ISBN 978-91-628-7556-5
LUNFD6/(NFFL-7227) 2008

Multiple Parton Interactions in Deep Inelastic ep-scattering at HERA

Thesis submitted for the degree of
Doctor of Philosophy
by

Sakar Osman

LUND
UNIVERSITY

DEPARTMENT OF PHYSICS
LUND, 2008

Organization **LUND UNIVERSITY** Department of Physics Lund University Box 118 SE-221 00 Lund SWEDEN	**Document name** **DOCTORAL DISSERTATION** **Date of issue** 30/10, 2008 **CODEN** LUNFD6/(NFFL-7227) 2008
Author(s) Sakar Osman	**Sponsoring organization**

Title and subtitle
Multiple Parton Interactions in Deep Inelastic ep-scattering at HERA

Abstract

The production of jets with low transverse momenta (mini-jets) in deep inelastic electron-proton scattering is studied. The analyses uses data taken with the H1 detector at HERA during the years 1999 to 2000. The events are required to contain either at least one leading jet of $P_T > 5$ GeV (the inclusive 1-jet sample) or at least two hard jets where one of them has to be at an angle larger than 140 degrees with respect to the leading jet (inclusive 2-jet sample). Mini-jet multiplicities and their average transverse momenta are presented as a function of Q^2, in two regions of psuedo-rapidity and for two bins in the hadronic mass, W for the inclusive 1-jet sample. For the inclusive 2-jet sample the results are shown for direct and resolved photon interactions in two bins of W. The results are compared to various QCD based models. A new method for calibrating jet energy measurements up to 10 GeV has been developed and its performance has been studied.

Key words:
Multiple, MPI, Mini-jets, Jets, Dijets, Parton Dynamics, Jet Calibration, Low-x, HERA, H1

Classification system and/or index terms (if any)

Supplementary bibliographical information:		**Language** English
ISSN and key title:		**ISBN** 978-91-628-7556-5
Recipient's notes	**Number of pages** 203	**Price**
	Security classification	

Distribution by (name and address)

Sakar Osman
Department of Physics
Div. of Experimental High-Energy Physics
Box 118
SE-221 00 Lund, SWEDEN

I, the undersigned, being the copyright owner of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature Sakar Osman **Date** 30/10, 2008

Contents

To my beloved family Ikhlas and Jarro,

and

to the memory of my mother
Zahida Mukhtar Ali (1939-2005)

Populärvetenskaplig Sammanfattning

Makroskopiska objekt beskrivs inom den klassiska fysiken av Newtons mekanik och den allmänna relativitetsteorin. Bland annat kan läget och rörelsemängden hos objektet bestämmas exakt i varje ögonblick. Detta är inte fallet för objekt i mikrokosmos där den klassiska mekaniken inte längre är tillämpbar. De objekt som vi pratar om i mikrokosmos är de partiklar som bygger upp den materia som vi består av och som vi ser runt omkring oss. För att beskriva sådana objekt får man istället tillgripa den relativistiska kvantmekniken, som emellertid förutsäger att läge och rörelsemängd inte är oberoende av varandra och bara kan mätas med en noggrannhet som ges av sambandet

$$\Delta x \cdot \Delta P \geq h/4\pi.$$

Här är ΔP och Δx osäkerheterna i rörelsemängd respektive läge, samt h ett tal, som kallas Planks konstant. En liknande relation gäller mellan tid och energi

$$\Delta t \cdot \Delta E \geq h/4\pi,$$

där Δt och ΔE är osäkerheterna i dessa observabler. Den senare relationen säger att energi kan skapas ur vakuum och existera under en tid som motsvarar Δt, dvs det är tillåtet att bryta fysikens lagar under detta lilla tidsrum. Partiklar som skapas under dessa förutsättningar brukar kallas *kvantfluktuationer*.

All materia är uppbyggd av atomer, som i sin tur består av protoner, neutroner och elektroner. Protonen och neutronen ansågs länge vara fundamentala beståndsdelar, men numera vet man att dessa i sin tur är uppbyggda av *kvarkar*. Elektronen däremot har liksom kvarkarna, så vitt vi vet, ingen inre struktur och de kallas därför för *punktformiga partiklar*. De fundamentala byggstenar som vi känner till idag är sex olika typer av kvarkar och lika många *leptoner*, varav elektronen är en. Byggstenarna hålls samman genom inverkan av naturkrafterna. Inom den gängse teorin, som kallas för *Standardmodellen*, anses krafterna verka genom utbyte av så kallade *kraftförmedlarpartiklar*. Således hålles t.ex. kvarkarna samman genom den starka kraften som förmedlas av *gluoner*. Den starka kraften har den speciella egenskapen att kvarkarna inte kan bryta sig ur den partikel de bygger upp, vilket har benämnts *innestängdhet*. Om en kvark ges hög rörelsemängd så att den avlägsnar sig från de andra kvarkarna så kommer gluonfältet mellan dem att brytas upp och nya partiklar bestående av kvarkar kommer att bildas. Vid tillräckligt höga rörelsemängder kommer dessa nybildade partiklar att bilda kollimerade partikelflöden, s.k. *partikeljettar*.

HERA acceleratorn vid forskningslaboratoriet DESY i Hamburg konstruerades för att studera egenskaperna hos kvarkarna och gluonerna inuti protonen. Metoden går ut på att kollidera elektroner och protoner vid höga energier, vid vilka den punktformiga elektronen utnyttjas som en sond för att titta in i protonen. Den fysikaliskt riktiga beskrivningen är att elektronen växelverkar med protonen via ett utbyte av fotoner, som förmedlar den elektromagnetiska kraften. Den utbytta fotonens våglängd bestämmer hur små strukturer som kan observeras inuti protonen. Små strukturer i detta sammanhang är ekvivalent med partoner som bär på små bråkdelar av protonens rörelsemängd. Ju högre kollisionsenergier desto mindre strukturer kan observeras. Resultat från HERA visar att protonen är en komplicerad soppa av kvarkar och gluoner och tätheten av partonerna blir större då protonen studeras med allt högre upplösning. Förutom de tre kvarkar (*valenskvarkarna*) som ger protonen dess egenskaper, så innehåler den även de gluoner som kontinuerligt utbytes mellan kvarkarna. Dessutom kan gluonerna fluktuera till kvark-antikvark par, vilka kallas *sjökvarkar*. Med högre upplösning hos den utbytta fotonen kommer en större andel av sjökvarkarna att kunna studeras.

På samma sätt som en gluon spontant kan bilda ett kvark-antikvark par, så kan också en foton spontant bilda ett kvark-antikvark par så att växelverkan under vissa omständigheter sker mellan en parton i fotonen och en parton i protonen genom gluonutbyte. Eftersom tätheten av partoner är stor vid höga kollisionsenergier så blir sannolikheten för att flera partoner växelverkar per kollision hög. Ju högre kollisionsenergier desto mer betydelsefulla blir dessa så kallade multipla partonväxelverkningar. En teoretisk beskrivning av växelverkningsprocesser mellan elektroner och protoner är generellt mycket komplicerad och blir än mer komplicerad i fallet då man måste ta hänsyn till flera växelverkningar i samma process. För närvarande är den teoretiska förståelsen för dessa processer dålig och därför är det viktigt att experimentalisterna förser teoretikerna med mätresultat som kan användas för att utveckla bättre modeller.

Denna analys har sökt efter signaler, som kan förväntas i en situation med multipla växelverkningar. Principen har varit att isolera de slutprodukter i form av partikeljettar, som bildas från den hårda spridningen mellan två partoner, vilka förväntas kunna beskrivas med de standardmodeller som utvecklats och finns tillgängliga för jämförelser mellan experimentella data och teoretiska förutsägelser. Områden av fasrummet, som inte antas påverkas i någon stor utsträckning av den hårda processen har därefter undersökts med avseende på partikeljettar. Analysen visar på avvikelser mellan det som modellerna förutsäger och resultaten från mätningarna, i dessa område skulle kunna förklaras med spridningsprocesserna som omfattar mer än en partonväxelverkan. En definitiv slutsats angående ursprunget till de jetaktiviteter som observerats, utöver dem som förväntas från den hårda processen, är svår att dra eftersom de existerande modellerna lämnar visst utrymme för andra tolkningar. Resultaten av analysen ger emellertid värdefull information som kan användas till att förbättra beskrivningen av dessa komplicerade spridningsprocesser och därmed ge en ökad förståelse för den grundläggande fysiken bakom multipla växelverkningar.

Introduction

In the HERA collider at DESY in Hamburg, electrons and protons are collided at high energies. The electrons are used as pointlike probes to investigate the proton constituents, the quarks and gluons. The scattering of the electrons against the partons of the proton proceeds via either electromagnetic or weak interaction through the exchange of virtual photons or the weak vector bosons, Z^0 or $W^{\pm}$. Since the partons can not exist as free particles and can therefore not be measured directly by the experiments, information about the dynamics of the partons has to be obtained from particle jets produced by the partons in the scattering process. Quantum Chromo Dynamics (QCD) is the theory of strong interaction, within which QCD processes are calculated using perturbative expansions in the coupling constant. Processes, where the exchanged boson is strongly virtual, are called Deep Inelastic Scattering (DIS), subdivided into neutral and charged current processes dependent on whether the exchanged boson is neutral or charged. The details of the proton structure, that can be probed, is related to the resolution power, or equivalently the four-momentum, of the exchanged boson. At low resolutions only partons carrying a large fraction of the proton momentum (valence quarks) can be investigated, whereas at high resolution mainly the sea-quarks are probed. The exchanged photon may fluctuate into quark-antiquark pairs and under certain kinematic conditions the scattering will take palce between the partons of the photon and the partons of the proton, similar to hadron-hadron scattering.

Among the most spectacular early observations at HERA was the strong rise of the proton structure function, $F_2(x, Q^2)$, which means that when the proton is probed at high resolution the density of partons is high and consequently each parton only carries a small fraction of the proton momentum, x. A class of events with a large rapidity gap close to the proton direction, produced by diffractive scattering was also observed, constituting about 10% of the total event sample. Since these early observations, the structure function of the proton has been measured at HERA to high precision in a wide range of the tranfered momentum squared, Q^2, of the exchanged photon and the fractional proton momentum, x. Also diffractive scattering has been studied extensively.

Although it is possible to describe the F_2 data by the exchange of a single gluon ladder, rapidity gaps can only be produced if the scattering takes place via the exchange of a system forming a colour singlet state, called the pomeron. Multi-gluon exchange, where a pair of gluons is the minimum to create a colour singlet, can describe diffractive scattering and can consequently be regarded as the new definition in QCD of the mysterious pomeron. At very high parton densities the probability of having more than one interaction per collision, between the partons of the photon and the proton, becomes significant, which means that more than one gluon ladder is exchanged in the scattering process, similar to the case in diffractive processes. Thus, it is reasonble to assume that Multiple Parton Interaction (MPI) and diffractive processes are related in some way.

Via a QCD reformulation of the so called Abramovasky- Griboov- Kanchelli (AGK) cutting rules a natural connection between saturation, diffractive processes and MPI is obtained by performing different cuts through the ladder diagrams in a multi-pomeron exchange. The contributions from MPI has turned out to be substantial in hadron collision at the TEVATRON energy scale. At the LHC the scattering will take place at even smaller x-values, where the parton density is extremely high and thus the influence of MPI will be even larger. So far MPI is theoretically not well understood and described. One reason is that existing QCD models rely upon the DGLAP parton evolution, which is not expected to be valid in the small x region. Another reason is that the exchange of several gluon ladders makes the description of the scattering process more involved, especially since it contains several scales. Direct evidence of MPI from HERA has been obtained in photoproduction, $Q^2 = 0$, whereas the present work is the first study of MPI in DIS. The modelling of MPI in DIS has the additional complication that not only the PDF of the proton but also of the photon has to be known, and so far the precision in the measurements of the photon PDF is not to the same precision as that of the proton.

The strategy of the analysis presented in this thesis is to identify the jet(s) of the hard scattering and define angular regions in azimuth, which are mainly populated by these. The regions in azimuth transverse to the hard scattering products are then investigated for additional jet activities, where the influence of the hard scattering is expected to be small.

The content of the thesis is structured as follows. In Chapter 1 a short description of QCD is given together with an overview of DIS, phenomenological QCD models and MPI. The HERA collider and the H1 detector are presented in Chapter 2. The event reconstruction and the event selection are explained in Chapter 3. In Chapter 4 a new method for calibrating measurements of jet energies in the region $P_T < 10$ GeV is described and the results from tests are presented. The physics analysis concerning MPI in DIS is discussed in Chapter 5 and the results are presented, followed by an interpretation of the results based on comparisons to various models.

Chapter 1

Theoretical Framework

In this chapter a short introduction to the theory of the strong force, Quantum Chromo Dynamics (QCD), is given. The perturbative and non-perturbative treatments of the strong interaction are discussed. Deep Inelastic Scattering (DIS) is introduced and different approaches to modeling DIS are presented. Finally, an overview of Multiple Parton Interaction (MPI) is given.

1.1 Quarks and Gluons

The concept 'atom', meaning indivisible, was introduced by the Greek in about 400 B.C. as the smallest constituent of matter. It took until 1897 before J.J. Thompson [1] proved that atoms are not indivisible by showing that electrons were produced when an electric field was applied across the electrodes in a cathode ray tube. In 1911 Ernest Rutherford [2, 3] used α-particles to scatter against a thin gold foil and discovered that essentially the entire mass of the atom was restricted to a very small volume inside the atom. This discovery led to the atomic model of Niels Bohr and it was realised that the nucleus of the atom must contain positively charged particles, protons. In 1932 James Chadwick [4] discovered a new particle with no charge and with a mass close to the proton mass, the neutron. The neutron provided the explanation to why, for example, helium is about four times as heavy as hydrogen and not just twice as heavy, as could be assumed if the nucleus contained only protons. It was not clear at that time whether the protons and neutrons are point-like objects or not. During the 1950s and 1960s, a variety of particles, called hadrons, were discovered in scattering experiments at particle accelerators and they were initially regarded as fundamental particles. Around 100 new particles were discovered during that period. This large number was in contradiction to the general belief that the number of fundamental consituents should be small. With the introduction of *quarks* in 1964 by Murray Gell-Manng [5] and George Zweig [6], it was possible to group the hadrons into multiplets, based on an SU(3) symmetry group, where the members of each multiplet have similar properties. Three quark flavours, the up-, down- and strange-quarks (u, d, s), with the electric charges 2/3, -1/3 and -1/3, were initially needed to account for all hadrons known at that time. The fundamental representation of the special unitary group SU(3) is thus the u, d, s-quarks or $\overline{u}, \overline{d}, \overline{s}$-antiquarks, from which the different hadron

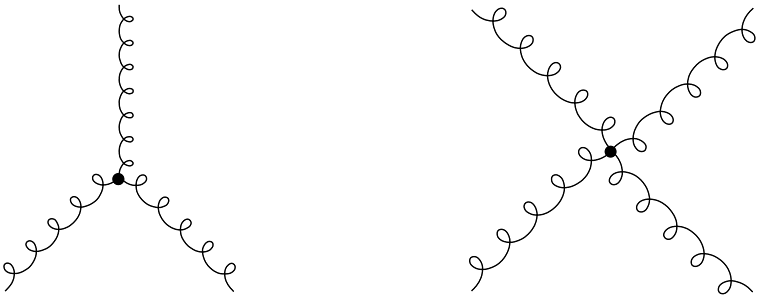


Figure 1.1: *Diagrams showing the self coupling between three and four gluons.*

multiplets can be constructed. The hadrons could be subdivided into *mesons* which were assumed to consist of a quark-antiquark pair and *baryons* described as a system of three quarks. The mesons could thus be assigned to either spin 0 or spin 1 flavour octets, and the baryons were included into either a spin 1/2 octet or a spin 3/2 decuplet. The quark content of the proton is *uud* and of the neutron *udd*. Since the quarks are fermions (spin = 1/2) they have to obey the Pauli exclusion principle. However, some of the baryons like the Δ^{++} with spin 3/2 have the same spin directions of all three u-quarks and would violate the exclusion principle. In the middle of 1960s, Oscar W. Greenberg [7], Yoichiro Nambu [8, 9] and Moo-Young Han [9] proposed that quarks carry an additional degree of freedom (quantum number), which later was called *colour*. Three colours, *red*, *green* and *blue*, had to be introduced in order to circumvent the problem with the Pauli exclusion principle. The combination of colour-anticolour and three different colours (or anticolours) both give colour neutral states, which is consistent with the fact that no experimental evidence of coloured particle states exists.

After the introduction of the quarks extensive effort was invested by the experimentalists in finding free quarks, but all attempts were in vain; no free quarks were found and still have not been found. An experiment at the Stanford Linear Accelerator Centre (SLAC) in 1968, where electrons were scattered against protons in deep inelastic processes, gave the first experimental evidence for a proton substructure. The constituents of the proton were called *partons* and they were later identified with the quarks. In order to explain the result from the SLAC experiment Richard P. Feynman, J. Bjorken and E. Paschos proposed, in 1969, that the electron interacts, via a photon exchange, with one of the three loosely bound point-like partons. The model was called the Quark Parton Model (QPM). This model predicted *scaling* which means that the cross section for scattering against a point-like object should not depend on the square of the momentum transferred by the exchanged photon, Q^2. As deep inelastic scattering experiments were extended to cover a larger Q^2 range, it was observed that scaling is violated. This scaling violation could be explained by the existence of the *gluon*, which is the mediator of the strong force.

