \rangle$ is shown together with that of the invariant mass of the three leading jets, M_{123}, in figure 9.34. $\langle P_{\mathrm{T}} \rangle$ reaches values of $5.5 < \langle P_{\mathrm{T}} \rangle < 40\,\mathrm{GeV}$, with a mean value of approximately 10 GeV. The invariant mass of the three jets peaks at 35 GeV and is always larger than 20 GeV. Both distributions are well described by the MC simulations.

9.11 Summary

The definitions of the phase space for the NC DIS, inclusive jet, dijet and trijet measurements are summarised in table 9.6. The larger phase space in η_{lab} compared to a previous jet measurement of H1 [H110c], leads to larger statistics. Whereas in the inclusive jet sample the number of observed jets is comparable with the one from the previous analysis, in the dijet and trijet samples more events are observed, even though a smaller dataset is used. In the dijet sample the gain in statistics is 10%. Statistically significant is the increase of the number of trijet events with about 40% more events than in the previous multijet analysis.

Sample	Phase space	Events
NC DIS	$150 < Q^2 < 15000\,\mathrm{GeV}^2$ $0.2 < y < 0.7$	337 048 events
Global jet selection	inclusive k_T-jets, $R_0 = 1$ P_T-recombination scheme $2.5 < \eta_\mathrm{lab} < -1.0$	
Inclusive jets	$7 < P_\mathrm{T} < 50\,\mathrm{GeV}$	133 621 jets (in 90 328 events)
Dijets	$5 < P_\mathrm{T} < 50\,\mathrm{GeV},\ M_{12} > 16\,\mathrm{GeV}$	52 028 events
Trijets	$5 < P_\mathrm{T} < 50\,\mathrm{GeV},\ M_{12} > 16\,\mathrm{GeV}$	10 351 events

Table 9.6: Phase space definition and statistics for NC DIS and jet samples.

Chapter 10

Measurement of Jet Cross Sections

The different data samples, obtained with the selection described in the previous chapter, are used to measure jet cross sections in bins of Q^2, P_{T} and ξ. In this chapter the chosen bin boundaries are introduced, and the migrations between bins are studied. In order to obtain cross sections independent from the experimental setup, the data are corrected for detector and QED radiation effects. The sources of systematic uncertainties and how they are estimated are given. This is followed by a summary of the statistical and systematic uncertainties for the different jet cross section measurements. The chapter concludes with a description of the corrections applied to the NLO calculations and the estimation of their uncertainties. These corrections are needed to account for effects not included in these calculations, namely Z^0 exchange and hadronisation effects.

10.1 Observables and Bin Definitions

The inclusive jet, dijet and trijet cross sections are denoted by σ_{jet}, $\sigma_{\text{2-jet}}$ and $\sigma_{\text{3-jet}}$, respectively. They are measured single and double differentially in Q^2, $P_{\mathrm{T,obs}}$ and ξ, where $P_{\mathrm{T,obs}}$ denotes P_{T} for inclusive jets and $\langle P_{\mathrm{T}} \rangle$ for di- and trijets. The binning in Q^2, $P_{\mathrm{T,obs}}$ and ξ is chosen to be identical with previous jet measurements from H1 [H107, H110c] in order to obtain directly comparable cross sections, where applicable. In the case of $P_{\mathrm{T,obs}}$ and ξ, the bin widths are approximately 4σ, where σ is the resolution of the measurement in a given observable. Four bins in $P_{\mathrm{T,obs}}$ and ξ are defined for every Q^2 bin. In the case of the trijet measurement the last bin in $\langle P_{\mathrm{T}} \rangle$ is not used, because of the smallness of the cross section, which results in too little statistics in this bin. In the case of Q^2, the bin widths are larger than 4σ to obtain high enough statistics in every range of Q^2. A summary of the bin labels and boundaries is given in table 10.1.

The normalised jet cross sections are defined as the ratio of the differential inclusive jet, dijet and trijet cross sections to the differential inclusive NC DIS cross section in a given Q^2 bin. Single differential normalised inclusive jet cross sections are given by

$$\frac{\sigma_{\text{jet}}}{\sigma_{\text{NC}}}(Q_i^2) = \frac{\sigma_{\text{jet}}(Q_i^2)}{\sigma_{\text{NC}}(Q_i^2)}\,, \tag{10.1}$$

where $\sigma_{\text{jet}}(Q_i^2)$ and $\sigma_{\text{NC}}(Q_i^2)$ denote the integrated inclusive jet and integrated in-

Bin label	Q^2 range (in GeV^2)
1	$150 \leq Q^2 < 200$
2	$200 \leq Q^2 < 270$
3	$270 \leq Q^2 < 400$
4	$400 \leq Q^2 < 700$
5	$700 \leq Q^2 < 5000$
6	$5000 \leq Q^2 < 15000$

Bin label	P_T or $\langle P_\mathrm{T} \rangle$ range (in GeV)
α	$7 \leq P_\mathrm{T} < 11$
β	$11 \leq P_\mathrm{T} < 18$
γ	$18 \leq P_\mathrm{T} < 30$
δ	$30 \leq P_\mathrm{T} < 50$

Bin label	Dijet ξ range
a	$0.006 \leq \xi < 0.020$
b	$0.020 \leq \xi < 0.040$
c	$0.040 \leq \xi < 0.080$
d	$0.080 \leq \xi < 0.316$

Bin label	Trijet ξ range
A	$0.01 \leq \xi < 0.04$
B	$0.04 \leq \xi < 0.08$
C	$0.08 \leq \xi < 0.50$

Table 10.1: Bin labels and boundaries for bins in Q^2 (top left), $P_\mathrm{T,obs}$ (top right), ξ for the dijet (bottom left) and ξ for the trijet (bottom right) measurements.

clusive NC cross sections in a given Q^2 bin i, respectively. The normalised dijet and trijet cross sections are defined analogously. In the case of a double differential measurement, the normalised inclusive jet cross section in Q^2 bin i and P_T bin j is defined as

$$\frac{\sigma_\mathrm{jet}}{\sigma_\mathrm{NC}}(Q_i^2, P_{\mathrm{T},j}) = \frac{\sigma_\mathrm{jet}(Q_i^2, P_{\mathrm{T},j})}{\sigma_\mathrm{NC}(Q_i^2)} , \tag{10.2}$$

with the variable $\sigma_\mathrm{jet}(Q_i^2, P_{\mathrm{T},j})$ denoting the integrated inclusive jet cross section in the given Q^2, P_T bin. Similar definitions are used for other observables and for the normalised dijet and trijet cross sections. The normalised inclusive jet cross sections can be considered as the average jet multiplicity in a given Q^2 region. The normalised dijet and trijet cross sections can be viewed as the average dijet and trijet event rates.

10.2 The Detector, Hadron and Parton Level

The Detector Level

Recorded events from ep collisions were measured with an imperfect detector with finite acceptance and intrinsic resolution. In order to estimate the effect of these imperfections, simulated events from MC event generators are passed through the detector simulation H1Sim. After this they are passed through the same reconstruction algorithms as the recorded data. At this stage, simulated events are said to be on detector level and are directly comparable to the reconstructed data (see also section 4.3).

The Hadron Level

Cross sections corrected for detector effects are termed hadron level cross sections. They are comparable to calculations including all known effects occuring in high energy collisions. In MC event generators hadron level cross sections are obtained by defining observables for the hadronic final state, i.e. after the hadronisation stage. Since many particles produced in high energy collisions are unstable and will decay soon after their production, the hadronic final state is chosen to be the list of all stable hadrons. Here stable refers to particles with an average decay length much larger than the dimensions of the beam pipe, which means that the lifetime τ is required to be larger than $3 \cdot 10^{-10}\,\mathrm{s}$. This definition renders all particles stable which have the chance to be detected within the active detector volume.
Usually one is not interested in the well known effects from QED radiation, which are generally not included in pQCD calculations. Hence, the hadron level cross sections are corrected for higher order QED effects as well. In order to obtain corrections leading from the detector level to a level corrected for QED radiation, a two-step procedure is used. The total experimental correction c_{exp} is defined as

$$c_{\mathrm{exp}} = c_{\mathrm{det}} \cdot c_{\mathrm{qed}} \,, \tag{10.3}$$

where c_{det} is the correction for detector effects, and c_{qed} corrects the hadron level for effects from QED radiation. The separation into contributions from detector and higher order QED effects leaves some ambiguity in the definition of the intermediate, radiative hadron level. In events with a radiative photon the event kinematics are not uniquely defined. If the photon has a small angle with respect to the scattered electron, it will not be resolved at the detector level. On the contrary, in simulated events on hadron level the scattered electron and the radiated photon can be distinguished. In this case the corrections c_{det} and c_{qed} will be artificially large. However, these effects cancel in the definition of c_{exp}, which remains correct. In order to obtain meaningful definitions already for the intermediate corrections c_{det} and c_{qed}, the radiative hadron level is defined in the following way:

- In an event with a radiated photon, the four-vector of the generated scattered electron p_{e} is combined with the four-vector of the radiated photon p_{γ}, if the opening angle between them is smaller than $\alpha_{\mathrm{e}\gamma}$. In this case the radiated photon is removed from the hadronic final state.

- All remaining photons are sufficiently isolated from the scattered electron and are treated as part of the hadronic final state.

The exact value of the resolution angle $\alpha_{\mathrm{e}\gamma}$ is not crucial, however a value close to the resolution power of the LAr calorimeter for the electron-photon system should be used. Studies showed that the choice $\alpha_{\mathrm{e}\gamma} = 5^{\circ}$ is reasonable for all scattering angles.

The Parton Level

Fixed-order pQCD calculations do not include long-range hadronisation effects. Instead, the obtained jet cross sections, after applying the same jet finder as for the

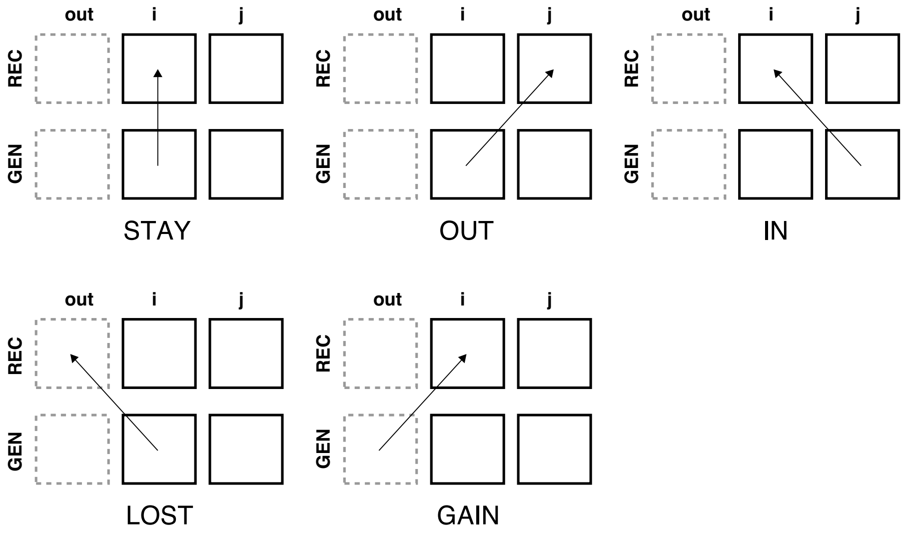


Figure 10.1: Schematic showing the possible migrations between reconstructed and generated level. Solid squares illustrate bins inside the phase space of the measurement, dashed squares illustrate regions outside the phase space.

data, are defined for the partonic final state and are said to be parton level jet cross sections. In order to compare these calculations to measured cross sections, they have to be corrected for the long-range hadronisation effects. To compute the size of these hadronisation corrections MC event generators are used, which is explained in section 10.6.2.

10.3 Acceptance, Purity and Stability

To study detector effects and the impact of migrations between individual bins on the measurement, observables on detector level (reconstructed MC events) and the intermediate radiative hadron level (generated MC events) are considered. For each measurement bin i, let N_{gen} and N_{rec} be the number of generated and reconstructed events, respectively. Then the following quantities may be defined:

- N_{stay} is the number of events generated and reconstructed in bin i.
- N_{out} is the number of events generated in bin i, but reconstructed in another bin inside the phase space of the measurement.
- N_{in} is the number of events reconstructed in bin i, but generated in another bin inside the phase space of the measurement.
- N_{lost} is the number of events generated in bin i, but reconstructed outside the phase space of the measurement.
- N_{gain} is the number of events reconstructed in bin i, but generated outside the phase space of the measurement.

These quantities are illustrated in figure 10.1. The quantities N_{stay}, N_{out} and N_{in} are related to migrations between the defined bins, whereas N_{gain} and N_{lost} describe migrations in and out of the phase space defined by the bin grid. These occur mainly because of the finite resolutions in the reconstruction of P_{T}, η, ξ and M_{12}. The quantities defined above are related to N_{gen} and N_{rec} via

$$\begin{aligned} N_{\mathrm{gen}} &= N_{\mathrm{stay}} + N_{\mathrm{out}} + N_{\mathrm{lost}}\,, \qquad (10.4) \\ N_{\mathrm{rec}} &= N_{\mathrm{stay}} + N_{\mathrm{in}} + N_{\mathrm{gain}}\,. \qquad (10.5) \end{aligned}$$

To quantify the relationship between reconstructed and generated variables as well as effects from smearing due to detector effects, the acceptance, purity and stability are studied.

Acceptance

The acceptance A is defined as

$$A = \frac{N_{\mathrm{rec}}}{N_{\mathrm{gen}}} \qquad (10.6)$$

and quantifies the number of events reconstructed in a bin with respect to the true number of events generated in this bin. It thus accounts for all reconstruction effects and detector inefficiencies.

Purity

The purity P is defined as

$$P = \frac{N_{\mathrm{stay}}}{N_{\mathrm{rec}}} \qquad (10.7)$$

and is the fraction of events generated and reconstructed in the same bin with respect to the total number of events reconstructed in this bin.

Stability

The stability S is defined as

$$S = \frac{N_{\mathrm{stay}}}{N_{\mathrm{gen}} - N_{\mathrm{lost}}} \qquad (10.8)$$

and describes the fraction of events generated and reconstructed in the same bin with respect to the number of events generated in this bin, which stay within the phase space of the measurement. The events migrating out of the visible phase space are ignored since they are already taken into account by the acceptance.

The values of the acceptance, stability and purity for each bin used in this measurement are required to be larger than 40% in order to ensure that migrations are well understood. In the bulk of the phase space of the measurement A, P and S are above 50%. Only for some of the lowest $P_{\mathrm{T,obs}}$ bins, the purity lies between 40–50%.

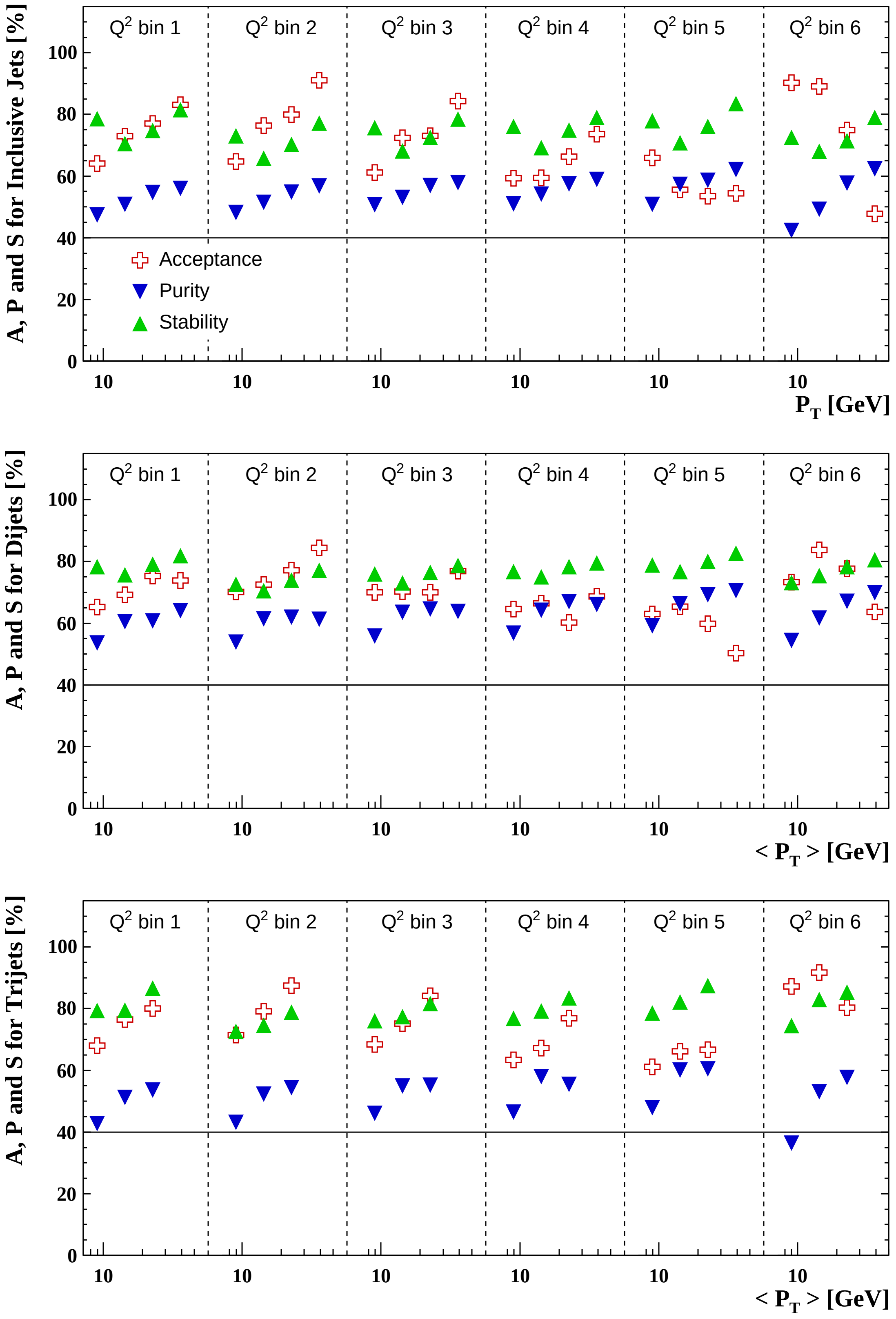


Figure 10.2: Acceptance, purity and stability for inclusive jets (top), dijets (middle) and trijets (bottom) for the Q^2, $P_{\mathrm{T,obs}}$ binning. The statistical errors are smaller than the histogram markers. The horizontal line represents the limit of 40%.

Q^2, P_T Binning

The values of A, P and S for the Q^2, $P_\mathrm{T,obs}$ binning are shown in figure 10.2 for the inclusive jet, dijet and trijet measurements.
In the inclusive jet sample the acceptance is between 60–90% in the first four Q^2 bins, where it rises as function of P_T. In the Q^2 bins 5 and 6 the acceptance shows the opposite behaviour but stays above 50% over the full range in P_T. This is due to the separation between detector and radiative correction factors. The radiative correction factors for inclusive jets show a more significant Q^2 dependence than for the dijet and trijet measurements. In the Q^2 bins 5 and 6 the radiative correction factors exhibit a slope in the opposite direction with respect to the acceptance. The total correction factors c_exp are relatively flat in P_T in all Q^2 bins, as shown below. The stability is relatively flat as function of P_T and lies between 70–80% in all Q^2 bins. The purity rises slowly as function of P_T. It takes values between 50–60% in all bins, except for bin 6α, where it is 43%. The reason for the purity to be smaller than the stability is the steeply falling P_T distribution. Only if the binning can be chosen such that the measured distribution is flat, will the purity and stability be approximately equal. Since the P_T distribution falls too steeply to achieve a flat measurement a compromise is made. On one hand the binning is chosen such that the bin widths are not smaller than 4σ for the migrations to be well controlled. On the other hand the bin widths are chosen to be not too large, such that the shape of the underlying distribution can still be measured.
In the dijet sample the acceptance and the stability lie between 60–80%. The purity is larger than in the inclusive jet sample and takes values of 55–70%. In the smallest $\langle P_\mathrm{T}\rangle$ bins the purity takes the lowest values of $\sim$55%, which is a consequence of the M_{12} cut. Since small values of M_{12} correspond to small $\langle P_\mathrm{T}\rangle$, the purity decreases due to events with small M_{12} smearing into the sample. An additional bin in the range $5 < \langle P_\mathrm{T}\rangle < 7\,\mathrm{GeV}$ has been studied, but was discarded due to too small values of the stability and purity.
In the trijet sample the values of the acceptance and stability are mostly between 70–80%, with the smallest values being 60% and the largest ones 90%. The purity shows the same behaviour as for the dijet sample, but is lower by about 10% due to the requirement of a third jet which has on average a relatively small P_T. This introduces migrations into the sample, which are larger than in the dijet case, where on average only one jet with low P_T is observed. In bin 6a, which is the bin with the highest values of Q^2 and the smallest values of $\langle P_\mathrm{T}\rangle$, the purity is smaller than 40%. This bin is removed from the measurement.

Q^2, ξ Binning

The values of A, P and S for the binning in Q^2 and ξ are shown in figure 10.3 for the dijet and trijet measurements. They show a similar behaviour to the binning in Q^2 and $\langle P_\mathrm{T}\rangle$. For the trijet measurement the purity is about 10% lower than for the dijet measurement. In bin 5a (5A), which is the lowest ξ bin in Q^2 bin 5, the purity is below 40% in the dijet (trijet) measurement. The reason is a small total cross section in this bin with respect to the next bin in ξ. This leads to significant migrations from bin 5b (5B) into bin 5a (5A). The bins 5a and 5A are discarded from both measurements.

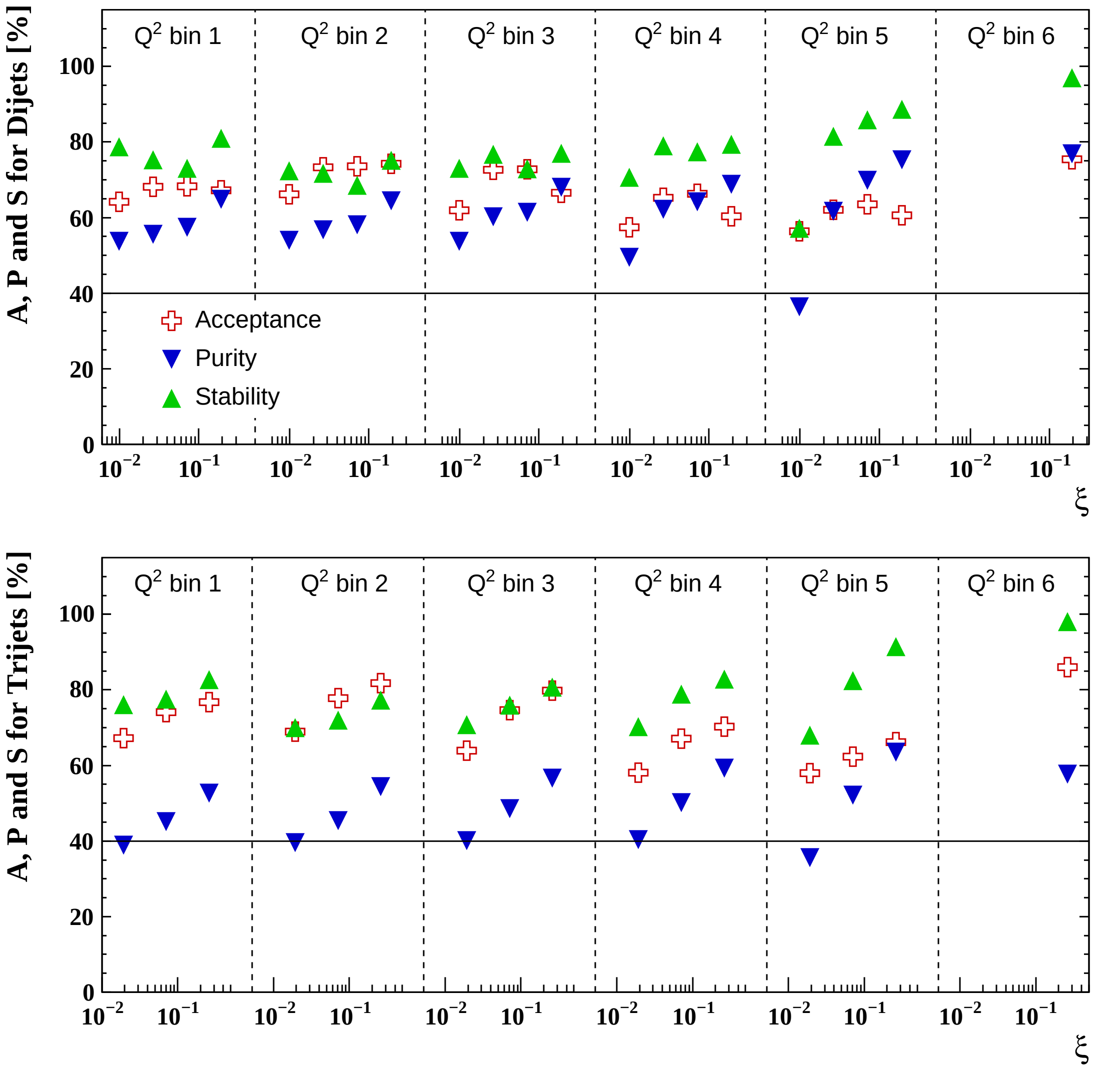


Figure 10.3: Acceptance, purity and stability for dijets (top) and trijets (bottom) for the Q^2, ξ binning. The statistical errors are smaller than the histogram markers. The horizontal line represents the limit of 40%.

In a previous measurement values of the purity between 70–80% were reported for the binning in $\langle P_\mathrm{T} \rangle$ in the case of the dijet measurement [Gou08]. For the trijet measurement, values of $P \sim 90\%$ were observed for the binning in Q^2. These high values of P are due to a different definition of P, which does not take migrations from outside of the defined phase space into account, $P_m = N_\mathrm{stay}/(N_\mathrm{stay} + N_\mathrm{in})$. Especially in the trijet measurement migrations from outside the defined phase space are the limiting factor for the purity. Using the definition of P_m for the purity, similar values as in the previous multi-jet measurement are observed.

10.4 Corrections to Data

In general, the cross section in a given bin is related to the number of corrected events in this bin for an integrated luminosity L via

$$\sigma = \frac{N}{L} \,. \tag{10.9}$$

Using a procedure where migrations and detector effects are corrected for bin-by-bin, the measured cross section in a given bin i is related to the number of observed events in this bin, N^i_data, through

$$\sigma^i = \frac{N^i_\mathrm{data} - N^i_\mathrm{bkgd}}{L \cdot A^i} \, c^i_\mathrm{qed} \,. \tag{10.10}$$

Here N^i_bkgd, A^i and c^i_qed are the number of events from *ep* background processes as estimated by MC simulations, the acceptance and the correction for QED effects in bin i, respectively. Equation (10.10) can be written as

$$\sigma^i = \sigma^i_\mathrm{meas} \, c^i_\mathrm{det} \, c^i_\mathrm{qed} \,, \tag{10.11}$$

with the definition of the background-subtracted cross section measured on detector level, $\sigma^i_\mathrm{meas} = (N^i_\mathrm{data} - N^i_\mathrm{bkgd})/L$, and the detector correction $c^i_\mathrm{det} = 1/A^i$.

10.4.1 Detector Correction

In this work a bin-by-bin correction is employed, which relies on a good simulation of migrations between bins as well as an adequate knowledge of migrations in and out of the samples used for the cross section measurements. The first requirement is connected to a good simulation of experimental effects and resolutions. These have been studied in great detail as shown in chapters 6–8. Additionally, the underlying cross sections have to be well modelled which is achieved by a reweighting of the MC cross sections. To achieve a good description of migrations into the samples, the data have to be well described by the simulations outside the defined phase space. This has been checked by relaxing the respective cuts, which has been partly shown in chapter 9.

Overall, the data are very well described by the MC simulations. In regions where the data are only reasonably well described, usually the RAPGAP and DJANGOH simulations differ by a larger amount than their deviations from the data. These differences may lead to different acceptance corrections, which are taken into account by using the arithmetic mean

$$c_{\mathrm{det}} = \frac{1}{2}\left(c_{\mathrm{det}}^{\mathrm{R}} + c_{\mathrm{det}}^{\mathrm{D}}\right) , \tag{10.12}$$

where $c_{\mathrm{det}}^{\mathrm{R}}$ and $c_{\mathrm{det}}^{\mathrm{D}}$ stand for the detector corrections obtained with Rapgap and Djangoh, respectively. The uncertainty of this procedure is estimated by the model dependence, which also takes the radiative correction into account and is described in more detail in section 10.5.

10.4.2 Radiative Correction

QED corrections can influence the measured jet cross sections through four mechanisms:

1. Real photon radiation will result in a shift of the kinematic variables, which leads to an additional smearing in the defined bins in Q^2.
2. The emission of real photons may lead to an error in the reconstruction of the boost vector. A badly reconstructed boost to the Breit frame can lead to an additional transverse component with respect to the photon-proton axis, which may be misinterpreted as an underlying QCD process.
3. Photons emitted under large angles with significant transverse momenta will not be identified during the reconstruction and will be reconstructed as jets, which enhances the measured jet cross sections.
4. Electroweak virtual corrections for γ and Z^0 exchange lead to a renormalisation of α_{em}, which becomes a function of Q^2, $\alpha_{\mathrm{em}} = \alpha_{\mathrm{em}}(Q^2)$. This changes the cross section as function of Q^2, with α_{em} being about 2.3% larger in the highest Q^2 bin than in the lowest Q^2 bin of this analysis.

It was already shown in a previous analysis [Wob00] that the $e\Sigma$-reconstruction method removes systematic biases in the reconstruction of the boost vector due to photons emitted collinear to the incoming electron. This is also verified by checking the transverse momentum components of the hadronic final state in the Breit frame. This shows that QED effects are well modelled in Heracles, which is interfaced to Rapgap and Djangoh. Heracles includes single photon emission from the incoming and scattered electron for γ and Z^0 exchange and the full electroweak virtual corrections to one loop accuracy, i.e. vertex correction, self energy and vacuum polarisation graphs, as shown in figure 10.4.