The Standard Model (SM) of particle physics is the theoretical framework that describes three of the four known fundamental interactions between the elementary particles: the electromagnetic, the weak and the strong forces. Models describing the forces as an exchange of force mediating particles (*bosons*) combine quantum mechanics and special relativity into a *quantum field theory*. Such theories were developed between 1954 and 1973. Perturbation calculations of for example a strong interaction process can be performed using Feynman diagrams, which describe all possible interactions between quarks and gluons leading to a specific final state, once the initial state has been defined. The Feynman rules, which have to be applied to such calculations can be derived from a Lagrangian density, which describes the interaction between coloured spin 1/2 quarks of mass m with massless spin 1 gluons. The various terms in the Lagrangian can be associated with a set of *propagators* and interaction *vertices*.

The gluons are the quanta of the colour field that binds the quarks together into hadrons. The exchange of a gluon between two quarks may change the colour of the quarks due to the fact that the gluons themselves carry a charge of colour-anticolour. Since there are three colours, the group of colour transformations is represented by an SU(3) symmetry group, which equivalently to the flavour states of a quark-antiquark pair, would lead to an octet of colour charge-anticharge combinations, i.e. there are eight gluons with different colour-anticolour content. The ninth combination is a colour singlet state which can not couple to colour charge. The first experimental evidence for the existence of gluons was presented in 1979 by experiments at the electron-positron collider PETRA [10] at DESY in Hamburg. By analogy with electromagnetic interactions, described by Quantum Electrodynamics (QED), where photons are exchanged between particles carrying electric charge, the strong interaction acts through the exchange of gluons between particles carrying colour charge. Consequently the theory describing the strong force can be constructed according to the same principles as QED and thus the theory of strong interactions has been named Quantum Chromo Dynamics (QCD) [11–13]. The fact that the photon has zero mass means that the electromagnetic interaction has an infinite range, which can easily be understood from the uncertainty principle of Heisenberg. It is well known that the strong force has a very short range, of the order of the extension of the nucleon (10^{-15} m), in spite of the fact that the gluon is massless. The explanation for the different behaviours of the two forces is given by the possibility for gluons to interact with (couple to) other gluons, since they carry coulour charge themselves, whereas photons can not interact with other photons, since the photon has no electric charge. The interaction between gluons is illustrated in Figure 1.1. The consequence of the gluon self-coupling is that, in contrast to the electromagnetic field, the colour field does not spread out in space as the interacting particles are moving apart. Instead it is confined to a colour tube, giving a constant force between the interacting quarks as they are separated. Thus, the quarks will notice the strong force at large distances, which prevents them from escaping the nucleon. This is called *confinement*. On the other hand, the quarks can essentially be treated as free particles when they are close together, which is called *asymptotic freedom*. Since short distances can be probed at high energies and large distances are probed at lower energies it has been found that the coupling strength of the strong force decreases with the energy of the probe.

Although perturbative methods can be used to describe the interactions between quarks and gluons at small distances, where the quarks can be regarded as free particles and thus the coupling strength is small, this is not true at the scale of hadrons ($\lesssim$ 1 GeV), where the coupling constant, α_s, is of the order of unity and perturbative expansions do not converge. The lattice formulation of QCD provides a non-perturbative tool for calculating the hadronic spectrum and matrix elements. Lattice QCD has also been used to demonstrate the confinement of quarks, which can not be understood perturbatively. In this theory QCD is formulated on a discrete space-time grid , which can be used to perform computerized calculations. The only tunable parameters entering into the theory are the strong coupling constant and the bare quark masses. Detailed predictions of quantities that depend on the strong coupling constant and the quark masses would then be possible. An alternative method of treating hadronic interactions at low energies is given by the theory of chiral perturbation theory, which relies on the symmetries of QCD alone, to construct an effective field theory. Chiral symmetry is a symmetry of QCD in the limit of vanishing quark masses. However, in comparison to hadronic scales the masses of the lightest quarks are small and the chiral symmetry may be considered as an approximate symmetry of strong interaction. Spontaneous breaking of the chiral symmetry predicts the existence of massless Goldstone bosons, which can be identified with the lightest pseudoscalar mesons.

1.2 Perturbative QCD

As long as the quarks can be regarded as free particles inside the hadrons, i.e. as long as they are probed at high enough energies, their interactions can be calculated analytically from QCD using perturbation theory (pQCD). Experimentally the cross section of a process with specific initial and final states can be measured, but the experimental information does not explain how the initial state has been turned into the final state. Theoretically, Feynman diagrams are used to write down the various possibilities for the two particles to interact via the exchange of virtual force mediators. The virtual force mediators are produced in a quantum fluctuation process and exist only for a short period of time, given by Heisenberg's uncertainty relation. Numerous diagrams can be drawn, describing more and more complex possibilities. The most basic diagram is the one with the smallest number of coupling vertices between the virtual force mediator and the interacting particles. This is called the lowest order diagram. More complicated diagrams are of higher orders in the coupling constant. Among the higher order diagrams are some that include quark or gluon loops as shown in Figure 1.2. In order to get the calculation correct, in principle all possible diagrams of higher orders have to be taken into account, which will, however, form an infinite power series in the coupling constant. The amplitudes for the different terms should be well behaved, which means that they should be non-divergent at high energies and for high orders in the coupling constant. It turns out that field theories, like QCD, in fact will contain divergent terms associated with integrals specified by the Feynman rules for calculating the amplitudes of the intermediate states. However, the masses and coupling constants that appear in the original Feynman diagrams are the so called 'bare' masses and coupling constants, which can never be measured. Instead we may use physical masses and coupling constants, which are the ones that are measurable, and are the same as the so called *renormalized* values. These are related to the 'bare' masses and couplings by:

$$m_{physical} = m_o + \delta m \qquad g_{physical} = g_0 + \delta g$$

Even if δm and δg are not well behaved it doesn't matter since they can not be measured anyhow.

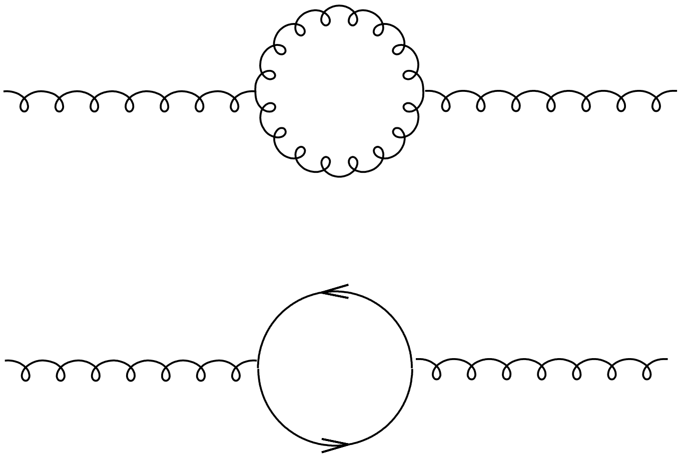


Figure 1.2: *QCD diagrams showing a gluon which fluctuates into a loop containing gluons and quarks.*

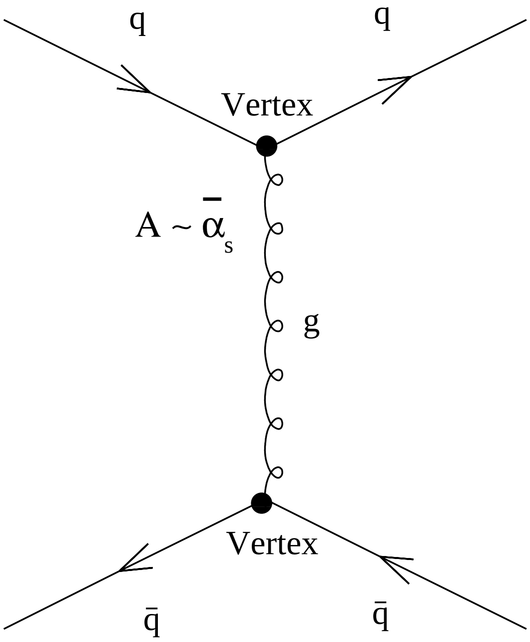


Figure 1.3: *The lowest order diagram for quark-antiquark scattering.*

1.2.1 The Strong Coupling Constant and Renormalization

So called ultraviolet divergences appear for example in the amplitude calculation of $q\overline{q} \rightarrow q\overline{q}$ scattering, where the lowest order process is described by the Feynman diagram shown in Figure 1.3. Here Q^2 is the four-momentum of the propagating gluon, which is used as the 'scale' in the evolution of the strong coupling constant, and in case Q^2 is large enough, α_s is small enough so that the perturbation expansion will converge. The coupling at the vertices of the Feynman diagram is given by the 'bare' coupling strength, $\overline{\alpha}_s$, which however is neither provided by QCD itself nor by measurements, and this is also true for the 'bare' mass. The lowest order diagram has two vertices but there are a number of diagrams in second order which have four vertices. In some of these the gluon propagator undergoes fluctuations into a quark-antiquark pair or alternatively into two gluons as illustrated in Figure 1.2. Although the four-momentum has to be conserved at each vertex, the momentum in the loop itself is unrestricted such that it can take any value between zero and infinity. Thus, for the calculation of the scattering amplitude, it is necessary to integrate over all possible momenta of the internal loop. The *regularization* and *renormalization* procedures used to deal with infinities in perturbative calculations, is discussed in some detail for QED processes by for example [14, 15]. Below only the major steps are given in the case of QCD, where the essential difference to QED comes from the gluon self coupling. Only 1-loop contributions, like the ones shown in Figure 1.2, are considered here. The integral over internal momenta of the gluon propagator can be written for large Q^2:

$$I(Q^2, \overline{\alpha}_s) = \frac{\overline{\alpha}_s}{4\pi} C_1 \int_{\Lambda^2}^{\infty} \frac{dp^2}{p^2} - \frac{\overline{\alpha}_s}{4\pi} C_1 log(\frac{Q^2}{\Lambda^2}) \tag{1.1}$$

where C_1 is a coefficient. Since the gluons are massless and the quarks are treated as massless, these loops lead to infrared divergences and consequently a lower cut-off, Λ^2, has to be introduced. The term $\int dp^2/p^2 = \log p^2$ in the integral gives rise to logarithmic divergences (ultraviolet divergences) as the momentum goes to infinity. In order to avoid this an upper cut M^2 is introduced. This is called momentum cut-off regularisation. The integral becomes:

$$I(Q^2, \overline{\alpha}_s, M) = \frac{\overline{\alpha}_s}{4\pi} C_1 log(\frac{M^2}{\Lambda^2}) - \frac{\overline{\alpha}_s}{4\pi} C_1 log(\frac{Q^2}{\Lambda^2}) = \frac{\overline{\alpha}_s}{4\pi} C_1 log(\frac{M^2}{Q^2}). \tag{1.2}$$

In general, each order n of the amplitude of the process $q\overline{q} \rightarrow q\overline{q}$ contains n loops as illustrated schematically in Figure 1.4. Their contributions are proportional to $I_n = \overline{\alpha}_s \ I(Q^2, \overline{\alpha}_s, M)^n$. If $|I_n| < 1$ the summation over all orders will converge. The relation between the physical, $\alpha_s(Q^2)$, and 'bare' couplings including higher orders is given by the following geometric series:

$$\alpha_s(Q^2) = \overline{\alpha}_s(1 + I(Q^2, \overline{\alpha}_s, M) + I(Q^2, \overline{\alpha}_s, M)^2 + ...) = \overline{\alpha}_s \ \frac{1}{1 - I(Q^2, \overline{\alpha}_s, M)} \ . \tag{1.3}$$

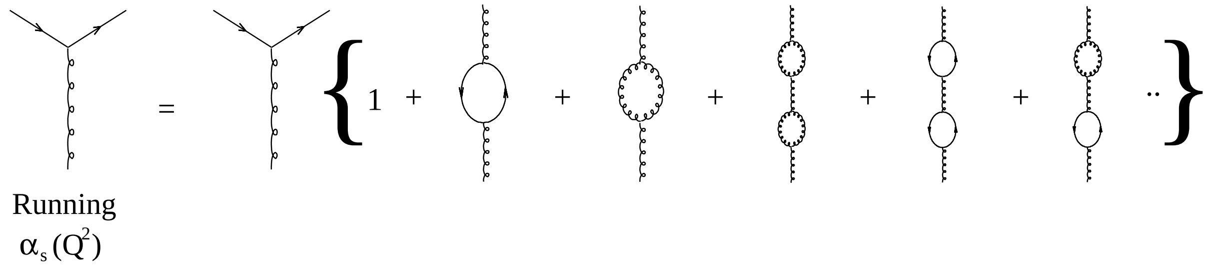


Figure 1.4: *A schematic illustration of the summation over 1-loop diagrams contributing to the running of the strong coupling constant.*

This relation tells us that the coupling that is experimentally measured depends on the scale Q^2 at which the measurement takes place and is referred to as the 'running' coupling. The running coupling describes how the effective coupling depends on the separation between the two interacting particles. The strong coupling can be calculated at a specific value of the gluon four-momentum, $Q^2 = \mu_r^2$, where μ_r is called the *renormalization scale* and is related to the experimental measurement:

$$\alpha_s(\mu_r^2) = \overline{\alpha}_s \; \frac{1}{1 - I(\mu_r^2, \overline{\alpha}_s, M)} \; . \tag{1.4}$$

Then, comparing the left hand side of Equations (1.3) and (1.4) and using the calculation of $I(Q^2, \overline{\alpha}_s, M)$ given in Equation (1.2), M^2 and $\overline{\alpha}_s$ can be eliminated (renormalized) in the expression of the strong coupling. The normalized strong coupling constant, in first order i.e. only including corrections from one-loop diagrams, becomes:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_r^2)}{1 + \frac{\alpha_s(\mu_r^2)}{4\pi} C_1 log(\frac{Q^2}{\mu_r^2})} \; . \tag{1.5}$$

The coefficient C_1 can be determined by calculating $I(Q^2)$ in QCD. This gives the value of the coefficient C_1 equal to:

$$11 - \frac{2}{3} n_f \tag{1.6}$$

where n_f is the number of quark flavours included in the calculation with a mass less than Q^2.

The theory of the strong force must not suffer from divergences caused by the loop calculations in QCD, since then it can not be used to make predictions. However, since these infinities can be controlled, by the renormalization procedure, the theory will still have a predictive power. Nevertheless, there is no reason to believe that QCD will still be a valid theory of the strong interaction at very high energies, since our ignorance of strong interaction theory in the physical limits is hidden in the measurable quantities.

The renormalized coupling $\alpha_s(\mu_r^2)$ depends on the value of the renormalization constant, which is an arbitrary parameter. Different choices of μ_r will lead to different expansions of the amplitude, which corresponds to different renormalization schemes. However, for any

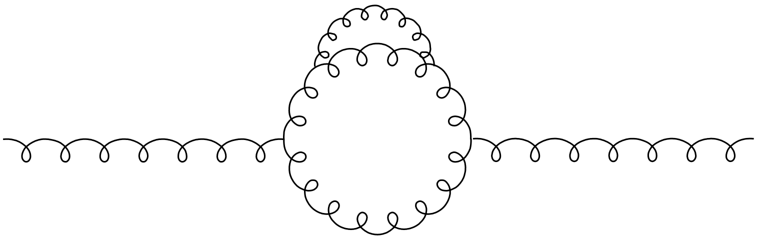


Figure 1.5: *An example of a 2-loop diagram.*

theory which is renormalizable a physical observable, R, must not depend on μ_r, implying that:

$$\mu_r^2 \frac{dR}{d\mu_r^2} \equiv \mu_r^2 \frac{\partial R}{\partial \mu_r^2} + \mu_r^2 \frac{\partial \alpha_s}{\partial \mu_r^2} \frac{\partial R}{\partial \alpha_s} = 0 \quad . \tag{1.7}$$

This is true if all diagrams, which contribute to the physical observable, are included in the calculation of R. The running coupling constant is determined by the renormalization group equation:

$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \beta(\alpha_s) = -\beta_0 \alpha_s^2 (1 + \beta_1 \alpha_s + ...) \tag{1.8}$$

where the loop diagrams at each order in α_s have to be calculated so as to determine:

$$\begin{aligned} \beta_0 &= (33 - 2n_f)/12\pi \\ \beta_1 &= (153 - 19n_f)/2\pi(33 - 2n_f) \\ ... \ . \end{aligned} \tag{1.9}$$

The second term β_1 describes contributions from gluon propagators with 2-loops like the one shown in Figure 1.5. From Equation (1.5) it can be seen that at large value of Q^2 the strong coupling gets small, leading to asymptotic freedom of the partons. This means that the partons are loosely bound at small distances.

Since α_s is increasing with decreasing Q^2 the value of the strong coupling constant will eventually get so large that perturbation technique in QCD cannot be used. This limit, at which the coupling would no longer converge, is given by the scale Λ_{QCD}. In this limit the quarks and gluons start forming hadrons. Experimental data suggest a value of $\Lambda_{QCD} \approx$ 0.200 GeV, which corresponds to about 1 fermi, i.e. the size of light hadrons.

To summarize, it has been shown that loop diagrams lead to divergencies in the calculation of the cross section, which however, can be removed by redefining the coupling constants and propagator mass in a renormalization procedure. From the renormalization procedure it becomes clear that the coupling depend on the energy scale at which the measurement is made. This dependence is given by the renomaliztion group equation. The different behaviour of the running coupling in QCD compared to QED is due to the gluon self-interaction.