The effects of QED radiation to the jet cross sections are corrected for by comparing cross sections on hadron level including QED radiation, σ_{rad}, with cross sections on hadron level without QED radiation σ_{norad}. Consequently, the jet cross sections are corrected for QED effects resulting from diagrams (a)–(c) in figure 10.4 and correspondingly to numbers 1–3 in the above list. The running of the electromagnetic coupling is not corrected for. It is taken into account when calculating σ_{norad}. Therefore, the corresponding diagrams (d)–(f) in figure 10.4, which lead to the running of

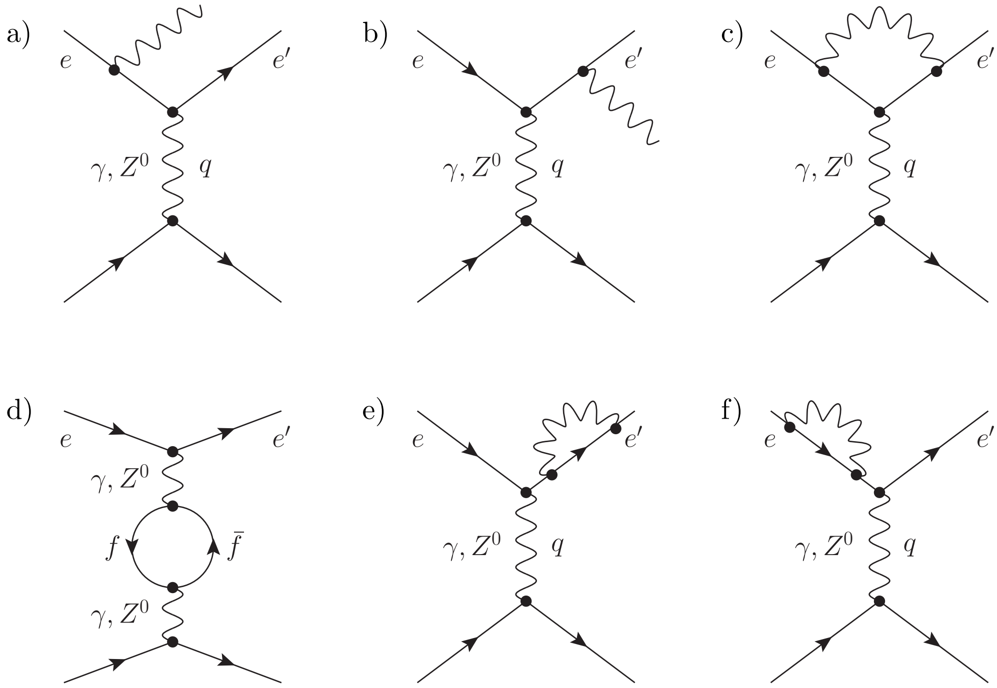


Figure 10.4: Diagrams of QED corrections to the process $e^{\pm}p \rightarrow e^{\pm}X$, included in HERACLES: real emission of photons (a, b), vertex correction (c), vaccum polarisation (d) and self energy (e, f) diagrams.

α_{em} and the renormalised electron mass, are taken into account when the final cross sections are given.
For every bin i, the radiative correction c^{i}_{qed} is given by

$$c^{i}_{\mathrm{qed}} = \frac{\sigma^{i}_{\mathrm{norad}}}{\sigma^{i}_{\mathrm{rad}}}\,, \tag{10.13}$$

where $\sigma^{i}_{\mathrm{norad}}$ and $\sigma^{i}_{\mathrm{norad}}$ are estimated using RAPGAP and DJANGOH.
The radiative correction factors are shown together with the detector and total correction factors for the dijet measurement as function of Q^2 and $\langle P_{\mathrm{T}} \rangle$ in figure 10.5. At low Q^2 and low $\langle P_{\mathrm{T}} \rangle$ the radiative corrections are small, becoming larger towards larger Q^2 and $\langle P_{\mathrm{T}} \rangle$. The detector correction factors are relatively flat around 1.5, leading to total experimental correction factors between 1.2 and 1.4.

10.4.3 Total Experimental Correction

The NC DIS sample is the basis for the jet selection, hence it is instructive to study its acceptance correction before looking at the jet samples. The requirements on the scattered electron, like the fiducial volume cut, rejection of ϕ and z-cracks and the requirement $E_{\mathrm{e}} > 11\,\mathrm{GeV}$ reduce the geometrical acceptance of the LAr calorimeter to about 0.7. Additional requirements for the background-rejection reduce the acceptance to approximately 0.6. None of these requirements exhibit a Q^2 dependence, which is why the acceptance is relatively flat over the full range in Q^2, except for the highest Q^2 bin with a higher acceptance of 0.75. These corrections are well

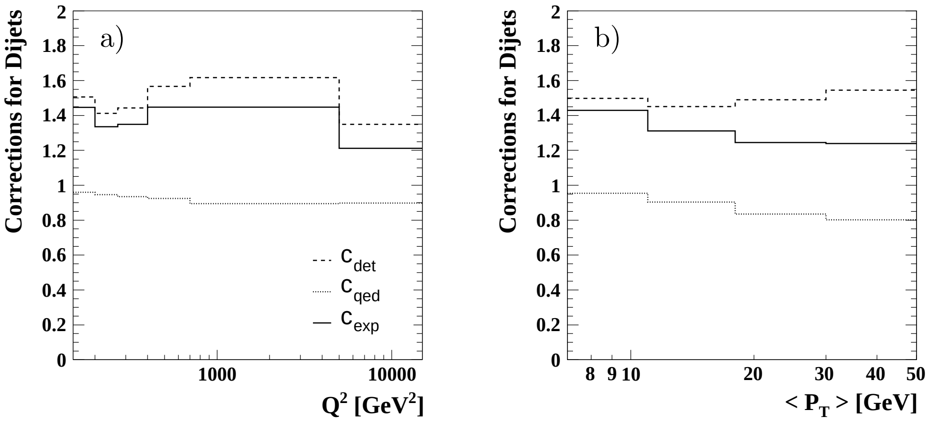


Figure 10.5: Correction factors for the dijet measurement as function of Q^2 (a) and $\langle P_\mathrm{T} \rangle$ (b). The detector (radiative) correction c_det (c_qed) is illustrated with a dashed (dotted) line, the total experimental correction $c_\mathrm{exp} = c_\mathrm{det} \cdot c_\mathrm{qed}$ is illustrated with a solid line. The statistical errors introduced by limited MC statistics are smaller than 0.3% and not shown here.

controlled, with most of them being geometrical corrections for the scattered electron only. The efficiencies of the remaining cuts have been studied in detail and are well understood, see chapter 9 and references therein. As a consequence of the flat acceptance, the experimental correction factors for the NC DIS are flat and take values around 1.5 (1.3 in the highest Q^2 bin).

In figure 10.6 the experimental correction factors for the inclusive jet, dijet and trijet measurements for the Q^2, $P_\mathrm{T,obs}$ binning are shown. In general, the acceptance rises with increasing $P_\mathrm{T,obs}$ (also see figure 10.2). At low values of $P_\mathrm{T,obs}$ it is close to the acceptance of the NC DIS sample, whereas at large values of $P_\mathrm{T,obs}$ it can take values of 0.8–0.9. Therefore the experimental correction factors for the jet measurement gets closer to unity for large $P_\mathrm{T,obs}$. For the inclusive jet sample the values of c_exp are closer to unity than for the dijet and trijet samples due to the requirement of a larger jet transverse momentum and no cut on M_{12}. In most of the bins the experimental correction factors take values between 1.0 and 1.5 for all jet samples. Only in the highest Q^2 bin for large $P_\mathrm{T,obs}$, the experimental correction factors can become smaller than unity due to radiative correction factors with values between 0.6 and 0.8.

In figure 10.6 the experimental correction factors for the dijet and trijet measurements for the Q^2, ξ binning are shown. The values of c_exp become closer to one for increasing values of ξ. This is due to large jet transverse momenta at high values of ξ, which leads to improved values of the acceptance compared to small values of ξ. Also, at small values of ξ the requirement on M_{12} leads to smaller values of the acceptance and therefore larger values of c_exp.

10.5 Experimental Uncertainties

Uncertainties in the performance of the detector, the normalisation of the background contribution and differences in the MC models used to obtain the experimental correction factors lead to systematic uncertainties on the cross section measure-

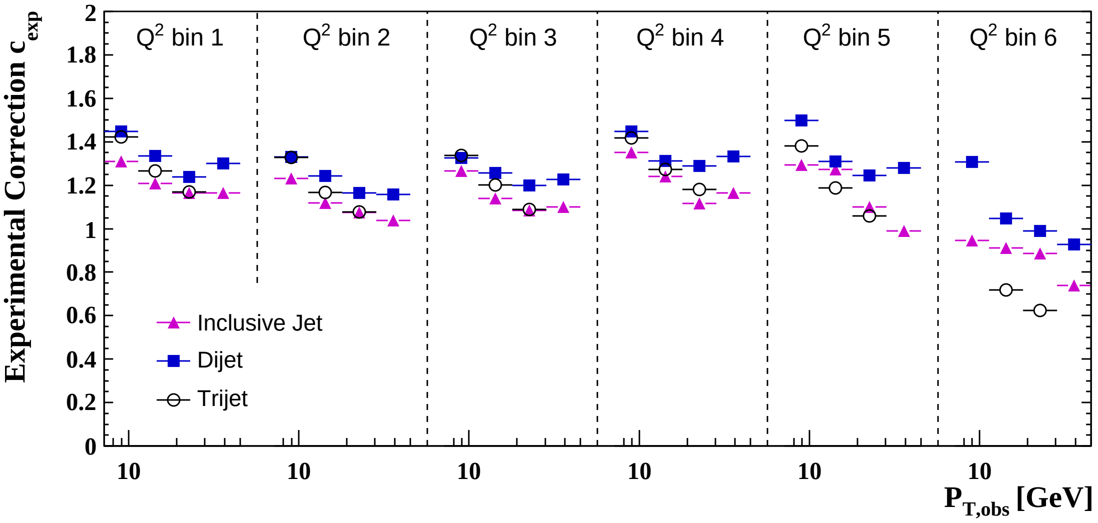


Figure 10.6: Experimental correction factors for the inclusive jet, dijet and trijet measurements as function of Q^2 and $P_{\mathrm{T,obs}}$. The statistical errors on the correction factors, introduced by limited MC statistics, are smaller than the histogram markers.

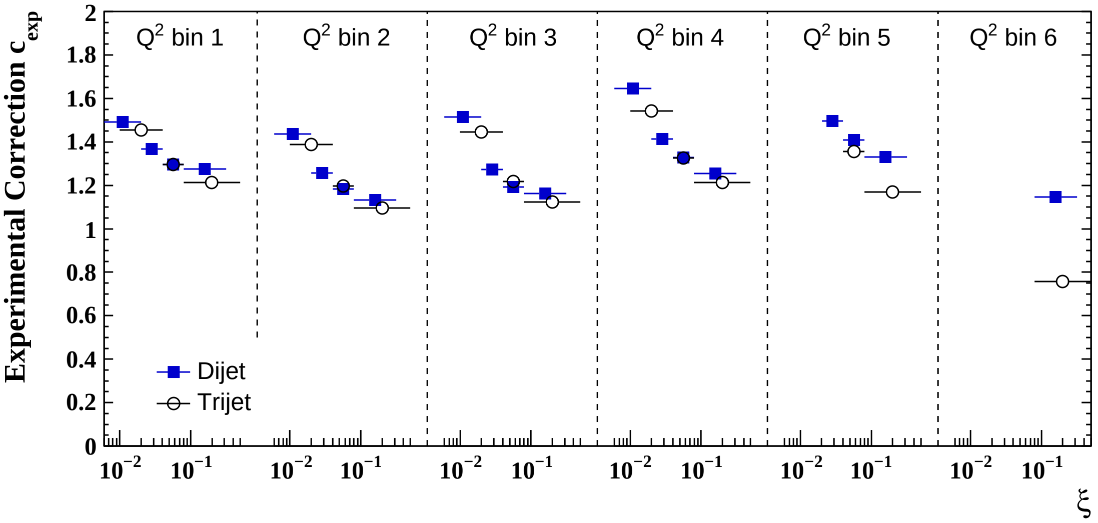


Figure 10.7: Experimental correction factors for the dijet and trijet measurements as function of Q^2 and ξ. The statistical errors on the correction factors, introduced by limited MC statistics, are smaller than the histogram markers.

ment. The uncertainties can have different sizes, depending on the observable and the region of phase space. A detailed examination of various sources of uncertainties has been performed and a summary is presented here. All systematic uncertainties are found to be symmetric to a good approximation and are assumed to be so in the following. The total experimental uncertainty is obtained by adding the individual systematic uncertainties with the statistical uncertainty in quadrature. The exact values of the statistical and systematic uncertainties from the various sources are given in the tables of the results in appendix A.

10.5.1 Statistical Uncertainty

Statistical uncertainties arise from a limited number of events used for the determination of a quantity. In general, not only the data but also the MC samples used for the determination of the detector and radiative correction have limited statistics. In practice, the statistical uncertainties from simulated events can be neglected if a large number of simulated events is used compared to the events in data. In this analysis the MC samples have approximately 25 times the luminosities of the HERA-2 dataset, resulting in a statistical uncertainty about five times smaller than in the data, which thus can be neglected safely.

However, migrations between bins can lead to an increase of the statistical uncertainty. This effect is taken into account when calculating the values of the acceptance, which enters the experimental correction factors c_{exp}. The migrations enter in the form of the uncorrelated variables N_{stay}, N_{out}, N_{lost}, N_{in} and N_{gain}, as defined in section 10.3. Expressing the acceptance A in terms of these uncorrelated variables, one obtains

$$A = \frac{N_{\mathrm{rec}}}{N_{\mathrm{gen}}} = \frac{N_{\mathrm{stay}} + N_{\mathrm{in}} + N_{\mathrm{gain}}}{N_{\mathrm{stay}} + N_{\mathrm{out}} + N_{\mathrm{lost}}} \,. \qquad (10.14)$$

The uncertainty on A is obtained via

$$\sigma_A^2 = \left(\frac{\partial A}{\partial N_{\mathrm{stay}}}\sigma_{\mathrm{stay}}\right)^2 + \left(\frac{\partial A}{\partial N_{\mathrm{in}}}\sigma_{\mathrm{in}}\right)^2 + \left(\frac{\partial A}{\partial N_{\mathrm{gain}}}\sigma_{\mathrm{gain}}\right)^2 + \ldots \qquad (10.15)$$

and propagated to the statistical uncertainty Δ_{stat} using

$$\Delta_{\mathrm{stat}} = \sqrt{\sigma_A^2 + \sigma_{\mathrm{data}}^2} \,. \qquad (10.16)$$

The statistical uncertainty on the measured number of events is obtained by assuming Possoinian statistics, $\sigma_{\mathrm{data}} = \sqrt{N_{\mathrm{data}}}$.

For the case of a measurement of normalised jet cross sections a similar procedure is used, with the normalised acceptance A_{norm} taking the more complicated form

$$A_{\mathrm{norm}} = \frac{N_{\mathrm{rec}}^{\mathrm{jet}} \big/ N_{\mathrm{rec}}^{\mathrm{NC}}}{N_{\mathrm{gen}}^{\mathrm{jet}} \big/ N_{\mathrm{gen}}^{\mathrm{NC}}} \,. \qquad (10.17)$$

This expression can be expanded in terms of uncorrelated variables similarly to equation (10.14). The error on A_{norm} is calculated similar to equation (10.15), with the assumption that the NC DIS and jet samples are statistically independent.

10.5.2 Normalisation Uncertainties

Normalisation uncertainties are independent of the measured observables. These uncertainties cancel for normalised cross section measurements.

Uncertainty on the Luminosity Determination

The run selection and corresponding integrated luminosity of the data set used is described in section 9.1. The uncertainty assigned due to the luminosity measurement is 2.5%.

Trigger Efficiency

The trigger efficiency is a combination of the efficiency of the LAr trigger element, the veto condition and the event timing requirement as described in section 9.2. The combined uncertainty is 1.2%.

Primary Vertex and Electron Track Finding Efficiency

The vertex finding efficiency together with the efficiency of the requirement of a track pointing to the electron cluster in the LAr calorimeter is studied in section 9.4. The efficiency is 95% with an uncertainty of 1%.

10.5.3 Model Dependence and Background Subtraction

The use of the bin-by-bin correction method and the statistical background subtraction introduce an uncertainty which depends on the measured observable. Whereas the background subtraction introduces only a small uncertainty due to the small amount of *ep* background, the model uncertainty is one of the dominant sources of uncertainties of this measurement. For normalised jet cross sections partial cancellations reduce the model dependence for some observables.

Model Dependence

The experimental correction factors c_{exp} are obtained as the arithmetic mean values of the correction factors $c^{\text{R}}_{\text{exp}}$ and $c^{\text{D}}_{\text{exp}}$, which are obtained from Rapgap and Djangoh, respectively. The values of $c^{\text{R}}_{\text{exp}}$ and $c^{\text{D}}_{\text{exp}}$ can differ if the underlying distributions are different. Especially at the boundaries of the defined phase space these differences can become significant due to migrations in and out of the jet phase space. The uncertainty assigned due to the experimental correction is chosen to be the difference between the mean value c_{exp} and the single model estimates which can also be written as

$$\Delta_{\text{model}} = \frac{1}{2}\left(c^{\text{R}}_{\text{exp}} - c^{\text{D}}_{\text{exp}}\right) . \tag{10.18}$$

Using this definition, in some measurement bins the model uncertainty may become very small due to artificial cancellations, like a crossing of the underlying distributions. In order to eliminate unphysically small values of Δ_{model}, a single exponential smoothing algorithm is applied [Bro63]. Let x_i denote the model dependence in bin i and s_i is a smoothed value of it. Then the values s_i are calculated as

$$s_i = \alpha x_i + (1 - \alpha) s_{i-1} \,. \tag{10.19}$$

The starting value s_0 is taken to be the total model uncertainty for a given observable, averaged over all bins. The parameter α is the smoothing factor with $0 < \alpha < 1$. Large values of α give larger weight to individual changes in the series of x_i, whereas in the limiting case of $\alpha = 1$ the smoothed series is just the input series. Small values of α have a large smoothing effect and result in a relatively flat model uncertainty. The value chosen for α is 0.5 with the additional constraint that the smoothed model uncertainty is not allowed to become smaller than 85% of its original value. This procedure leaves the model uncertainty, averaged over all bins, unchanged within 0.5% for all observables.
The effect of the smoothing algorithm is shown in figure 10.8 for the inclusive jet measurement. The dashed lines correspond to the original model uncertainty, obtained with equation (10.18). The green area is the result of the exponential smoothing and corresponds to the assigned model uncertainty. Before the smoothing, the model uncertainty is unphysically small in the second P_{T} bin. The smoothing cures these fluctuation by preserving the overall structures of the model uncertainty. Similar results are obtained for the other observables.

Background Subtraction

The *ep* background is estimated with dedicated MC simulations, as described in section 9.6.1. The relative background is estimated to be 1.2%, 1.1%, 0.4% and 0.4% in the NC DIS, inclusive jet, dijet and trijet samples, respectively. Only a small fraction of the background is due to photoproduction, which has the largest normalisation uncertainty. Therefore, a conservative estimate for the uncertainty due to the background subtraction is obtained by varying the total background contribution by ±30%. The resulting overall uncertainty is 0.4% for the NC DIS, 0.3% for the inclusive jet and 0.1% for the dijet and trijet measurements.

10.5.4 Measurement Uncertainties

Uncertainties due to the reconstruction of electron and hadronic final state variables are estimated by varying the corresponding reconstructed quantities in simulated events. The effect of this variation on the measured cross section is assigned as experimental uncertainty. In figure 10.9 the ratios $\sigma_{\mathrm{jet}}/\sigma_{\mathrm{jet}}^{(+\delta)}$ and $\sigma_{\mathrm{jet}}/\sigma_{\mathrm{jet}}^{(-\delta)}$ are shown for the inclusive jet measurement as function of Q^2 and P_{T}. The up and down variations of the corresponding reconstructed quantity are denoted by $+\delta$ and $-\delta$. It can be observed that the effect of the variations on the cross section is to a good approximation symmetric. The dominant systematic uncertainty comes from the hadronic energy scale.

Electron Energy

The electron energy is measured with a precision of 0.5% in the central region of the detector ($z_{\mathrm{impact}} \leq 100\,\mathrm{cm}$) and 1% in the forward region ($z_{\mathrm{impact}} > 100\,\mathrm{cm}$), as

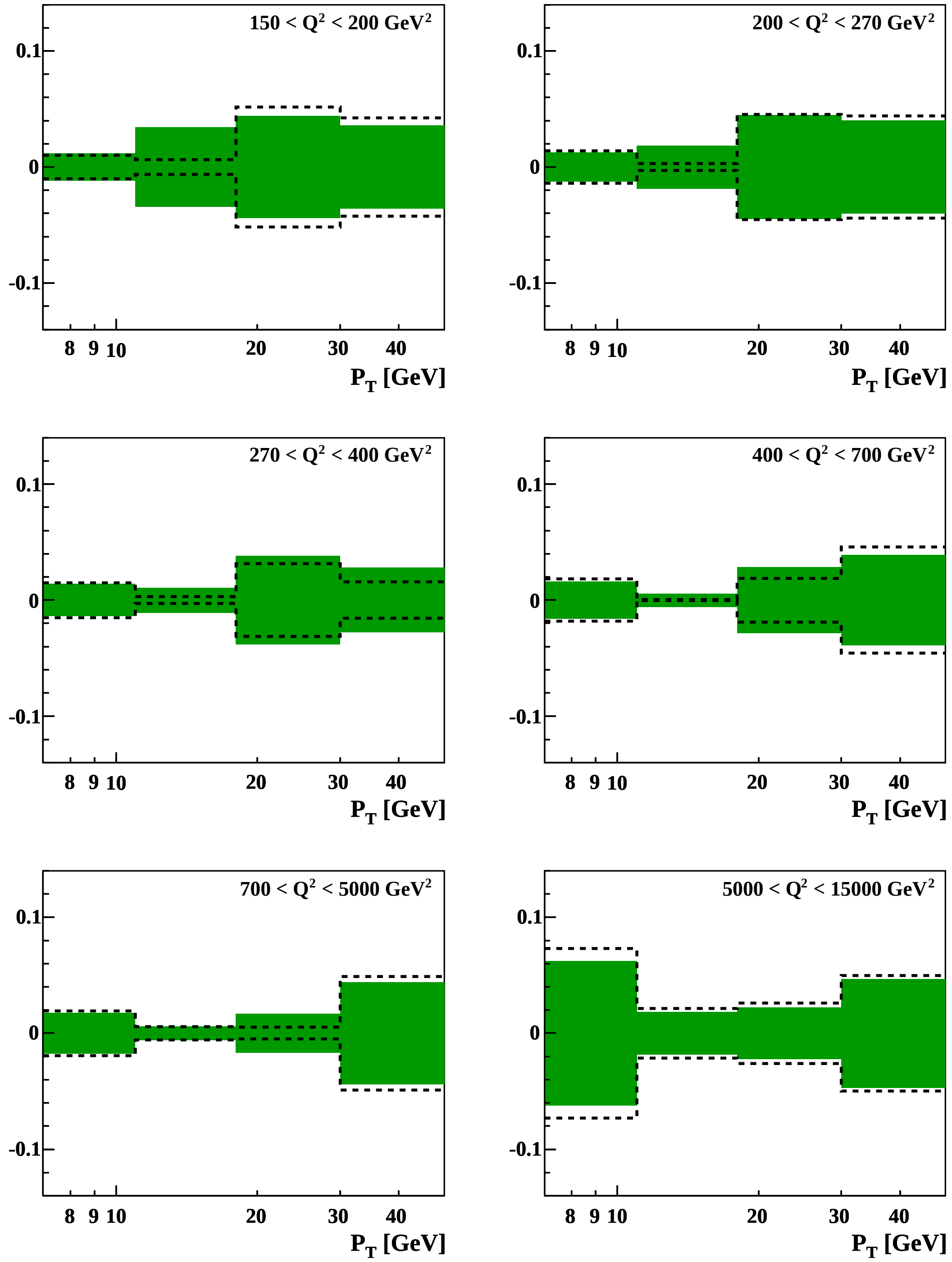


Figure 10.8: Uncertainty due to the model dependence for the inclusive jet measurement as function of P_{T} in different Q^2 bins. The dashed line corresponds to the model uncertainty calculated with equation (10.18) and the green area is the model uncertainty after the application of the single exponential smoothing algorithm. The statistical errors due to limited MC statistics are neglible and not shown.

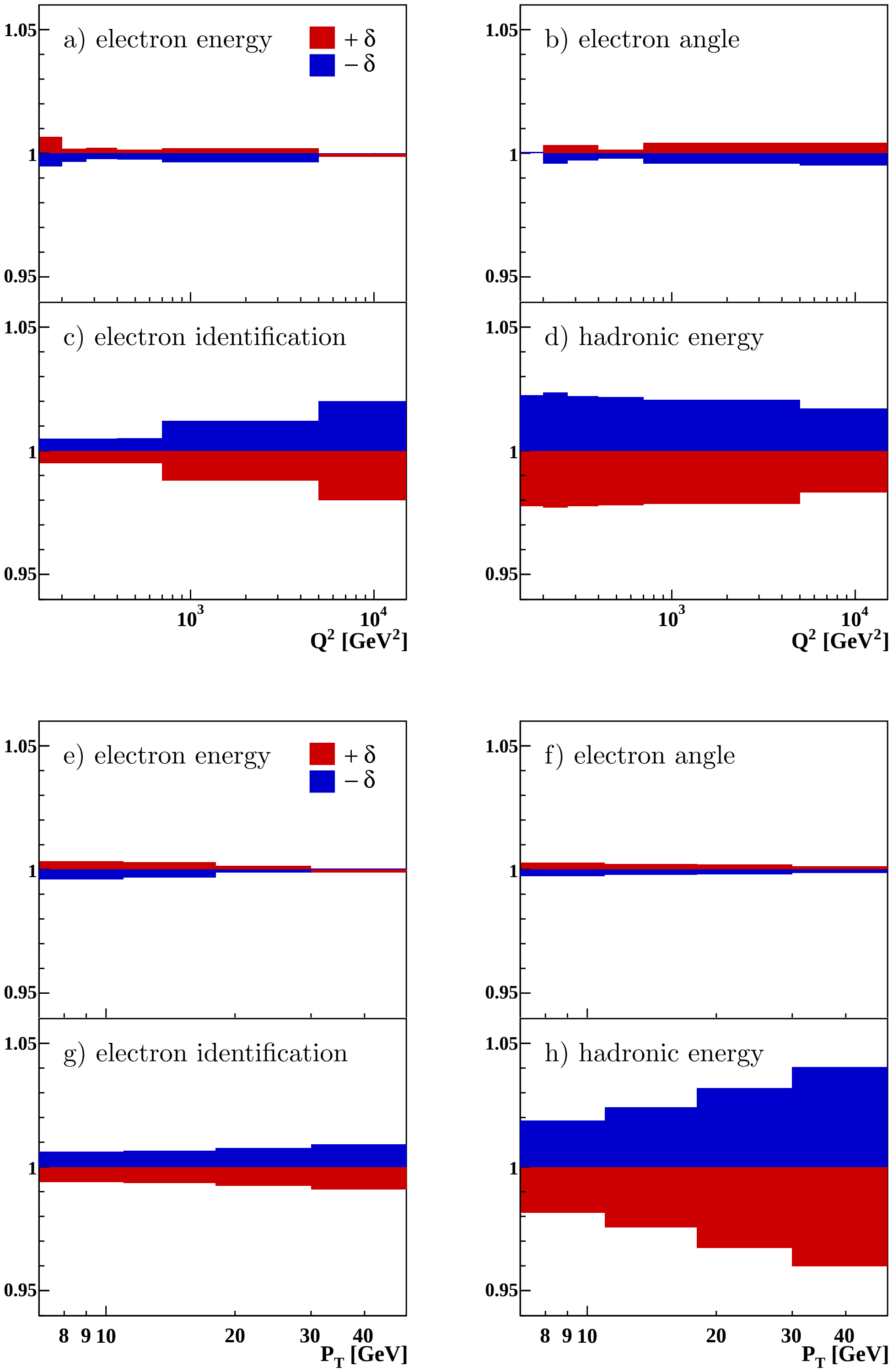


Figure 10.9: Systematic uncertainties obtained by variations of the corresponding reconstructed quantity for the inclusive jet measurement as function of Q^2 (top) and P_{T} (bottom). Shown are the ratios $\sigma_{\text{jet}}/\sigma_{\text{jet}}^{(\pm\delta)}$, where the plus (minus) sign corresponds to an upwards (downwards) shift of the reconstructed quantity.

shown in section 8.1. This results in an uncertainty of approximately 0.3% for jet cross sections and 0.5% for normalised jet cross sections. This increase of the latter is due to an opposite effect that the variation of E_{e} has on the NC DIS cross section compared to the jet cross sections.

Electron Polar Angle

The polar angle of the scattered electron is measured with the LAr calorimeter, which has been aligned with the CTD to a precision of 1 mrad as described in chapter 7. The variation of θ_{e} by ± 1 mrad leads to an uncertainty of 0.2–0.3%.

Electon Identification

The uncertainty on the electron finding efficiency is 0.5% for $z_{\mathrm{impact}} \leq 0$ and 2% for $z_{\mathrm{impact}} > 0$ (see section 7.2). This leads to an uncertainty which is Q^2 dependent, as shown in figure 10.9c. Since the uncertainty is relatively flat as function of P_{T} (figure 10.9g), it largely cancels for normalised jet cross sections.

Hadronic Energy Scale

The hadronic energy scale uncertainty is one of the dominating uncertainties in every jet measurement. A large part of this work has been dedicated to study the detector response to jets and to improve the uncertainty on the energy measurement of the hadronic final state. The separation of electromagnetic showers in the LAr calorimeter based on statistical methods, and the subsequent reconstruction and calibration of the hadronic final state are described in chapters 6 through 8. One of the major achievements of this work is the reduction of the uncertainty on the hadronic energy scale to 1%.

The effect of this uncertainty on the measured jet cross sections is studied by varying the energy of all calibrated HFS objects by ±1%. The shifted objects are then boosted to the Breit frame where the jet finding is performed. No distinction between tracks and calorimeter clusters is made for the variations, since the employed energy flow algorithm results in combined objects. The uncertainty of 1% thus applies to the fullly reconstructed and calibrated hadronic final state and takes the matching of energy from tracks and clusters into account.

The effect of the hadronic energy scale uncertainty on the inclusive jet measurement is shown in figures 10.9d and 10.9h as function of Q^2 and P_{T}. The uncertainty caused on the cross section is flat in Q^2 at values of approximately 2%. As function of P_{T}, the uncertainty increases with increasing P_{T}, taking values of about 2% in the lowest P_{T} bin and 4% in the highest one.

10.5.5 Summary of Uncertainties

A summary of the contributions from the individual sources of uncertainty is given in table 10.2. The normalisation uncertainty Δ_{norm} is a combination of the uncertainties due to the luminosity measurement and the trigger and vertex-track-link efficiencies, added in quadrature. Δ_{model} and Δ_{bkgd} denote the uncertainties due to the model dependence and background subtraction. The uncertainty due to the

Measurement	Δ_{norm}	Δ_{model}	Δ_{bkgd}	Δ_{elec}	Δ_{HFS}	Δ_{sys}	Δ_{stat}	Δ_{tot}
σ_{NC}	2.9%	1.9%	0.4%	0.7%	0.6%	3.7%	0.2%	**3.7%**
σ_{jet}	2.9%	1.3%	0.3%	0.7%	2.2%	4.0%	0.3%	**4.0%**
$\sigma_{\text{jet}}/\sigma_{\text{NC}}$	-	0.4%	0.2%	0.6%	1.4%	1.6%	0.9%	**1.9%**
$\sigma_{\text{2-jet}}$	2.9%	3.7%	0.1%	0.7%	2.0%	5.2%	0.5%	**5.2%**
$\sigma_{\text{2-jet}}/\sigma_{\text{NC}}$	-	1.6%	0.1%	0.2%	1.2%	2.0%	1.3%	**2.4%**
$\sigma_{\text{3-jet}}$	2.9%	5.6%	0.1%	0.8%	3.0%	7.1%	1.0%	**7.2%**
$\sigma_{\text{3-jet}}/\sigma_{\text{NC}}$	-	5.5%	0.1%	0.3%	2.2%	6.1%	2.1%	**6.5%**

Table 10.2: Summary of the experimental uncertainties for the NC DIS and jet cross section measurements. A detailed explanation is given in the text.

electron measurement, Δ_{elec}, combines the uncertainties due to the electron energy and polar angle measurements with the uncertainty of the electron finding efficiency. Δ_{HFS} is the uncertainty due to the hadronic energy scale. Δ_{sys} denotes the total systematic uncertainty, with the individual sources added in quadrature. The total uncertainty Δ_{tot} is obtained by adding the systematic uncertainty with the statistical uncertainty Δ_{stat} in quadrature.

For the NC DIS measurement σ_{NC} the normalisation and model uncertainties give the major contributions to the experimental uncertainty of 3.7%.

The uncertainty of the inclusive jet measurement σ_{jet} is dominated by the normalisation and the hadronic energy scale uncertainties. The model uncertainty is smaller than in the NC DIS case. The normalised inclusive jet measurement $\sigma_{\text{jet}}/\sigma_{\text{NC}}$ is independent of the overall normalisation. In addition it benefits from cancellations which reduce the model and hadronic energy scale uncertainties. This reduces the uncertainty from 4.0% for the σ_{jet} measurement to 1.9% for the $\sigma_{\text{jet}}/\sigma_{\text{NC}}$ measurement.

In the case of the dijet measurement $\sigma_{\text{2-jet}}$, the normalisation, model and hadronic energy scale uncertainties are of comparable size and lead to a total uncertainty is 5.2%. Again, the normalised measurement $\sigma_{\text{2-jet}}/\sigma_{\text{NC}}$ benefits from cancellations resulting in an overall uncertainty of 2.4% only.

The normalised and non-normalised trijet measurements are dominated by the model uncertainty, which becomes nearly twice as large as the hadronic energy scale uncertainty. Both MC event generators used to obtain the experimental correction factors do not include the LO diagrams for trijet production, which results in a poor description of the trijet data sample. The MC weighting described in section 9.7.4 improves the description of the trijet data significantly. This reduces the model dependence already by a factor of two and results in an uncertainty of 5.5%. A further reduction of the model dependence for the trijet measurement proved difficult.

For the normalised measurements the statistical uncertainty becomes more important than for the corresponding cross section measurements. This is mostly due to the more complicated form of the acceptance in the case of normalised jet cross

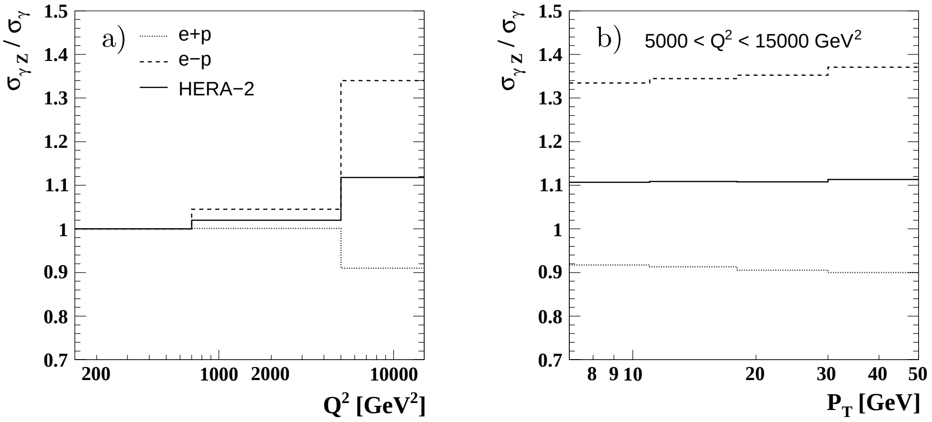


Figure 10.10: EW correction factors for the inclusive jet cross section as function of Q^2 (a) and P_{T} in the highest Q^2 bin (b). The values of c_{EW} are shown for e^+p and e^-p scattering as well as for a luminosity weighted mixture, according to the HERA-2 dataset.

sections. The uncertainty on the acceptance is related to the uncertainty due to migrations between bins and is propagated to the statistical uncertainty, as shown in section 10.5.1.

10.6 Corrections to NLO Calculations

The NLOJet++ program, employed in this work for the calculation of jet cross sections in NLO in α_s, does not include effects from Z^0 exchange. Also, the jet cross sections calculated with NLOJet++ are defined on the parton level. In order to compare the measured cross sections with the predictions, the NLO calculations are corrected for the effects of Z^0 exchange and hadronisation.

10.6.1 Electroweak Correction

In the fixed-order pQCD calculation for multi-jet production employed in this work Z^0 exchange and γ/Z^0 interference effects are not included. These effects are negligible as long as $Q^2 \ll M_Z^2$, but in the two highest Q^2 bins these effects cannot be neglected anymore. Therefore a correction, c_{EW}, to simulate electroweak effects, is applied to the NLO calculations. The correction c_{EW} is obtained with the LEPTO MC event generator and is defined by $c_{\mathrm{EW}} = \sigma_{\gamma Z}/\sigma_{\gamma}$ [H107]. Here $\sigma_{\gamma Z}$ denotes a cross section obtained with all EW effects included and σ_{γ} is the same cross section calculated with pure γ exchange only. Since the cross section for Z^0 exchange depends on the charge of the incoming lepton, c_{EW} is different for e^+p and e^-p scattering.
The correction c_{EW} is shown in figure 10.10 for e^+p and e^-p scattering and for a mixture weighted according to the HERA-2 luminosities. Whereas the EW corrections are sizeable in Q^2 bins 5 and 6 for e^+p and e^-p scattering, the corrections for the HERA-2 average are only 2% and 12% in bins 5 and 6, respectively. In the highest Q^2 bin c_{EW} depends on the jet transverse momentum for e^+p and e^-p scattering, because the phase space in Q^2 depends slightly on P_{T} (figure 10.10b).

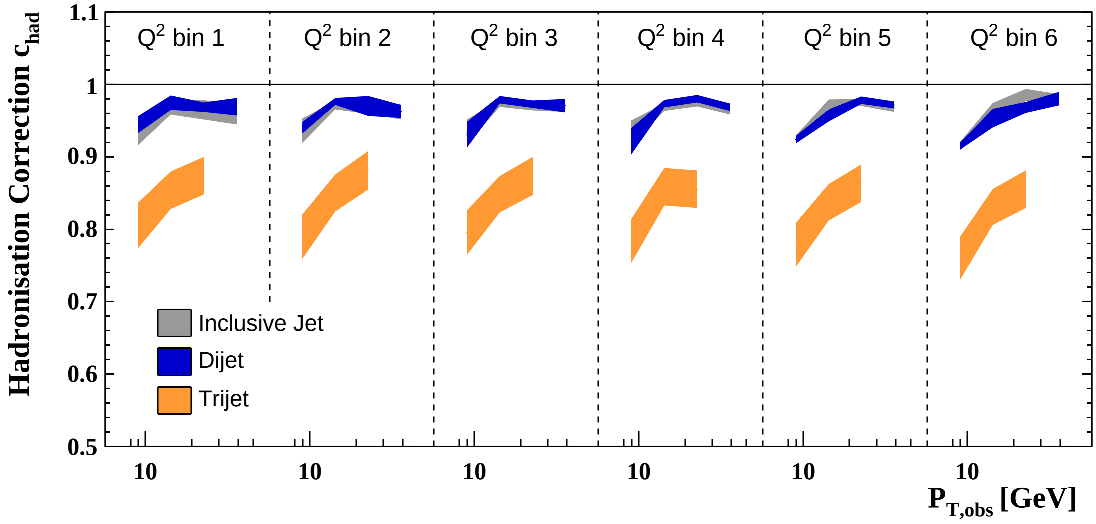


Figure 10.11: Hadronisation correction factors for the inclusive jet, dijet and trijet measurements as function of Q^2 and $P_{\mathrm{T,obs}}$.

For the HERA-2 mixture this dependence largely cancels and c_{EW} is flat within 1%. The EW corrections for dijet and trijet measurements show a similar behaviour as for the inclusive jet case and are not shown here.

10.6.2 Hadronisation Correction

At present NLO predictions for jet cross sections do not include non-perturbative contributions to account for the fragmentation of partons into stable hadrons. Before measured cross sections on hadron level can be compared to NLO predictions on parton level, the size of these hadronisation effects has to be estimated. One possibility is to use phenomenological fragmentation models like the Lund string model as implemented in JETSET and used in RAPGAP and DJANGOH. The hadronisation correction c_{had} can then be estimated by comparing cross sections on hadron level σ_{had} with parton level cross sections σ_{part},

$$c_{\mathrm{had}} = \frac{\sigma_{\mathrm{had}}}{\sigma_{\mathrm{part}}} \,. \tag{10.20}$$

The correction c_{had} is applied as multiplicative factor to NLO cross sections. The values of c_{had} are determined as the average of values obtained from RAPGAP and DJANGOH.

The hadronisation corrections for the inclusive, dijet and trijet measurements are shown in figure 10.11. The values of c_{had} are observed to be smaller than unity over the whole phase space considered. This is generally expected, since hadronisation effects alter the final state used for the jet finding and cause particles to migrate out of the phase space defined for a given jet. This leads to smaller transverse momenta for hadron jets than for parton jets and consequently $c_{\mathrm{had}} < 1$. The values of c_{had} are very similar for the inclusive jet and dijet measurements, where they differ from unity only by 4% on average. In the lowest $P_{\mathrm{T,obs}}$ bin the hadronisation correction is largest with values between 0.92–0.94. For the trijet measurement the hadronisation corrections are larger than for the inclusive and dijet case, with average values of c_{had} of approximately 0.83.

The uncertainty on c_{had} is estimated by taking the difference between the correction factors obtained using DJANGOH and RAPGAP. This is a conservative estimate compared to some previous jet cross section measurements [H107, H110b, ZEUS10a], which is motivated by a comparison to hadronisation corrections obtained with an analytic approach [DM+08, H110c] and a variation of model parameters in MC event generators which influence the hadronisation [WW99]. These studies showed differences in the obtained hadronisation corrections of a few percent, which is still smaller than the uncertainties of NLO calculations due to missing higher orders, but show our present inability to model non-perturbative effects very accurately.

For the inclusive jet and dijet measurements the uncertainty on c_{had} lies typically between 1 and 1.5%. For trijet observables c_{had} is on average 0.83 with uncertainties between 2–3%.

Chapter 11

Results

In this chapter the measured cross sections for inclusive and multi-jet production in NC DIS at high Q^2 are presented and compared to NLO calculations. The NC DIS phase space is defined through the virtuality of the exchanged boson, $150 < Q^2 < 15000\,\mathrm{GeV}^2$ and the inelasticity $0.2 < y < 0.7$. The measured cross sections are based on the full HERA-2 data set with an integrated luminosity of $351.6\,\mathrm{pb}^{-1}$. The data are corrected for detector effects and QED corrections at the leptonic vertex. The cross sections are corrected to the hadron level and not to the parton level, meaning that they are not corrected for non-perturbative hadronisation effects. No corrections for the running of the electromagnetic coupling or effects from Z^0 exchange are applied to the data. For the illustration of single and double differential jet cross sections, the corrected cross sections are divided by the respective bin width, which is not done for normalised cross sections. The statistical uncertainty of the measurement is always illustrated by the inner error bars, while the outer error bars represent the quadratic sum of the statistical and the systematic uncertainties. The numerical values of the cross sections and normalised cross sections presented in this chapter can be found in appendix A.

11.1 The Neutral Current Cross Section

Since the jet samples are based on the NC DIS sample, the extraction of the inclusive NC cross section provides an important consistency check of the analysis. The inclusive NC cross sections σ_{NC} for the unpolarised neutral current reactions $e^{\pm}p \to e^{\pm}X$ are shown in figure 11.1 as function of Q^2. The cross sections are measured in the inelasticity range $0.2 < y < 0.7$. They are obtained from the full e^-p and e^+p HERA-2 datasets with integrated luminosities of $159.6\,\mathrm{pb}^{-1}$ and $192.0\,\mathrm{pb}^{-1}$, respectively. The measurements are compared to NLO calculations using QCDNUM [Bot11] with the CTEQ6.6 PDF [CTEQ08] and $\alpha_s(M_Z) = 0.118$. The factorisation and renormalisation scales are chosen to be $\mu_{\mathrm{f}} = \mu_{\mathrm{r}} = Q$. No corrections are applied to the predictions from QCDNUM, since EW effects from Z^0 exchange are included, and because the inclusive NC cross section is insensitive to hadronisation corrections. The NC cross sections show the expected $1/Q^4$ behaviour (2.32) and are well described by the theoretical predictions. A charge dependent asymmetry due to effects from Z^0 exchange can be observed in the highest Q^2 bin, where the e^-p NC cross section is larger than the e^+p NC cross section. At low values of Q^2 the NLO

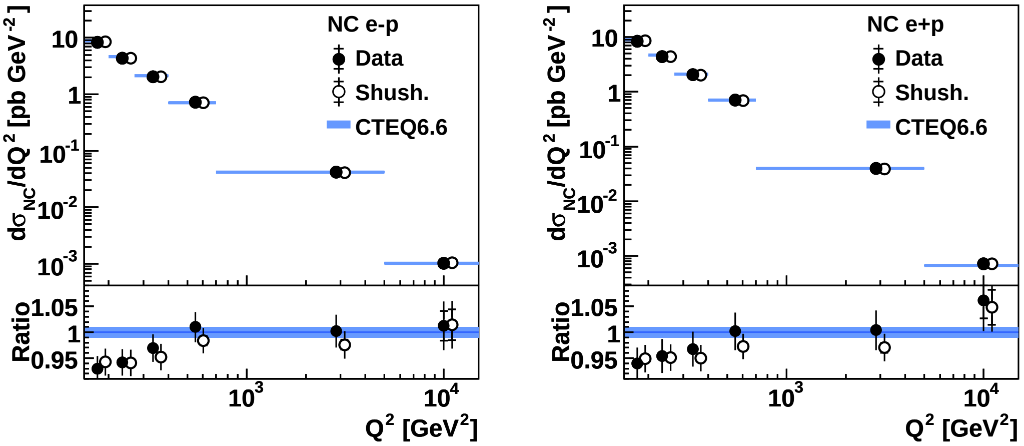


Figure 11.1: The inclusive NC $e^-p \to e^-X$ (left) and $e^+p \to e^+X$ (right) cross sections in the inelasticity range $0.2 < y < 0.7$. The data are compared to NLO calculations from QCDNUM using the CTEQ6.6 PDF and $\alpha_s = 0.118$. Also shown are cross sections obtained by an independent analysis on HERA-2 data [H109a], where the phase space and binning have been adjusted to this analysis [Shu10], illustrated by open circles and shifted to the right for better visibility. The uncertainty due to the luminosity measurement of 2.5%, which is common to both analyses, is not included in the experimental uncertainties. The ratio of data with respect to the NLO calculation is shown at the bottom.

calculation predicts approximately 4–5% larger cross sections than are measured. A different choice of PDF, as for example the use of the NNPDF set [BD$^+$10], leads to an improved description of the data.

Also shown in figure 11.1 are inclusive NC cross sections obtained from an independent analysis of the HERA-2 data [H109a], which has been adjusted to the phase space and binning used in this work [Shu10]. Good agreement between the two measurements is observed, with the central values agreeing within 1.5% in most bins. Only in two Q^2 bins larger deviations are observed, with values of about 2–2.5% in Q^2 bins 4 and 5, respectively. These deviations are well within the systematic uncertainties of the two measurements. Excluding the statistical uncertainty and the uncertainty due to the luminosity measurement, the inclusive NC cross sections obtained in this work have an uncertainty of 3.1% in these bins. This is larger than the uncertainty of the independent, inclusive NC analysis, which is 2.5%. These differences can be attributed to the different goals of these analyses. The selection criteria in this work are optimised for a precise jet measurement. In terms of the scattered electron these selection criteria ensure an accurate reconstruction of the boost to the Breit frame. The inclusive NC analysis aims for a precise measurement of the inclusive NC cross sections, without additional requirements. The selection criteria employed in this work render the acceptance for the NC measurement smaller compared to the acceptance of the NC analysis. This results in larger experimental correction factors with larger uncertainties, which is reflected in an uncertainty due to the model dependence of nearly 2%. Furthermore, the different selection criteria make the NC cross section measurement more susceptible to experimental uncertainties. For example, the cut on $E - p_z$ results in an uncertainty of 0.6% due the hadronic energy scale uncertainty. In the case of the NC analysis a much weaker requirement is imposed on $E - p_z$, which renders the cross section nearly independent of the hadronic energy scale in the phase space considered.

11.2 Jet Cross Sections

One of the central physics results of this analysis are the measurements of the inclusive jet, dijet and trijet cross sections in NC DIS. The DIS phase space is given by

$$150 < Q^2 < 15000\,\mathrm{GeV}^2, \qquad 0.2 < y < 0.7. \tag{11.1}$$

Jets are defined in the Breit frame of reference with the longitudinally invariant k_T algorithm. The phase space for the inclusive jet measurement is

$$-1.0 < \eta_\mathrm{lab} < 2.5, \qquad 7 < P_\mathrm{T} < 50\,\mathrm{GeV}\,. \tag{11.2}$$

The dijet and trijet cross sections are defined through

$$-1.0 < \eta_\mathrm{lab} < 2.5, \qquad 5 < P_\mathrm{T} < 50\,\mathrm{GeV}, \qquad M_{12} > 16\,\mathrm{GeV} \tag{11.3}$$

as discussed in chapter 9. The jet cross section measurement is based on the full HERA-2 data set recorded by the H1 collaboration in the years 2003–2007. The total integrated luminosity of the data used is $351.6\,\mathrm{pb}^{-1}$.
The jet cross sections are compared to pQCD calculations performed in NLO in the strong coupling. The calculations are carried out in the $\overline{\mathrm{MS}}$ renormalisation and factorisation schemes as implemented in the NLOJet++ program. In these calculations massless quarks are assumed and the number of active flavours is set to five. CTEQ6.6 is chosen as the parametrisation of the proton PDF. The value of the strong coupling at the mass of the Z^0 boson M_Z, is taken to be $\alpha_s(M_Z) = 0.118$, which is in accordance with the value of α_s used for the determination of the PDF. The strong coupling is evolved as function of the renormalisation scale in the two-loop approximation. The factorisation scale μ_f is chosen to be Q, and the renormalisation scale μ_r is taken to be $\sqrt{(Q^2 + P^2_\mathrm{T,obs})/2}$. The variable $P_\mathrm{T,obs}$ denotes the transverse momentum P_T of the jet in case of inclusive jet cross sections and the average transverse momentum $\langle P_\mathrm{T}\rangle$ of the two (three) leading jets in the dijet (trijet) case. This choice of the renormalisation scale is motivated by the presence of the two hard scales Q and $P_\mathrm{T,obs}$ in jet production in DIS. The theoretical predictions are corrected for non-perturbative hadronisation effects, as described in section 10.6.2. Corrections due to Z^0 exchange are taken into account by the EW correction, which differs from unity only in the two highest Q^2 bins (section 10.6.1).
The uncertainty of the NLO predictions due to missing higher orders is conventially estimated by varying the chosen renormalisation and factorisation scales in the range between 0.5 and 2. The uncertainty due to a scale variation f is then the change of the cross section with respect to the nominal scale, $\sigma_\mathrm{NLO}(f\mu)/\sigma_\mathrm{NLO}(\mu)$, where f is either 0.5 or 2. In regions where the cross section is not a monotonical function of μ_r, the theoretical uncertainty is estimated by taking the maximum and minimum values that the cross section takes in the interval $[0.5\mu_\mathrm{r}, 2\mu_\mathrm{r}]$ in order not to underestimate the scale dependence. The uncertainties obtained by a variation of the renormalisation scale are generally by a factor of two to three larger than the

uncertainties from variations of the factorisation scale. The total theoretical uncertainties are calculated by adding in quadrature the uncertainties estimating missing higher orders in the perturbative series and the uncertainties of the hadronisation corrections.

The single differential inclusive jet, dijet and trijet cross sections as function of Q^2 and $P_{\mathrm{T,obs}}$ are presented in figure 11.2. The cross sections are steeply falling as function of Q^2 and $P_{\mathrm{T,obs}}$. They are well described by the NLO calculations over more than four orders of magnitude. The uncertainties of the inclusive and dijet measurements are typically of 4% and 5%, respectively. These are in most bins by a factor of two smaller than the uncertainties of the theoretical predictions, which are dominated by the uncertainties due to missing higher orders. The measured trijet cross sections have uncertainties of 6% to 7% at low Q^2 and low $\langle P_{\mathrm{T}}\rangle$, which are also about a factor of two smaller than the theoretical uncertainties. At high Q^2 and high $\langle P_{\mathrm{T}}\rangle$ the experimental uncertainties increase because of a larger model dependence and the increasing uncertainty due to the hadronic energy scale, whereas the theoretical uncertainties decrease. In these regions the theoretical and experimental uncertainties of the trijet cross sections become comparable.

The double differential inclusive jet cross sections $\mathrm{d}^2\sigma_{\mathrm{jet}}/\mathrm{d}Q^2\mathrm{d}P_{\mathrm{T}}$ are shown in figure 11.3 together with the NLO predictions. The data are well described by the calculations over the full phase space of the measurement. Only in two bins, 4δ and 5δ (the nomenclature can be found in table 10.1), are the data below the uncertainty band of the NLO predictions. In both cases the deviation from the NLO prediction is less than two standard deviations of the experimental and theoretical uncertainties. The uncertainty of the measurement is dominated either by the normalisation uncertainty or the uncertainty due to the hadronic energy scale, depending on the region in phase space. Except for the bin 6δ, where the statistical uncertainty dominates, the experimental uncertainty is always smaller than the theoretical uncertainty. In regions with $P_{\mathrm{T}} < 30\,\mathrm{GeV}$ the experimental uncertainty is by a factor of two smaller than the NLO uncertainty. Only in regions of high jet transverse momenta, $P_{\mathrm{T}} > 30\,\mathrm{GeV}$, the uncertainties become comparable.

The double differential dijet cross sections measured as function of Q^2 and $\langle P_{\mathrm{T}}\rangle$ are presented in figure 11.4. The cross section falls steeply as function of $\langle P_{\mathrm{T}}\rangle$, with a smaller slope for high Q^2 values. The experimental precision is around 5% at low $\langle P_{\mathrm{T}}\rangle$ and increases to about 12% at the highest values of $\langle P_{\mathrm{T}}\rangle$. The uncertainties due to the model dependence, the hadronic energy measurement and the luminosity normalisation contribute in roughly equal parts to the experimental precision. Only at the highest values of $\langle P_{\mathrm{T}}\rangle$ and Q^2 are statistical uncertainties between 7–14% dominating the experimental uncertainty. The double differential dijet cross sections as function of Q^2 and ξ are presented in figure 11.5. Kinematic constraints due to the required jet transverse momenta, as well as the demand of a minimal invariant mass M_{12} of the two jets leads to a reduction of the cross section at small values of ξ. At large values of ξ the decreasing quark and gluon densities lead to a decrease of the cross section. The cross section has a maximum at small values of ξ at low Q^2, which shifts to higher values as Q^2 increases with a fixed range in inelasticity. The uncertainties on the theoretical predictions are approximately 10% at low Q^2 and become smaller with increasing Q^2, where they are about 6–7%. Similar to the inclusive jet measurement, the theoretical uncertainties are larger than

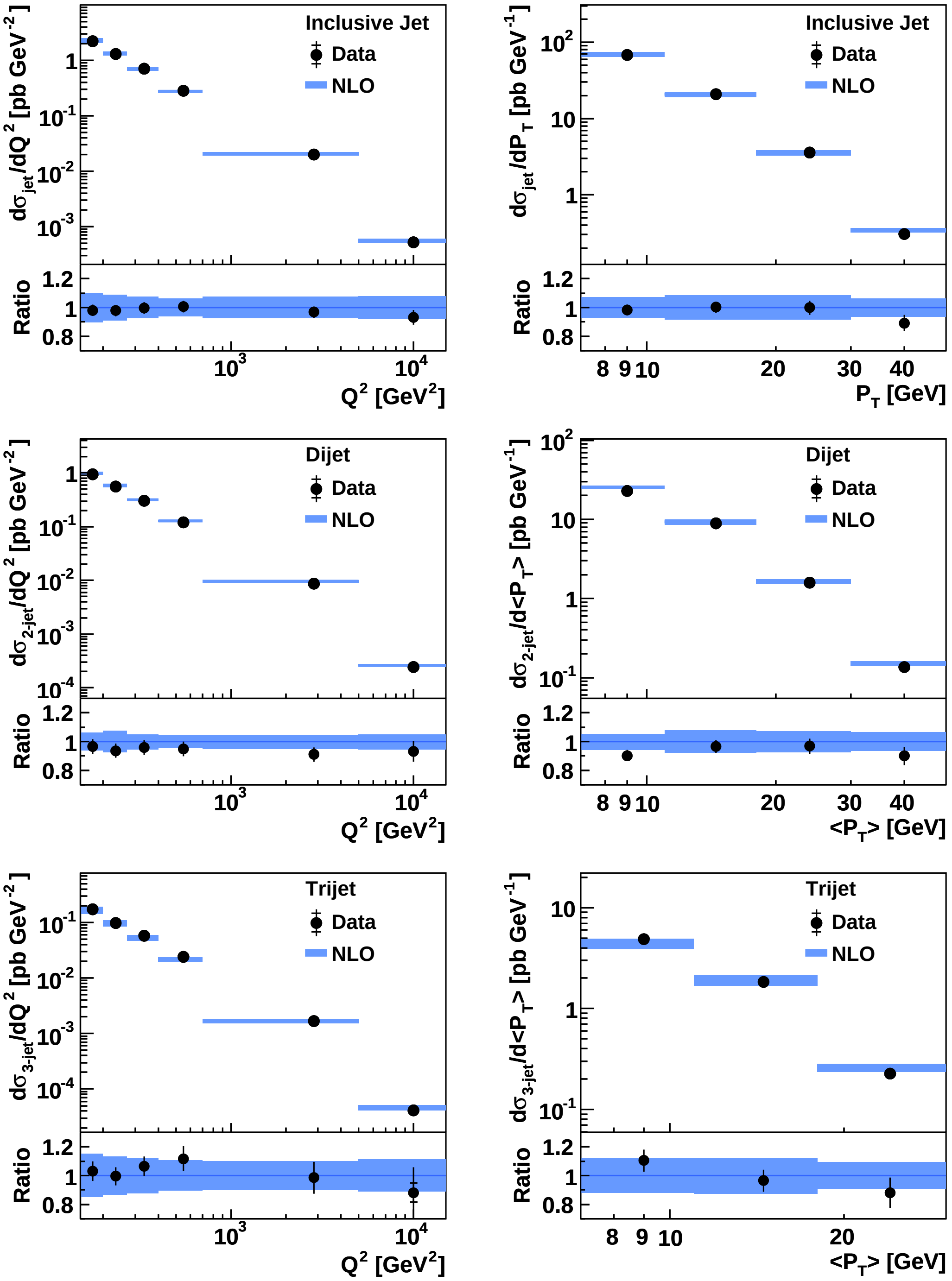


Figure 11.2: Single differential inclusive, dijet and trijet cross sections (from top to bottom) as function of Q^2 and $P_{\mathrm{T,obs}}$. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