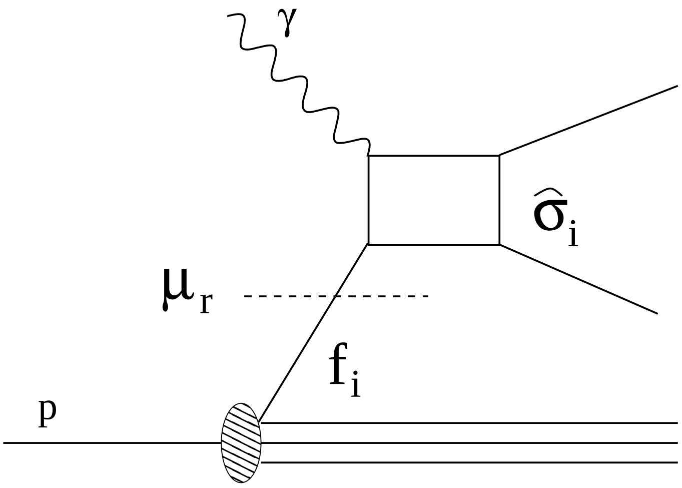


Figure 1.6: *A schematic illustration of factorization in γp scattering.*

1.2.2 Factorization

The predictions of QCD can be tested experimentally by comparing them to results obtained by collider experiments at high energies. In collisions between electrons and protons, the electron is used as a point-like probe to study the partonic content of the proton. At not too high energies (below the masses of the weak vector bosons $W^{\pm}$ and Z^0) the electron-proton interaction is dominated by an exchange of virtual photons. Effectively the scattering takes place between the virtual photon and one of the partons inside the proton. Thus, the most basic scattering process is a pure electromagnetic interaction, whereas higher order processes include strong vertices as well.

The electron-proton cross section can be obtained by convoluting the analytically calculable partonic cross section, $\hat{\sigma}_i$, with a "bare" parton density function, f, that provides the probability of scattering against a parton carrying a certain four-momentum k.

$$\sigma_{\gamma p}(k,q) = \sum f_i(k) \otimes \hat{\sigma}_i(k,q,\alpha_s) \tag{1.10}$$

where q is the four-momentum of the exchanged photon. The "bare" parton density function can, by analogy with the "bare" coupling constant $\overline{\alpha}_s$ introduced in section 1.2.1, not be calculated or measured. The partonic cross section $\hat{\sigma}_i$ of the hard interaction, between the photon and the parton i, is obtained by summing over all possible final states given by QED and QCD:

$$\hat{\sigma}_i = \hat{\sigma}(\gamma i \to i) + \hat{\sigma}(\gamma i \to q\overline{q}) + \hat{\sigma}(\gamma i \to qg) + \hat{\sigma}(\gamma i \to gq\overline{q}) + \ldots \tag{1.11}$$

The higher order terms of $\hat{\sigma}_i$, suffer from logarithmic divergences in various limits of the particle kinematic phase space. These divergences appear for example when the transverse momentum of a final state gluon approches zero as its energy becomes very small (soft divergence) or as it is emitted at very small angles (collinear divergence). As in the case of the strong coupling where divergences are absorbed into the definition of α_s, the singular terms in $\hat{\sigma}_i$ can be separated (factorized) and resummed into the parametrization of the *Parton Distribution Functions (PDF).* This introduces a new arbitrary scale called *factorization scale*, μ_f. Thus, Equation (1.10) now becomes

$$\sigma_{\gamma p}(k,q) = \sum f_i(k,\mu_f^2) \otimes \hat{\sigma}_i^{fact}(k,q,\alpha_s,\mu_f^2), \tag{1.12}$$

where $f_i(k,\mu_f^2)$ and $\hat{\sigma}_i^{fact}$ are now the factorized parton distribution function and the factorized partronic cross section, respectively. Although factorization provides a prescription for handling the logarithmic divergencies, there is an arbitrariness how the finite contributions should be treated. How much of the finite contribution is factored out depends on the factorization scheme used. In the *DIS scheme* all the gluon contributions are included into the parton distribution, whereas in the $\overline{MS}$ *scheme* only the divergent terms are hidden in the PDF.

There are two main approaches to factorizing DIS cross sections: *collinear factorization* and k_t*-factorization*. In the collinear case, (used in the $\overline{\text{MS}}$ and the DIS schemes) the transverse momenta of the incoming partons involved in the hard interaction are neglected (collinear with the proton). Collinear factorization works well in a large fraction of the phase space covered by HERA. As the energy of the probe gets very large, the scattering mainly takes place against the sea-quarks, which means that essentially the gluon density is probed. In the region where the partons carry a very small fraction of the proton momentum, the collinear factorization scheme will no longer reproduce the data, since the transverse momenta of the partons are no longer small compared to their longitudinal components. In this region the k_t-factorization scheme has to be used, describing the scattering cross section as a convolution of the k_t-dependent partonic cross section, where the incoming partons are treated off-mass shell, and *un-integrated Parton Distribution Functions (uPDF)*, where the transverse momenta of the partons are taken into account.

The cross section $\sigma_{\gamma p}(k,q)$ should not depend on the choice of μ_f^2, which means that the product of the partonic cross section and the PDF should be independent of the factorization scale, if all contributing processes are included in the calculations of the partonic cross section $\hat{\sigma}$. Although the PDF's themselves can not be analytically calculated, the dependence of the quark and gluon distribution functions on the factorization scale can be calculated perturbatively. Taking the derivative of Equation (1.12) with respect to μ_f^2 gives the μ_f^2 dependence of $f_i(k,\mu_f^2)$ as a differential equation, which is known as the DGLAP equation. In kinematic regions where k_t-factorization is valid, the so called BFKL equation is used to evolve the un-integrated PDF.

Only a few higher order terms in the expression of $\hat{\sigma}_i$ have been calculated analytically due to the increasing complexity of the calculations. Contributions from the missing higher orders are included via phenomenological treatment of parton emissions before (initial state radiation) and after (final state radiation) the hard sub-process. Depending on the kinematic region the DGLAP and BFKL evolution equations are used to describe initial and final state radiation, as explained in section 1.4.3. This concept is used in Monte Carlo generators in order to simulate higher order effects, see section 1.6.

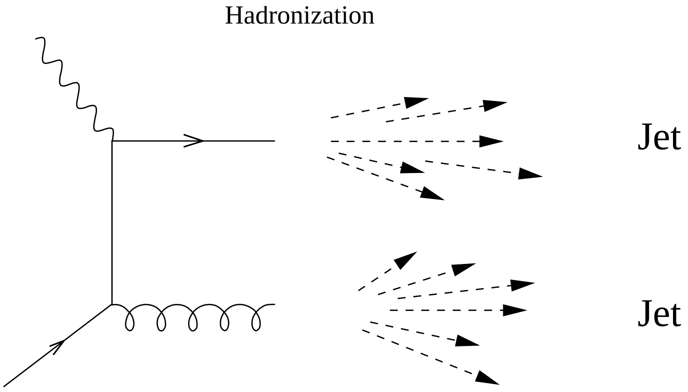


Figure 1.7: *The QCDC diagram indicating the production of collimated flows of hadrons, jets, along the direction of the partons.*

1.3 Non-perturbative QCD

As discussed in the previous section, higher order scattering processes involve the emission of additional partons, which if they are soft can not be described by the parton evolution equations. At the end of the parton showering process, the partons will be converted into hadrons, in a process called *hadronization.* This process will happen at large distances between the partons and can also not be described by perturbative calculations. Similar to the case with PDF's, into which all non-perturbative contributions are absorbed, fragmentation functions are introduced to describe the probability for a parton to hadronize into a particular hadron, carrying a certain fraction of the parton's energy. As in the case of PDF's, the parameters of the fragmentation functions have also to be determined through fits to experimental data. This has been done for an example by comparisons to e^+e^--data.

1.3.1 Hadronization

The Lund string fragmentation model and the cluster fragmentation model are widely used in MC generators to provide the link between the partons (at the parton level) and the measurable hadrons (at the hadron level). Below, both models are described briefly.

The Lund string fragmentation model

The basic idea of the Lund string fragmentation model [16, 17] is that the colour field, connecting a quark with an antiquark or a quark with a diquark, is contained in a narrow colour tube, due to the gluon self-interaction. Mathematically, the colour field is approximated by a massless relativistic one dimensional string, described by a Coulomb potential plus a term, which gives a linear increases with the distance r between the partons,

$$V(r) = -\frac{A}{r} + \kappa r.$$

The constant κ ($\approx$ 1 GeV/fm) is the energy per unit length of the colour tube, called the string tension. If a quark and an antiquark are forced to move apart, which is what

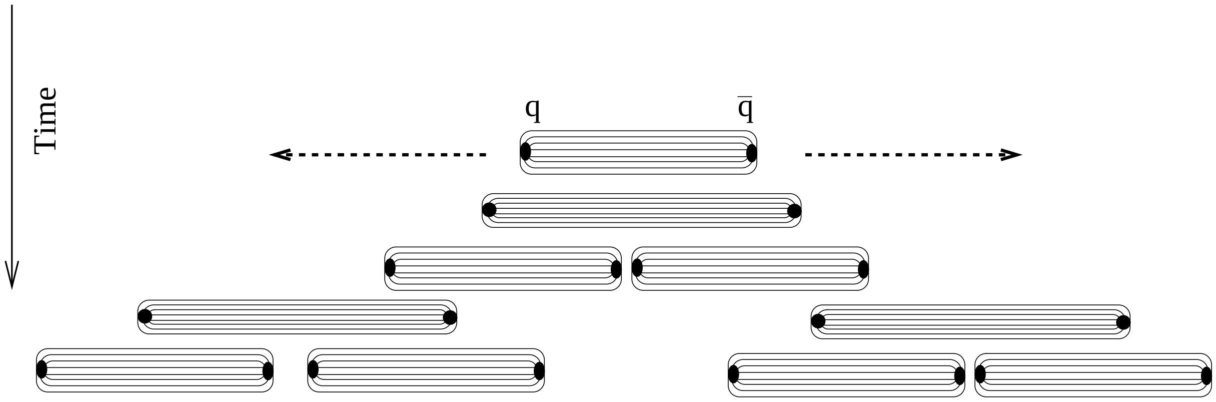


Figure 1.8: *Illustration of the hadronization process according to the Lund string model.*

happens in a scattering experiment, a colour tube is formed as the distance between the quarks exceeds ≈ 1 GeV^{-1}, and it will be stretched out with time. A $q\overline{q}$-pair, which is produced out of vacuum from a quantum fluctuation process, may tunnel through the barrier presented by the constant field inside the tube, with a probability given by

$$exp(-\frac{\pi(m^2 + P_T^2)}{\kappa}),$$

where m is the mass of the quark-pair and P_T is the transverse momentum of the quark and the anti-quark relative to the string. The new $q\overline{q}$-pair will be pulled apart by the field of the original quarks and the field that is built up between them will at some point cancel the original field in that region, and cause the tube to split up in two parts of lower energy. If the initial energy in the string is very high a number of new string fragments will be created. Since the energy of the string is higher at the ends than in the center, due to the kinetic energy carried by the initial quark and antiquark, the string will primarily break at the ends and the produced particles will essentially move in the same directions as the original quark and antiquark, respectively. This gives rise to collimated flows of particles, called *jets*. Figure 1.8 illustrates how mesons are produced through consecutive string breaking.

In case a diquark and an antidiquark is produced as the string breaks, the original quark will connect to the diquark via one colour string and the original antiquark will be colour connected to an antidiquark, i.e. a baryon and an antibaryon have been created. The probability of creating baryons is lower than that of creating mesons due to the higher masses of the baryons. If the quark or antiquark emits a gluon this will cause a kink in the colour string, which will gain kinetic energy in the direction of the gluon. In the hadronization process this leads to an additional jet so that a three jet event will be observed.

The cluster fragmentation model

The cluster fragmentation model [18] contains two steps. In the first step primary clusters are created which, in the second step, decay into secondary clusters or directly into hadrons.

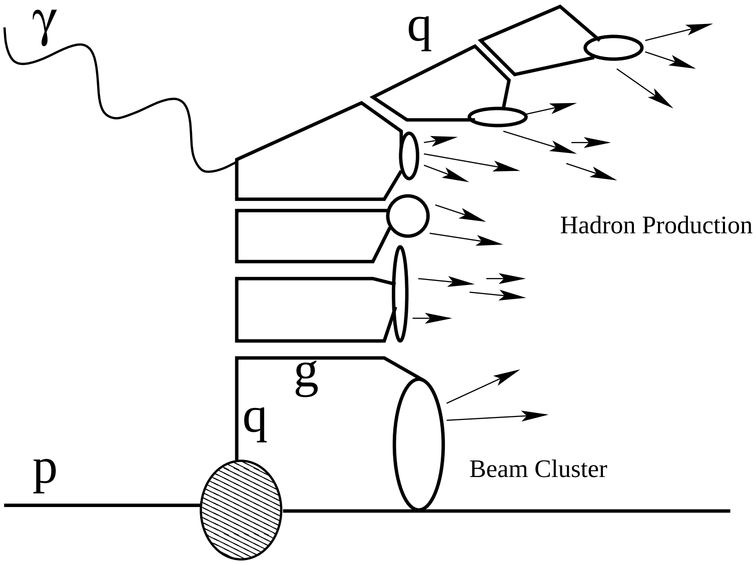


Figure 1.9: *Illustration of the hadronization process according to the cluster fragmentation model.*

When the showering process comes to an end (at some scale, typically ≈ 1 GeV) the quarks are brought to their constituent masses and the gluons are assigned a finite mass. The creation of colour-singlet clusters is performed through a non-perturbative splitting of the gluons into quark-antiquark pairs. Neighbouring quarks and antiquarks, not coming from the same gluon, are then combined into clusters. Large mass clusters may decay isotropically into secondary clusters, whereas clusters of lower masses decay directly into hadrons. This hadronization procedure is illustrated for a DIS event in Figure 1.9.

The cluster mass spectrum is universal and follows a steeply falling behavior with increasing mass. This gives most of the clusters masses of a few GeV, and they are treated as superpositions of meson resonances. These may decay isotropically, in their rest frames, into pair of hadrons, with probabilities given by the density of states. Decays into heavier mesons or baryons are less frequent due to the reduced phase space, in accordance with experimentally measured multiplicity distributions can be described. It turns out that the experimentally obtained hadron energy and transverse momentum distributions can be reproduced quite well without introducing any adjustable fragmentation functions.

1.4 Deep Inelastic Scattering

The dynamics of the partons, confined inside the hadrons, might be explored by using a probe of sufficiently high energy. There is an obvious advantage if the probe itself is a particle without any substructure i.e. a point-like particle. Electrons and positrons have been chosen to probe the partonic structure of the proton in the HERA collider. Such a scattering process where the electron interacts with the proton via the exchange of a photon is illustrated in Figure 1.10. If the photon is quasi-real the process is called photoproduction, whereas virtual photon exchange gives rise to Deep Inelastic Scattering

(DIS). The resolution power of the probe is given by its wavelength, or equivalently by the four-momentum. The higher the four-momentum, the smaller the structures can be resolved. Smaller structures in this context means partons carrying smaller fractions of the proton momentum. DIS processes can also proceed through weak interactions, but due to the large masses of the weak vector bosons (Z^0 and $W^{\pm}$), the photon exchange dominates at low energies. For the energy range covered in this analysis the contributions from weak interaction can be neglected.

1.4.1 Event Kinematics and the DIS Cross Section

A schematic diagram of the DIS process $ep \rightarrow eX$ is shown in Figure 1.10 where the incoming electron, with a four-momentum k, interacts with the proton, of four-momentum P, via a photon, carrying a four-momentum q. The final state contains the scattered electron, having four-momentum $k^{'}$, and a hadronic final state X, of four-momentum W. The kinematic variables of a DIS event can be defined using the four-momenta of the incoming particles and the scattered electron. Due to the rotational symmetry of the event with respect to the incident beam directions, a DIS event can be determined by three Lorentz invariant variables, which are:

The virtuality of the exchanged photon:

$$Q^2 \equiv -q^2 = -(k - k^{'})^2,$$

The Bjorken scaling variable:

$$x_{Bj} \equiv \frac{Q^2}{2P \cdot q} \ , \ x_{Bj} \in [0, 1],$$

The inelasticity:

$$y \equiv \frac{P \cdot q}{P \cdot k} \ , \ y \in [0, 1].$$

Since k and $k^{'}$ are time-like the resulting photon four-momenta is space-like, $q^2 < 0$, and thus Q^2 is always positive. The resolution power of the photon by which the proton is probed, is $\sim 1/\sqrt{Q^2}$. In general the fraction of the longitudinal proton momentum taken by the parton is denoted x. For the lowest order scattering process the x_{Bj} variable gives the fractional proton momentum carried by the interacting quark i.e. $x = x_{Bj}$. In the proton rest frame the inelasticity is the energy fraction of the incoming electron transferred by the photon to the proton. Further, the center-of-mass energy $\sqrt{s}$ of the collision is given by:

$$s = (P + k)^2.$$

The invariant mass of the hadronic final state X is:

$$W^2 = (P + q)^2.$$

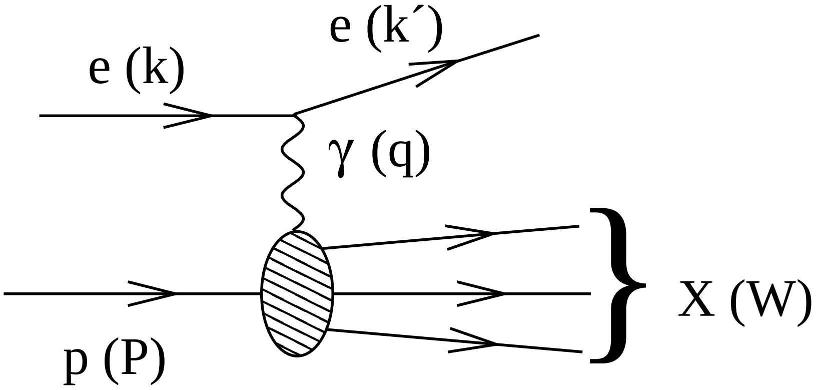


Figure 1.10: *A generic diagram showing a DIS event. The four-momentum notation of the involved particles are given in brackets.*

At HERA collision energies the proton and the electron masses can be neglected. The three variables Q^2, x_{Bj} and y are related by:

$$Q^2 = x_{Bj} y s \approx W^2 \frac{x_{Bj}}{(1 - x_{Bj})}$$

and since s is known, only two independent variables are needed to describe the full kinematics of a DIS event.