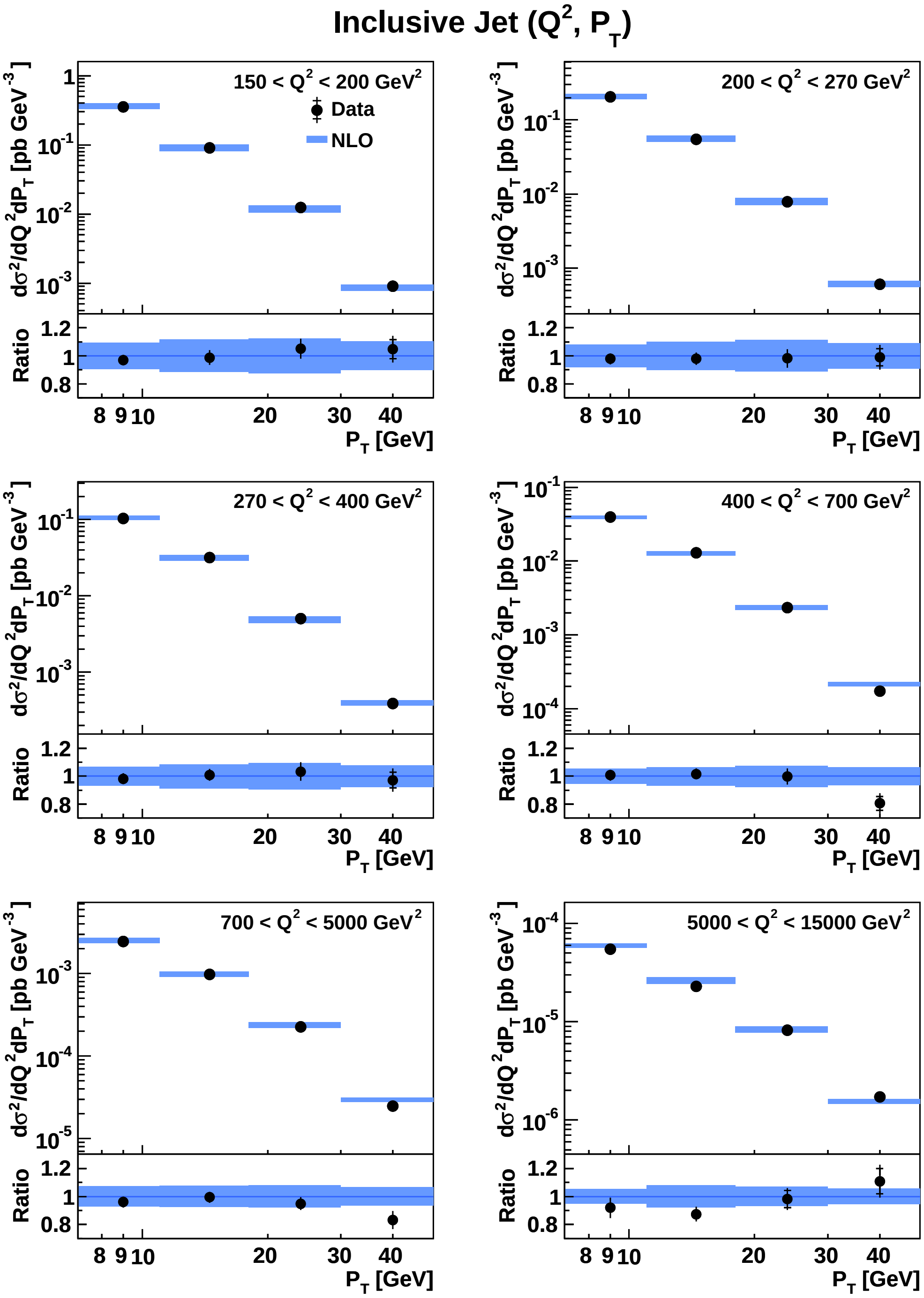


Figure 11.3: Double differential inclusive jet cross sections as function of Q^2 and P_{T}. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

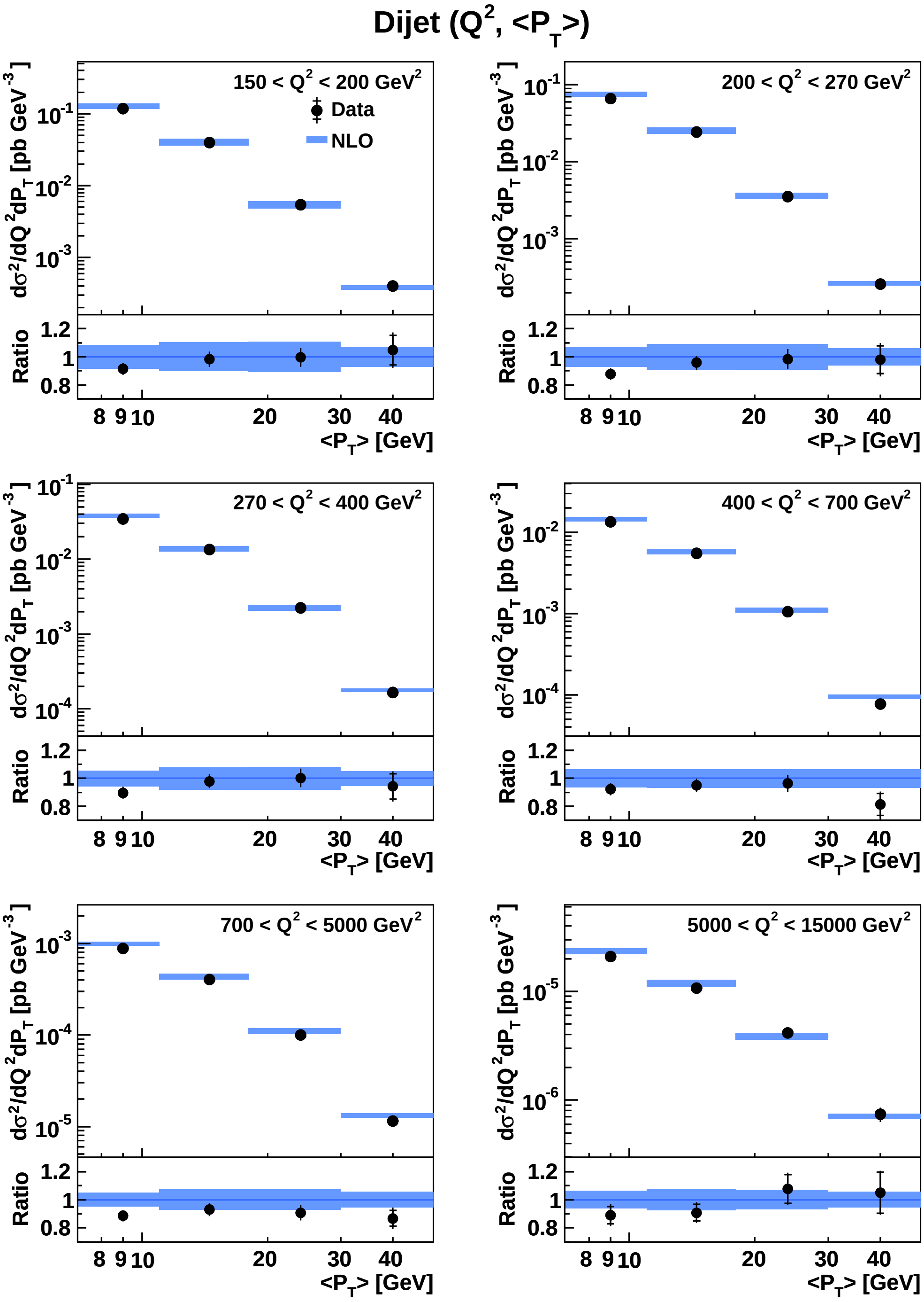


Figure 11.4: Double differential dijet cross sections as function of Q^2 and $\langle P_\mathrm{T} \rangle$. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

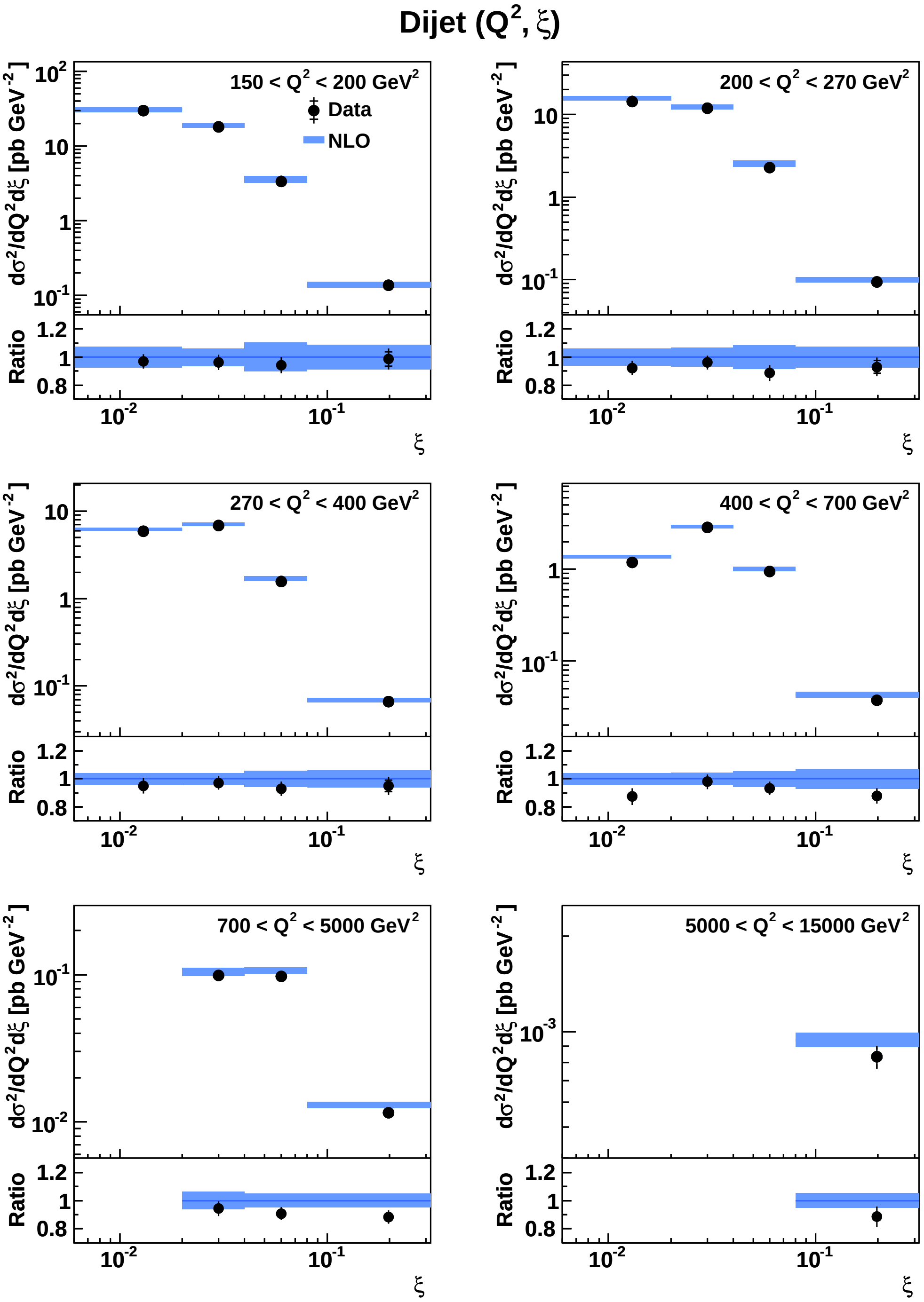


Figure 11.5: Double differential dijet cross sections as function of Q^2 and ξ. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

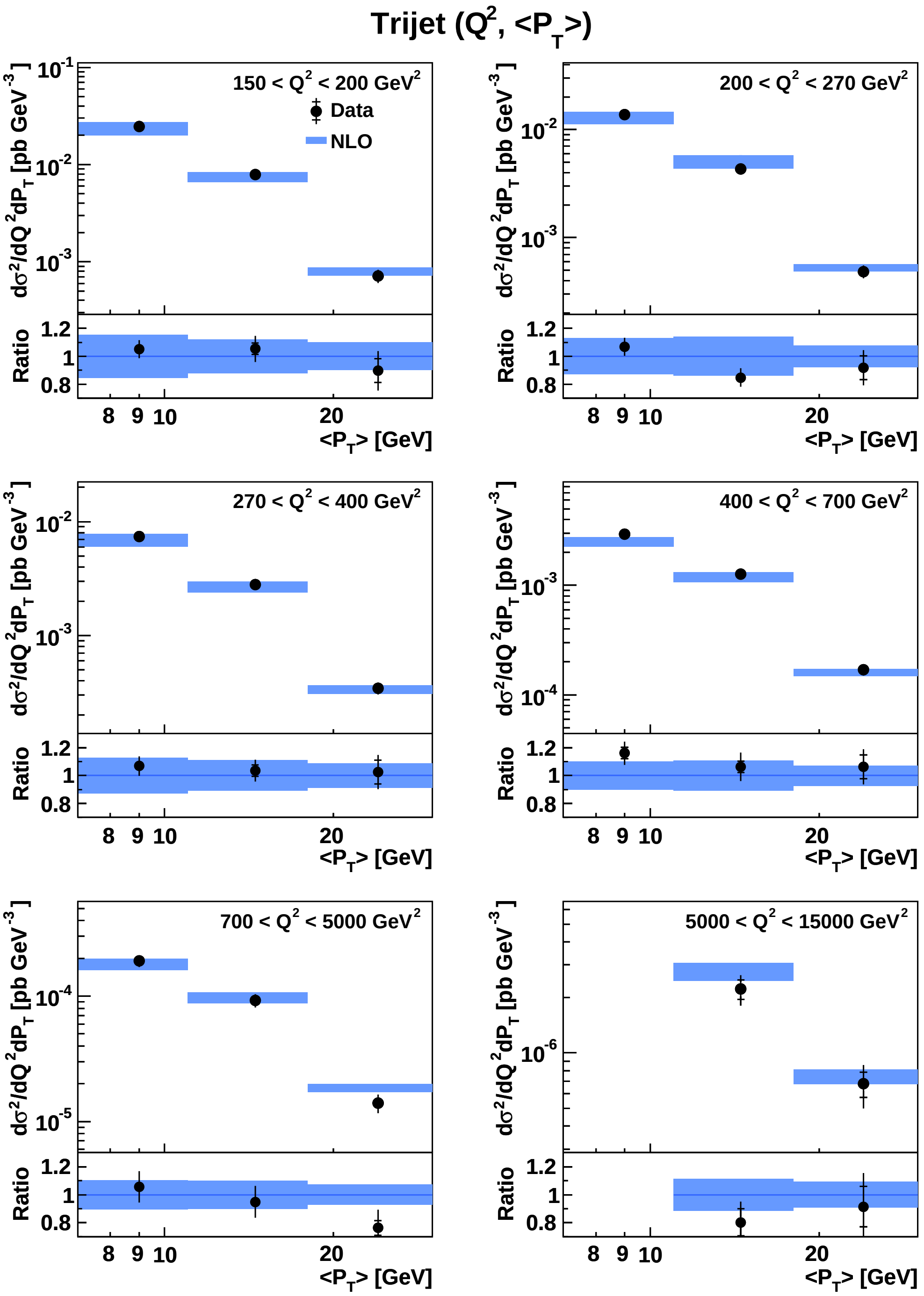


Figure 11.6: Double differential trijet cross sections as function of Q^2 and $\langle P_{\mathrm{T}} \rangle$. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

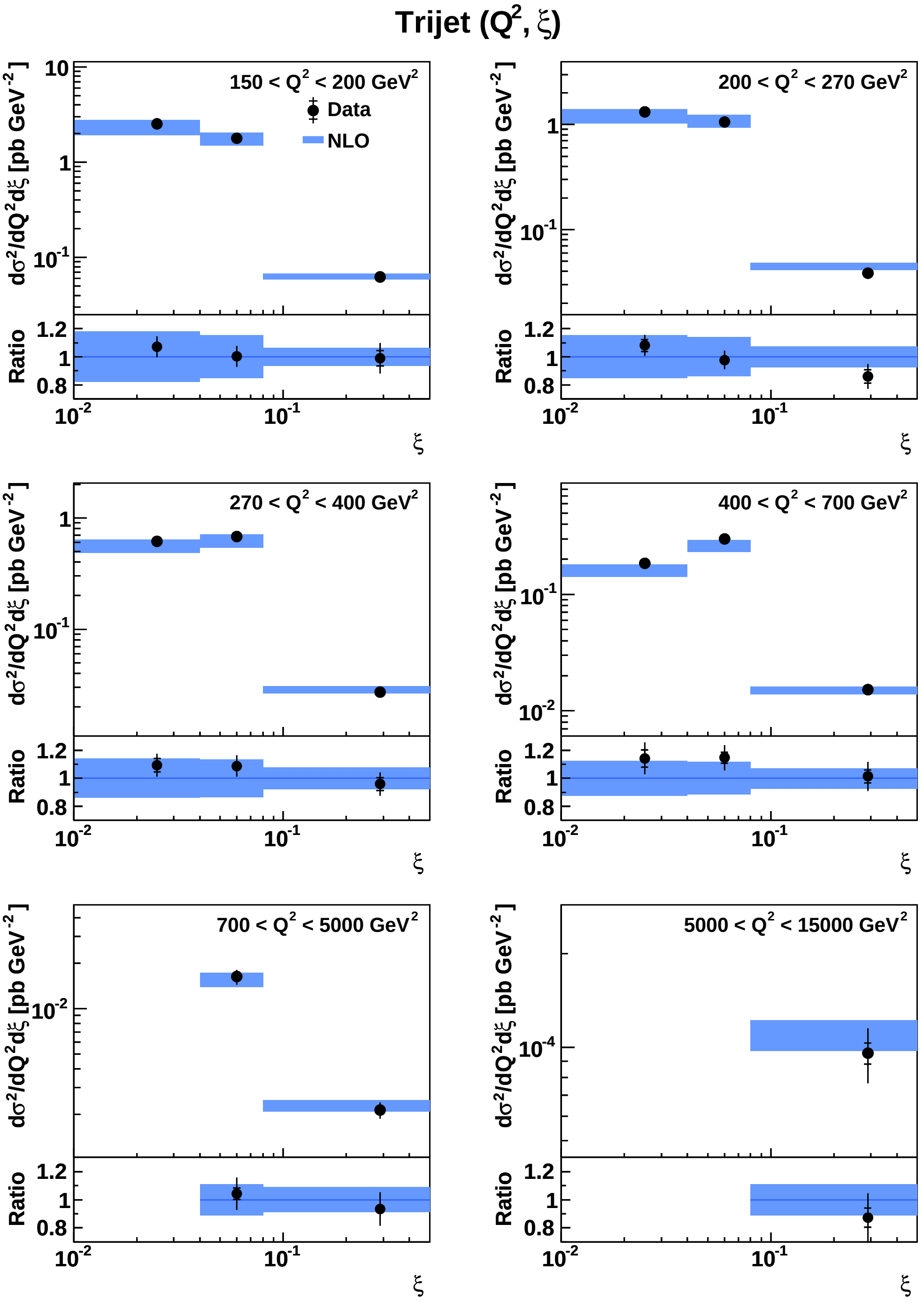


Figure 11.7: Double differential trijet cross sections as function of Q^2 and ξ. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation and EW effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

the experimental ones over most of the phase space. The data are well described by the NLO calculations over more than five orders of magnitude, with deviations no larger than 2σ. However, the central value of the NLO calculation is systematically higher than the data. Especially at small values of $\langle P_{\mathrm{T}} \rangle$ and ξ the NLO calculation shows the tendency to lie above the data, which indicates a preference for a smaller value of α_s. Tests showed that NLO calculations using a value of $\alpha_s(M_Z) = 0.116$ provide an improved description of the data.

The trijet cross sections are already at the Born level proportional to $\mathcal{O}(\alpha_s^2)$ and the NLO calculations include terms of $\mathcal{O}(\alpha_s^3)$. This means that already at LO the triple gluon vertex contributes to the cross section and the four gluon vertex is included in NLO calculations. The trijet cross sections thus provide an important test of pQCD and the running of the strong coupling. The double differential trijet cross sections measured as function of Q^2 and $\langle P_{\mathrm{T}} \rangle$ and Q^2 and ξ are presented in figures 11.6 and 11.7. They show a similar behaviour as the dijet cross sections, but are about five times smaller. The cross section as function of ξ is shifted towards larger values of ξ, due to the higher invariant mass of the trijet system compared to that of the dijet system. The experimental uncertainties are to a large part dominated by the model dependence. They are about 7% at small values of $\langle P_{\mathrm{T}} \rangle$ and ξ and between 10–15% at large values of these variables. The uncertainties on the theoretical predictions are between 10–15%, with smaller uncertainties at high Q^2 and $\langle P_{\mathrm{T}} \rangle$ and larger uncertainties at small Q^2 and $\langle P_{\mathrm{T}} \rangle$. The data are well described by the NLO calculations over the full phase space of the measurement.

11.3 Normalised Jet Cross Sections

The measurement of normalised jet cross sections reduces the experimental uncertainties for some observables nearly by a factor of 2. This is achieved by a measurement of the ratio of jet events to the number of NC DIS events on detector level. This ratio is then corrected for experimental and QED effects. The determination of the experimental correction factors for this ratio leads to a cancellation of the overall normalisation uncertainty. Additionally, the ratio benefits from partial cancellations of the model dependence and the uncertainty due to the hadronic energy scale, which are two of the dominating uncertainties of the jet measurements.

The NLO calculations for the normalised jet cross sections are performed with the NLOJet++ program with the settings as described in the previous section. The inclusive NC cross sections are calculated with the QCDNUM program in NLO, i.e. including terms $\mathcal{O}(\alpha_s)$. The same PDF, CTEQ6.6, and the same value of $\alpha_s(M_Z) = 0.118$ are used for the calculation of the inclusive NC cross sections as for the jet cross sections. For the calculation of inclusive NC cross sections the renormalisation and factorisation scales μ_{r} and μ_{f} are chosen to be Q. The uncertainties due to missing higher orders are much smaller in the case of inclusive NC cross sections than for jet cross sections and can be neglected. The NLO calculations of the normalised jet cross sections are corrected for hadronisation effects similar to the jet cross sections. Therefore, the uncertainties of the NLO predictions for jet production, added in quadrature with the uncertainties of the hadronisation corrections, are taken as the uncertainties of the normalised jet cross sections. The dependence of the EW

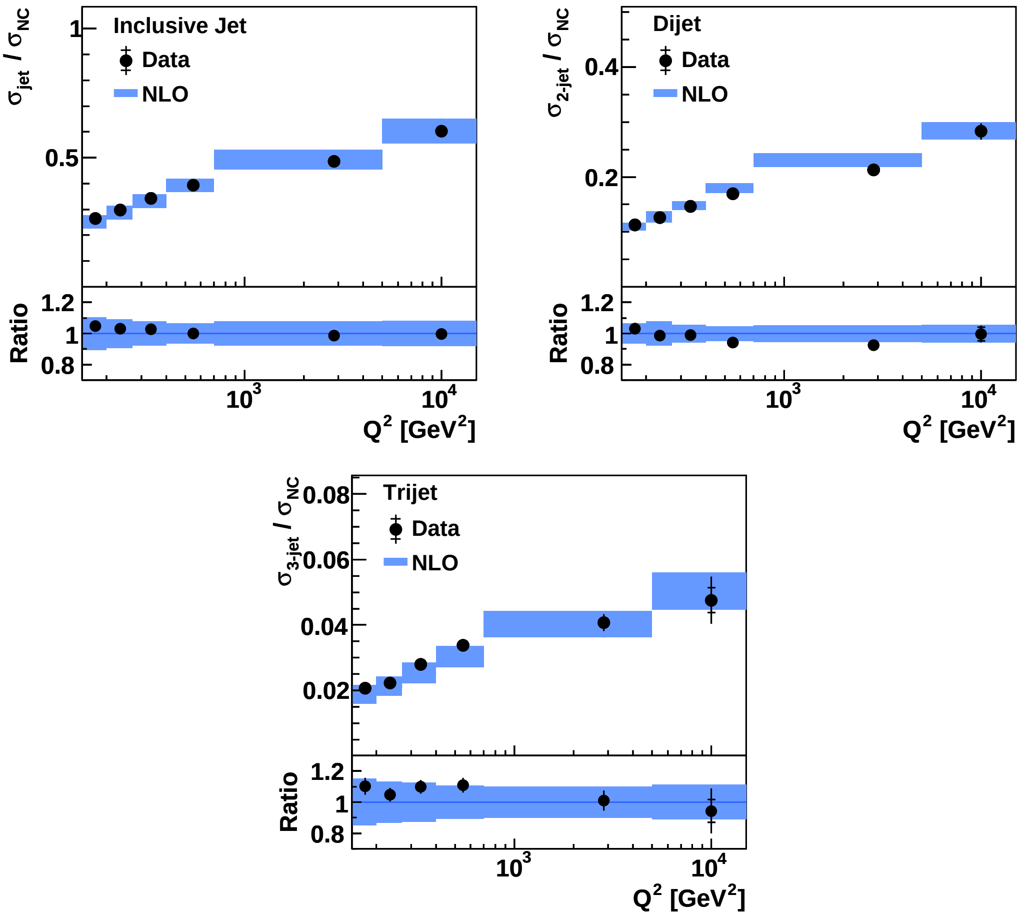


Figure 11.8: Normalised single differential inclusive jet, dijet and trijet cross sections as function of Q^2. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

corrections on P_T and ξ is found to be smaller than 1% for the mixture of e^-p and e^+p scattering data of the full HERA-2 sample (see section 10.6.1). Hence, the EW corrections cancel in the ratio and no corrections need to be applied to the NLO calculations.

The normalised, single differential inclusive jet, dijet and trijet cross sections as function of Q^2 are shown in figure 11.8. The cross sections increase with increasing Q^2 as the available phase space for jet production opens. The rate of dijet production is about a factor two smaller than the average jet multiplicity in NC events $\sigma_\mathrm{jet}/\sigma_\mathrm{NC}$, in agreement with the intuitive expectation. The rate of trijet production is about a factor of five smaller than the rate for dijet production. The data are well described by the NLO calculations. The total experimental uncertainties are approximately 2%, 2.5% and 5% for normalised inclusive jet, dijet and trijet cross sections, respectively. These are about a one third of the theoretical uncertainties. Only in the highest Q^2 region do the statistical uncertainties become important, which leads to experimental uncertainties about twice as large as in regions with smaller values of Q^2.

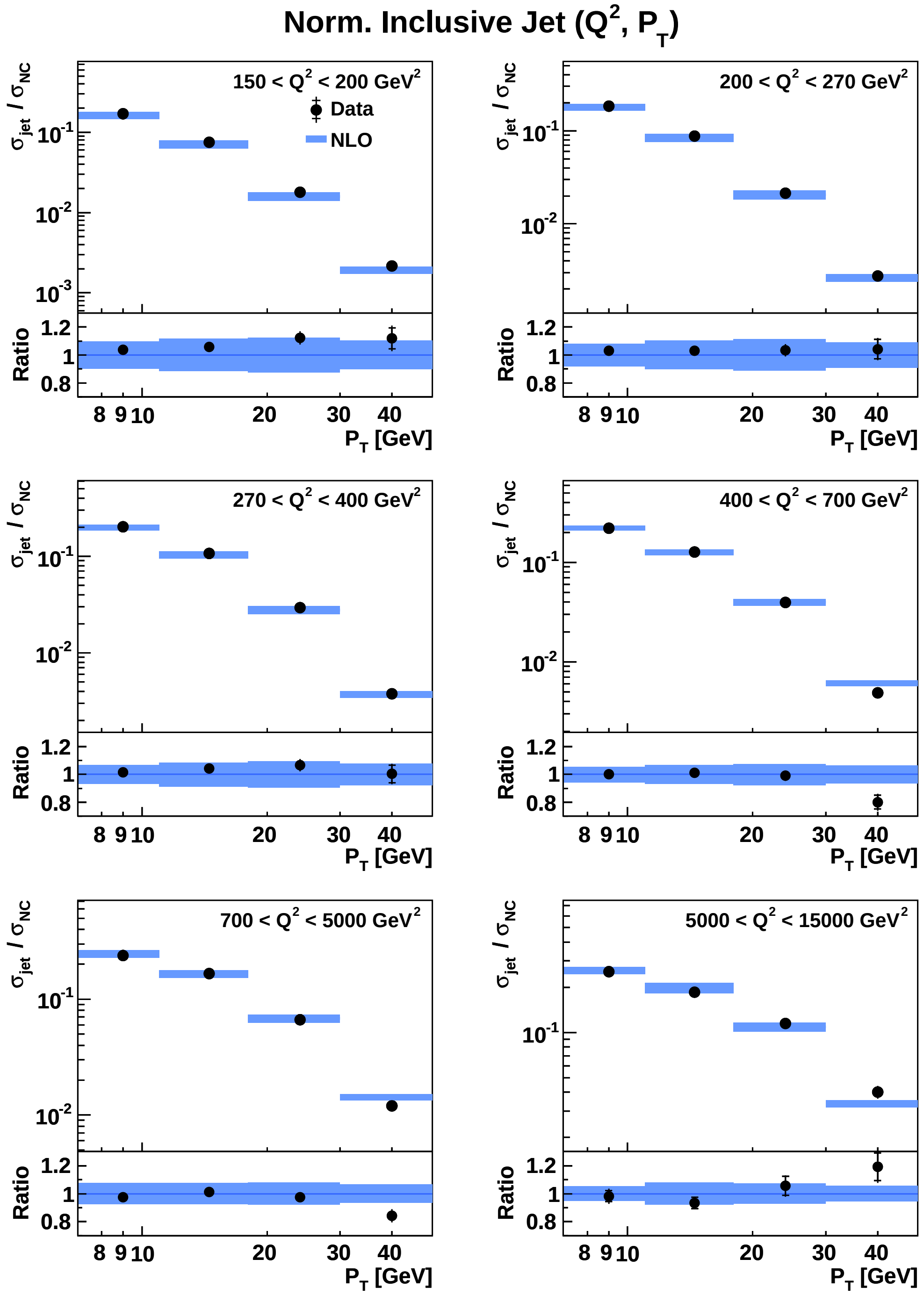


Figure 11.9: Normalised double differential inclusive jet cross sections as function of Q^2 and P_{T}. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

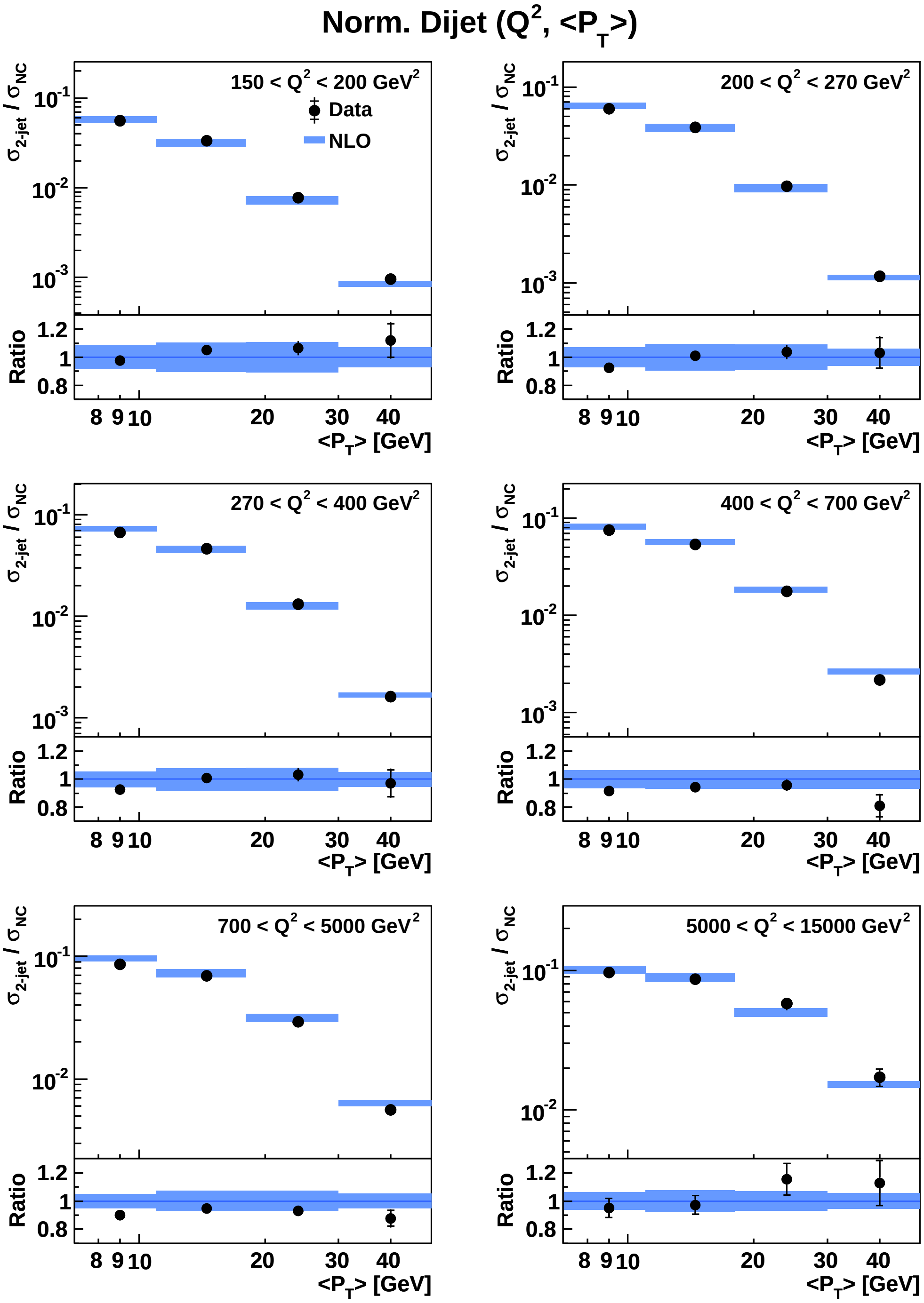


Figure 11.10: Normalised double differential dijet cross sections as function of Q^2 and $\langle P_\mathrm{T} \rangle$. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

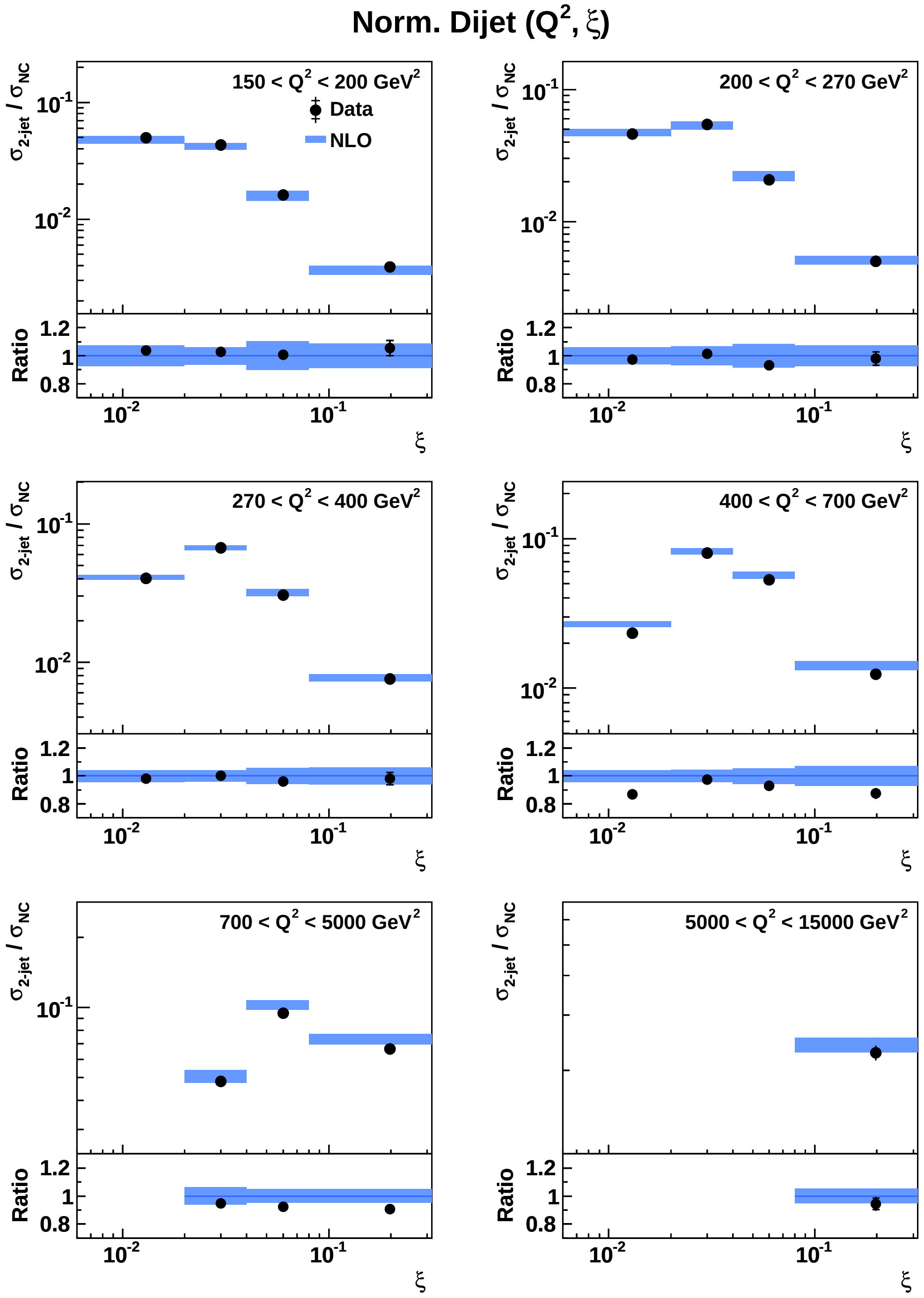


Figure 11.11: Normalised double differential dijet cross sections as function of Q^2 and ξ. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

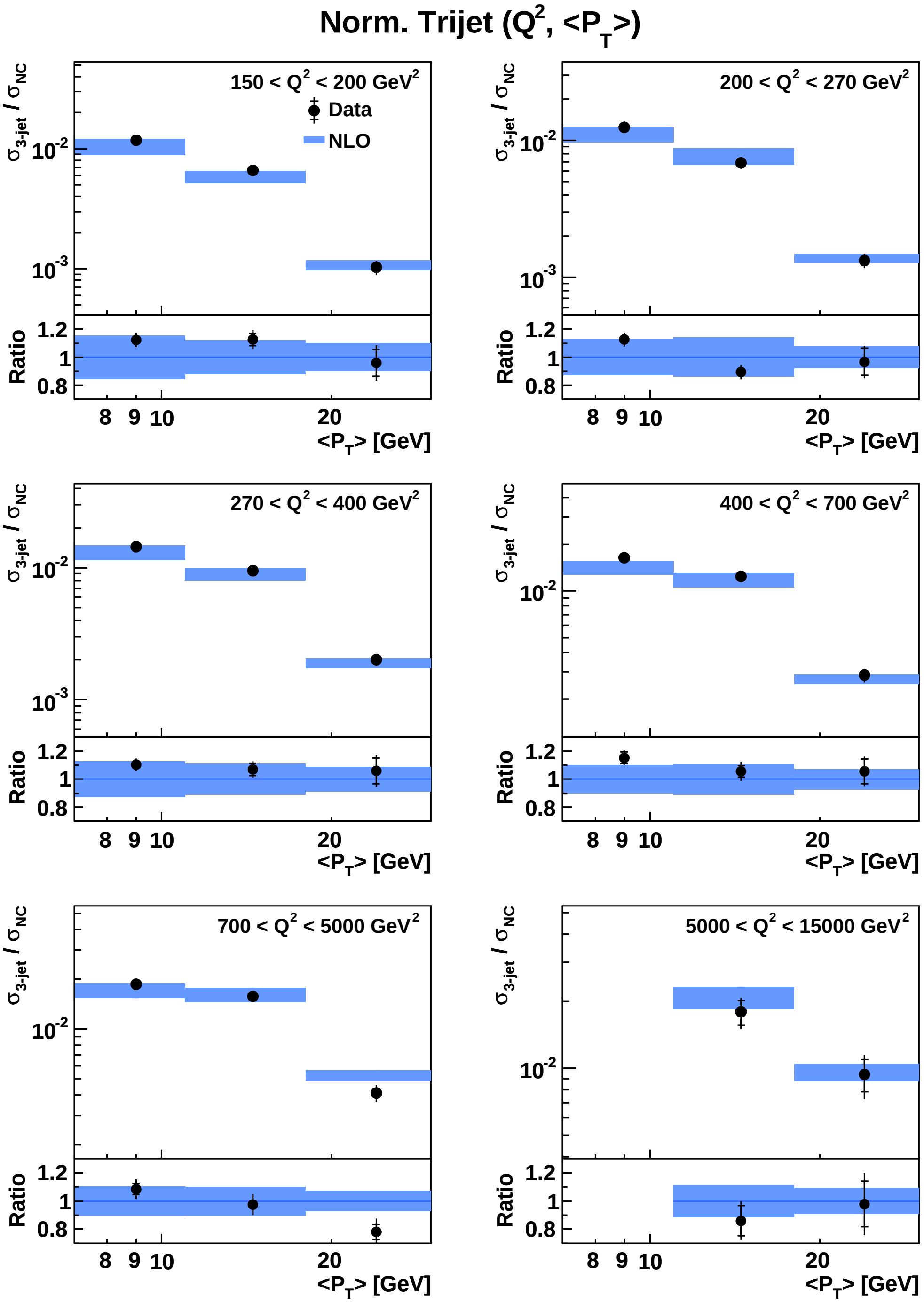


Figure 11.12: Normalised double differential trijet cross sections as function of Q^2 and $\langle P_\mathrm{T} \rangle$. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

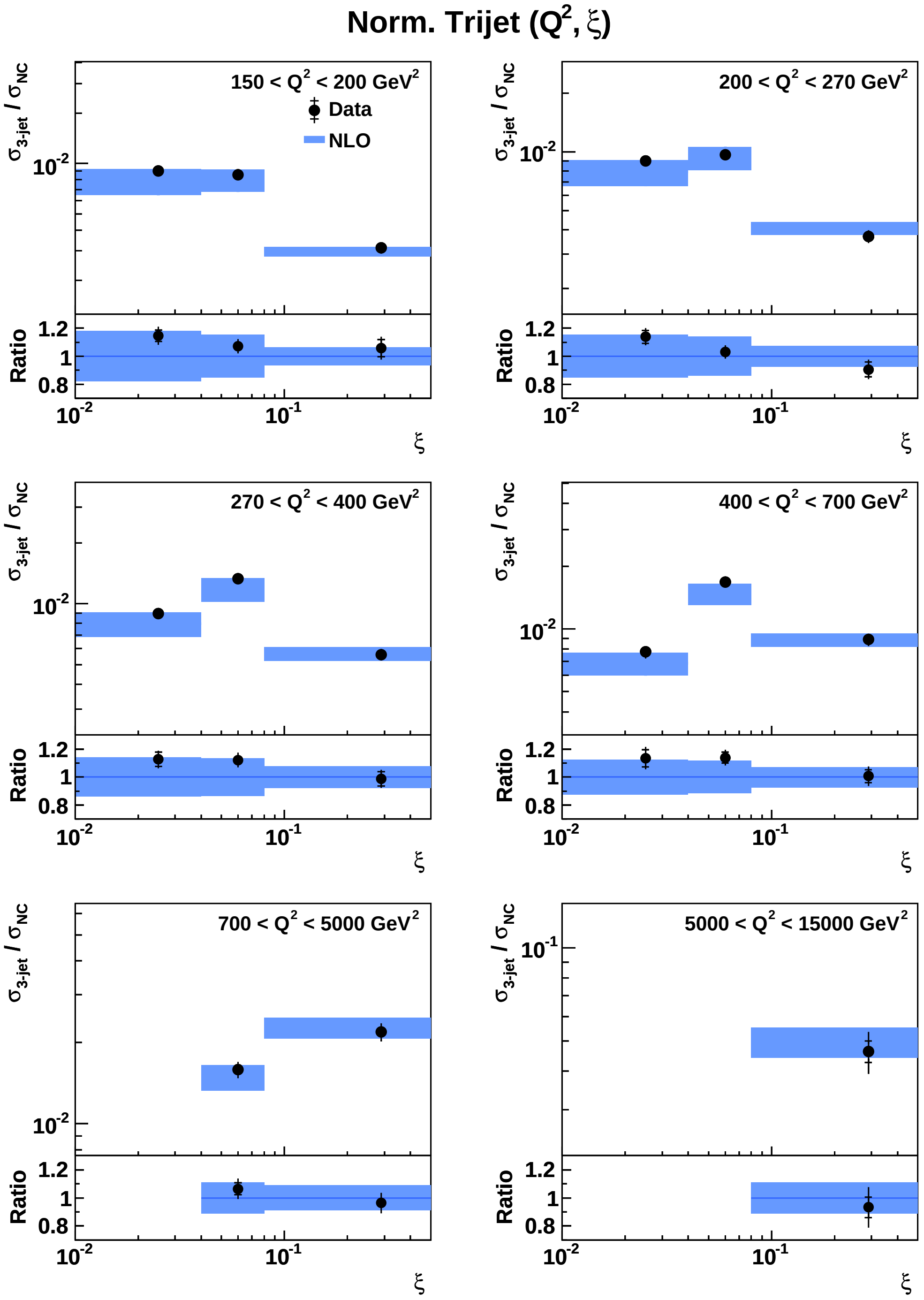


Figure 11.13: Normalised double differential trijet cross sections as function of Q^2 and ξ. The NLO calculations are shown together with their uncertainties as blue bands. The predictions are corrected for hadronisation effects. Further details on the theoretical predictions are given in the text. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

In figures 11.9, 11.10 and 11.11 the normalised double differential inclusive jet and dijet cross sections as function of P_T and ξ are presented for different regions of Q^2. The cross sections are well described by the NLO calculations over the full phase space. The largest deviations between the data and the NLO calculations are found in the bins 4δ and 5δ, where the data lie below the theoretical predictions by about 2.5σ (1.5σ) in the case of the normalised inclusive jet (dijet) cross sections. Preliminary tests for the determination of $\alpha_s(M_Z)$ utilising the normalised inclusive jet and dijet data showed a preference for smaller values of $\alpha_s(M_Z)$ than the value of 0.118. A value of $\alpha_s(M_Z) = 0.116$ leads to an improved description of the normalised inclusive jet and dijet cross sections in bins 4δ and 5δ, as well as at low values of $\langle P_\mathrm{T}\rangle$ in the case of the normalised dijet cross sections.

In figures 11.12 and 11.13 the normalised double differential trijet cross sections as function of $P_\mathrm{T,obs}$ and ξ are presented for different regions of Q^2. The normalised trijet measurement mostly benefits from partial cancellations in the model dependence, which reduces the total experimental uncertainty to 5–6% in most bins, which can be compared to theoretical uncertainties between 10% and 15%.

11.4 Comparison to Other Measurements

The inclusive jet cross sections in NC DIS at high Q^2 have been measured previously by the H1 [H107] and ZEUS [ZEUS07a] collaborations. These measurements are based on HERA-1 data with integrated luminosities of $65.4\,\mathrm{pb}^{-1}$ and $82\,\mathrm{pb}^{-1}$, respectively. The ZEUS measurement is restricted to a different phase space with a requirement on the cosine of the hadronic angle, $|\cos\gamma_\mathrm{h}| < 0.65$, which makes a direct comparison difficult. The H1 measurement on the contrary is performed in the same phase space as this analysis, and the cross sections can be directly compared. The comparison of the double differential inclusive jet cross sections measured as function of P_T in different Q^2 bins is shown in figure 11.14. The new measurement is compatible with the previous H1 data, with the largest deviations observed being less than 2σ in bins 3δ and 4δ. The reduction of the experimental uncertainty by a factor of about two with respect to the HERA-1 measurement is mainly achieved because of the decreased uncertainty due to the hadronic energy measurement. This reduction, from 2% as estimated for the HERA-1 analysis, to 1% as achieved in this work, leads to a reduction of the uncertainty on jet the cross section by about 30%. At high values of Q^2 and P_T the reduced statistical uncertainty and an improved model dependence helped to improve the experimental uncertainty further. The measurement of the inclusive jet cross sections in NC DIS presented here is the most precise measurement up to now.

In two analyses of HERA-1 data the H1 [H101a] and ZEUS [ZEUS07a] collaborations measured the dijet cross sections in NC DIS at high Q^2 single and double differentially. The published cross sections have uncertainties between 10% to 15%, which are twice as large as the uncertainties obtained in this work. In a more recent analysis by the ZEUS collaboration of the combined HERA-1+2 data set, the dijet cross sections are measured with improved experimental precision [ZEUS10a]. In the ZEUS analysis the inelasticity is required to be within $0.2 < y < 0.6$, and the jet phase space is defined through a higher jet transverse momentum of $P_\mathrm{T} > 8\,\mathrm{GeV}$ and

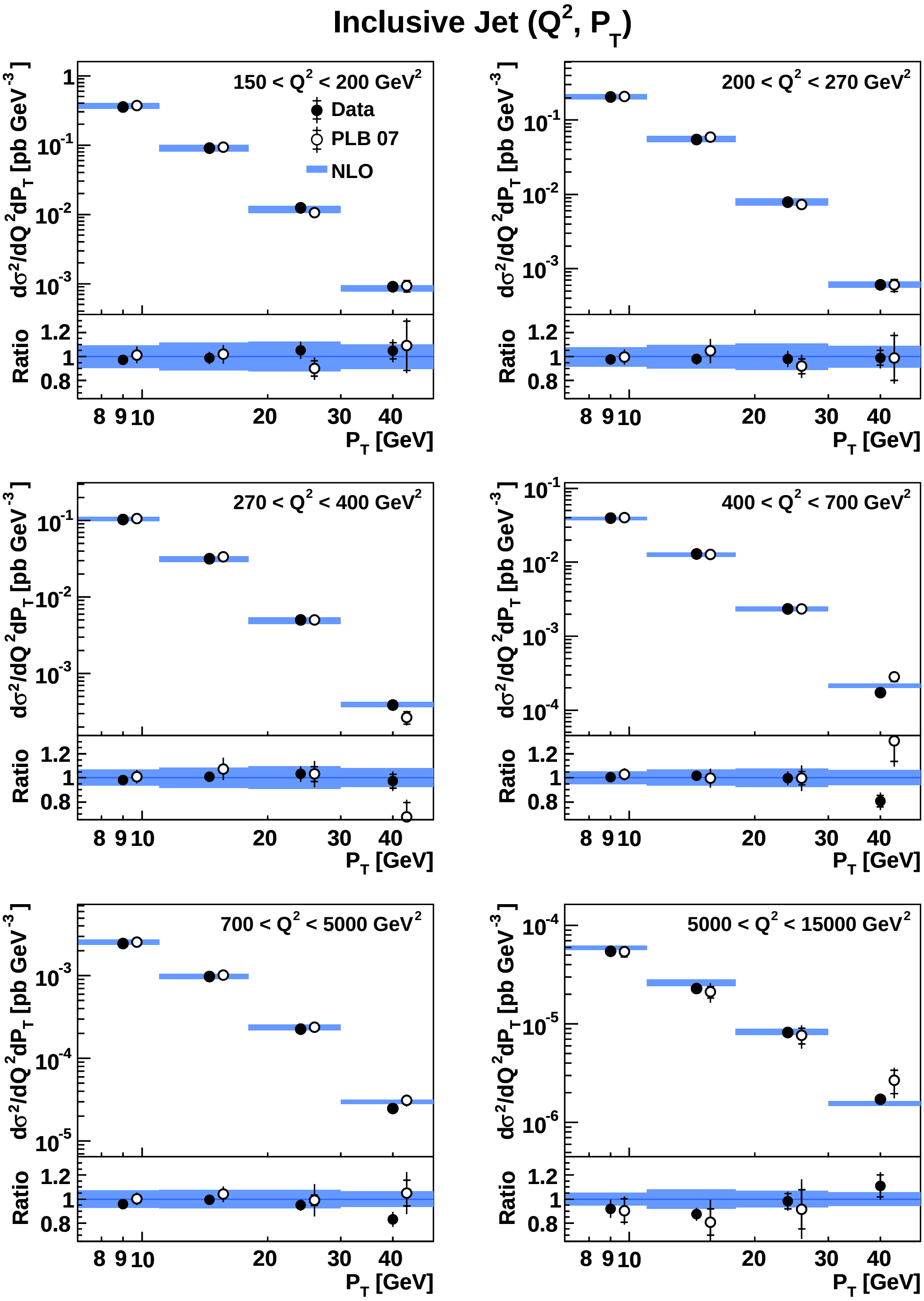


Figure 11.14: Double differential inclusive jet cross sections as function of Q^2 and P_T compared to NLO calculations and to data from a previous H1 measurement using HERA-1 data [H107]. The HERA-1 measurement, exploiting e^+p scattering data only, is corrected for EW effects in order to be directly comparable to the mixture of e^-p and e^+p scattering data from the HERA-2 data set. The open circles are shifted to the right for better visibility. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

a higher invariant mass of the two leading jets, $M_{12} > 20\,\mathrm{GeV}$, than in this analysis. Even though the phase space of the ZEUS analysis is quite similar compared to the one used in this work, a direct comparison of the measured cross sections would still need an extrapolation. The experimental uncertainties of the ZEUS measurement are comparable to the experimental uncertainties of the dijet cross sections presented in this work. A relatively small adjustment of the phase space by the H1 and ZEUS analyses would allow for a first combination of jet data between two experiments. This could help to improve the experimental precision further, as prominently demonstrated by the combined measurement of inclusive NC and CC cross sections in DIS [AA^{+}10a].

The trijet cross sections in NC DIS at high Q^2 have been measured by the H1 [H101b] and ZEUS [ZEUS05a, ZEUS08] collaborations single differentially. In these analyses 2–3 data points at high Q^2 have been measured with uncertainties of 15–20%. The single differential trijet cross sections presented in this work have uncertainties of about half this size. In a recent dissertation by Jörg Behr [Beh10], the trijet cross sections at high Q^2 are measured single and double differentially, exploiting HERA-1+2 data recorded with the ZEUS detector. The phase space of the ZEUS analysis is different, most noticeably a higher jet transverse momentum of $P_{\mathrm{T}} > 8\,\mathrm{GeV}$ and a higher invariant mass of the two hardest jets of $M_{12} > 20\,\mathrm{GeV}$ are required. The more restrictive phase space results in smaller trijet cross sections compared to the ones measured in this work. The experimental uncertainties are comparable between the two analyses, which makes the trijet measurement another possible candidate for a combination between H1 and ZEUS results.

Analyses of normalised inclusive jet cross sections in NC DIS at high Q^2 have been performed by the H1 collaboration on HERA-1 data [H107] and on a combined HERA-1+2 data set [H110c]. In the latter analysis also normalised dijet and trijet cross sections are measured, but the range of pseudorapidity of jets in the laboratory rest frame is more restricted than in this work, with $-0.8 < \eta_{\mathrm{lab}} < 2.0$. The data from the HERA-1 analysis can be directly compared to the normalised inclusive jet cross sections presented here. For a comparison with the normalised jet cross sections from the HERA-1+2 analysis, the measurement is repeated with the more restrictive jet pseudorapidity range. The normalised inclusive jet cross sections as measured in this work, as well as the normalised inclusive jet, dijet and trijet cross sections, obtained with the more restrictive jet pseudorapidity range, are compared to data from the HERA-1 and HERA-1+2 analyses in figure 11.15. The measurements are compatible. A large reduction of the experimental uncertainties is observed when comparing to the HERA-1 analysis. A detailed comparison with the HERA-1+2 analysis reveals a reduction of the experimental uncertainties of about 15% on average. It can be observed that the normalised dijet cross sections are not as well described by the NLO calculations as in the published analysis [H110c], where they are compared to NLO calculations using the CTEQ6.5 PDF set and $\alpha_s(M_Z) = 0.1168$. The NLO calculations presented here use the PDF set CTEQ6.6 and $\alpha_s(M_Z) = 0.118$. This choice of PDF and $\alpha_s(M_Z)$ results in larger normalised jet cross sections at NLO.

In the previous H1 analysis of HERA-1+2 data the normalised inclusive and dijet cross sections are also measured double differentially as function of $P_{\mathrm{T,obs}}$ and ξ in different regions of Q^2. A comparison between these and the normalised jet

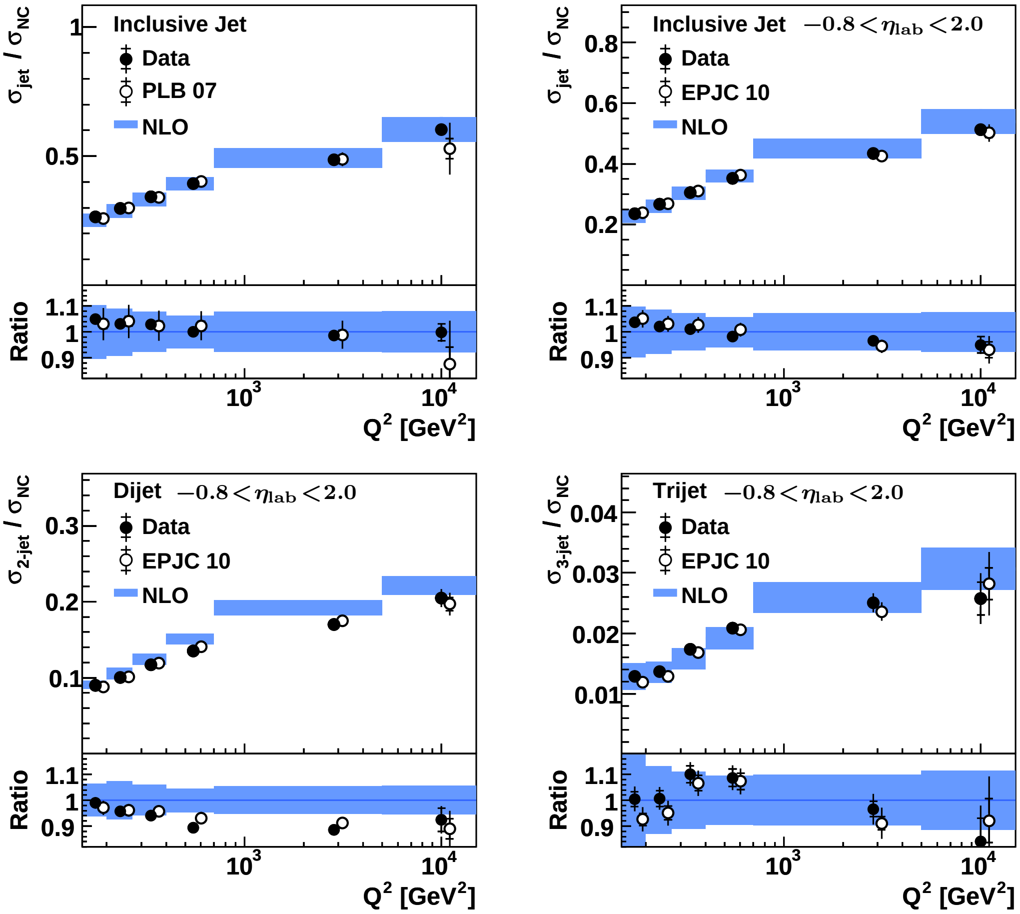


Figure 11.15: Normalised inclusive jet (top), dijet (bottom left) and trijet (bottom right) cross sections compared to data from previous H1 measurements [H107, H110c]. For the comparison with the HERA-1+2 data the range of pseudorapidity of jets in the laboratory rest frame is restricted to $2.0 < \eta_{\text{lab}} < -0.8$. The open circles are shifted to the right for better visibility. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

cross sections obtained in this work with the requirement $-0.8 < \eta_{\text{lab}} < 2.0$, shows differences of less than one standard deviation in all measurement bins. An example of this comparison is shown in figure 11.16 for normalised dijet cross sections as function of $\langle P_{\text{T}} \rangle$ and Q^2. It can be observed that the data are compatible, but are not as well described by the NLO calculations due to the different choice of PDF and $\alpha_s(M_Z)$, as described above. This demonstrates the sensitivity of the data to the choice of the proton PDFs and the value of $\alpha_s(M_Z)$.