Already before the introduction of the quark-parton model, the one photon approximation was used to describe deep inelastic scattering. This model took into account that the proton is not a point-like object and described the cross section as being proportional to the leptonic and hadronic tensors:

$$\sigma \approx L_{\mu\nu} W^{\mu\nu}$$

Here $W^{\mu\nu}$ is an unknown quantity describing the photon proton scattering, which can be expressed in terms of two scalar structure functions, F_1 and F_2. A fundamental problem of the theory is to determine these functions, which at this stage are completely arbitrary and have to be measured experimentally.

In Section 1.2.2 it was discussed that the cross section of $ep \rightarrow eX$ processes can be factorized into two parts, where one is the calculable partonic cross section and the other the parton density function. The parton density functions have mainly been determined from measurements of the proton structure functions, which in leading order QCD are related to the PDF's. In the quark parton model the inclusive double differential neutral current electron-proton cross section can be expressed in terms of the structure functions, $F_1(x_{Bj}, Q^2)$ and $F_2(x_{Bj}, Q^2)$ according to:

$$\frac{d^2\sigma}{dx_{Bj}dQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[[1 + (1-y)^2]F_1 + \frac{(1-y)}{x_{Bj}}(F_2 - 2x_{Bj}F_1)\right]. \qquad (1.13)$$

where α is the fine structure constant.

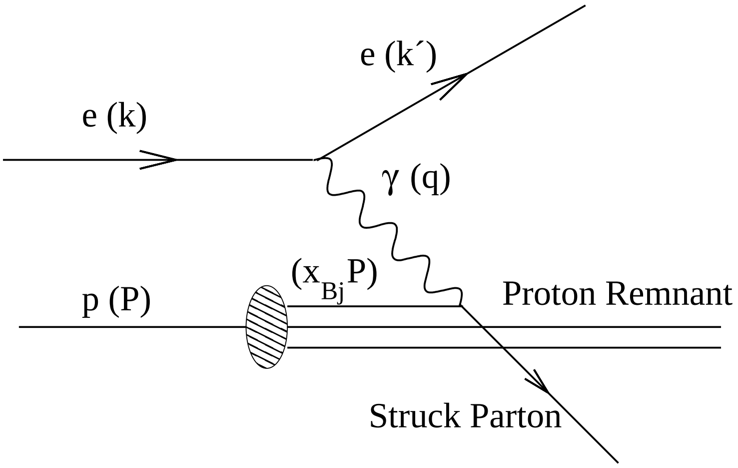


Figure 1.11: *A lowest order DIS event (a so called QPM event), where the photon interacts with a parton carrying a longitudinal momentum fraction* $x_{Bj}P$.

Due to the fact that the interacting photon in DIS is not real but virtual, in addition to being transversely polarised it also has a longitudinal component. The longitudinal structure function is given by $F_L = F_2 - 2xF_1$. The structure functions can be expressed in terms of the cross sections σ_T and σ_L, which are related to the coupling with transversely and longitudinally polarised photons, respectively, such that $\sigma^{\gamma p}_{tot} = \sigma_L + \sigma_T$. At energies where photon exchange dominates and at small x_{Bj}, the cross section is dominated by contributions from F_2, which can be written in the DIS scheme as:

$$F_2(x_{Bj}, Q^2) = \frac{Q^2}{4\pi\alpha^2}(\sigma_L + \sigma_T) = \sum_q e_q^2(x_{Bj} f_q(x_{Bj}, Q^2) + x_{Bj}\overline{f}_q(x_{Bj}, Q^2)) \tag{1.14}$$

i.e. the sum of the momentum weighted quark and antiquark densities, $x_{Bj} f_q(x_{Bj}, Q^2)$ and $x_{Bj}\overline{f}_q(x_{Bj}, Q^2)$, and the electromagnetic couplings to the photon given by the electric charge of the quarks.

The longitudinal structure function, F_L, describes the interaction with longitudinally polarised photons.

$$F_L(x_{Bj}, Q^2) = \frac{Q^2}{4\pi\alpha^2}\sigma_L \sim x f_g(x_{Bj}, Q^2) \tag{1.15}$$

with $x f_g(x_{Bj}, Q^2)$ being the momentum weighted gluon distribution. At high energies the contribution from longitudinally polarised photons is normally neglected due to helicity conservation suppression.

1.4.2 The Quark Parton Model

The Quark Parton Model (QPM) [19, 20] is valid for electron-proton scattering in the 'infinite momentum frame' ($M_p \ll |\bar{P}|$) where the quarks, confined inside the proton, move parallel with the proton direction. A schematic diagram of a QPM event is shown in Figure 1.11, where the virtual photon interacts with a quark carrying a momentum fraction x. The quarks not involved in the hard interaction, which constitute the proton remnant, are also shown. In the QPM process the fractional proton momentum carried by the scattered quark can be identified with x_{Bj}, as was already mentioned.

If the exchanged photon scatters against an object with no extension, like the point-like quarks in the QPM, then it is expected that the structure functions should not depend on the momentum (scale) of the probe, since there is no structure to be resolved, i.e. scaling should be valid. In the QPM, interactions of the quarks via gluon exchange are not considered, which means that the cross section given by Equation (1.13) is reduced to:

$$\frac{d^2\sigma}{dx_{Bj}dQ^2} = \frac{4\pi\alpha^2}{Q^4 x_{Bj}}[1 + (1 - y)^2]F_2(x_{Bj}). \tag{1.16}$$

In Figure 1.12 the structure function F_2 is shown as a function of Q^2 for different values of x_{Bj}. It can be noted that the structure function is independent of the scale Q^2 for x_{Bj} values around 0.13. At low values of x_{Bj}, F_2 is increasing with increasing Q^2, whereas at high values of x_{Bj}, F_2 is decreasing with increasing Q^2. Thus, experimental data from HERA clearly demonstrate that scaling is violated. This is related to the QCD structure of the hadrons. Thus, with high resolution power of the exchanged photon (high Q^2) smaller distances are probed i.e. the sensitivity to partons carrying a small fraction of the proton momentum (small x) is high, whereas with low resolution of the photon only partons having a large fraction of the proton momentum (large x) can be probed. Due to the fact that the structure functions depend on Q^2 and the strong coupling, α_s, the scaling violation effect can be used to extract the parton density functions and the strength of the strong coupling.

At very high parton densities the wave functions of the partons will start to overlap and cause interference effects, such that the partons interact and recombine, which is called *saturation* of the parton density. Such an effect will damp the strong rise of the structure function at very small x-values. In Figure 1.13 the structure function F_2 is shown as a function of x_{Bj} for different values of Q^2. The data in the lowest x range are from the H1 experiment at HERA. The strong rise of F_2 with decreasing x is clearly seen but there is no clear evidence for saturation from the F_2 data.

1.4.3 QCD in DIS

The Feynman diagram shown in Figure 1.11 corresponds to the lowest order DIS diagram. This is a pure electromagnetic process and thus of order zero in the strong coupling constant. Processes of higher orders in α_s involve the emission of gluons. Each additional gluon emission or virtual parton loop increases the order in α_s by one unit. A virtual correction to the zeroth order diagram is shown in Figure 1.14a, whereas the scattering processes of first order in α_s are the QCD-Compton (QCDC) and Boson-Gluon Fusion (BGF) processes, shown in Figure 1.14(b, c). In Figure 1.15 a higher order diagram is drawn, where the soft parton activities, which can not be perturbatively calculated, are

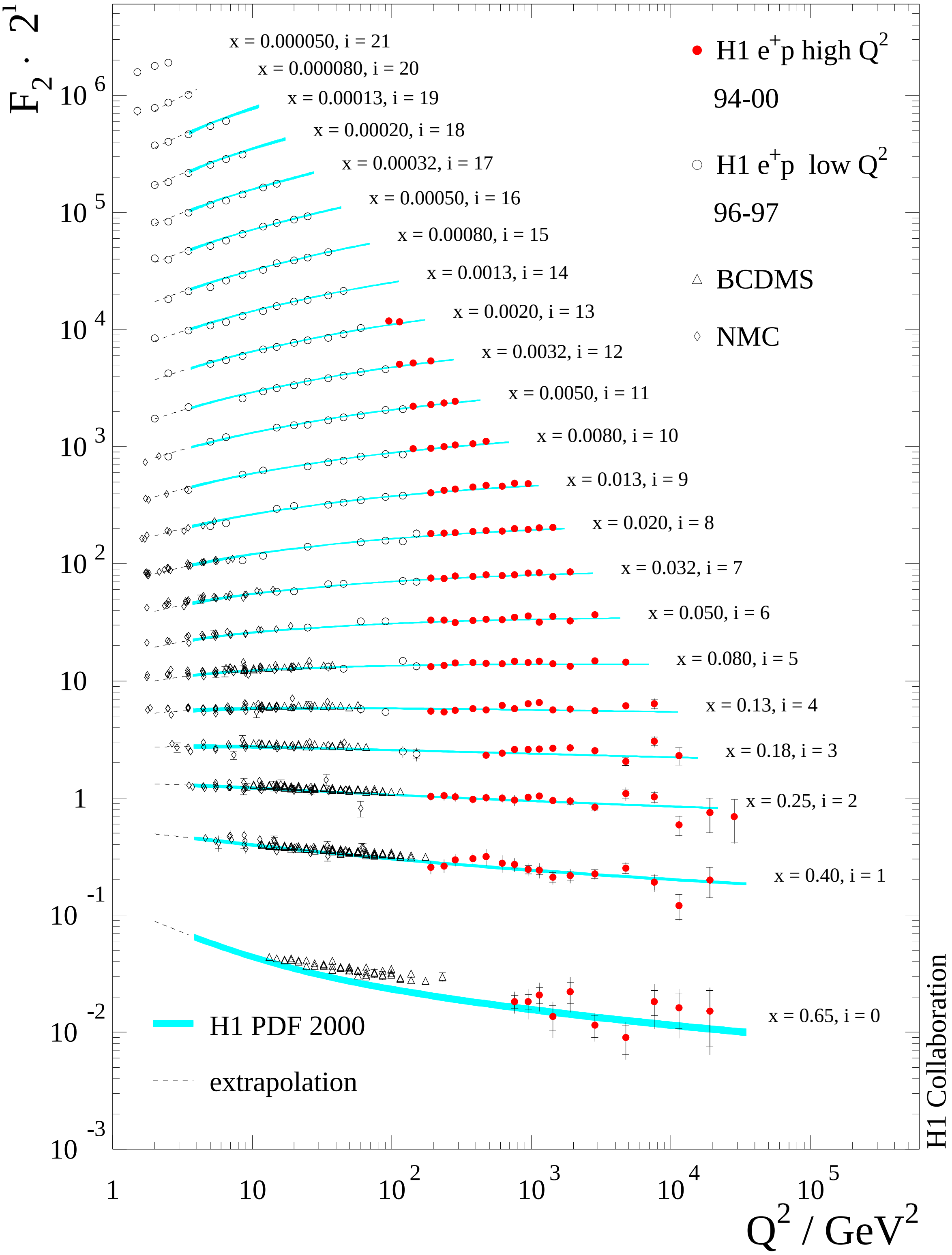


Figure 1.12: *The proton structure function $F_2(x, Q^2)$ as a function of Q^2 at fixed values of x_{Bj}, obtained from measurements by the H1, BCDMS and NMC experiments [21].*

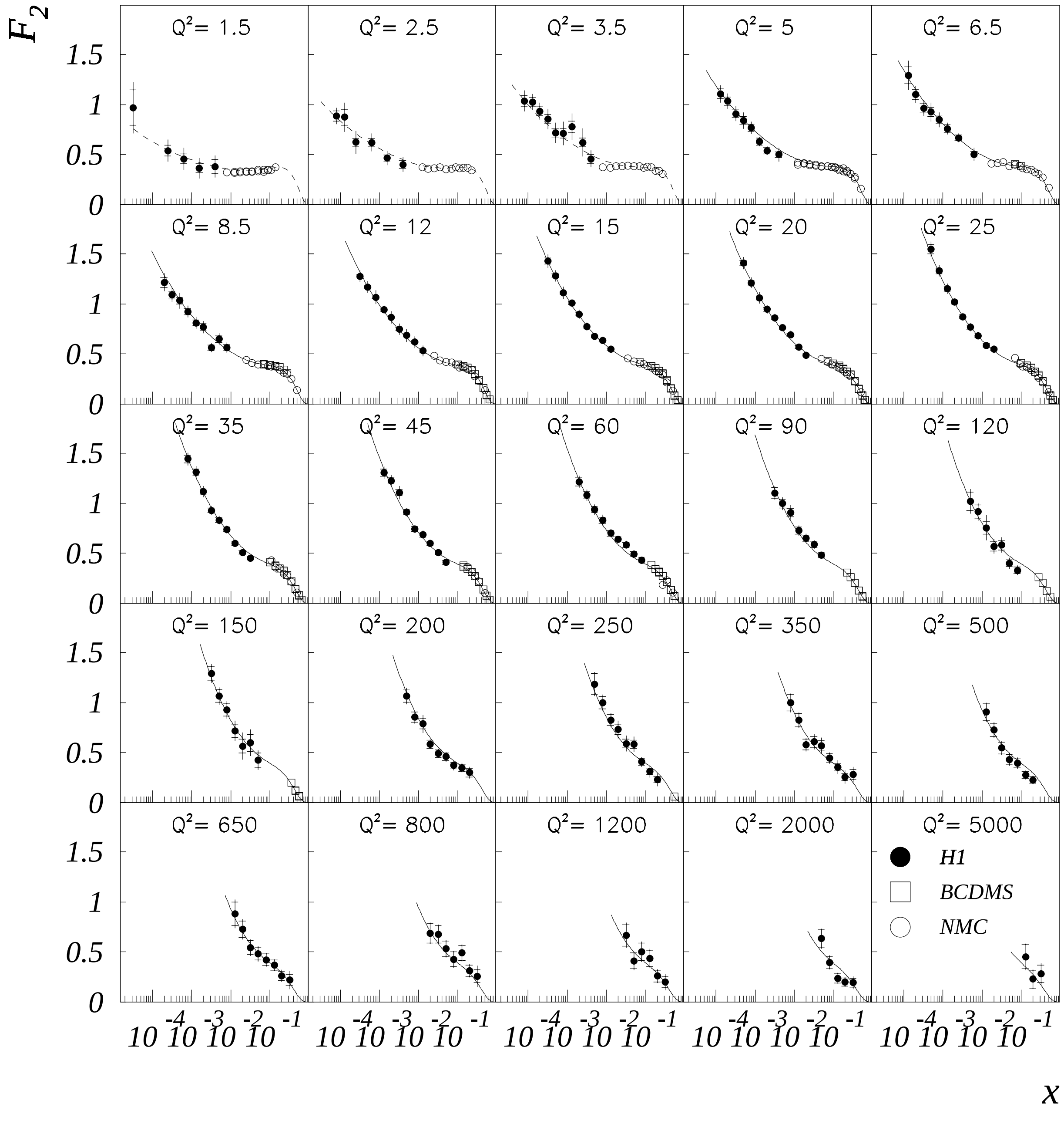


Figure 1.13: *The proton structure function $F_2(x, Q^2)$ as a function of x_{Bj} at fixed values of Q^2, obtained from measurements by the H1, BCDMS and NMC experiments [22].*

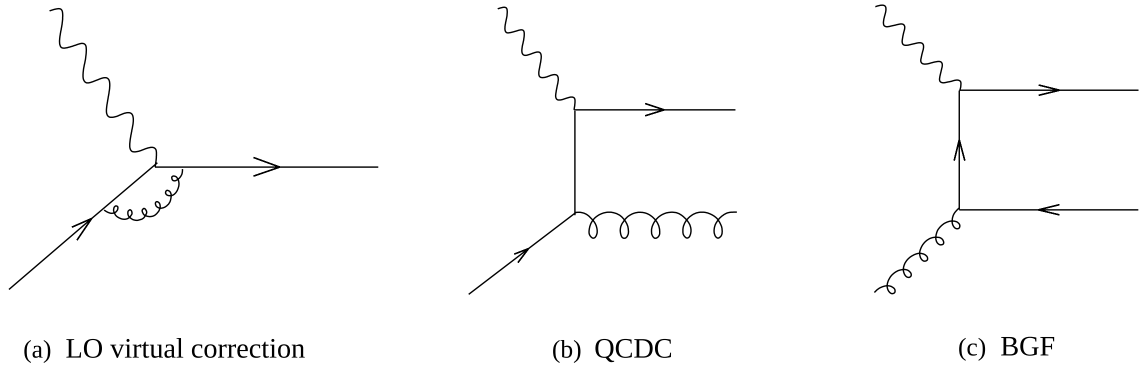

(a) LO virtual correction (b) QCDC (c) BGF

Figure 1.14: *Diagrams showing a) the virtual correction to the lowest order DIS diagram, b) the QCDC and c) the BGF processes.*

hidden in the PDF, represented by the shaded blob. The semi-hard emissions which are indicated in the figure, are treated phenomenologically using the parton evolution equations and the hard scattering is calculated analytically using matrix elements. There are three different approaches to perform the parton evolution depending on which kinematic region is probed. These are the evolution equations of the DGLAP, BFKL and CCFM models.