11.5 Summary

In this chapter the central physics results of this work have been presented. These consist of a high precision measurement of inclusive jet, dijet and trijet cross sections in NC DIS at high Q^2. The presented measurements span over more than six orders of magnitude between the inclusive jet cross sections at low Q^2 and low

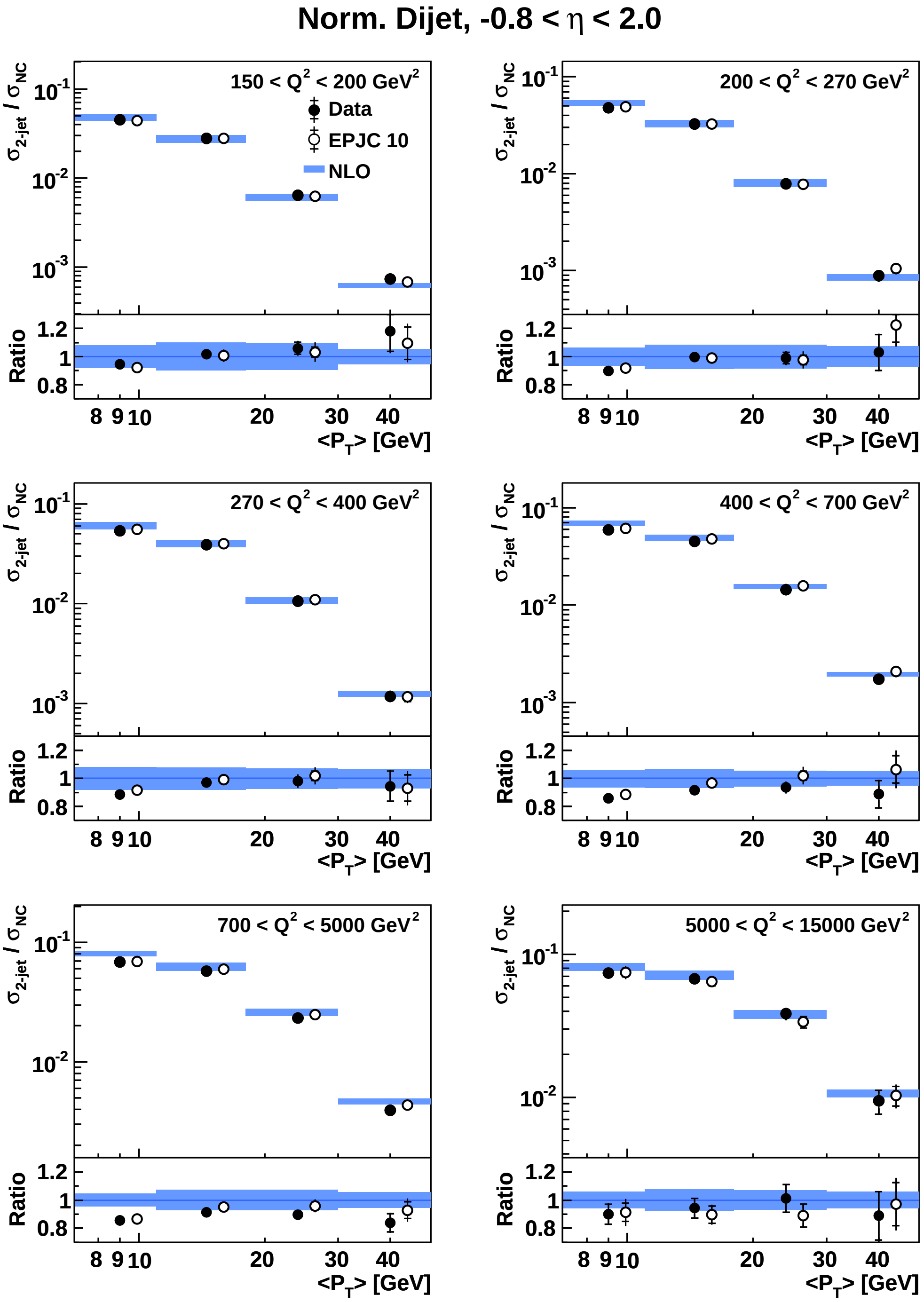


Figure 11.16: Normalised double differential dijet cross sections compared to data from a previous H1 measurement [H110c] and to NLO calculations. The range of pseudorapidity of jets in the laboratory rest frame is restricted to $-0.8 < \eta_{\mathrm{lab}} < 2.0$. The open circles are shifted to the right for better visibility. The ratio of data with respect to the theoretical prediction is shown at the lower part of each plot.

P_{T} and the trijet cross sections at high Q^2 and high $\langle P_{\mathrm{T}} \rangle$. The data are well described by NLO calculations within the uncertainties over the full phase space considered, demonstrating the validity of pQCD for multijet production in NC DIS at high Q^2. Furthermore it has been shown in this chapter how the measurement of normalised jet cross sections improves the experimental precision further due to partial cancellations of the experimental uncertainties.

The jet cross sections presented in this work are sensitive to the strong coupling at LO and can be used to determine the value of α_s at a given scale with high precision. Additionally, the jet cross sections provide an independent test of the proton PDFs. At low values of Q^2 jet production in NC DIS is dominated by gluon induced processes, whereas at high values of Q^2 quark induced processes dominate the cross section. Jet data thus allows to disentangle the value of the gluon PDF from the chosen value of $\alpha_s(M_Z)$. The dijet and trijet cross sections measured as function of ξ, which is in leading order the proton momentum fraction carried by the struck parton, shows the sensitivity of the data to medium to large values of ξ. Since in this region of ξ the gluon PDF is not strongly constrained by inclusive scattering data, jet data can provide an important constraint on the gluon at this range of ξ. This will make the presented jet cross sections valuable input for the extraction of proton PDFs with a simultaneous determination of $\alpha_s(M_Z)$.

The experimental uncertainties of the (normalised) jet cross sections are usually smaller by a factor of (three) two than the theoretical uncertainties. Since the theoretical uncertainties are dominated by missing terms of higher orders, NNLO calculations are necessary in order to make full use of the data. Recently, the antenna subtraction method [GG$^+$05] was extended to include one hadron in the initial state [DG$^+$10a]. This work allows the construction of a parton level event generator program for the calculation of jet production in DIS at NNLO.

Chapter 12

Conclusion

Jet production in neutral current deep-inelastic scattering provides an important test of the validity of pQCD and offers the possibility of an independent test of proton PDFs, which are extracted using predominantly inclusive $e^{\pm}p$ scattering data. Conversely, jet data can be used to determine the value of the strong coupling at a given scale, and it can provide valueable constraints for the determination of the proton PDFs. In order for jet data to make an appreciable contribution in these respects, high precision jet measurements are indespensible.

In this work a considerable effort was undertaken to improve the jet energy scale uncertainty, which is the dominating uncertainty for every jet measurement. As a first step, hadronic and electromagnetic showers measured with the highly segmented LAr calorimeter have been separated on a statistical basis by employing a feed-forward backpropagation neural network. It has been shown that the thereby obtained probability of a shower to be induced by purely electromagnetically interacting particles improves the performance of the energy flow algorithm, which is used for the reconstruction of the hadronic final state. Based on these results a new method to calibrate jets has been developed. This calibration, which has been shown to yield a jet energy scale uncertainty of 1% over a wide range of energies and the full geometrical acceptance of the LAr calorimeter, has been tested by several independent analyses and has recently become the standard used by H1.

The data set used for the measurement of jet cross sections comprises the complete H1 HERA-2 data set, corresponding to a total integrated luminosity of $351.6\,\mathrm{pb}^{-1}$. This improves the statistical uncertainty of the measurement considerably with respect to previous jet cross section measurements based on HERA-1 data. Besides the reconstruction of the hadronic final state, other possible sources of systematic uncertainties have been studied in detail. Efforts by the H1 collaboration to improve the track reconstruction, calibration of the energy of the scattered electron and an independent determination of the luminosity with QED Compton events have improved the uncertainties due to the scattered electron reconstruction and the luminosity measurement. The uncertainty due to the model dependence of the employed acceptance correction is reduced by reweighting the MC simulations. After the application of dedicated non-*ep* and *ep* background finders the uncertainty due to subtraction of background became negligible.

The presented measurements of inclusive jet, dijet and trijet cross sections profited significantly from these developments. Jet cross sections have been measured in the kinematic range $150 < Q^2 < 15000\,\mathrm{GeV}^2$ and $0.2 < y < 0.7$. The inclusive jet, dijet and trijet cross sections have been measured with an average precision of 4%, 5.2% and 7.2%, respectively. The calibration of the hadronic final state up to large values of pseudorapidity made the extension of the phase space to values of η_{lab} up to 2.5 possible, without a deterioration of the jet energy scale uncertainty. This resulted in an increase in statistics in the trijet sample, making the first double-differential measurement of trijet production at high Q^2 at H1 possible. The measurement of jet cross sections normalised to the inclusive NC cross sections provided a possibility to further reduce the experimental uncertainties due to partial cancellations of systematic effects. The obtained normalised inclusive jet, dijet and trijet cross sections have average uncertainties of 1.9%, 2.4% and 6.5%, respectively. The measured jet cross sections are found to be consistent with previous measurements by the H1 collaboration. The measurement of jet production presented in this thesis results in the most precise jet cross sections in neutral current DIS at high Q^2 from HERA to date.

Calculations in NLO in the strong coupling using the CTEQ6.6 parameterisations of the proton PDFs and $\alpha_s(M_Z) = 0.118$ have been compared to the data. Good overall agreement with the data is found for all jet observables. The theoretical uncertainties, being mostly due to missing higher orders in the perturbative expansion, are found to be two to three times larger than the experimental uncertainties on average. A systematic difference between the NLO calculations and the data is seen for the dijet cross sections at small values of $\langle P_{\mathrm{T}} \rangle$, where the NLO predictions are about 10% too high. This suggests a preference for a smaller value of $\alpha_s(M_Z)$ than the used value of 0.118 or a less steep gluon PDF.

The normalised and non-normalised inclusive jet, dijet and trijet cross sections offer the possibility for a precise determination of $\alpha_s(M_Z)$. They will also provide valueable input for the determination of proton PDFs, where they can help to constrain the gluon at medium to large values of x. However, in order to fully profit from the precision of the measured jet cross sections NNLO calculations are required, which are expected to become available in the near future.

Appendix A

Tables of the Results

In this chapter tables with numerical values of the jet cross sections and their uncertainties are given. The phase space of the measurement is summarised in table 9.6 and in section 11.2. The bin labels and boundaries are given in table A.1.
In the following, the normalisation uncertainty denotes the uncertainties due to the luminosity measurement, the trigger efficiency and the vertex-track link efficiency added in quadrature. The electron uncertainty combines the uncertainties due to the measurement of the energy of the scattered electron, the electron polar angle and the electron identification efficiency.

Bin label	Q^2 range (in GeV^2)
1	$150 \leq Q^2 < 200$
2	$200 \leq Q^2 < 270$
3	$270 \leq Q^2 < 400$
4	$400 \leq Q^2 < 700$
5	$700 \leq Q^2 < 5000$
6	$5000 \leq Q^2 < 15000$

Bin label	P_T or $\langle P_\mathrm{T}\rangle$ range (in GeV)
α	$7 \leq P_\mathrm{T} < 11$
β	$11 \leq P_\mathrm{T} < 18$
γ	$18 \leq P_\mathrm{T} < 30$
δ	$30 \leq P_\mathrm{T} < 50$

Bin label	Dijet ξ range
a	$0.006 \leq \xi < 0.02$
b	$0.020 \leq \xi < 0.04$
c	$0.040 \leq \xi < 0.08$
d	$0.080 \leq \xi < 0.316$

Bin label	Trijet ξ range
A	$0.01 \leq \xi < 0.04$
B	$0.04 \leq \xi < 0.08$
C	$0.08 \leq \xi < 0.50$

Table A.1: Bin labels and boundaries for bins in Q^2 (top left), $P_{\mathrm{T,obs}}$ (top right), ξ for the dijet (bottom left) and ξ for the trijet (bottom right) measurements.

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
Inclusive jet cross section in bins of Q^2										
1	$1.10\ 10^{2}$	0.6	4.0	2.9	0.4	1.3	2.2	0.8	0.95	1.5
2	$9.14\ 10^{1}$	0.7	4.1	2.9	0.4	1.3	2.3	0.7	0.95	1.4
3	$9.12\ 10^{1}$	0.7	4.0	2.9	0.4	1.3	2.2	0.6	0.95	1.4
4	$8.42\ 10^{1}$	0.7	4.0	2.9	0.3	1.3	2.1	0.6	0.95	1.3
5	$8.58\ 10^{1}$	0.7	4.1	2.9	0.1	1.4	2.0	1.3	0.95	1.0
6	$5.19\ 10^{0}$	2.4	5.3	2.9	0.0	2.6	1.7	2.1	0.95	0.7
Inclusive jet cross section in bins of P_{T}										
α	$2.73\ 10^{2}$	0.4	3.8	2.9	0.3	1.4	1.8	0.8	0.93	1.4
β	$1.46\ 10^{2}$	0.5	4.0	2.9	0.3	0.8	2.4	0.8	0.97	1.1
γ	$4.28\ 10^{1}$	0.9	4.9	2.9	0.1	1.8	3.3	0.8	0.97	0.9
δ	$6.11\ 10^{0}$	2.3	6.2	2.9	0.2	2.4	4.2	1.0	0.97	0.7

Table A.2: Single differential inclusive jet cross sections measured as function of Q^2 (top) and P_{T} (bottom). The bin labels of column 1 are defined in table A.1. The statistical and total uncertainties on the cross section are given in columns 3 and 4. The uncertainties due to the normalisation, the background subtraction, the model dependence, the hadronic energy scale and the electron measurement are given in columns 5–9. In column 10 and 11 the hadronisation correction factors applied to the NLO calculations and their uncertainties are shown.

Inclusive jet cross section in bins of Q^2 and P_{T}

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$7.08\ 10^{1}$	0.8	3.9	2.9	0.4	1.2	1.8	0.8	0.94	1.8
1β	$3.15\ 10^{1}$	1.1	5.4	2.9	0.3	3.4	2.6	0.7	0.97	1.0
1γ	$7.48\ 10^{0}$	2.2	6.7	2.9	0.2	4.4	3.4	0.5	0.97	1.2
1δ	$9.02\ 10^{-1}$	6.5	9.2	2.9	0.1	3.6	4.5	0.6	0.96	1.1
2α	$5.70\ 10^{1}$	0.9	3.9	2.9	0.4	1.3	1.9	0.7	0.94	1.8
2β	$2.70\ 10^{1}$	1.2	4.6	2.9	0.3	1.9	2.6	0.7	0.97	0.9
2γ	$6.58\ 10^{0}$	2.3	6.9	2.9	0.1	4.5	3.6	0.6	0.97	1.1
2δ	$8.47\ 10^{-1}$	6.3	8.9	2.9	0.0	4.0	3.8	0.8	0.96	1.1
3α	$5.36\ 10^{1}$	0.9	3.9	2.9	0.4	1.4	1.8	0.7	0.93	1.7
3β	$2.86\ 10^{1}$	1.1	4.2	2.9	0.4	1.1	2.5	0.7	0.97	0.7
3γ	$7.87\ 10^{0}$	2.2	6.4	2.9	0.2	3.8	3.6	0.5	0.97	0.9
3δ	$1.00\ 10^{0}$	5.9	8.4	2.9	0.0	2.8	4.4	0.7	0.97	0.8
4α	$4.74\ 10^{1}$	1.0	3.9	2.9	0.3	1.6	1.7	0.6	0.93	1.7
4β	$2.73\ 10^{1}$	1.2	4.0	2.9	0.4	0.6	2.2	0.5	0.97	0.6
4γ	$8.46\ 10^{0}$	2.0	5.8	2.9	0.1	2.8	3.4	0.6	0.97	0.8
4δ	$1.04\ 10^{0}$	6.0	9.2	2.9	0.4	3.9	5.0	0.6	0.97	0.8
5α	$4.21\ 10^{1}$	1.0	4.1	2.9	0.1	1.8	1.6	1.3	0.92	1.2
5β	$2.94\ 10^{1}$	1.2	4.0	2.9	0.2	0.6	2.0	1.3	0.97	0.6
5γ	$1.17\ 10^{1}$	1.7	4.9	2.9	0.1	1.7	2.8	1.3	0.97	0.6
5δ	$2.13\ 10^{0}$	3.8	7.9	2.9	0.2	4.4	4.1	1.5	0.97	0.6
6α	$2.19\ 10^{0}$	3.7	8.1	2.9	0.0	6.2	1.1	2.1	0.92	1.0
6β	$1.61\ 10^{0}$	4.2	6.1	2.9	0.0	1.8	1.8	2.1	0.96	0.9
6γ	$9.88\ 10^{-1}$	6.3	7.9	2.9	0.0	2.2	2.2	2.1	0.99	0.7
6δ	$3.45\ 10^{-1}$	8.2	10.5	2.9	0.0	4.7	3.0	2.1	0.98	0.5

Table A.3: Double differential inclusive jet cross sections measured as function of Q^2 and P_{T}. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Dijet cross section in bins of Q^2

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1	$4.74\ 10^{1}$	1.0	5.3	2.9	0.1	3.7	2.1	0.7	0.94	1.1
2	$3.90\ 10^{1}$	1.0	5.2	2.9	0.1	3.6	2.1	0.6	0.94	1.0
3	$3.91\ 10^{1}$	1.0	5.3	2.9	0.1	3.7	2.0	0.6	0.94	1.0
4	$3.64\ 10^{1}$	1.1	5.3	2.9	0.1	3.8	1.9	0.6	0.94	1.0
5	$3.77\ 10^{1}$	1.1	5.3	2.9	0.0	3.7	1.7	1.3	0.94	0.7
6	$2.44\ 10^{0}$	4.0	7.8	2.9	0.0	5.4	1.6	2.1	0.94	1.6
Dijet cross section in bins of $\langle P_{\mathrm{T}} \rangle$										
α	$9.19\ 10^{1}$	0.7	4.3	2.9	0.1	2.6	1.6	0.7	0.93	1.0
β	$6.27\ 10^{1}$	0.8	4.6	2.9	0.1	2.5	2.3	0.7	0.97	0.8
γ	$1.91\ 10^{1}$	1.4	5.4	2.9	0.1	2.8	3.2	0.8	0.97	0.7
δ	$2.73\ 10^{0}$	3.7	7.0	2.9	0.1	2.7	4.2	1.0	0.97	0.5
Dijet cross section in bins of ξ										
a	$5.12\ 10^{1}$	0.9	5.5	2.9	0.2	3.8	2.4	0.6	0.94	1.3
b	$7.83\ 10^{1}$	0.7	5.1	2.9	0.1	3.6	1.8	0.6	0.94	1.1
c	$4.93\ 10^{1}$	0.9	5.2	2.9	0.1	3.7	1.8	0.7	0.94	0.9
d	$2.14\ 10^{1}$	1.4	5.6	2.9	0.0	3.8	2.2	1.2	0.94	0.6

Table A.4: Single differential dijet cross sections measured as function of Q^2 (top), $\langle P_{\mathrm{T}} \rangle$ (middle) and ξ (bottom). The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Dijet cross section in bins of Q^2 and $\langle P_{\mathrm{T}}\rangle$

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$2.35\ 10^{1}$	1.4	4.6	2.9	0.1	2.7	1.7	0.7	0.94	1.1
1β	$1.40\ 10^{1}$	1.7	5.4	2.9	0.1	3.4	2.4	0.6	0.97	0.9
1γ	$3.23\ 10^{0}$	3.5	6.9	2.9	0.1	3.9	3.4	0.5	0.97	1.0
1δ	$3.99\ 10^{-1}$	10.2	12.1	2.9	0.0	3.7	4.5	0.7	0.97	1.1
2α	$1.85\ 10^{1}$	1.5	4.6	2.9	0.1	2.6	1.7	0.7	0.94	1.0
2β	$1.19\ 10^{1}$	1.8	5.1	2.9	0.1	2.9	2.4	0.6	0.98	0.7
2γ	$2.99\ 10^{0}$	3.5	7.0	2.9	0.0	4.0	3.4	0.6	0.97	1.3
2δ	$3.62\ 10^{-1}$	10.1	12.1	2.9	0.0	4.1	4.4	0.8	0.96	1.0
3α	$1.79\ 10^{1}$	1.5	4.7	2.9	0.1	2.9	1.6	0.6	0.93	1.6
3β	$1.23\ 10^{1}$	1.8	5.0	2.9	0.2	2.6	2.3	0.6	0.98	0.5
3γ	$3.50\ 10^{0}$	3.3	6.6	2.9	0.1	3.5	3.4	0.6	0.97	0.9
3δ	$4.32\ 10^{-1}$	9.5	11.3	2.9	0.0	2.8	4.5	0.7	0.97	1.0
4α	$1.61\ 10^{1}$	1.7	4.7	2.9	0.1	2.9	1.4	0.6	0.92	1.7
4β	$1.15\ 10^{1}$	1.9	4.9	2.9	0.1	2.6	2.3	0.5	0.97	0.5
4γ	$3.80\ 10^{0}$	3.3	6.4	2.9	0.1	3.2	3.2	0.6	0.98	0.7
4δ	$4.64\ 10^{-1}$	9.4	11.3	2.9	0.1	2.3	4.8	0.7	0.97	0.8
5α	$1.52\ 10^{1}$	1.7	4.3	2.9	0.0	2.1	1.1	1.3	0.92	1.2
5β	$1.22\ 10^{1}$	1.8	4.9	2.9	0.0	2.6	1.9	1.3	0.96	0.7
5γ	$5.16\ 10^{0}$	2.7	5.9	2.9	0.1	2.9	2.9	1.4	0.98	0.6
5δ	$9.85\ 10^{-1}$	6.4	8.8	2.9	0.1	3.3	3.9	1.5	0.97	0.7
6α	$8.37\ 10^{-1}$	6.9	9.0	2.9	0.0	4.3	0.8	2.2	0.91	0.9
6β	$7.50\ 10^{-1}$	6.6	8.1	2.9	0.0	2.6	1.6	2.2	0.95	1.1
6γ	$4.98\ 10^{-1}$	9.5	10.6	2.9	0.0	1.8	2.3	2.2	0.97	0.7
6δ	$1.48\ 10^{-1}$	14.0	14.9	2.9	0.0	2.1	2.7	2.1	0.98	0.8

Table A.5: Double differential dijet cross sections measured as function of Q^2 and $\langle P_{\mathrm{T}}\rangle$. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Dijet cross section in bins of Q^2 and ξ										
Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1a	$2.08\ 10^1$	1.5	5.4	2.9	0.2	3.6	2.4	0.5	0.94	1.7
1b	$1.81\ 10^1$	1.5	5.6	2.9	0.1	4.0	1.7	0.9	0.94	1.2
1c	$6.73\ 10^0$	2.5	6.1	2.9	0.1	4.2	2.0	0.9	0.94	0.9
1d	$1.62\ 10^0$	5.0	7.6	2.9	0.0	3.9	2.7	0.8	0.93	1.5
2a	$1.42\ 10^1$	1.8	5.3	2.9	0.2	3.3	2.3	0.6	0.94	1.5
2b	$1.67\ 10^1$	1.5	5.2	2.9	0.1	3.6	1.8	0.7	0.94	1.1
2c	$6.38\ 10^0$	2.4	6.2	2.9	0.1	4.4	2.1	0.7	0.95	0.7
2d	$1.54\ 10^0$	4.8	7.1	2.9	0.0	3.2	2.8	0.8	0.94	1.3
3a	$1.08\ 10^1$	2.1	6.0	2.9	0.2	4.0	2.5	0.6	0.93	1.4
3b	$1.79\ 10^1$	1.5	5.0	2.9	0.1	3.3	1.8	0.6	0.94	1.0
3c	$8.20\ 10^0$	2.1	5.6	2.9	0.1	3.7	2.0	0.6	0.94	0.6
3d	$2.02\ 10^0$	4.3	6.7	2.9	0.0	3.3	2.6	0.6	0.95	1.3
4a	$4.99\ 10^0$	3.2	6.9	2.9	0.3	4.7	2.6	0.5	0.92	1.4
4b	$1.71\ 10^1$	1.6	5.3	2.9	0.1	3.7	1.8	0.7	0.93	1.2
4c	$1.13\ 10^1$	1.9	5.1	2.9	0.1	3.3	1.7	0.6	0.94	0.6
4d	$2.65\ 10^0$	3.8	6.2	2.9	0.1	2.6	2.7	0.6	0.96	1.1
5b	$8.49\ 10^0$	2.3	5.6	2.9	0.1	3.5	1.8	1.1	0.93	0.9
5c	$1.67\ 10^1$	1.6	5.1	2.9	0.0	3.3	1.6	1.2	0.94	0.5
5d	$1.17\ 10^1$	1.9	5.5	2.9	0.1	3.5	1.9	1.5	0.94	0.9
6d	$1.97\ 10^0$	4.2	8.2	2.9	0.0	5.9	1.6	2.1	0.94	1.6

Table A.6: Double differential dijet cross sections measured as function of Q^2 and ξ. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
Trijet cross section in bins of Q^2										
1	$8.67\ 10^{0}$	2.2	6.6	2.9	0.1	4.5	3.1	0.6	0.82	2.4
2	$6.89\ 10^{0}$	2.4	6.3	2.9	0.1	3.7	3.2	0.7	0.80	2.4
3	$7.46\ 10^{0}$	2.3	6.4	2.9	0.2	4.1	3.2	0.6	0.81	2.5
4	$7.23\ 10^{0}$	2.4	7.8	2.9	0.1	6.0	3.0	0.7	0.81	2.5
5	$7.19\ 10^{0}$	2.3	11.1	2.9	0.1	9.9	2.8	1.4	0.80	2.7
6	$4.13\ 10^{-1}$	7.5	19.9	2.9	0.0	17.9	2.3	2.2	0.80	2.9
Trijet cross section in bins of $\langle P_{\mathrm{T}} \rangle$										
α	$1.96\ 10^{1}$	1.5	6.9	2.9	0.1	5.5	2.6	0.7	0.79	2.5
β	$1.29\ 10^{1}$	1.7	8.0	2.9	0.1	6.4	3.4	0.8	0.85	2.6
γ	$2.73\ 10^{0}$	3.6	12.0	2.9	0.1	10.0	4.6	1.0	0.87	2.8
Trijet cross section in bins of ξ										
A	$1.11\ 10^{1}$	2.0	7.2	2.9	0.2	5.1	3.5	0.7	0.81	2.4
B	$1.64\ 10^{1}$	1.6	7.0	2.9	0.1	5.5	2.8	0.7	0.81	2.5
C	$1.00\ 10^{1}$	1.9	10.3	2.9	0.1	9.1	3.1	1.0	0.81	2.5

Table A.7: Single differential trijet cross sections measured as function of Q^2 (top), $\langle P_{\mathrm{T}} \rangle$ (middle) and ξ (bottom). The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Trijet cross section in bins of Q^2 and $\langle P_T \rangle$

Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$4.94\ 10^{0}$	3.0	6.2	2.9	0.1	3.6	2.7	0.6	0.81	2.4
1β	$2.75\ 10^{0}$	3.9	8.9	2.9	0.1	6.6	3.4	0.6	0.85	2.2
1γ	$4.30\ 10^{-1}$	9.5	15.8	2.9	0.1	11.4	4.4	0.7	0.87	2.5
2α	$3.86\ 10^{0}$	3.3	6.1	2.9	0.1	2.9	2.9	0.7	0.79	2.1
2β	$2.11\ 10^{0}$	4.2	7.9	2.9	0.1	4.8	3.6	0.6	0.85	2.5
2γ	$4.08\ 10^{-1}$	9.3	13.8	2.9	0.0	8.3	4.8	1.4	0.88	2.5
3α	$3.85\ 10^{0}$	3.3	6.5	2.9	0.2	3.9	2.6	0.6	0.80	2.4
3β	$2.54\ 10^{0}$	3.9	7.6	2.9	0.2	4.7	3.5	0.5	0.85	2.5
3γ	$5.36\ 10^{-1}$	8.3	12.0	2.9	0.0	6.1	5.4	0.6	0.87	2.7
4α	$3.51\ 10^{0}$	3.5	7.2	2.9	0.1	5.0	2.4	0.7	0.78	2.2
4β	$2.66\ 10^{0}$	3.9	9.7	2.9	0.2	7.6	3.5	0.6	0.86	2.5
4γ	$6.12\ 10^{-1}$	8.0	11.9	2.9	0.1	6.5	5.0	0.8	0.86	2.8
5α	$3.28\ 10^{0}$	3.6	10.6	2.9	0.0	9.1	2.3	1.3	0.78	2.4
5β	$2.78\ 10^{0}$	3.6	12.1	2.9	0.0	10.7	3.0	1.5	0.84	2.7
5γ	$7.25\ 10^{-1}$	6.9	17.0	2.9	0.4	14.6	4.3	1.5	0.86	2.6
6β	$1.55\ 10^{-1}$	12.1	18.8	2.9	0.1	13.7	2.2	2.3	0.83	2.7
6γ	$8.16\ 10^{-2}$	15.7	26.4	2.9	0.0	20.6	3.4	2.3	0.86	3.2

Table A.8: Double differential trijet cross sections measured as function of Q^2 and $\langle P_T \rangle$. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Trijet cross section in bins of Q^2 and ξ										
Bin label	Cross section [pb]	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1A	$3.77\ 10^{0}$	3.5	7.0	2.9	0.2	4.0	3.4	0.7	0.82	2.2
1B	$3.57\ 10^{0}$	3.4	7.5	2.9	0.0	5.3	2.6	0.6	0.81	2.4
1C	$1.32\ 10^{0}$	5.5	11.0	2.9	0.2	8.4	3.2	0.6	0.81	2.4
2A	$2.77\ 10^{0}$	3.9	6.9	2.9	0.1	3.4	3.4	0.7	0.80	2.3
2B	$2.97\ 10^{0}$	3.6	6.9	2.9	0.1	4.0	3.0	0.7	0.81	2.4
2C	$1.14\ 10^{0}$	5.6	10.3	2.9	0.1	7.3	3.5	0.7	0.81	2.4
3A	$2.39\ 10^{0}$	4.3	7.5	2.9	0.3	3.9	3.5	0.7	0.80	2.1
3B	$3.54\ 10^{0}$	3.3	7.2	2.9	0.1	4.8	2.9	0.6	0.81	2.5
3C	$1.49\ 10^{0}$	4.9	8.8	2.9	0.1	5.8	3.3	0.6	0.81	2.7
4A	$1.66\ 10^{0}$	5.4	10.0	2.9	0.3	7.1	3.5	0.7	0.80	2.3
4B	$3.60\ 10^{0}$	3.4	7.8	2.9	0.1	5.8	2.7	0.7	0.81	2.4
4C	$1.91\ 10^{0}$	4.5	10.4	2.9	0.1	8.2	3.2	0.7	0.82	2.6
5B	$2.79\ 10^{0}$	3.9	11.1	2.9	0.0	9.6	2.7	1.2	0.80	2.4
5C	$3.85\ 10^{0}$	3.1	12.6	2.9	0.1	11.4	2.8	1.5	0.81	2.8
6C	$4.02\ 10^{-1}$	7.7	20.0	2.9	0.0	18.0	2.4	2.1	0.80	3.1