The DGLAP evolution equation

At Q^2, large enough for perturbative calculation too be valid, and for not too small values of x_{Bj}, the DGLAP [23–26] evolution equations can be used to evolve the density function of quarks, $q_i(x,\mu^2)$, and gluons, $g(x,\mu^2)$, from the starting scale μ_o to the scale of the interaction. The DGLAP equation can be described by the matrix equation:

$$\mu^2\frac{\partial}{\partial\mu^2}\begin{pmatrix} q_i(x,\mu^2) \\ g(x,\mu^2) \end{pmatrix} = \frac{\alpha_s}{2\pi}\sum_{q_j,\overline{q}_j}\int_x^1 \frac{dx'}{x'}\begin{pmatrix} P_{q_i,q_j}(\frac{x}{x'},\alpha_s) & P_{q_i,g}(\frac{x}{x'},\alpha_s) \\ P_{g,q_j}(\frac{x}{x'},\alpha_s) & P_{g,g}(\frac{x}{x'},\alpha_s) \end{pmatrix}\begin{pmatrix} q_j(x',\mu^2) \\ g(x',\mu^2) \end{pmatrix}. \tag{1.17}$$

Here, $P_{ij}(z)$, $z = x/x'$, are the splitting functions, which describe the probability of finding a parton of type i with an energy fraction x in a parton of type j with an energy fraction z. The splitting functions depend on the renormalization and the factorization scheme (e.g. the $\overline{\text{MS}}$ and DIS schemes). They can be expressed as a perturbative expansion in the strong coupling constant:

$$P(z,\alpha_s) = P^0(z) + \frac{\alpha_s}{2\pi}P^{(1)}(z) + \frac{\alpha_s^2}{(2\pi)^2}P^{(2)}(z)... \,. \tag{1.18}$$

Recently, the splitting functions have been calculated up to order α_s^2 [27]. In $\mathcal{O}(\alpha_s^0)$ processes the splitting functions, P_{ij}^0, derived from 2 → 2 scattering processes, describe branchings of the types $q \to qg$, $q \to gq$, $g \to q\overline{q}$ and $g \to gg$, and are given by:

$$P^0_{qq}(z) = \frac{4}{3}\,\frac{1+z^2}{1-z} \tag{1.19}$$

$$P^0_{gq}(z) = \frac{4}{3}\,\frac{1+(1-z)^2}{z} \tag{1.20}$$

$$P^0_{qg}(z) = \frac{1}{2}\left[z^2+(1-z)^2\right] \tag{1.21}$$

$$P^0_{gg}(z) = 6\left(\frac{z}{1-z}+\frac{1-z}{z}\right)+z(1-z) \tag{1.22}$$

where the terms $1/(1-z)$ and $1/z$ are called singular terms, since they give infinite contributions as $z \to 1$ and $z \to 0$, respectively.

As mentioned previously the choice of the scale μ^2 is arbitrary. In the *Leading Logarithmic Approximation* (LLA), only the leading term of the splitting functions, $P^0(z)$, is taken into account. The DGLAP equation, Equation 1.18, in LLA contains a *resummation* over all terms of the type $(\alpha_s ln(\mu^2/\mu_0^2))^n$, where n denotes the order of the term in the expansion of the cross section. It can be shown that the leading log terms $(\alpha_s ln(\mu^2/\mu_0^2))^n$ corresponds to an evolution, in which the parton virtuality is strongly ordered,

$$Q^2 \gg |k_n^2|.. \gg |k_{i+1}^2| \gg |k_i^2|.. \gg |k_0^2|, \tag{1.23}$$

where k_i is the four-momentum of parton i. Also, the parton momentum fraction has to be ordered, $x_{i+1} < x_i$, since momentum conservation has to be fulfilled in each splitting. The condition $k_{t,i}^2 = (1-x_i)|k_i^2|$ is fulfilled in each splitting, which leads to a strong ordering of the transverse momenta of the propagating partons,

$$Q^2 \gg |k_{t,n}^2|.. \gg |k_{t,i+1}^2| \gg |k_{t,i}^2| \gg |k_{t,0}^2|. \tag{1.24}$$

Since the virtualities and transverse momenta of all the parton propagators are small compared to the hard scale, given by Q^2, the partons can be treated as massless and assumed to move in the same direction as the incoming proton. Due to the fact that DGLAP evolution resums $(\alpha_s ln(\mu^2/\mu_0^2))^n$ terms, it will only be valid at large enough values of Q^2, where these terms are dominating. The DGLAP evolution generally gives a good description of data in this region but it should be remembered that higher order contributions to the splitting functions lead to smaller x values for the partons, which at the end might spoil the convergence.

The BFKL evolution equation

At high four-momenta of the probe, more of the gluon content in the proton will be resolved. The sea quarks, which are created from quantum fluctuations of gluons, will carry small fractions x of the proton momentum. The smaller the x-values that are probed the more important terms of the type $\alpha_s ln(1/x)$, in the evolution, become and eventually they will dominate over the terms $\alpha_s ln(\mu^2/\mu_0^2)$. On entering the small x region, it is thus expected that the fixed order DGLAP approximation will break down and has to be replaced by the BFKL evolution scheme [28–30], which resums terms of the type $(\alpha_s ln(1/x))^n$ and

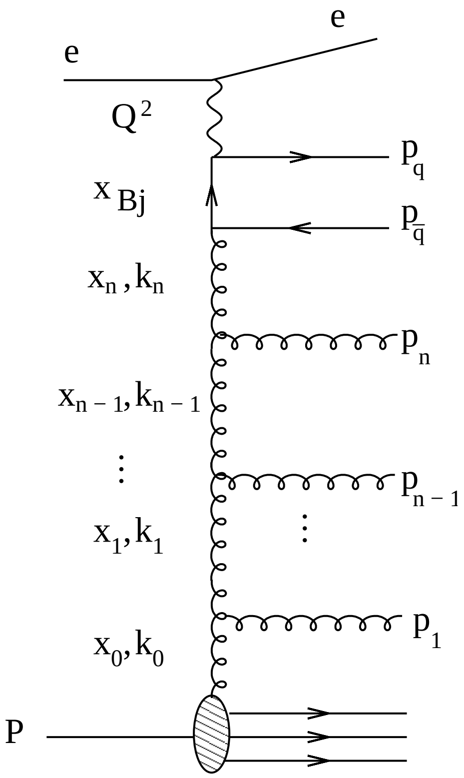


Figure 1.15: *A higher order DIS diagram with a gluon ladder, where the four-momenta of the propagator gluons and the emitted gluons are denoted* k_i *and* p_i*, respectively.* Q^2 *is the four-momentum of the virtual photon and* x_{Bj} *is the parton momentum fraction at the interaction vertex.*

consequently is only valid at small x. A strong ordering in the longitudinal momentum fraction of the parton propagators is assumed in this evolution scheme i.e.

$$x_o^2 \gg x_1^2 .. \gg x_i^2 \gg x_{i+1}^2 .. \gg x_n^2 \gg x_{Bj}^2 \tag{1.25}$$

which means that the evolution is made in increasing $ln(\frac{1}{x})$. This implies that the emitted partons will take a large fraction of the propagator momentum. Since there are no restrictions in k^2 or k_t^2, the virtualities and transverse momenta of the propagators can take any kinematically allowed value and need not be smaller than the photon virtuality as was the case in the DGLAP description. Consequently, the collinear approach is no longer applicable and therefore the parton masses and the transverse momenta of the parton propagators can no longer be neglected. Instead k_t-factorization has to be applied. Here the matrix elements, which are used to calculate the partonic cross section, must be taken off mass shell, by giving the partons a virtual mass, and un-integrated parton densities have to be used, in which the transverse momenta of the propagators are taken into account. The cross section can be written as the convolution of a k_t-dependent partonic cross section and a k_t-dependent parton density function:

$$\sigma = \sum_i \int_0^1 dx \int d^2k_t \ \mathcal{F}_i(x, k_t^2, \mu_f^2) \hat{\sigma}_i(x, k_t^2). \tag{1.26}$$

Here $\mathcal{F}_i(x, k_t^2, \mu_f^2)$ are the Un-integrated Parton Density Functions (UPDF), and $\hat{\sigma}_i(x, k_t^2)$ the off-shell parton cross section. Integrating out the k_t dependence of the partons and neglecting the k_t dependence in σ, the UPDF is reduced to the collinear PDF by the relation:

$$f(x, \mu_f^2) \simeq \int_0^{\mu_f^2} dk_t^{'2} \ \mathcal{F}(x, k_t^{'2}, \mu_f^2). \tag{1.27}$$

The BFKL evolution is of the form:

$$\frac{d\mathcal{F}(x, k_T^2)}{d\ln(\frac{1}{x})} = \int dk_T^{'2} \ \mathcal{F}(x, k_T^{'2}) K(k_T^2, k_T^{'2}), \tag{1.28}$$

where the function K is the splitting function (also called the BFKL splitting kernel). One consequence of non ordering in k_t of the emitted partons, is the absence of a lower limit in the transverse momentum which is received by the propagator after each branching and thus the calculations may enter into non perturbative regions. The cross section (1.26) is dominated by the gluon distribution at very small x and thus it is a good approximation to only consider the gluon evolution in Equation (1.28).

The CCFM evolution equation

The CCFM [31–34] evolution equation is valid both at large and small x-values, since it resums terms of the types $(\alpha_s ln(\frac{1}{x}))^n$ and $(\alpha_s ln(\frac{1}{1-x}))^n$. This means that the CCFM evolution will be DGLAP-like at large x-values and BFKL-like at small values of x. The CCFM evolution introduces strong angular ordering of the parton emissions, which gives

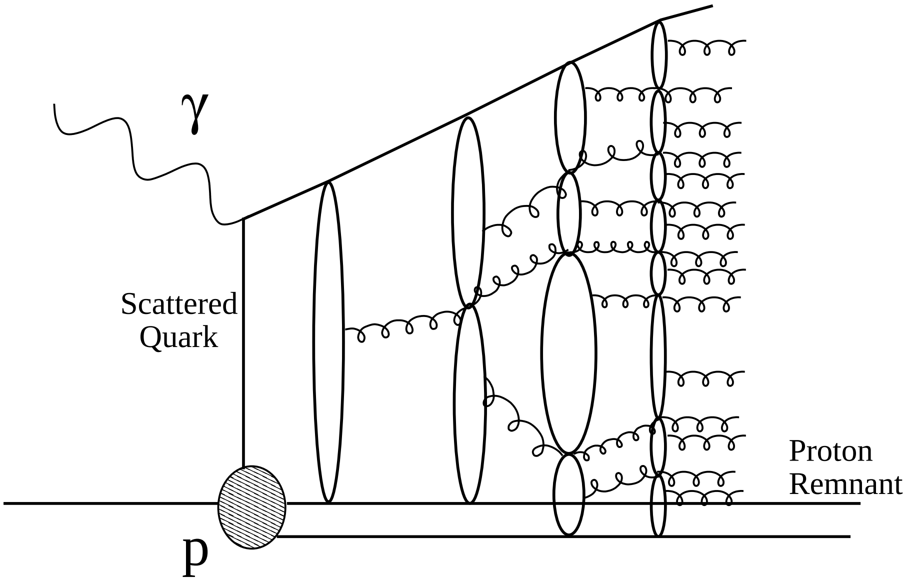


Figure 1.16: *Parton showering according to the Colour Dipole Model (CDM).*

a correct treatment of colour coherence effects [35]. The emission angles, ξ, are increasing according to:

$$\Xi \gg \xi_n \gg ..\xi_i \gg \xi_{i+1} \gg \xi_o \tag{1.29}$$

where the maximum angle Ξ allowed in the emission is given by the hard scattering quark box. The original CCFM splitting function is given by:

$$\tilde{P}_g(z_i, \overline{q}_i, k_{ti}^2) = \frac{\overline{\alpha}_s(\overline{q}_i^2(1-z_i)^2)}{1-z_i} + \frac{\overline{\alpha}_s(k_{ti}^2)}{z_i}\Delta_{ns}(z_i, \overline{q}_i^2, k_{ti}^2) \tag{1.30}$$

were $\overline{\alpha}_s = 3\alpha_s/\pi$ and Δ_{ns} is the non-Sudakov form factor (includes all higher order virtual corrections to the gluon vertex), $z_i = x_i/x_{i-1}$ is the ratio of the energy fractions in the branching $(i-1) \rightarrow i$, and k_{ti} is the transverse momentum of the gluon in the initial state of the gluon cascade. The rescaled transverse momenta of the emitted gluons, $\overline{q}_i$, is defined as:

$$\overline{q}_i = \frac{p_{ti}}{1-z_i} = x_{i-1}\sqrt{s\xi_i}. \tag{1.31}$$

1.4.4 The Colour Dipole Model

The DGLAP, BFKL and CCFM parton evolution schemes have in common that the parton emissions are described by splitting functions. In the Colour Dipole Model (CDM) [36–40] gluons are emitted from colour dipoles, which are stretched between the emitted partons in an avalanche process. Dipoles can be stretched between pairs of $q\bar{q}$, qg and gg, where the probability for an emission is:

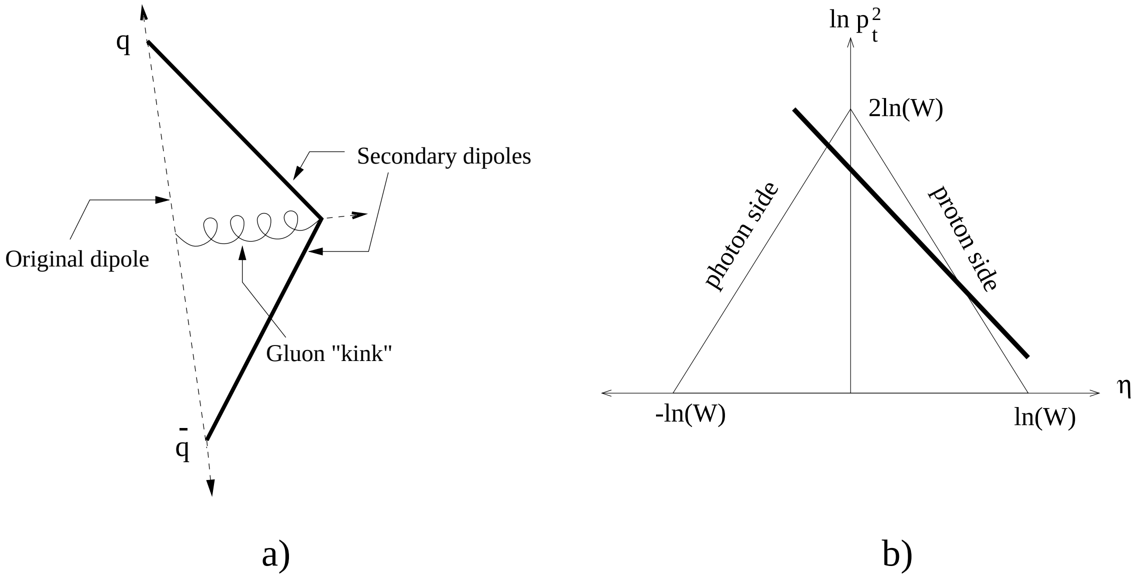


Figure 1.17: *a) The emission of a gluon, which according to the colour string model causes the original colour dipole between the quark and the antiquark to break up into two secondary dipoles, stretched between the quark and the gluon and between the antiquark and the gluon. The phase space for gluon emission is shown in (b), expressed in terms of* $ln(p_t^2)$ *and* η*, i.e. the momentum and rapidity of the gluon, where the gluon is emitted from a dipole of mass* W*. The thick line indicates the reduction in phase space due to the extension of the proton [40].*

$$d\sigma = \frac{\alpha_s}{4\pi} N_c \frac{dP_T^2}{P_T^2} dy d\phi, \tag{1.32}$$

with y, P_T and ϕ being the rapidity, transverse momentum and the azimuth angle, respectively, of the emitted gluon in the dipole center of mass system. A schematic picture of a parton cascade from a DIS event as described by the CDM is shown in Figure 1.16. The first dipole, spanned between the proton remnant and the scattered quark, radiates a gluon. Two new dipoles are stretched between the emitted gluon and the scattered quark on one hand , and the proton remnant on the other hand. These dipoles independently radiate one gluon each. Thereby, four new dipoles are created and so on. This continues until no dipole gives a transverse momentum above a certain cut-off value, which defines the point at which the hadronization starts. The independent radiation of the dipoles leads to non-ordering in the transverse momenta of the parton emissions, similar to the BFKL evolution scheme.

Figure 1.17 illustrates how the emission of a gluon from the colour dipole, formed by a quark and an antiquark, causes a kink in the original dipole, which results in two new dipoles. The available phase space for gluon emissions can be represented by a triangular area with η, on the horizontal axis and ln p_t^2 on the vertical axis, where η and p_t are the pseudo rapidity and transverse momentum, respectively, of the emitted gluon in the dipole center-of-mass system. This is illustrated in Figure 1.17b. For dipoles stretched between point-like particles the phase space boundaries are given by $|\eta| < ln(W/p_t)$. However, the proton remnant is an extended object, from which radiation of wavelengths smaller than its extension is suppressed. This corresponds to an extra restriction of the phase space, which is indicated by the thick line in Figure 1.17b.

1.4.5 Fixed Order α_s Calculations

QCD predictions in leading order (LO) of some measurable observable normally suffer from large higher order corrections and scale uncertainties. The corresponding order expressed in the strong coupling constant is related to the measured observable, such that LO inclusive jet production means $\mathcal{O}(\alpha_s^0)$, whereas LO dijet production means $\mathcal{O}(\alpha_s^1)$. Generally speaking LO is the lowest order on which the observable obtains a non-zero contribution, and next-to leading order (NLO) is the next higher order in α_s. In LO calculations all higher order contributions are absorbed into the parametrization of the parton density function. In NLO one additional real or virtual gluon is taken into account and in next-to-next-to leading order (NNLO) two additional gluons are added. The additional real gluons in next orders are restricted kinematicly. The more terms that are included into the perturbative expansion of the parton evolution the smaller the dependence on the choice of the factorization and renormalization scales will be. However, the degree of complication in the calculations increases dramatically with each order that is added.

The matrix element calculations diverge if one of the partons produced in the hard sub-process become either collinear with the incoming or outgoing parton or if they are soft. However, these divergencies in principle are cancelled by the negative contributions from the virtual gluon corrections. There are two different techniques to perform the cancellation of soft and collinear singularities in the computations. One is the so called phase space slicing method [41,42], in which an invariant mass resolution cut is introduced to isolate the soft and collinear singularities. The other is the subtraction method [43], in which the virtual loop corrections are subtracted by defining an approximate cross section that regularise the real corrections. However, depending on how the physical observable is defined it might happen that the virtual contributions are still not completely cancelled since there are not enough soft or collinear emissions included in the calculations due to kinematic restrictions [44–46]. In order to make a relevant comparison between higher order calculations and the experimental results, the observables must be either infrared-safe, i.e. safe with respect to collinear and soft singularity, or kinematic cuts have to be introduced in order avoid the problematic regions of the phase space.

LO and NLO 2-jet processes

The BGF and QCDC processes, shown in Figure 1.14(b-c), are of order α_s with two partons of high transverse momenta in the final state, which hadronize into two jets. Since the cross section of these processes depend on both the strong coupling constant and the parton densities of the proton, they have been used to experimentally measure these quantities. Since QCDC is dominating in the high x_{Bj}-region, where the dependence of the gluon PDF is small, it has been used to determine $\alpha_s(\mu_r)$. The gluon initiated BGF process, on the other hand, dominates at low x_{Bj} and has therefore been measured to directly determine the gluon density. Due to momentum conservation, the two final state partons of the QCDC and BGF processes will have equal and opposite momenta. Such a process can be

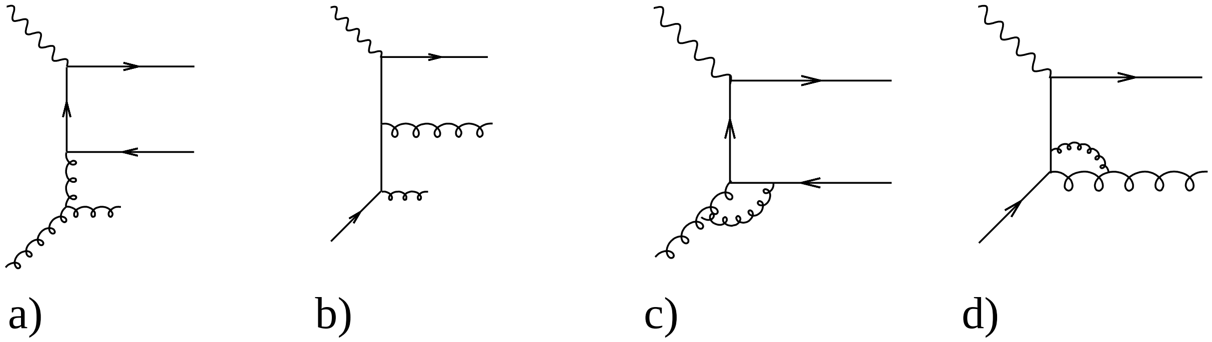


Figure 1.18: *Diagrams showing NLO contributions to dijet production, where* $(a - b)$ *show the real, and* $(c - d)$ *the virtual emissions.*

calculated analytically using so called matrix elements. Calculations to LO order, however, do not reproduce the experimental data.

The inclusion of NLO corrections, significantly reduces the dependence of the renormalization and factorization scales, present in LO calculations, and much more reliable cross sections are obtained. However, with the experimental requirement of two back-to-back jets with equal momenta, the phase space for additional gluon radiation is limited to what can be allowed within the measuring accuracy. This means that in the symmetric jet situation the gluon emission will be either very soft or essentially collinear with the outgoing parton. As was discussed above, such a situation will cause errors in the NLO calculations due to the effects of infrared divergencies. The insufficient cancellation of soft gluons, Figure 1.18(a-b), with virtual corrections, Figure 1.18(c-d), will introduce large logarithms in the cross section calculations, which give the cross section an unphysical behavior. A reliable perturbative expansion can be obtained if the large logarithms are resummed to all orders in α_s. Such calculations have recently been performed [45]. The alternative is to introduce experimental cuts like asymmetric jet energies, which opens up the phase space for additional radiation and makes the dijet cross section infrared stable.

LO and NLO 3-jet processes

The diagrams in Figure 1.19(a-b) would be regarded as NLO corrections to dijet production in case the emitted gluon is soft enough not to produce a reconstructed jet (compare Figure 1.18a. However, if the gluon is hard the diagram represents a 3-jet event, which is of order α_s^2 in LO. The real and virtual NLO corrections to the BGF and QCDC processes are shown in Figure 1.19(c-d) and Figure 1.19(e-d), respectively. Also calculations of 3-jet cross sections encounter problems with singularities due to the emission of soft or collinear gluons as discussed above.

1.4.6 The photon in DIS

The time taken for the electron and proton to interact in DIS is related to the scale of the hard sub-process (μ) via the Heisenberg uncertainty principle and is given by $\tau_{int.} \sim 1/\sqrt{\mu^2}$.

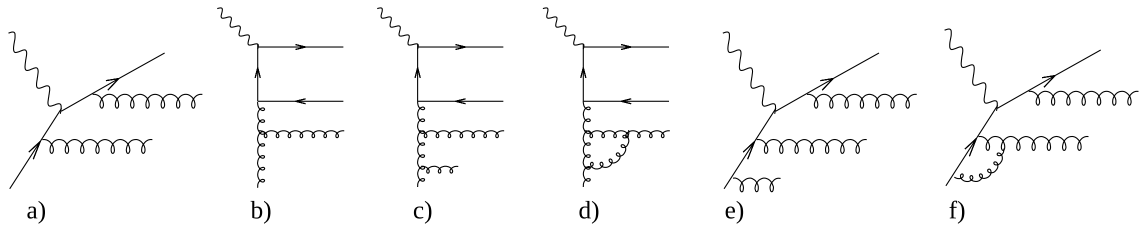


Figure 1.19: *Diagrams showing $(a-b)$ the LO 3-jet production, and the the NLO contributions $(c-f)$.*

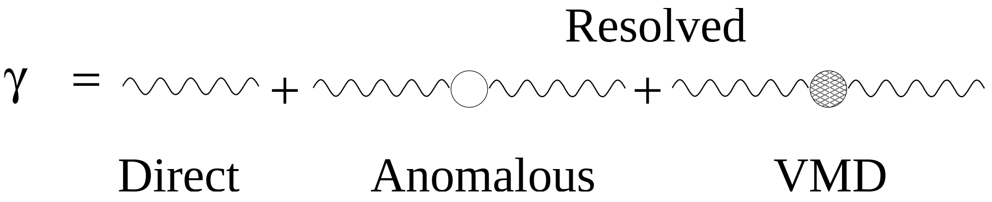


Figure 1.20: *The three quantum states of the photon given by the 'bare' photon providing direct interactions, the anomalous contribution where the photon fluctuates into a quark-antiquark pair, and the VDM (Vector Meson Dominance) contribution where the photon fluctuates into a bound quark-antiquark state.*

If the sea-quark structure of the proton is to be resolved, $\tau_{int.}$ has to be smaller than the life time of the parton fluctuation in the proton, $1/R_p \sim \Lambda_{QCD}$. Similarly, if $\tau_{int.}$ is smaller than the lifetime of the photon fluctuation into partons, $1/R_\gamma \sim 1/\sqrt{Q^2}$, the photon can be treated as an object with a hadronic structure. This means that the photon interacts either as a point-like particle or as a resolved object via its partonic content. The photon can fluctuate into either a bound $q\bar{q}$-state, creating a vector meson, or into a decoupled quark and antiquark, provided the p_t of the partons is large enough. The latter is called the anomalous contribution (see Figure 1.20). Soft processes like the creation of vector mesons can not be calculated using pQCD. Instead the Vector Dominance Model (VDM) has been used to describe how the photon fluctuates into a vector meson $(\rho^0, \omega, \phi, J/\psi, ...)$ before it interacts with the proton. Contributions of the VMD state are negligible in regions where Q^2 is larger than the vector meson masses ($Q^2 > 4$ GeV^2).

A quark (antiquark) produced in the anomalous process can interact either directly with a gluon from the proton ($\gamma \rightarrow q\bar{q}$) or it may first emit a gluon ($\gamma \rightarrow q\bar{q}g$), which then interacts with a parton from the proton. The partonic processes can be described perturbatively by calculating the various splittings in the evolution of the partons. The perturbative calculations have to be complemented by a photon structure function, F_2^γ, describing the parton momentum distribution in the photon at the starting scale of the evolution. Although the photon structure function has been determined to good precision in photoproduction ($Q^2 = 0$), its dependence on Q^2 has not yet been extracted to very high accuracy. The total cross section is given by the sum of the three contributions:

$$\sigma_{\gamma p}^{tot} = \sigma_{\gamma p}^{Dir} + \sigma_{\gamma p}^{Anom} + \sigma_{\gamma p}^{VMD} \tag{1.33}$$

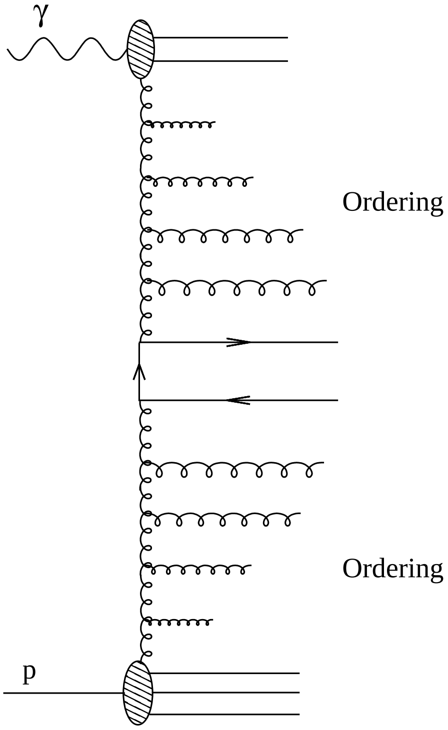


Figure 1.21: *A diagram of higher order, illustrating the interaction between a resolved photon and the proton, including two DGLAP ladders.*

Including the possibility of having interactions with resolved photons the total cross section has to be factorized into partonic cross sections and structure functions of both the proton and the photon. The scale dependence of the photon structure functions can, as for the proton, be described by the DGLAP evolution function in the collinear factorization approximation. A resolved photon event thus will contain two ladders where one describes the parton evolution from the photon PDF to the hard scattering scale and the other describes the evolution from the proton PDF to the hard scattering scale, as illustrated in Figure 1.21.

The NLO BGF process can be described either as a gluon from the proton, which after having emitted another gluon, splits up into a quark-antiquark pair and the hard interaction takes place with a point-like photon, as shown in Figure 1.22a, or alternatively the photon splits up into a quark-antiquark pair (resolved photon) and the interaction takes place via a gluon exchange, shown in Figure 1.22b. This means that a fraction of the NLO corrections in direct processes are naturally included via the resolved contribution. This is true if the photon virtuality, Q^2, is of the same order as the scale of the hard sub-process $\mu^2 = P_T^2$ of the hardest parton.

1.5 Multiple Parton Interactions in DIS

Due to the fact that hadrons contain quarks and gluons, collisions between two hadrons may involve more than one parton interaction. The probability for Multiple Parton In-

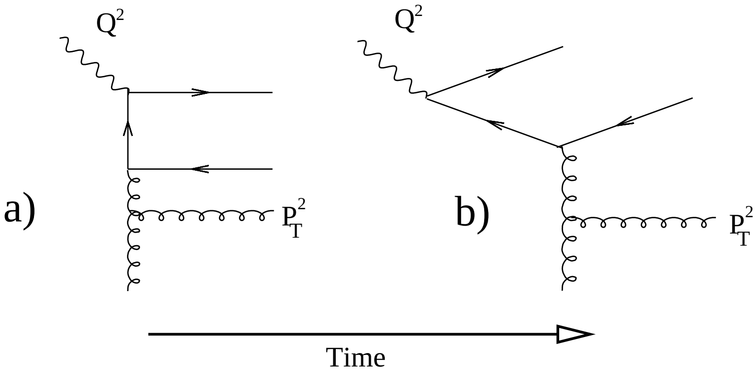


Figure 1.22: *Diagrams with a three parton final state, $\mathcal{O}(\alpha_s^2)$, where a) represents a NLO 2-jet diagram whereas b) represents a LO 2-jet diagram with a photon remnant.*

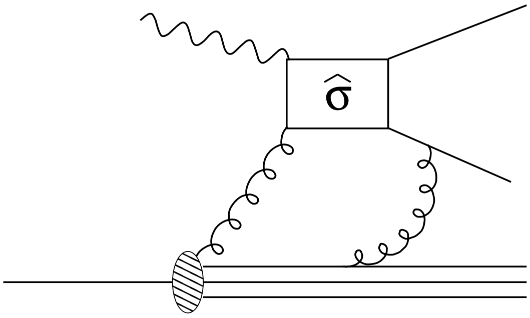


Figure 1.23: *An example of multiple parton interaction in ep scattering.*

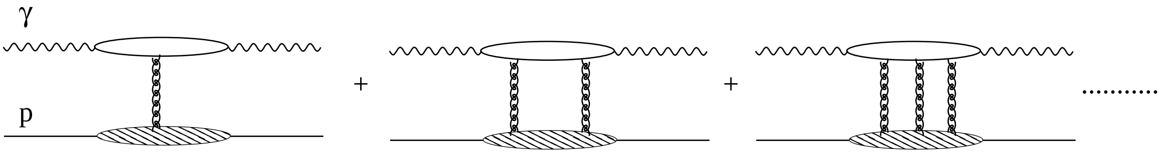


Figure 1.24: *The lowest contributions to the γp elastic scattering.*

teractions (MPI) increases with increasing collision energy because of the strongly rising parton density at low x-values. MPI may also accur in electron-proton collisions. One possibility is that the exchanged photon fluctuates into a $q\bar{q}$-pair and thus interacts through its hadronic component with the proton. A second contribution may come from gluon exchange between the scattered quark and the quarks of the proton remnant (rescattering). The additional interaction may be hard (can be described by pQCD) or soft (part of non-perturbative QCD). Figure 1.23 illustrates a scattering process with the exchange of an additional gluon between the scattered quark and the proton. In the collinear factorization scheme, with momentum ordered gluon ladders, MPI can be regarded as so-called higher-twist contributions, which provide power corrections to the structure functions, given by:

$$F_2(x,Q^2) = \sum_{n=0}^{\infty} \frac{F_2^{(n)}(x,Q^2)}{Q^{2n}}, \tag{1.34}$$

where the superscript n indicates the 'order of the twist' and thus $F_2^{(0)}(x,Q^2)$ is the contribution from the lowest order DIS process (the leading-twist diagram).

At small x, higher order twist diagrams may become important due to the large logarithms $ln(1/x)$. According to the so called cutting rules of Abramovasky-Gribov-Kanchelli (AGK) [47–49], MPI is related to diffraction, i.e. events where the exchanged particle is in a colour singlet state, and to saturation of the parton density.

There is a general consensus that all final state particles not produced directly by the hard scattering are included into the so called *underlying event* (UE). Thus, the UE will have contributions from initial and final state radiation, from additional soft or semi-hard parton interactions and from the beam remnant. Experimentally it is far from obvious how these contributions can be separated and although agreement with Monte Carlo models, including all contributions, can be obtained, it is not guaranteed that each contribution to the UE is correctly described by the model.

1.5.1 The Optical Theorem and Regge Theory

A scattering processes can be described as an initial incident state, followed by an interaction between the components of the system, resulting in a final scattered state. The scattering matrix (or S-matrix) describes how the initial state is transformed into the final state. Conservation of probability in the transformation requires that the scattering matrix to be unitary. By summing over all intermediate states it can be shown that the imaginary part of the forward scattering amplitude is related to the total scattering cross section, via the optical theorem:

$$\sigma_{\gamma p}^{tot} = \frac{1}{s} Im[\ A^{\gamma p \to \gamma p}(s, t = 0)\ ] \tag{1.35}$$

Here, Im[A] is the imaginary part of the amplitude, s and t are the Mandelstam variables defined as $s = (P_\gamma + P_p)^2$ and $t = (P_p - P_p')^2$, where P_γ, P_p and P_p' are the four-momenta of the γ, the incoming proton and the scattered proton, respectively. In Figure 1.24 the lowest order contributions to $\gamma p \to \gamma p$ scattering are shown. After the interaction, the proton remains intact and therefore the exchanged virtual particles are in a colour singlet state. It can be shown that the total cross section cannot rise faster than $ln^2(s)$ at very large value of s if all contributions to the scattering are summed over.