Table A.9: Double differential trijet cross sections measured as function of Q^2 and ξ. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
Normalised inclusive jet cross section in bins of Q^2										
1	$2.64\ 10^{-1}$	0.7	1.9	-	0.4	0.2	1.7	0.5	0.95	1.5
2	$2.96\ 10^{-1}$	0.8	1.9	-	0.4	0.2	1.6	0.5	0.95	1.4
3	$3.42\ 10^{-1}$	0.8	1.8	-	0.4	0.2	1.5	0.4	0.95	1.4
4	$3.93\ 10^{-1}$	0.9	1.8	-	0.3	0.4	1.5	0.4	0.95	1.3
5	$4.86\ 10^{-1}$	0.8	1.8	-	0.1	0.4	1.4	0.5	0.95	1.0
6	$6.02\ 10^{-1}$	3.2	3.6	-	0.0	0.8	1.1	0.9	0.95	0.7
Normalised dijet cross section in bins of Q^2										
1	$1.13\ 10^{-1}$	1.0	2.5	-	0.1	1.6	1.5	0.4	0.94	1.1
2	$1.26\ 10^{-1}$	1.1	2.3	-	0.1	1.4	1.4	0.3	0.94	1.0
3	$1.46\ 10^{-1}$	1.1	2.3	-	0.1	1.4	1.4	0.2	0.94	1.0
4	$1.70\ 10^{-1}$	1.2	2.1	-	0.1	1.3	1.3	0.1	0.94	1.0
5	$2.14\ 10^{-1}$	1.2	2.0	-	0.0	1.1	1.1	0.1	0.94	0.7
6	$2.83\ 10^{-1}$	4.5	5.3	-	0.0	2.6	1.0	0.4	0.94	1.6
Normalised trijet cross section in bins of Q^2										
1	$2.07\ 10^{-2}$	2.2	5.0	-	0.1	3.7	2.5	0.3	0.82	2.4
2	$2.23\ 10^{-2}$	2.4	4.2	-	0.1	2.3	2.6	0.1	0.80	2.4
3	$2.79\ 10^{-2}$	2.3	4.0	-	0.2	2.1	2.5	0.2	0.81	2.5
4	$3.37\ 10^{-2}$	2.5	4.4	-	0.1	2.8	2.3	0.2	0.81	2.5
5	$4.07\ 10^{-2}$	2.4	6.5	-	0.1	5.6	2.2	0.3	0.80	2.7
6	$4.76\ 10^{-2}$	7.8	15.3	-	0.0	13.0	1.7	0.3	0.80	2.9

Table A.10: Normalised single differential inclusive jet (top), dijet (middle) and trijet (bottom) cross sections measured as function of Q^2. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Normalised inclusive jet cross section in bins of Q^2 and P_T

Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$1.69\ 10^{-1}$	0.8	1.8	-	0.4	0.7	1.3	0.6	0.94	1.8
1β	$7.52\ 10^{-2}$	1.2	3.1	-	0.3	2.0	2.1	0.5	0.97	1.0
1γ	$1.79\ 10^{-2}$	2.4	4.3	-	0.2	2.2	2.9	0.2	0.97	1.2
1δ	$2.15\ 10^{-3}$	6.8	8.2	-	0.1	2.5	3.9	0.4	0.96	1.1
2α	$1.85\ 10^{-1}$	0.9	1.7	-	0.4	0.4	1.2	0.5	0.94	1.8
2β	$8.76\ 10^{-2}$	1.2	2.7	-	0.3	1.4	2.0	0.5	0.97	0.9
2γ	$2.13\ 10^{-2}$	2.4	4.4	-	0.1	2.2	2.9	0.3	0.97	1.1
2δ	$2.75\ 10^{-3}$	6.6	7.7	-	0.0	2.3	3.2	0.4	0.96	1.1
3α	$2.01\ 10^{-1}$	0.9	1.6	-	0.4	0.3	1.1	0.5	0.93	1.7
3β	$1.07\ 10^{-1}$	1.2	2.5	-	0.4	1.2	1.8	0.5	0.97	0.7
3γ	$2.95\ 10^{-2}$	2.2	4.1	-	0.2	1.7	3.0	0.4	0.97	0.9
3δ	$3.76\ 10^{-3}$	6.2	7.3	-	0.0	1.2	3.7	0.3	0.97	0.8
4α	$2.21\ 10^{-1}$	1.0	1.6	-	0.3	0.3	1.1	0.4	0.93	1.7
4β	$1.27\ 10^{-1}$	1.3	2.4	-	0.4	1.3	1.5	0.4	0.97	0.6
4γ	$3.95\ 10^{-2}$	2.1	3.6	-	0.1	0.9	2.8	0.4	0.97	0.8
4δ	$4.86\ 10^{-3}$	6.2	7.7	-	0.4	1.6	4.3	0.3	0.97	0.8
5α	$2.39\ 10^{-1}$	1.0	1.6	-	0.1	0.2	1.0	0.6	0.92	1.2
5β	$1.67\ 10^{-1}$	1.2	2.3	-	0.2	1.1	1.4	0.6	0.97	0.6
5γ	$6.64\ 10^{-2}$	1.7	3.0	-	0.1	0.9	2.2	0.4	0.97	0.6
5δ	$1.20\ 10^{-2}$	3.9	5.6	-	0.2	1.9	3.5	0.5	0.97	0.6
6α	$2.54\ 10^{-1}$	4.0	5.4	-	0.0	3.6	0.5	0.8	0.92	1.0
6β	$1.86\ 10^{-1}$	4.4	4.7	-	0.0	0.7	1.2	1.0	0.96	0.9
6γ	$1.15\ 10^{-1}$	6.4	7.4	-	0.0	3.0	1.6	1.0	0.99	0.7
6δ	$4.00\ 10^{-2}$	8.3	9.7	-	0.0	4.3	2.4	0.7	0.98	0.5

Table A.11: Normalised double differential inclusive jet cross sections measured as function of Q^2 and P_T. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Normalised dijet cross section in bins of Q^2 and $\langle P_\mathrm{T} \rangle$										
Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$5.61\ 10^{-2}$	1.4	2.0	-	0.1	0.8	1.2	0.4	0.94	1.1
1β	$3.33\ 10^{-2}$	1.8	2.8	-	0.1	1.1	1.8	0.4	0.97	0.9
1γ	$7.72\ 10^{-3}$	3.6	4.9	-	0.1	1.8	2.8	0.2	0.97	1.0
1δ	$9.53\ 10^{-4}$	10.6	11.5	-	0.0	1.9	3.9	0.5	0.97	1.1
2α	$6.01\ 10^{-2}$	1.5	2.0	-	0.1	0.7	1.1	0.3	0.94	1.0
2β	$3.86\ 10^{-2}$	1.8	2.7	-	0.1	0.8	1.7	0.3	0.98	0.7
2γ	$9.68\ 10^{-3}$	3.6	4.9	-	0.0	1.8	2.8	0.3	0.97	1.3
2δ	$1.17\ 10^{-3}$	10.6	11.4	-	0.0	2.0	3.7	0.5	0.96	1.0
3α	$6.70\ 10^{-2}$	1.5	2.0	-	0.1	0.8	1.0	0.2	0.93	1.6
3β	$4.61\ 10^{-2}$	1.8	2.6	-	0.2	0.6	1.7	0.3	0.98	0.5
3γ	$1.31\ 10^{-2}$	3.4	4.6	-	0.1	1.4	2.8	0.3	0.97	0.9
3δ	$1.62\ 10^{-3}$	9.8	10.6	-	0.0	1.0	3.9	0.5	0.97	1.0
4α	$7.50\ 10^{-2}$	1.7	2.0	-	0.1	0.6	0.8	0.1	0.92	1.7
4β	$5.38\ 10^{-2}$	1.9	2.5	-	0.1	0.5	1.6	0.2	0.97	0.5
4γ	$1.77\ 10^{-2}$	3.3	4.3	-	0.1	1.0	2.5	0.2	0.98	0.7
4δ	$2.17\ 10^{-3}$	9.7	10.5	-	0.1	0.6	4.2	0.3	0.97	0.8
5α	$8.63\ 10^{-2}$	1.8	2.0	-	0.0	0.7	0.5	0.1	0.92	1.2
5β	$6.93\ 10^{-2}$	1.9	2.3	-	0.0	0.4	1.3	0.2	0.96	0.7
5γ	$2.93\ 10^{-2}$	2.8	3.7	-	0.1	0.7	2.3	0.3	0.98	0.6
5δ	$5.58\ 10^{-3}$	6.4	7.3	-	0.1	1.1	3.3	0.6	0.97	0.7
6α	$9.71\ 10^{-2}$	7.1	7.3	-	0.0	1.5	0.6	0.6	0.91	0.9
6β	$8.69\ 10^{-2}$	6.7	6.9	-	0.0	0.4	1.1	1.1	0.95	1.1
6γ	$5.78\ 10^{-2}$	9.6	10.0	-	0.0	1.8	1.8	0.4	0.97	0.7
6δ	$1.72\ 10^{-2}$	14.2	14.4	-	0.0	0.9	2.1	0.6	0.98	0.8

Table A.12: Normalised double differential dijet cross sections measured as function of Q^2 and $\langle P_\mathrm{T} \rangle$. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Normalised dijet cross section in bins of Q^2 and ξ										
Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1a	$4.96\ 10^{-2}$	1.5	2.8	-	0.2	1.5	1.8	0.1	0.94	1.7
1b	$4.32\ 10^{-2}$	1.6	2.6	-	0.1	1.5	1.2	0.7	0.94	1.2
1c	$1.61\ 10^{-2}$	2.6	3.6	-	0.1	2.0	1.4	0.7	0.94	0.9
1d	$3.88\ 10^{-3}$	5.2	5.9	-	0.0	1.8	2.1	0.7	0.93	1.5
2a	$4.60\ 10^{-2}$	1.8	2.7	-	0.2	1.2	1.7	0.1	0.94	1.5
2b	$5.41\ 10^{-2}$	1.6	2.3	-	0.1	1.3	1.2	0.2	0.94	1.1
2c	$2.07\ 10^{-2}$	2.5	3.6	-	0.1	2.0	1.4	0.5	0.95	0.7
2d	$5.01\ 10^{-3}$	5.0	5.6	-	0.0	1.2	2.1	0.7	0.94	1.3
3a	$4.03\ 10^{-2}$	2.1	3.2	-	0.2	1.6	1.8	0.2	0.93	1.4
3b	$6.72\ 10^{-2}$	1.5	2.2	-	0.1	1.1	1.1	0.1	0.94	1.0
3c	$3.07\ 10^{-2}$	2.2	3.0	-	0.1	1.5	1.3	0.4	0.94	0.6
3d	$7.58\ 10^{-3}$	4.4	5.0	-	0.0	1.2	2.0	0.5	0.95	1.3
4a	$2.33\ 10^{-2}$	3.2	4.1	-	0.3	1.7	1.9	0.3	0.92	1.4
4b	$7.99\ 10^{-2}$	1.6	2.3	-	0.1	1.1	1.1	0.1	0.93	1.2
4c	$5.29\ 10^{-2}$	1.9	2.4	-	0.1	1.0	1.0	0.2	0.94	0.6
4d	$1.23\ 10^{-2}$	3.9	4.5	-	0.1	0.6	2.0	0.2	0.96	1.1
5b	$4.81\ 10^{-2}$	2.3	2.8	-	0.1	0.9	1.2	0.1	0.93	0.9
5c	$9.49\ 10^{-2}$	1.6	2.1	-	0.0	0.9	1.0	0.1	0.94	0.5
5d	$6.63\ 10^{-2}$	1.9	2.6	-	0.1	1.2	1.3	0.4	0.94	0.9
6d	$2.28\ 10^{-1}$	4.4	5.5	-	0.0	3.0	1.1	0.6	0.94	1.6

Table A.13: Normalised double differential dijet cross sections measured as function of Q^2 and ξ. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Normalised trijet cross section in bins of Q^2 and $\langle P_{\mathrm{T}} \rangle$

Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1α	$1.18\ 10^{-2}$	3.1	4.6	-	0.1	2.6	2.2	0.3	0.81	2.4
1β	$6.57\ 10^{-3}$	4.0	6.0	-	0.1	3.4	2.9	0.3	0.85	2.2
1γ	$1.03\ 10^{-3}$	10.1	13.1	-	0.1	7.4	3.9	0.5	0.87	2.5
2α	$1.25\ 10^{-2}$	3.4	4.3	-	0.1	1.5	2.2	0.2	0.79	2.1
2β	$6.85\ 10^{-3}$	4.3	5.7	-	0.1	2.2	3.0	0.2	0.85	2.5
2γ	$1.32\ 10^{-3}$	9.9	11.9	-	0.0	5.1	4.1	1.1	0.88	2.5
3α	$1.44\ 10^{-2}$	3.4	4.3	-	0.2	1.7	2.0	0.2	0.80	2.4
3β	$9.53\ 10^{-3}$	4.0	5.4	-	0.2	2.1	2.8	0.3	0.85	2.5
3γ	$2.01\ 10^{-3}$	8.8	10.6	-	0.0	3.4	4.8	0.4	0.87	2.7
4α	$1.64\ 10^{-2}$	3.6	4.5	-	0.1	2.1	1.7	0.2	0.78	2.2
4β	$1.24\ 10^{-2}$	4.0	6.3	-	0.2	3.9	2.9	0.2	0.86	2.5
4γ	$2.85\ 10^{-3}$	8.4	10.1	-	0.1	3.4	4.3	0.6	0.86	2.8
5α	$1.86\ 10^{-2}$	3.7	6.4	-	0.0	4.9	1.7	0.2	0.78	2.4
5β	$1.57\ 10^{-2}$	3.7	7.6	-	0.0	6.1	2.4	0.4	0.84	2.7
5γ	$4.10\ 10^{-3}$	7.1	12.0	-	0.4	8.9	3.7	0.5	0.86	2.6
6β	$1.79\ 10^{-2}$	12.5	16.0	-	0.1	9.8	1.6	0.6	0.83	2.7
6γ	$9.39\ 10^{-3}$	16.5	22.7	-	0.0	15.3	2.9	0.8	0.86	3.2

Table A.14: Normalised double differential trijet cross sections measured as function of Q^2 and $\langle P_{\mathrm{T}} \rangle$. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Normalised trijet cross section in bins of Q^2 and ξ										
Bin label	Norm. cross section	Stat. unc. (%)	Total unc. (%)	Norm. unc. (%)	Bkgd. unc. (%)	Model. unc. (%)	HFS. unc. (%)	Elec. unc. (%)	Had. corr.	Had. unc. (%)
1A	$9.01\ 10^{-3}$	3.6	5.7	-	0.2	3.3	2.9	0.4	0.82	2.2
1B	$8.53\ 10^{-3}$	3.6	4.8	-	0.0	2.4	2.1	0.3	0.81	2.4
1C	$3.14\ 10^{-3}$	5.8	7.9	-	0.2	4.6	2.7	0.2	0.81	2.4
2A	$8.97\ 10^{-3}$	4.0	5.3	-	0.1	2.0	2.8	0.2	0.80	2.3
2B	$9.64\ 10^{-3}$	3.7	4.7	-	0.1	1.6	2.3	0.2	0.81	2.4
2C	$3.69\ 10^{-3}$	5.8	7.6	-	0.1	4.0	2.8	0.3	0.81	2.4
3A	$8.95\ 10^{-3}$	4.4	5.6	-	0.3	1.8	2.9	0.3	0.80	2.1
3B	$1.33\ 10^{-2}$	3.4	4.7	-	0.1	2.3	2.2	0.2	0.81	2.5
3C	$5.60\ 10^{-3}$	5.2	6.5	-	0.1	2.9	2.7	0.5	0.81	2.7
4A	$7.74\ 10^{-3}$	5.4	7.0	-	0.3	3.3	2.9	0.3	0.80	2.3
4B	$1.68\ 10^{-2}$	3.5	4.9	-	0.1	2.7	2.0	0.2	0.81	2.4
4C	$8.89\ 10^{-3}$	4.7	6.9	-	0.1	4.4	2.6	0.2	0.82	2.6
5B	$1.58\ 10^{-2}$	3.9	6.9	-	0.0	5.3	2.1	0.2	0.80	2.4
5C	$2.18\ 10^{-2}$	3.2	7.7	-	0.1	6.7	2.2	0.4	0.81	2.8
6C	$4.64\ 10^{-2}$	8.0	15.4	-	0.0	13.1	1.8	0.4	0.80	3.1

Table A.15: Normalised double differential trijet cross sections measured as function of Q^2 and ξ. The bin labels of column 1 are defined in table A.1. Other details are given in the caption of table A.2.

Appendix B

List of Used Acronyms

BBE Backward Barrel, Electromagnetic (calorimeter wheel)
BCDMS Bologna-Cern-Dubna-Munich-Saclay
BGF Boson Gluon Fusion
BPC Backward Proportional Chamber
BST Backward Silicon Tracker
CC charged current
CDM Colour Dipole Model
CB Central Barrel (calorimter wheel)
CIP Central Inner Proportional Chamber
CJC Central Jet Chamber
COP Central Outer Proportional Chamber
COZ Central Outer z-Chamber
CST Central Silicon Tracker
CTD Central Track Detector
DESY Deutsches Elektronen Synchrotron
DIS deep-inelastic scattering
DGLAP Dokshitzer, Gribov, Lipatov, Altarelli, Parisi
DREAM Dual-Readout Module
DST Data Summary Tape
DVCS Deeply Virtual Compton Scattering
EW electroweak
EMC Electromagnetic Calorimeter
ET Electron Tagger
FB Forward Barrel (calorimeter wheel)
FMD Forward Muon Detector
FST Forward Silicon Tracker
FTD Forward Track Detector
FTT Fast Track Trigger
H1OO H1 Object Oriented Analysis Software
HAC Hadronic Calorimeter
HAT H1 Event Tag
HERA Hadron-Elektron-Ring-Anlage
HFS Hadronic Final State
IF Inner Forward (calorimeter wheel)

IP Interaction Point
IR infrared
LAr Liquid Argon
LEP Large Electron Positron Collider
LHC Large Hadron Collider
LO leading order
MC Monte Carlo (event generator)
MEPS matrix elements and parton shower
NC neutral current
NMC New Muon Collaboration
NNLO next-to-next-to-leading order
NLO next-to-leading order
OF Outer Forward (calorimeter wheel)
PETRA Positron-Elektron-Ring-Anlage
PD Photon Detector
PDF Parton Distribution Function
POT Production Output Tape
pQCD perturbative QCD
QCD Quantum Chromodynamics
QCDC QCD Compton
QED Quantum Electrodynamics
QEDC QED Compton
QPM Quark Parton Model
RGE Renormalisation Group Equation
SLAC Stanford Linear Accelerator Center
SM Standard Model
SpaCal Spaghetti Calorimeter
TC Tail Catcher
ToF Time-of-Flight
UV ultraviolet

References

[AA$^+$02] M. Albrow, S. Aota et al., *Intercalibration of the longitudinal segments of a calorimeter system*, Nucl. Instr. and Meth. A **487**, 381–395 (2002).

[AA$^+$03] S. Agostinelli, J. Allison et al., *GEANT4—a simulation toolkit*, Nucl. Instr. and Meth. A **506**, 250–303 (2003).

[AA$^+$10a] F. Aaron, H. Abramowicz et al., *Combined measurement and QCD analysis of the inclusive $e^{\pm}p$ scattering cross sections at HERA*, JHEP **01**, 109 (2010).

[AA$^+$10b] P. Adragna, C. Alexa et al., *Measurement of pion and proton response and longitudinal shower profiles up to 20 nuclear interaction lengths with the ATLAS Tile calorimeter*, Nucl. Instr. and Meth. A **615**(2), 158–181 (2010).

[AB$^+$92] D. Acosta, S. Buontempo et al., *Lateral shower profiles in a lead / scintillating fiber calorimeter*, Nucl. Instr. and Meth. A **316**, 184–201 (1992).

[AB$^+$02] V. Andreev, V. Boudry et al., *The new H1 luminosity system for HERA II*, Nucl. Instr. and Meth. A **494**, 45–50 (2002).

[Abe01] T. Abe, *GRAPE-Dilepton (Version 1.1): A generator for dilepton production in ep collisions*, Comput. Phys. Commun. **136**, 126–147 (2001).

[AC$^+$05a] N. Akchurin, K. Carrell et al., *Comparison of high-energy electromagnetic shower profiles measured with scintillation and Cherenkov light*, Nucl. Instr. and Meth. A **548**, 336–354 (2005).

[AC$^+$05b] N. Akchurin, K. Carrell et al., *Hadron and jet detection with a dual-readout calorimeter*, Nucl. Instr. and Meth. A **537**, 537–561 (2005).

[AC$^+$08] N. Akchurin, K. Carrell et al., *Comparison of high-energy hadronic shower profiles measured with scintillation and Cherenkov light*, Nucl. Instr. and Meth. A **584**, 304–318 (2008).

[AD$^+$81] H. Abramowicz, J. De Groot et al., *The response and resolution of an iron-scintillator calorimeter for hadronic and electromagnetic* *showers between 10 GeV and 140 GeV*, Nucl. Instr. and Meth. **180**, 429–439 (1981).

[AG$^+$83] B. Andersson, G. Gustafson et al., *Parton fragmentation and string dynamics*, Physics Reports **97**, 31–145 (1983).

[AI$^+$77] G. Akopdjanova, A. Inyakina et al., *Determination of photon coordinates in a hodoscope Cherenkov spectrometer*, Nucl. Instr. and Meth. **140**, 441–445 (1977).

[AP77] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126**, 298–318 (1977).

[BB$^+$78] W. Bardeen, A. Buras et al., *Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories*, Phys. Rev. D **18**, 3998–4017 (1978).

[BB$^+$87] R. Brun, F. Bruyant et al., *GEANT: Simulation Program for Particle Physics Experiments. User Guide and Reference Manual*, CERN-DD/EE 84-1 (1987).

[BB$^+$93] A. Bernstein, T. Bienz et al., *Beam tests of the ZEUS barrel calorimeter*, Nucl. Instr. and Meth. A **336**, 23–52 (1993).

[BB95] U. Bassler and G. Bernardi, *On the kinematic reconstruction of deep inelastic scattering at HERA*, Nucl. Instr. and Meth. A **361**, 197–208 (1995).

[BB99] U. Bassler and G. Bernardi, *Structure function measurements and kinematic reconstruction at HERA*, Nucl. Instr. and Meth. A **426**, 583–598 (1999).

[BB$^+$06] F. Ballarini, G. Battistoni et al., *The FLUKA code: an overview*, J. Phys. Conf. Ser. **41**, 151–160 (2006).

[BB$^+$08] J. Becker, K. Bösiger et al., *A vertex trigger based on cylindrical multiwire proportional chambers*, Nucl. Instr. and Meth. A **586**, 190–203 (2008).

[BC$^+$69] E. Bloom, D. Coward et al., *High-energy inelastic ep scattering at* 6° *and* 10°, Phys. Rev. Lett. **23**, 930–934 (1969).

[BC$^+$03] J. Butterworth, J. Couchman et al., *KtJet: A C++ implementation of the* $K_\perp$ *clustering algorithm*, Comput. Phys. Commun. **153**, 85–96 (2003).

[BCDMS89] A. C. Benvenuti, D. Bollini et al. (BCDMS), *A high statistics measurement of the proton structure functions* $F_2(x, Q^2)$ *and R from deep-inelastic muon scattering at high* Q^2, Phys. Lett. B **223**, 485–489 (1989).

[BD$^+$10] R. D. Ball, L. D. Debbio et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, Nucl. Phys. B **838**, 136–206 (2010).

[Beh10] J. Behr, *Jets at High Q^2 at HERA and Test Beam Measurements with the EUDET Pixel Telescope*, Dissertation, Universität Hamburg, DESY-THESIS-2010-038, 2010.

[Bet09] S. Bethke, *The 2009 world average of α_s*, Eur. Phys. J. C **64**, 689–703 (2009).

[BF$^+$69] M. Breidenbach, J. Friedman et al., *Observed behavior of highly inelastic electron-proton scattering*, Phys. Rev. Lett. **23**, 935–939 (1969).

[BHK$^+$81] R. Bock, T. Hansl-Kozanecka et al., *Parametrization of the longitudinal development of hadronic showers in sampling calorimeters*, Nucl. Instr. and Meth. **186**, 533–539 (1981).

[Bjo69] J. Bjorken, *Asymptotic sum rules at infinite momentum*, Phys. Rev. **179**, 1547–1553 (1969).

[Blo04] V. Blobel, *Central Track Reconstruction*, Talk at H1 tracking group meeting, March 11, 2004 (available at `https://www-h1.desy.de/idet/itracker/TrackingGroup/AgnMin/Meet041103/blobel041103.pdf`) (2004).

[Blo06a] V. Blobel, *A new fast track-fit algorithm based on broken lines*, Nucl. Instr. and Meth. A **566**, 14–17 (2006).

[Blo06b] V. Blobel, *Software alignment for tracking detectors*, Nucl. Instr. and Meth. A **566**, 5–13 (2006).

[Bot11] M. Botje, *QCDNUM: Fast QCD Evolution and Convolution*, Comput. Phys. Commun. **182**, 490–532 (2011).

[Bro63] R. G. Brown, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, Englewood Cliffs, 1963.

[Bru98] P. Bruel, *Recherche d'interactions au-dela du Modele Standard a HERA*, Dissertation, Universite Paris XI, Orsay, 1998.

[BS88] M. Bengtsson and T. Sjöstrand, *Parton showers in leptoproduction events*, Z. Phys. C **37**, 465–476 (1988).

[BSC$^+$01] A. Baird, H. Schultz-Coulon et al., *A fast track trigger for the H1 Collaboration*, Nucl. Instr. and Meth. A **461**, 461–464 (2001).

[Cas74] W. Caswell, *Asymptotic behavior of non-Abelian gauge theories to two-loop order*, Phys. Rev. Lett. **33**, 244–246 (1974).

[CC^{+}91] T. Carli, A. Courau et al., *Quasireal QED Compton Monte Carlo*, Proceedings of the Workshop on Physics at HERA, eds. W. Buchmüller, G. Ingelmann, Vol. 2 , 902 (1991).

[CD^{+}93] S. Catani, Y. Dokshitzer et al., *Longitudinally invariant k_T clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406**, 187–224 (1993).

[CD10] CDF and D0, *Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with up to 6.7 fb-1 of Data*, arXiv:hep-ex/1007.4587 (2010).

[CDF08] T. Aaltonen, J. Adelman et al. (CDF), *Measurement of the inclusive jet cross section at the Fermilab Tevatron $p\bar{p}$ collider using a cone-based jet algorithm*, Phys. Rev. D **78**, 52006 (2008).

[CG69] C. G. Callan and D. Gross, *High-energy electroproduction and the constitution of the electric current*, Phys. Rev. Lett. **22**, 156–159 (1969).

[CK92] A. Courau and P. Kessler, *QED Compton scattering in high-energy electron-proton collisions*, Phys. Rev. D **46**, 117–124 (1992).

[CK^{+}97] K. Chetyrkin, B. Kniehl et al., *Strong Coupling Constant with Flavor Thresholds at Four Loops in the Modified Minimal-Subtraction Scheme*, Phys. Rev. Lett. **79**, 2184–2187 (1997).

[CK^{+}00] K. Chetyrkin, J. Kühn et al., *RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses*, Comput. Phys. Commun. **133**, 43–65 (2000).

[CS^{+}88] J. Collins, D. E. Soper et al., *Perturbative QCD*, Adv. Ser. Direct. High Energy Physics **5**, 1 (1988).

[CS^{+}94] K. Charchula, G. Schuler et al., *Combined QED and QCD radiative effects in deep inelastic lepton-proton scattering: The Monte Carlo generator DJANGO6*, Comput. Phys. Commun. **81**, 381–402 (1994).

[CS97] S. Catani and M. Seymour, *A general algorithm for calculating jet cross sections in NLO QCD*, Nucl. Phys. B **485**, 291–419 (1997).

[CS06] M. Cacciari and G. Salam, *Dispelling the N^3 myth for the k_T jet-finder*, Phys. Lett. B **641**, 57–61 (2006).

[CS^{+}08] M. Cacciari, G. Salam et al., *The anti-k_T jet clustering algorithm*, JHEP **04**, 063 (2008).

[CTEQ08] P. Nadolsky, H. Lai et al. (CTEQ), *Implications of CTEQ global analysis for collider observables*, Phys. Rev. D **78**, 013004 (2008).

[CZ99] J. Cao and Z. Zhang, *Towards an Unbiased Measurement of Kinematic Variables at Low y Region*, H1-IN-580 (1999).

[D008] V. Abazov, B. Abbott et al. (D0), *Measurement of the Inclusive Jet Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **101**, 062001 (2008).

[DCS04] R. Devenish and A. M. Cooper-Sarkar, *Deep Inelastic Scattering*, Oxford University Press, 2004.

[DG$^+$10a] A. Daleo, A. Gehrmann-De Ridder et al., *NNLO Antenna Subtraction with One Hadronic Initial State*, arXiv:hep-ph/1001.2397 (2010).

[DG$^+$10b] G. Dissertori, A. Gehrmann-De Ridder et al., *Precise Determination of the Strong Coupling Constant at NNLO in QCD from the Three-Jet Rate in Electron–Positron Annihilation at LEP*, Phys. Rev. Lett. **104**, 072002 (2010).

[DL$^+$97] Y. Dokshitzer, G. Leder et al., *Better jet clustering algorithms*, JHEP **08**, 001 (1997).

[DM$^+$08] M. Dasgupta, L. Magnea et al., *Non-perturbative QCD effects in jets at hadron colliders*, JHEP **02**, 055 (2008).

[Dok77] Y. L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, Sov. Phys. JETP **46**, 641–653 (1977).