In a space-like representation, i.e. the exchanged particle is space-like, s is the square of the center-of-mass energy, whereas t is the squared four-momentum transfer at the proton vertex. In the crossed diagram, which is the time-like representation, the two variables are interchanged so that t is the square of the center-of-mass energy, while s is the momentum transfer squared. The time-like representation of $\gamma p \to \gamma p$ scattering, via one particle exchange, is shown in Figure 1.25. The analytic properties of the S-matrix ensures that $A^{\gamma p \to \gamma p}$ can be calculated from $A^{\gamma\gamma \to p\bar{p}}$ scattering after exchanging s with t. The scattering process can thus be interpreted in either of two representations, a fact exploited in Regge theory.

In the Regge theory, formulated long before the introduction of QCD, the exchanged particles leading to inelastic scattering are identified as mesons, in the time-like representation. The scattering amplitude is analytically extended to complex angular momenta, which gives an interpolating function $A(l, t)$. This is reduced to $A_l(t)$ for integer values of the angular momentum and for energies t, which give rise to singularities, poles, in the scattering matrix. These so called *Regge poles* can be identified with the exchanged objects of masses M_l. A sequence of poles with $l = L_1$ at $t = t_1$, $l = L_2$ at $t = t_2$ and so on, are found to fall on a straight trajectory, called the *Regge trajectory*, in a representation where the angular momentum is plotted against t, where $t = M^2$ in the t-channel representation. Such plots are called Chew-Frautschi plots. The Regge trajectories can be parametrized as:

$$\alpha(t) = \alpha(0) + \alpha^{'} t, \tag{1.36}$$

where $\alpha(0)$ is the intercept at $t = 0$ and α' is the slope of the trajectory. It can be shown that the total cross section is proportional to $s^{\alpha(0)-1}$.

Figure 1.26 shows the Regge trajectories of the pion and of the Reggeon for $M^2 < 6$ GeV^2, which fitted to data give $\alpha(t) = 0.0 + 0.7t$ and $\alpha(t) = 0.5 + 0.9t$, respectively. All trajectories associated with meson exchange have intercepts smaller than 1 and thus a decrease of the total cross section is expected as a function of s.

Empirically, the total cross section of a scattering process $AB \to CD$ can be described by:

$$\sigma = X^{AB} s^{\epsilon} + Y^{AB} s^{-\eta}, \tag{1.37}$$

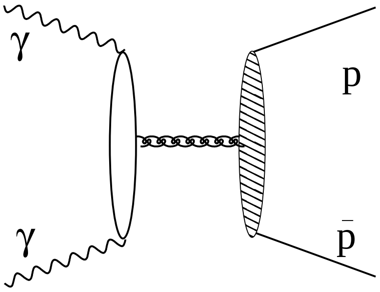


Figure 1.25: *The t-channel process of the 1-particle exchange of $\gamma p \rightarrow \gamma p$ scattering.*

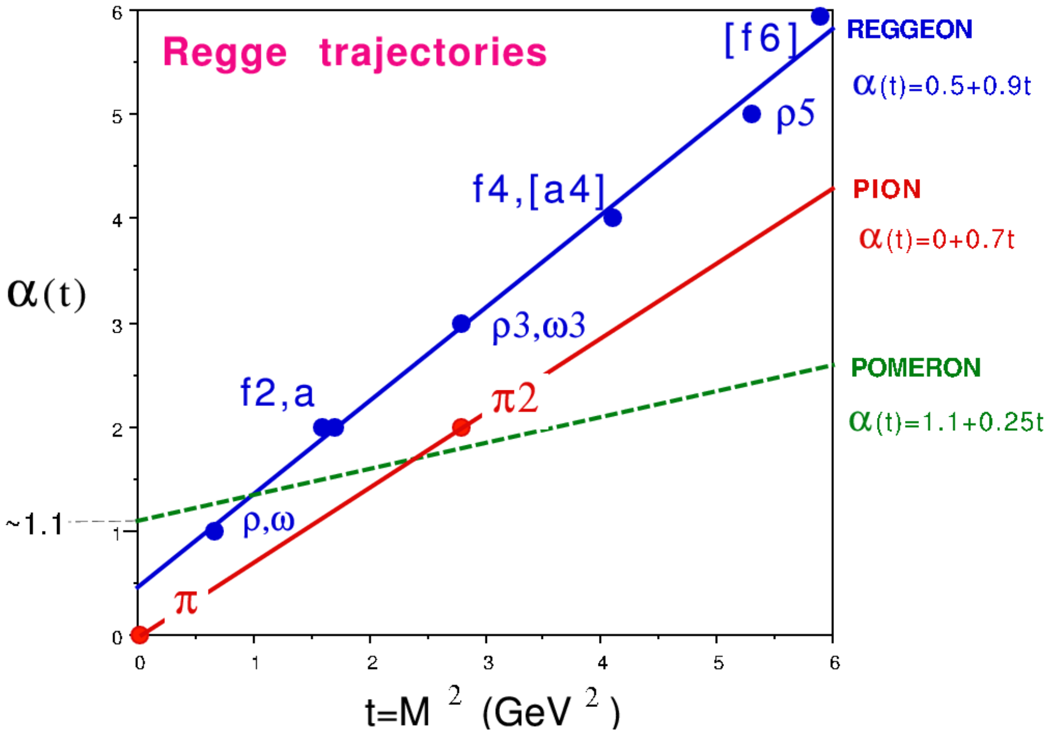


Figure 1.26: *A Chew-Frautschi plot showing Regge trajectories of vector mesons and the pomeron [50].*

where A and B denote the interacting particles and X, Y, ϵ and η are constants extracted from fits to data. Results from fits to pp, $p\bar{p}$ and γp scattering data, are all nicely described by the parametrization of Equation (1.37), as shown in Figure 1.27. It should be noted that the fit results predict universal values of the fit parameters ϵ and η, independent of the interacting particle types.

At low center-of-mass energies the s dependence of the cross section is dominated by the second term in Equation (1.37), with $\eta = 0.4525$ from the fits. This value agrees well with the intercept of the pion trajectory, since $s^{\alpha(0)-1} = s^{-\eta}$ gives $\eta = -0.4525 + 1 = 0.5475$.

However, in going to higher energies the cross section first reaches a minimum and then starts rising. In the high energy region the first term of the fit is dominating such that

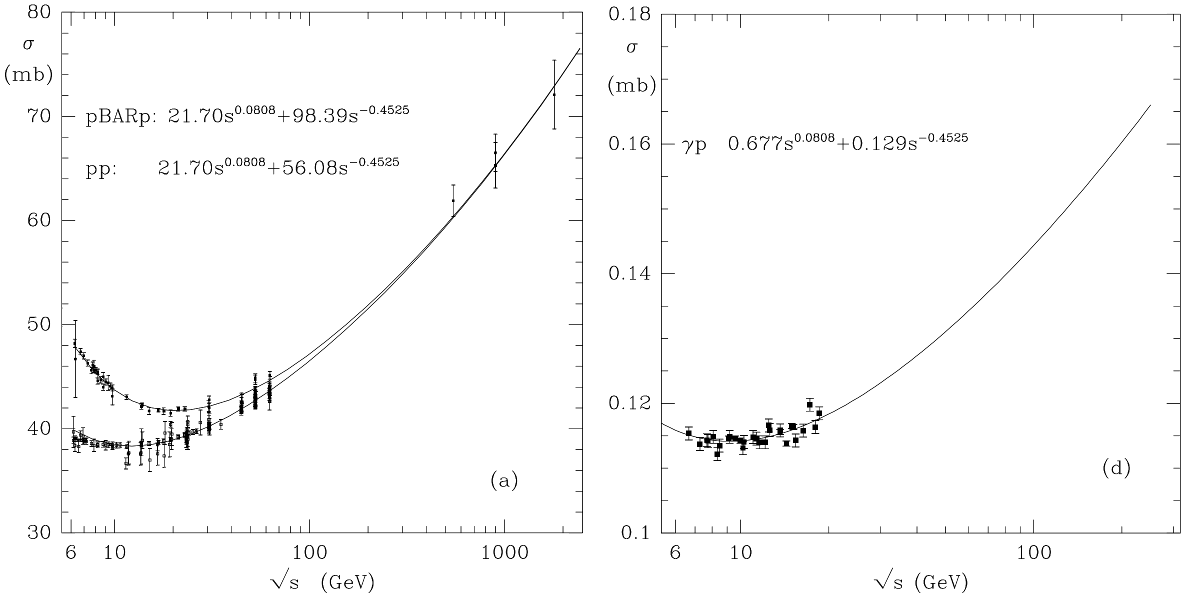


Figure 1.27: *The total cross section of pp, $p\bar{p}$ and γp scattering as a function of the center-of-mass energy. The fits according to $\sigma = X^{AB} * s^{\epsilon} + Y^{AB} * s^{-\eta}$ are shown and the resulting parameter values are given [53].*

$s^{\alpha(0)-1} = s^{0.0808}$ gives an intercept of $\alpha(0) = 1 + 0.0808 \approx 1.1$. In order to explain this behaviour of the total cross section a trajectory with the quantum numbers of vacuum with the intercept of 1.1 was introduced, called the *Pomeron* trajectory, which is also indicated in Figure 1.26. So far no particles belonging to the pomeron trajectories have been found. It is obvious that single pomeron exchange will cause the cross section to grow faster than $\ln s^2$ and thus violate unitarity. However, by introducing multiple gluon exchange this problem of the model prediction can be solved [51, 52].

1.5.2 Diffraction, Saturation and MPI in DIS

Regge theory and the optical theorem provide a powerful tool to understand how the total cross section of $AB \to CD$ scattering can be calculated and interpreted. The Regge theory is however not applicable to calculate multi-jet cross sections, for which QCD is needed.

From a closer investigation of the multi-pomeron exchange diagram, which contributes to the total cross section via the optical theorem, final states, f, with specific characteristics can be identified:

$$\frac{1}{s} Im[\ A^{\gamma p \to \gamma p}(s,t)\ ] = \sum_f \int d\Omega^f |A^{\gamma p \to f}|^2 \ . \tag{1.38}$$

In lowest order QCD the pomeron is described as two connected parton ladders, which can be described by either of the DGLAP, BFKL or CCFM evolution equations depending on the kinematic region. The specific types of final state f are recognized by applying different cuts through the exchanged ladders.

Figure 1.28 shows some amplitude diagrams of single and double ladder exchange. A cut through the amplitude diagram, as specified by the dashed line in Figure 1.28a, will split the diagram into a real part (A) and an imaginary part (A^*).

Three different ways of applying cuts to a double gluon ladder exchange are shown in the Figure 1.28(b-d). From the different possibilities of introducing the AGK cuts through these amplitude diagrams, different contributions to the imaginary part of the scattering amplitude can be identified, which through the optical theorem correspond to different contributions to the total cross section. A cut through either of the two gluon ladders in Figure 1.28b leads to single gluon exchange with the same final state as the diagram in Figure 1.28a, but including a higher order correction from the second gluon ladder. Even if the additional gluon ladder does not effect the final state multiplicity it gives a negative contribution to the cross section. A cut through two overlapping gluon ladders, as shown in Figure 1.28c, gives rise to two independent gluon chains, corresponding to multiple parton interaction. A cut between two separated gluon ladders, as shown in Figure 1.28d, corresponds to a ladder exchange leading to diffractive scattering (scattering via an exchange of a colourless object). Thus, multi-ladder exchange provides natural connections between inclusive scattering, diffractive scattering and multiple interactions via the QCD reformulation of the AGK cutting rules.

From photoproduction results ($Q^2 = 0$) at HERA evidence for soft and hard MPI [54] have been reported. The large fraction of diffractive processes also provides a strong indication of multi-gluon chain exchanges. Such events, which are usually identified by a large rapidity gap, with no particle activity close to the incoming proton direction, are only described by models where the interaction involves pomeron and Reggeon exchange. The structure function $F_2(x, Q^2)$, for example, has been measured at HERA in the transition region between the non-perturbative and perturbative domains ($Q^2 \sim 1$ GeV^2) in order to investigate its x-dependence. According to simple Regge phenomenology a $F_2(x, Q^2) \sim x^{\lambda(Q^2)}$ dependence is expected, where $\lambda = \alpha_{IP}(0) - 1$ is given by the pomeron intercept. From the results shown in Figure 1.29, a clear break in a linear dependence is observed at around Q^2 of 1 GeV^2, below which the λ-parameter becomes essentially constant at a level of 0.1. This value is consistent with the intercept $\alpha_{IP}(0) = 1.1$ of the pomeron trajectory [55, 56]. Models which are based on multiple gluon ladder exchange, provide a good description of the cross section dependence at low Q^2 and x [57–59]. However, so far there is no well established model, which can provide a detailed description of MPI in DIS. There are some calculations [60], which predict that the fraction of events with hard MPI in the kinematic region covered by HERA is about 14% at Q^2 =4 GeV^2 and 6% at Q^2 =40 GeV^2.

Properties of the UE in hadron-hadron collisions have been studied extensively in 1- and 2-jet events at the Tevatron [62, 63]. Global event variables like charged particle multiplicity and the scalar sum of the charge particle transverse momenta have been investigated as a function of the leading jet momentum in angular regions where the jets from the hard interaction have been excluded. The basic observations are that the average charged particle multiplicity and the total scalar sum of the transverse momentum grow rapidly with the leading jet transverse momentum but flattens out above 5 GeV to become independent of the leading jet transverse momentum above 50 GeV. The corresponding measurements from minimum bias events (without requiring hard jets) exhibit activities which are twice as low at the same collision energies. The experimental data have been compared to the

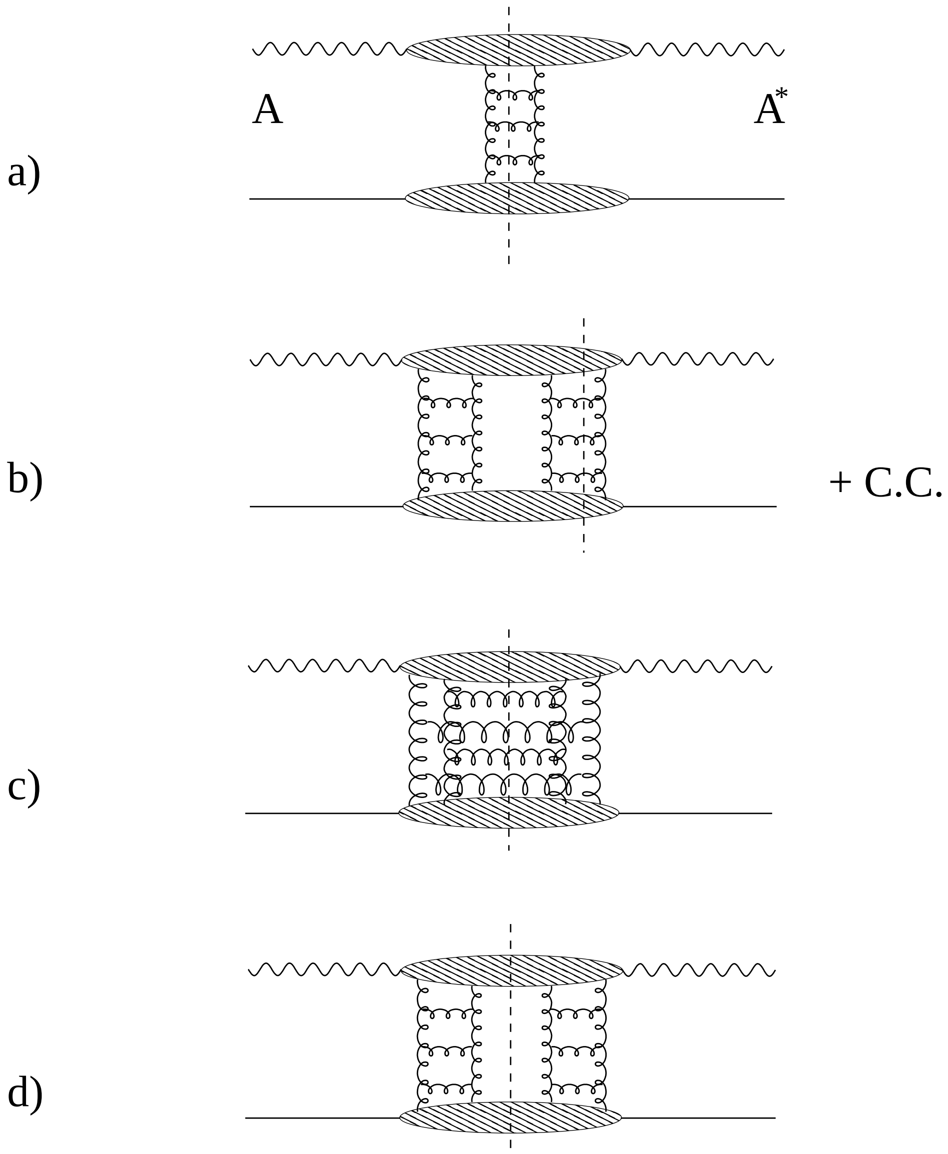


Figure 1.28: *Diagrams showing the exchange of a) single and (b-d) double gluon ladders in* γp *scattering. The dashed line illustrates various cuts through the ladders leading to different final state configurations. The cut in a) gives a single gluon exchange, whereas the cut in b) corresponds to a higher order single ladder exchange. The cut in c) represents multiple parton interaction and d) gives an exchange of a colourless object leading to diffraction.*

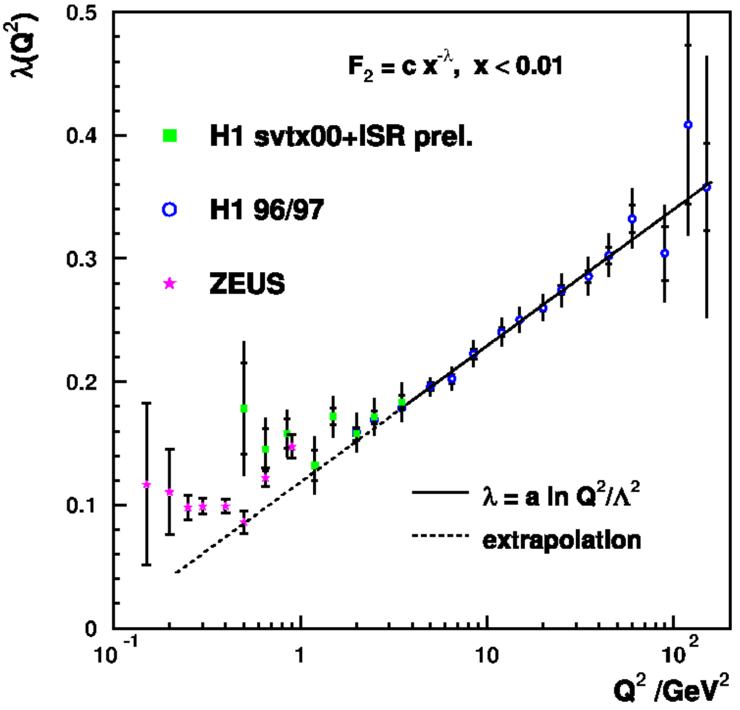


Figure 1.29: *The slope (λ) of F_2 as a function of Q^2 in the region above 0.1 GeV^2 [61].*

MC models of Pythia (with MPI) and Herwig. Only after elaborate tuning of Pythia could good agreement with data be obtained, whereas Herwig without tuning was less successful in reproducing the data.