[Dup09] A. Duperrin, *Review of searches for Higgs bosons and beyond the Standard Model Physics at the Tevatron*, Eur. Phys. J. C **59**, 297–334 (2009).

[EB64] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13**, 321–323 (1964).

[EG$^+$79] R. K. Ellis, H. Georgi et al., *Perturbation theory and the parton model in QCD*, Nucl. Phys. B **152**, 285–329 (1979).

[EM$^+$05a] M. Ellerbrock, J. Marks et al., *Description of Hadronic Showers in the Optimized Version of H1FAST*, H1-IN-623 (2005).

[EM$^+$05b] M. Ellerbrock, J. Marks et al., *Improved Description of Electromagnetic Showers in H1FAST*, H1-IN-617 (2005).

[ES93] S. Ellis and D. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48**, 3160–3166 (1993).

[ES$^+$96] R. K. Ellis, W. J. Stirling et al., *QCD and Collider Physics*, Cambridge University Press, 1996.

[Fei04] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, arXiv:physics/0402093 (2004).

[Fes85] H. C. Fesefeldt, *Simulation of hadronic showers, physics and application*, PITHA 85–02, RWTH Aachen (1985).

[Fey69] R. Feynman, *Very high-energy collisions of hadrons*, Phys. Rev. Lett. **23**, 1415–1417 (1969).

[Fey72] R. Feynman, *Photon – Hadron Interactions*, Benjamin, New York, 1972.

[FG03] C. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. **75**, 1243–1286 (2003).

[FGM+73] H. Fritzsch, M. Gell-Mann et al., *Advantages of the color octet gluon picture*, Phys. Lett. B **47**, 365–368 (1973).

[FK72] J. Friedman and H. Kendall, *Deep inelastic electron scattering*, Annual Review of Nuclear Science **22**, 203–254 (1972).

[FK06] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, Nucl. Instr. and Meth. A **559**, 190–194 (2006).

[GG+94] T. Gabriel, D. Groom et al., *Energy dependence of hadronic activity*, Nucl. Instr. and Meth. A **338**, 336–347 (1994).

[GG+05] A. Gehrmann-De Ridder, T. Gehrmann et al., *Antenna subtraction at NNLO*, JHEP **9**, 056 (2005).

[GH+64] G. Guralnik, C. Hagen et al., *Global conservation laws and massless particles*, Phys. Rev. Lett. **13**, 585–587 (1964).

[GL72] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15**, 438–450 (1972).

[Gla61] S. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22**, 579–588 (1961).

[GM64] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett **8**, 214–215 (1964).

[Gou08] M. Gouzevitch, *Mesure de la constante de couplage forte α_s avec les jets hadroniques en Diffusion Inélastique Profonde*, Dissertation, Ecole Polytechnique Palaiseau, DESY-THESIS-2008-047, 2008.

[GP00] G. Grindhammer and S. Peters, *The Parameterized Simulation of Electromagnetic Showers in Homogeneous and Sampling Calorimeters*, arXiv:hep-ex/0001020 (2000).

[GR+90] G. Grindhammer, M. Rudowicz et al., *The fast simulation of electromagnetic and hadronic showers*, Nucl. Instr. and Meth. A **290**, 469–488 (1990).

[Gro07] D. Groom, *Energy flow in a hadronic cascade: Application to hadron calorimetry*, Nucl. Instr. and Meth. A **572**, 633–653 (2007).

[GS+62] J. Goldstone, A. Salam et al., *Broken symmetries*, Phys. Rev. **127**, 965–970 (1962).

[GS08] C. Grupen and B. Shwartz, *Particle Detectors*, Cambridge University Press, 2008.

[Gus86] G. Gustafson, *Dual description of a confined colour field*, Phys. Lett. B **175**, 453–456 (1986).

[GW73] D. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30**, 1343–1346 (1973).

[GW91] L. Görlich and H. Wellisch, *Documentation of the LAr clustering*, H1-IN-204 (1991).

[H193a] B. Andrieu et al. (Calorimeter Group of H1), *The H1 liquid argon calorimeter system*, Nucl. Instr. and Meth. A **336**, 460–498 (1993).

[H193b] B. Andrieu, Y. Feng et al. (Calorimeter Group of H1), *Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations*, Nucl. Instr. and Meth. A **336**, 499–509 (1993).

[H194a] B. Andrieu, J. Ban et al. (Calorimeter Group of H1), *Electron/pion separation with the H1 LAr calorimeters*, Nucl. Instr. and Meth. A **344**, 492–506 (1994).

[H194b] B. Andrieu, E. Barrelet et al. (Calorimeter Group of H1), *Beam tests and calibration of the H1 liquid argon calorimeter with electrons*, Nucl. Instr. and Meth. A **350**, 57–72 (1994).

[H196a] R. Appuhn, C. Arndt et al. (SPACAL Group of H1), *Hadronic response and e/pi separation with the H1 lead/fibre calorimeter*, Nucl. Instr. and Meth. A **382**, 395–412 (1996).

[H196b] T. Nicholls, L. Hajduk et al. (SPACAL Group of H1), *Performance of an electromagnetic lead/scintillating-fibre calorimeter for the H1 detector*, Nucl. Instr. and Meth. A **374**, 149–156 (1996).

[H197a] I. Abt et al. (H1), *The H1 detector at HERA*, Nucl. Instr. and Meth. A **386**, 310–347 (1997).

[H197b] I. Abt et al. (H1), *The Tracking, calorimeter and muon detectors of the H1 experiment at HERA*, Nucl. Instr. and Meth. A **386**, 348–396 (1997).

[H197c] R. Appuhn, C. Arndt et al. (SPACAL Group of H1), *The H1 lead / scintillating fiber calorimeter*, Nucl. Instr. and Meth. A **386**, 397–408 (1997).

[H101a] C. Adloff, V. Andreev et al. (H1), *Measurement and QCD analysis of jet cross sections in deep-inelastic positron-proton collisions at* $\sqrt{s}$ *of 300 GeV*, Eur. Phys. J. C **19**, 289–311 (2001).

[H101b] C. Adloff, V. Andreev et al. (H1), *Three Jet production in deep inelastic scattering at HERA*, Phys. Lett. B **515**, 17–29 (2001).

[H104] A. Aktas, V. Andreev et al. (H1), *Measurement of the proton structure function* F_2 *at low* Q^2 *in QED compton scattering at HERA*, Phys. Lett. B **598**, 159–171 (2004).

[H107] A. Aktas, C. Alexa et al. (H1), *Measurement of inclusive jet production in deep-inelastic scattering at high* Q^2 *and determination of the strong coupling*, Phys. Lett. B **653**, 134–144 (2007).

[H109a] F. D. Aaron, M. Aldaya Martin et al. (H1), *Neutral Current Interactions in* $e^{\pm}p$ *Scattering with Longitudinally Polarised Leptons*, H1prelim-09-042, available at `http://www-h1.desy.de/publications/H1preliminary.short_list.html` (2009).

[H109b] F. D. Aaron, C. Alexa et al. (H1), *A general search for new phenomena at HERA*, Phys. Lett. B **674**, 257–268 (2009).

[H109c] F. Aaron, C. Alexa et al. (H1), *A precision measurement of the inclusive ep scattering cross section at HERA*, Eur. Phys. J. C **64**, 561–587 (2009).

[H110a] F. D. Aaron, M. Aldaya Martin et al. (H1), *Prompt photons in photoproduction at HERA*, Eur. Phys. J. C **66**, 17–33 (2010).

[H110b] F. D. Aaron, C. Alexa et al. (H1), *Jet production in ep collisions at low* Q^2 *and determination of* α_s, Eur. Phys. J. C **67**, 1–24 (2010).

[H110c] F. D. Aaron et al. (H1), *Jet production in ep collisions at high* Q^2 *and determination of* α_s, Eur. Phys. J. C **65**, 363–383 (2010).

[Hab09] S. Habib, *Unpolarized Neutral Current* $e^{\pm}p$ *Cross Section Measurements at the H1 Experiment, HERA*, Dissertation, Universität Hamburg, DESY-THESIS-2009-039, 2009.

[Hei99] B. Heinemann, *Measurement of Charged Current and Neutral Current Cross Sections in Positron-Proton Collisions at* $\sqrt{s}$*=300 GeV*, Dissertation, Universität Hamburg, 1999.

[Hig64] P. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13**, 508–509 (1964).

[HL00] H. Henschel and R. Lahmann, *The backward silicon tracker of the H1 Experiment at HERA*, Nucl. Instr. and Meth. A **453**, 93–97 (2000).

[HM84] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, 1984.

[Hol08] B. J. Holzer, *HERA: Lessons learned from the HERA upgrade*, Final CARE-HHH Workshop on Scenarios for the LHC Upgrade and FAIR (CERN-2009-004) , 30–35 (2008).

[Höp97] M. Höppner, *Anwendung neuronaler Netze in der Analyse hadronischer Schauer in feinsegmentierten Flüssigargon-Kalorimetern*, Dissertation, Universität Dortmund, 1997.

[HW$^+$90] J. Huth, N. Wainer et al., *Toward a standardization of jet definitions*, Research directions for the decade. Summer Study on HEP, Snowmass, Colorado (1990).

[HW97] M. Höppner and D. Wegener, *Applications of neural networks to shower analysis in a highly segmented LAr calorimeter*, Nucl. Instr. and Meth. A **389**, 154–156 (1997).

[IE$^+$97] G. Ingelman, A. Edin et al., *LEPTO 6.5–A Monte Carlo generator for deep inelastic lepton-nucleon scattering*, Comput. Phys. Commun. **101**, 108–134 (1997).

[JB79] F. Jacquet and A. Blondel, *Study of an ep facility for Europe*, ECFA Proceedings, DESY 79/48 (1979).

[JL$^+$94] J. Janoth, A. Lindner et al., *Response of mesh-type photomultiplier tubes in strong magnetic fields*, Nucl. Instr. and Meth. A **350**, 221–225 (1994).

[JR75] F. James and M. Roos, *Minuit: A system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun **10**, 343–367 (1975).

[Jun95] H. Jung, *Hard diffractive scattering in high energy ep collisions and the Monte Carlo generator RAPGAP*, Comput. Phys. Commun. **86**, 147–161 (1995).

[JZ$^+$99] M. Jacquet, Z. Zhang et al., *Absolute Hadronic Jet Calibration of the H1 Liquid Argon Calorimeter*, H1-IN-571 (1999).

[KD$^+$01] C. Kiesling, B. Denby et al., *The h1 neural network trigger project*, Advanced computing and analysis techniques in physics research **583**, 36–44 (2001).

[Kel04] N. Keller, *A Measurement of the QED Compton Cross Section in Electron-Proton Scattering with the H1 Experiment at HERA*, Dissertation, Universität Zürich, 2004.

[Kin62] T. Kinoshita, *Mass singularities of Feynman Amplitudes*, J. Math. Phys. **3**, 650–677 (1962).

[Kle07] C. Kleinwort, *H1 Alignment Experiences*, CERN report 2007-004 (2007).

[KM+89] A. Kalinovskii, N. Mokhov et al., *Passage of high-energy particles through matter*, American Institute of Physics, 1989.

[Kog10] R. Kogler, *Jet Production at Low and High Q^2 and Determination of the Strong Coupling α_s at H1*, Proceedings of the 18th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2010), PoS DIS2010:116 (2010), arXiv:hep-ex/1006.4184 (2010).

[Kos02] M. Kossov, *Approximation of photonuclear interaction cross-sections*, Eur. Phy. J. A **14**, 377–392 (2002).

[KS+92] A. Kwiatkowski, H. Spiesberger et al., *HERACLES: an event generator for ep interactions at HERA energies including radiative processes*, Comput. Phys. Commun. **69**, 155–172 (1992).

[Kuh92] M. Kuhlen, *The Fast H1 detector Monte Carlo*, Proceedings of the XXVI International Conference on High Energy Physics, vol. 2 , 1787–1790 (1992).

[LEP07] J. Alcaraz, A. Bajo-Vaquero et al. (LEP), *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:hep-ex/0712.0929 (2007).

[Lis05] B. List, *The H1 silicon tracker*, Nucl. Instr. and Meth. A **549**, 33–36 (2005).

[LN64] T. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, Phys. Rev. **133**, 1549–B1562 (1964).

[Loc92] P. Loch, *Kalibration des H1 Flüssig-Argon Kalorimeters unter Berücksichtigung der Gewichtungsmethode für Teilchenjets*, Dissertation, Universität Hamburg, 1992.

[Lön92] L. Lönnblad, *Ariadne version 4 – A program for simulation of QDC cascades implementing the colour dipole model*, Comput. Phys. Commun. **71**, 15–31 (1992).

[Lön95] L. Lönnblad, *Rapidity gaps and other final state properties in the colour dipole model for deep inelastic scattering*, Z. Phys. C **65**, 285–291 (1995).

[LR00] C. Leroy and P.-G. Rancoita, *Physics of cascading shower generation and propagation in matter: principles of high-energy, ultrahigh-energy and compensating calorimetry*, Rep. Prog. Phys. **63**, 505–606 (2000).

[LS75] E. Longo and I. Sestili, *Monte Carlo calculation of photon-initiated electromagnetic showers in lead glass*, Nucl. Instr. and Meth. **128**, 283–307 (1975).

[LS78] S. Libby and G. Sterman, *Jet and lepton-pair production in high-energy lepton-hadron and hadron-hadron scattering*, Phys. Rev. D **18**, 3252–3268 (1978).

[LS02] K. Long and K. Schüler, *Polarisation measurements on $e^{\pm}$ beams*, Nucl. Instr. and Meth. A **494**, 75–80 (2002).

[Lüs03] M. Lüscher, *Lattice QCD–from Quark Confinement to Asymptotic Freedom*, Annales Henri Poincare **4**, 197–210 (2003).

[Mai07] A. Maire, *Forward LAr electron calibration*, Talk at H1 calorimeter task force meeting, June 15, 2007 (available at `https://www-h1.desy.de/icgi-h1wiki/moin.cgi/CalorimetersAnalysisTaskForce/Meeting2007-06-15`) (2007).

[Man08] M. L. Mangano, *Understanding the Standard Model, as a bridge to the discovery of new phenomena at the LHC*, Int. J. Mod Phys A **23**, 3833–3848 (2008).

[Mav04] G. Mavromanolakis, *Quartz fiber calorimetry and calorimeters*, arXiv:physics/0412123 (2004).

[MS$^+$09] A. Martin, W. Stirling et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63**, 189–285 (2009).

[MV$^+$04] S. Moch, J. Vermaseren et al., *The three-loop splitting functions in QCD: The non-singlet case*, Nucl. Phys. B **688**, 101–134 (2004).

[Nik07] A. Nikiforov, *Measurements of the Neutral Current $e^{\pm}p$ Cross Sections Using Longitudinally Polarised Lepton Beams At HERA II*, Dissertation, LMU München, 2007.

[Nik10] A. Nikiforov, *LAr EM calibration*, Talk at H1 ELAN group meeting, February 4, 2010 (available at `https://www-h1.desy.de/icgi-h1wiki/moin.cgi/ElanInclusive/HighQ2NC-CC`, 2010-02-04) (2010).

[NMC97] M. Arneodo, A. Arvidson et al. (NMC), *Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T*, Nucl. Phys. B **483**, 3–43 (1997).

[Nož03] M. Nožička, *The forward and backward silicon trackers of H1*, Nucl. Instr. and Meth. A **501**, 54–59 (2003).

[NT98] Z. Nagy and Z. Trocsanyi, *Next-to-leading order calculation of four-jet observables in electron-positron annihilation*, Phys. Rev. D **59**, 14020 (1998).

[NT01] Z. Nagy and Z. Trocsanyi, *Multijet Cross Sections in Deep Inelastic Scattering at Next-to-Leading Order*, Phys. Rev. Lett. **87**, 82001 (2001).

[OS09] S. Osman and D. Salek, *The Low Pt HFS and Jet Energy Calibration*, H1-IN-631 (2009).

[PB$^+$00] D. Pitzl, O. Behnke et al., *The H1 silicon vertex detector*, Nucl. Instr. and Meth. A **454**, 334–349 (2000).

[PDG10] K. Nakamura (Particle Data Group PDG), *Review of Particle Physics*, J. Phys. G: Nucl. Part. Phys. **37**, 075021 (2010).

[Pee03] M. Peez, *Recherche de déviations au Modèle Standard dans les Processus de grande énergie transverse sur le collisionneur électron-proton HERA*, Dissertation, Univ. Claude Bernard, Lyon, DESY-THESIS-2003-023, 2003.

[Per06] A. Perieanu, *The Structure of Charm Jets and the Dead Cone Effect in Deep-Inelastic Scattering at HERA*, Dissertation, Universität Hamburg, DESY-THESIS-2006-002, 2006.

[Pet92] S. Peters, *Die parametrisierte Simulation elektromagnetischer Schauer*, Dissertation, Universität Hamburg, Max-Planck-Institut für Physik, München, MPI-PhE/92-13, 1992.

[Pic95] A. Pich, *QUANTUM CHROMODYNAMICS*, arXiv:hep-ph/9505231 (1995).

[Pit09] D. Pitzl, *Tracking Group Report: DST7*, Talk at H1 software plenary, June 10, 2009 (available at `https://www-h1.desy.de/h1/iww/iminutes/talks.cm609/Pitzl.pdf`) (2009).

[Pla06] R. Plačakytė, *First Measurement of Charged Current Cross Sections with Longitudinally Polarised Positrons at HERA*, Dissertation, LMU München, MPP-2006-148, 2006.

[Pol73] H. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30**, 1346–1349 (1973).

[PP$^+$05] M. Peez, B. Portheault et al., *An energy flow algorithm for Hadronic Reconstruction in OO: Hadroo2*, H1-IN-616 (2005).

[PR80] R. D. Peccei and R. Rückl, *Energy flow and energy correlations in deep inelastic scattering*, Nucl. Phys. B **162**, 125–167 (1980).

[Qui09] C. Quigg, *Unanswered Questions in the Electroweak Theory*, Annual Review of Nuclear and Particle Science **59**, 505–555 (2009).

[Rud92] M. Rudowicz, *Hadronische Schauersimulation für den H1-Detektor*, Dissertation, Universität Hamburg, Max-Planck-Institut für Physik, München, MPI-PhE/92-14, 1992.

[RV$^+$97] T. V. Ritbergen, J. Vermaseren et al., *The four-loop β-function in quantum chromodynamics*, Phys. Lett. B **400**, 379–384 (1997).

[Sau09] M. Sauter, *Measurement of Beauty Photoproduction at Threshold using Di-Electron Events with the H1 Detector at HERA*, Dissertation, ETH Zürich, 2009.

[Sch85] P. Schmüser, *The electron-proton colliding beam facility HERA*, Nucl. Instr. and Meth. A **235**, 201–208 (1985).

[Sch98] U. Schneekloth, *Recent HERA Results and Future Prospects*, arXiv:hep-ex/9806010 (1998).

[Sch02] C. Schwanenberger, *The Jet Calibration in the H1 Liquid Argon Calorimeter*, arXiv:physics/0209026 (2002).

[Sch06] A. Schöning, *The Fast Track Trigger at the H1 experiment design concepts and algorithms*, Nucl. Instr. and Meth. A **566**, 130–132 (2006).

[Sei99] M. Seidel, *Luminosity Upgrade of HERA*, Proceedings of the 1999 Particle Accelerator Conference , 34–36 (1999).

[Shu10] S. Shushkevic, *private communication, see also: Measurement of Neutral Current Reactions in the H1 Experiment at HERA-2*, Dissertation, LMU München, in preparation, 2010.

[SM$^+$06] T. Sjöstrand, S. Mrenna et al., *PYTHIA 6.4 physics and manual*, JHEP **2006**, 026 (2006).

[Sop97] D. E. Soper, *Basics of QCD Perturbation Theory*, arXiv:hep-ph/9702203 (1997).

[Sou10] D. M. South, *High Q^2 Measurements from HERA*, arXiv:hep-ex/1005.5729 (2010).

[SS$^+$95] G. Sterman, J. Smith et al., *Handbook of perturbative QCD*, Rev. Mod. Phys. **67**, 157–248 (1995).

[SS07] G. Salam and G. Soyez, *A practical seedless infrared-safe cone jet algorithm*, JHEP **05**, 086 (2007).

[SS09] S. Schmitt and E. Sauvan, *QED Compton analysis update*, Talk at H1 physics plenary, September 17, 2009 (available at `https://www-h1.desy.de/h1/iww/iminutes/talks.20090917/Schmitt.pdf`) (2009).

[Ste78] G. Sterman, *Mass divergences in annihilation processes. I. Origin and nature of divergences in cut vacuum polarization diagrams*, Phys. Rev. D **17**, 2773–2788 (1978).

[Sud56] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **3**, 65–71 (1956).

[SW64] A. Salam and J. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13**, 168–171 (1964).

[tHV72] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44**, 189–213 (1972).

[Tol10] T. Toll, *MC@NLO for Heavy Flavour Photoproduction at HERA*, Dissertation, Universität Hamburg, 2010.

[Tra10] T. H. Tran, *Mesures de précision de la section efficace des processus courant chargé et courant neutre à grand Q^2 à HERA avec le faisceau d'électrons polarisés*, Dissertation, Universite Paris XI, Orsay, LAL10-28, 2010.

[Tsa74] Y. Tsai, *Pair production and bremsstrahlung of charged leptons*, Rev. Mod. Phys. **46**, 815–851 (1974).

[TV$^+$80] O. Tarasov, A. Vladimirov et al., *The Gell-Mann-Low function of QCD in the three-loop approximation*, Phys. Lett. B **93**, 429–432 (1980).

[Urb04] M. Urban, *The new CIP2k z-Vertex Trigger for the H1 Experiment at HERA*, Dissertation, Univ. Zürich, DESY-THESIS-2004/044, 2004.

[Vel02] C. Velkeen, *H1NonepBgFinder – Rejection of cosmic muon and beam-halo events in the H1OO framework*, H1-IN-603 (2002).

[VM$^+$04] A. Vogt, S. Moch et al., *The three-loop splitting functions in QCD: the singlet case*, Nucl. Phys. B **691**, 129–181 (2004).

[W$^+$81] B. H. Wiik et al., *HERA, A Proposal for a Large Electron Proton Colliding Beam Facility at DESY*, DESY-HERA **81/10** (1981).

[War50] J. Ward, *An identity in quantum electrodynamics*, Phys. Rev. **78**, 182–182 (1950).

[Web84] B. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238**, 492–528 (1984).

[Web93] B. Webber, *Factorization and jet clustering algorithms for deep-inelastic scattering*, J. Phys. G: Nucl. Part. Phys. **19**, 1567–1575 (1993).

[Wei67] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19**, 1264–1266 (1967).

[Wei73] S. Weinberg, *New Approach to the Renormalization Group*, Phys. Rev. D **8**, 3497–3509 (1973).

[Wes00] L. West, *How to use the Heavy Flavour Working Group Track, Muon and Electron Selection Code*, H1 Software Manual (available at https://www-h1.desy.de/icas/imanuals/h1phan/track_manual30106.ps) (2000).

[Wig87] R. Wigmans, *On the energy resolution of uranium and other hadron calorimeters*, Nucl. Instr. and Meth. A **259**, 389–429 (1987).

[Wig00] R. Wigmans, *Calorimetry: energy measurement in particle physics*, Oxford University Press, 2000.

[Wig02] R. Wigmans, *Sampling calorimetry*, Nucl. Instr. and Meth. A **494**, 277–287 (2002).

[Wil74] K. Wilson, *Confinement of quarks*, Phys. Rev. D **10**, 2445–2459 (1974).

[WK$^+$94] H. Wellisch, J. Kubenka et al., *Hadronic Calibration of the H1 LAr Calorimeter using Software Weighting Techniques*, H1-IN-346, MPI-PhE/94-03 (1994).

[Wob00] M. Wobisch, *Measurement and QCD Analysis of Jet Cross Sections in Deep-Inelastic Positron-Proton Collisions at $\sqrt{s} = 300$ GeV*, Dissertation, RWTH Aachen, DESY-THESIS-2000-049, 2000.

[WW99] M. Wobisch and T. Wengler, *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*, Proceedings of the HERA Monte Carlo Workshop, eds. G. Grindhammer, G. Ingelman, H. Jung, T. Doyle, DESY-PROC-02-1999 , 270–280 (1999).

[YM54] C. Yang and R. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Phys. Rev. **96**, 191–195 (1954).

[Yud69] T. Yuda, *Electron-induced cascade showers in inhomogeneous media*, Nucl. Instr. and Meth. **73**, 301–312 (1969).

[ZEUS05a] S. Chekanov, M. Derrick et al. (ZEUS), *Multijet production in neutral current deep inelastic scattering at HERA and determination of α_s*, Eur. Phys. J. C **44**, 183–193 (2005).

[ZEUS05b] S. Chekanov, M. Derrick et al. (ZEUS), *An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment*, Eur. Phys. J. C **42**, 1–16 (2005).

[ZEUS07a] S. Chekanov, M. Derrick et al. (ZEUS), *Inclusive-jet and dijet cross sections in deep inelastic scattering at HERA*, Nucl. Phys. B **765**, 1–30 (2007).

[ZEUS07b] S. Chekanov, M. Derrick et al. (ZEUS), *Jet-radius dependence of inclusive-jet cross sections in deep inelastic scattering at HERA*, Phys. Lett. B **649**, 12–24 (2007).

[ZEUS08] S. Chekanov, M. Derrick et al. (ZEUS), *Angular correlations in three-jet events in ep collisions at HERA*, arXiv:hep-ex/0808.3783 (2008).

[ZEUS10a] H. Abramowicz, I. Abt et al. (ZEUS), *Inclusive dijet cross sections in neutral current deep inelastic scattering at HERA*, arXiv:hep-ex/1010.6167, to be published in Eur. Phys. J C, DESY-10-170 (2010).

[ZEUS10b] H. Abramowicz, I. Abt et al. (ZEUS), *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**, 127–137 (2010).

Acknowledgements / Danksagung

Zuallererst möchte ich meinem Doktorvater Günter Grindhammer danken, der diese Arbeit erst ermöglichte. Lieber Günter, Du hast meine Liebe zur Teilchenphysik geweckt und ich empfinde es als große Bereicherung Dich als Mentor zu haben. Es war mir eine große Freude von Deinem reichen physikalischen Wissen zu lernen, aber auch Deine Gesellschaft als Freund und Kollege schätze ich sehr und ich freue mich schon auf unsere weitere Zusammenarbeit.

Ich bin Robert Klanner sehr dankbar, dass der sich bereit erkärt hat Gutachter dieser Dissertation zu sein. Besonders gefreut habe ich mich über die schnelle Durchsicht und die hilfreichen Kommentare und Anmerkungen. Außerdem möchte ich mich bei Hannes Jung und Johannes Haller für die Beurteilung der Disputation bedanken.

Die Hilfe und Unterstützung die ich im Laufe meiner Arbeit von meinen Kollegen und Kolleginnen in der H1 Kollaboration bekam, war mir von großem Wert. Es ist nicht möglich euch allen hier zu danken, deswegen ein herzliches “Thank you, Heinz!” an die H1 Kollaboration. Ganz besonders möchte ich allerdings Stefan Schmitt danken, der mir bei der χ^2 Minimierung mit Rat und Tat zur Seite stand und mir auch seinen Computercode dafür zur Verfügung stellte.

Ahoj Juraj, who taught me the art of tea making and consumption, with our meaningful discussions not always being about physics and computing. Adil, ohne Dich wäre ich noch immer ein `C++` Analphabet und Du hast mit jedem Besuch in unserem Büro nicht nur Ablenkung, sondern auch Freude gebracht. Meine Bürokollegen Zuzka, Andrej, Biljana, Aziz, Daniel und Stas haben mit so einigen vergnüglichen Plaudereien bei Kaffee oder Tee eine angenehme Arbeitsatmosphäre geschaffen. Aber auch für die vielen wertvollen Diskussionen und eure Unterstützung bin ich sehr dankbar. Ganz besonderer Dank geht an Daniel für seine Hilfe bei den NLO Berechnungen. Dr. Shiraz, without you I would still be Mr. Roman and wonder how to calculate my trigger efficiencies. Boy, thank you for all your help and don’t stop limin’ with me.

Ein Hoch auf Dave, Michael und Philipp alias die Software-Gang, die mir zwar immer wieder neue Arbeit geschafft hat, aber mir auch immer den Rücken frei hielt wenn es brenzlig wurde. Dear Dave, even though you can’t beat me at Badminton, you can still fire me (especially after reading the first part of this sentence). I’m very grateful for your support and for reading parts of the manuscript, especially when you did not really have the time to do it. Lieber Michael, es freut mich sehr, dass Du es bist, mit dem ich gemeinsam die Software Gruppe leite. Es macht mir großen Spaß mit Dir zu arbeiten und bei jeder Änderung von H1OO warte ich gespannt ob Du all die Fehler findest, die ich eingebaut habe. Ich kann es kaum erwarten, mit Dir endlich die “very, very, very final” Version von H1OO herauszubringen. Danke Philipp, für all die spannenden 2×45 Minuten die wir gemeinsam verbracht haben. Genauso habe ich die Zeit genossen die wir einfach nur “rumhängten” oder uns bei eigenen sportlichen Aktivitäten von der Arbeit erholten.

Ganz besonderer Dank geht an Dich, meine liebe Jennifer! Du hast mir immer Rückhalt gegeben und an mich geglaubt, auch wenn ich es selbst nicht tat. Du hast es immer wieder geschafft mich aufzumuntern und meine Gedanken in eine neue Richtung zu lenken. Ich danke Dir für jeden gemeinsamen Tag!

Liebe Mutter, ich möchte Dir dafür danken, dass Du mich während meines Studiums und auch während meiner Zeit in Hamburg immer unterstützt hast. Du hast nicht nur die Voraussetzungen für diese Arbeit geschaffen, sondern bist mir auch währenddessen zur Seite gestanden.
All meinen Freunden und Verwandten möchte ich danken dass sie mich nie aufgaben, auch wenn ich während des Schreibens dieser Arbeit schon als verschollen galt. Egal ob ihr in Österreich, Hamburg oder sonstwo seid, die gemeinsame Zeit mit Euch und auch all die interessierten Fragen über meine Arbeit haben mir sehr geholfen.