The investigation of MPI in DIS at HERA is interesting from several aspects:

- Although evidence for MPI and diffraction events has been reported from measurements at the TEVATRON and from photoproduction data obtained at HERA there is still no established signal of MPI from DIS despite the fact that diffraction events correspond to as much as 10% of the total cross section.
- Models involving multiple gluon ladder exchange have been successful in describing data on diffractive scattering at low Q^2 from HERA. It would be interesting to investigate whether these models are also able to describe MPI data from DIS in kinematic regions which are expected to contain a significant signal.
- The dependence of the MPI rate on Q^2 and x_{Bj} can only be measured at HERA.
- Due to the asymmetric beam conditions at HERA with 27.5 GeV electrons scattering against 920 GeV protons, it can be expected that the soft interactions of MPI will be boosted along the proton direction. Thus, experimentally a separation between the soft physics and the hard scattering process might be possible. Such a separation is not possible at the TEVATRON where the laboratory system is identical to the $p\overline{p}$ center-of-mass system.

In Chapter 5 an anlysis dedicated to study MPI effects in DIS will be presented.

1.6 Monte-Carlo Generators

Frequently the main goal of an experiment is to test existing theoretical models and possibly contribute to an increased understanding of the physics involved, which can be used to improve the theory. Analytic calculations are complemented with phenomenological models where the calculations are too complicated or where perturbative methods can not be applied. Various Monte Carlo (MC) event generators use in addition to matrix element calculations different modelings of parton showering and fragmentation in order to produce the final state particles on the so called *hadron level*.

Below the MC generators used in this thesis are described briefly.

Rapgap

The Rapgap [64] Monte Carlo program was originally developed to model diffractive events from electron-proton interactions. However, it was later extended into a multi-purpose event generator also covering non-diffractive processes in ep-scattering both for photoproduction and for deep inelastic scattering. The latest version of Rapgap includes calculations of standard LO matrix element calculations of the cross sections for direct and resolved photon interactions together with initial and final state radiation obtained using the DGLAP evolution equations in the leading log approximation. In order to provide a smooth transition between direct and resolved processes the renormalization scale has been chosen to be $\mu^2 = Q^2 + P_T^2$ where P_T is the transverse momentum of the final state partons produced in the hard scattering. The hadronization is performed using the Lund string fragmentation model as given by the Jetset Monte Carlo program [65, 66]. Initial photon radiation, usually called QED radiation, is simulated via an interface to the Heracles [67] program. Rapgap does not include any model for multiple parton interactions.

Ariadne

In Ariadne [68] the CDM is implemented. In QCD-Compton processes the gluon emission is generated by a colour dipole stretched between the scattered quark and the diquark of the proton remnant. The hard scattering of the BGF process cannot be generated by the colour dipole approximation but is calculated from LO matrix element calculations. In the BGF process the emitted gluon splits up into a quark-antiquark pair, either of which interacts with the exchanged photon. Two colour dipoles are formed in the BGF event, one between the hard antiquark and a quark in the proton remnant, and the other between the hard quark and the remaining diquark system in the proton remnant. The dipoles radiate independently to create a QCD parton Cascade with no ordering in the transverse momenta of the emitted partons. Ariadne is used in combination with other MC programs to calculate the hard BGF interaction, and to describe hadronization and particle decays.

Lepto

The original Lepto [69] event generator, just like Rapgap, uses LO matrix elements to calculate the hard scattering process, followed by DGLAP evolution in the leading log approximation to account for higher order emissions and the Lund string fragmentation to describe the hadronization. In this thesis, however, Lepto is used in connection with Ariadne so that the hard scattering is calculated by Lepto and the higher order emissions are generated according to the CDM. Initial QED radiation is simulated with Heracles via an interface provided by the Django program [70].

Pythia

The Pythia generator [66,71] can provide complete events from e^+e^-, pp and ep collisions. It uses leading order α_s matrix elements supplemented by initial and final state radiation generated according to the DGLAP evolution scheme in the leading logarithm approximation. The renormalization and factorization scales are both chosen to be the transverse momentum, P_T, of the partons produced in the hard scattering, which always causes the scale to be higher than Q^2. A lower cut in the transverse momenta of the final partons from the $2 \rightarrow 2$ diagrams, $P_T \geq P_T^{min}$ is necessary in order to avoid the region where the QCD cross section of hard parton scattering diverges and to ensure that perturbative calculations can be applied. For low values of P_T^{min} it may happen that the partonic hard scattering cross section, σ_{hard}, becomes larger than the total cross section, σ_{tot}. However, if the virtual photon and the proton are regarded as colliding beams of partons, there is a certain probability for more than one interaction per beam crossing. In that respect it should be perfectly allowed that $\sigma_{hard} > \sigma_{tot}$. The interactions are assumed to occur independently of each other with the number of interactions given by a Poisson distribution. The number of additional interactions is sensitive to the P_T^{min} cut-off such that an increase in P_T^{min} decreases the number of multiple interactions. The hadronization is performed according to Lund string fragmentation model within the JETSET program. In this thesis Pythia6224 is used with the default parameter setting except for the following parameters:

- MSEL(1)= 1 and MSTP(14)= 30 to provide good mixing of different contributions of the resolved and non resolved virtual photon.
- MSTP(32)= 9 will ensure that μ_r and μ_f are always higher than Q^2, which is the case for resolved virtual photon processes.
- PARP(67)= 4 to increase the contribution from initial state radiation.

Two options are considered, Pythia without MPI, referred to as Pythia, and Pythia including MPI, denoted Pythia+MPI.

Herwig

Herwig [72] is a general purpose event generator with the options to simulate Lepton-Lepton, Lepton-hadron as well as hard and soft hadron-hadron collisions. In addition to first order α_s matrix elements it uses the parton-shower approach for initial- and final-state QCD radiation, including colour coherence effects and azimuthal correlations both within jets and between jets. Cluster fragmentation is used for the parton hadronization.

Soft additional interactions between the resolved photon and the proton remnant, in ep-scattering, can also be generated. In the language of Herwig, this is called Soft Underlying Events (SUE) and cannot be calculated in pQCD. The underlying event is implemented by superimposing on the hard emission an underlying event structure similar to that of a minimum-bias collision, where the amount of SUE has to be specified by hand. In this thesis the Herwig version 65 in the direct and resolved photon mode is used. The two cases of 0% and 10% probability for SUE are simulated in the resolved photon interactions, and are referred to as Herwig and Herwig+SUE10%, respectively.

Cascade

Cascade [73, 74] is a full hadron level generator using k_t-factorization of the cross section into an off-shell matrix element and an unintegrated gluon density function. Higher order emissions are generated by the CCFM evolution equations. In the present versions, only gluons are used in the parton Cascade. Different parametrizations of the unintegrated gluon density are available, where the parameter values have been obtained from fits to structure function data. The UPDF set 'A0' [75] is evolved using only the singular terms in the splitting function (see section 1.4.3), while for the 'J2003 set2' [76, 77] also non-singular terms are also included in the splitting function.

1.6.1 Detector Simulation

To allow for comparisons between the measured data and the predictions of theoretical models, detector effects such as resolution, acceptance, dead material and inefficient detector components, have to be considered. Therefore a full simulation of the detector performance has to be applied to events generated on hadron level by the MC programs. From the ratio of the detector simulated MC events and the generated hadron level MC data, corrections due to the influence of the detector can be obtained. These corrections are then applied to the experimental data to receive the measured signal on hadron level. Information on the performance of the various detector components have been obtained from measurements at test beams and has been implemented, together with a detailed description of the detector geometry and the material properties, into the H1SIM program [78]. The H1SIM package is based on the GEANT program [79].

Chapter 2

Experimental Setup

This chapter provides an introduction to the HERA accelerator and the H1 experiment. An overall description of the general construction of the H1 detector is given but only the detector components essential to this thesis are presented in more detail as of their status during the data taking period 1999/2000.

2.1 The HERA Accelerator

The Hadron-Electron Ring Accelerator (HERA) situated at DESY in Hamburg was originally designed to collide 820 GeV protons with 30 GeV electrons. However, as it came into operation in 1992 the electron beam energy was limited to 27.5 GeV. In the year 1998 the proton beam was upgraded to 920 GeV. After more than 15 years of successful operation HERA was closed down in the summer of 2007. Initial technical problems in storing high electron currents could be circumvented by using positrons instead of electrons. Although these problems were solved later on, the major part of the data were collected with positrons[1].

A schematic view of the accelerator facilities at DESY can be seen in Figure 2.1, which also contains a magnification of the pre-accelerators necessary for the injection of protons and electrons into HERA. Four experiment are located along the HERA ring, out of which two are collider experiments H1 and ZEUS, and the other two are the fixed target experiments HERMES and HERA-B. Their locations at the HERA ring are marked in the figure. The H1 and ZEUS detectors were designed to cover the same type of *ep*-physics, and consequently results from the two experiments can be used not only for cross checks, but also for combining data to reduce the experimental errors. The HERA-B experiment used the proton beam to collide with thin wires of Carbon in order to produce a large sample of B-mesons. The main purpose of the experiment was to measure the CP-violation in B-meson decay and to study rare B-meson decays. The HERMES experiment made use of the electron beam, which was polarized using the Siberian snake procedure, and brought to collide with polarized protons from a gas jet target. The main physics goal was to measure the spin structure of the nucleons.

---

[1] In the following, electrons are used as a generic name for both electrons and positrons

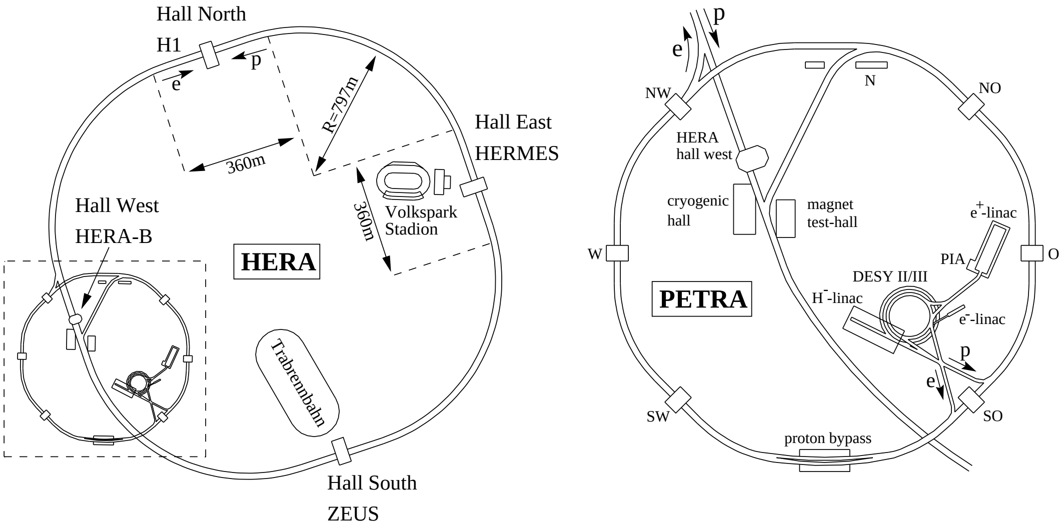


Figure 2.1: *Overview of the DESY accelerator facilities shown to the left with the pre-accelerator system magnified to the right.*

On the right hand side of Figure 2.1 the pre-accelerators used to prepare the beam particles for the HERA ring are shown. H^- ions with an energy of 50 MeV, accelerated in H^--LINAC, are passed through a stripper foil to produce the proton beams. The protons are then sent to the DESY-II accelerator in bunches. The proton bunches are accelerated to an energy of 7.5 GeV, at which point they are injected into the PETRA ring where they are accelerated further to 40 GeV. After injection into the HERA-ring, the protons are accelerated to their final energy of 920 GeV. The electron beams are produced in the $e^{\pm}$-LINAC and brought to an energy of 500 MeV. In DESY II the electron bunches are accelerated to 7 GeV and in PETRA to 12 GeV. Finally, the electrons are accelerated in HERA to 27.5 GeV.

The HERA collider consists of two separate storage rings, one for the electrons and one for the protons. It is placed in a tunnel of about 6.4 km circumference, 10-15 m under the ground level. The large difference in beam energies which can be reached for protons and electrons is related to the large difference in mass between protons and electrons. A charged particle which is forced to deviate from a straight path will lose energy by emitting photons, called synchrotron radiation. In a circular ring with radius R, the radiative energy loss per turn, ΔE, of a charged particle with energy E and mass m is:

$$\Delta E = \frac{E^4}{m^4 R}$$

Acceleration fields are provided by radio frequency (RF) cavities and as long as the energy delivered by the cavities is larger than the energy lost from synchrotron radiation, the particles will be accelerated to higher energies. Strong acceleration fields have been obtained by using superconducting RF cavities. Due to the $1/m^4$ dependence electrons will suffer a lot from synchrotron radiation, which will be the limiting factor for its maximum energy, whereas for protons the radiative losses are negligible. Instead the maximum proton beam energy is dictated by the strength of the magnetic field, that is needed to keep the protons in their orbit. This is given by the relation:

$$p = B \cdot e \cdot R,$$

where p is the particle momentum, B is the magnetic field strength, e is the electric charge of the particle and R is the bending radius. In HERA superconducting magnets have been used, providing a field of 4.2 T. From 1998 the magnetic field of the bending magnets could be raised to 4.7 T leading to an increase in the proton energy from 820 GeV to 920 GeV. The proton and electron energies of 920 GeV and 27.5 GeV, respectively, were used during the 1999-2000 run period, and this corresponds to a centre-of-mass energy $\sqrt{s} \approx 318$ GeV.

At HERA the time between two bunch crossings is 96 ns and the transverse size of the bunches is $\sigma_x \approx 300$ μm and $\sigma_y \approx 80$ μm for the protons and $\sigma_x \approx 300$ μm and $\sigma_y \approx 40$ μm for the electrons. The luminosity L is defined as:

$$L = \int f \frac{N_1 N_2}{A} dt$$

were A is the transverse area of the collision, N_1 and N_2 are the number of particles in each bunch and f is the collision frequency of the particles. In the years 1999 and 2000 the delivered luminosity was 45.045 pb^{-1} and 67.889 pb^{-1} respectively.

2.2 The H1 Detector

In Figure 2.2 a schematic view of the H1 detector is given. The electrons enter the detector from the left hand side in the figure and the protons from the right. The reference frame of the H1 detector is described by a Cartesian coordinate system with its origin placed at the nominal interaction point. The proton beam direction defines the positive z direction. The y-axis points upwards and the x-axis is directed towards the center of the HERA ring. The polar angle θ is the angle with respect to the positive z-axis. Thus, for an unscattered electron, θ is equal to 180^o. The region of small θ-angles is referred to as the forward region. The azimuthal angle ϕ is measured with respect to the positive x-axis in the xy-plane

The central region of the H1 detector, closest to the nominal collision point, is surrounded by silicon detectors. The silicon trackers provide excellent position resolution which gives a precise reconstruction of particle tracks. Thus, they are used to give an accurate determination of the interaction vertex and are able to resolve secondary vertices from short lived particle decays. The central $\boxed{2}$ and forward tracking $\boxed{3}$ detectors provide together with the silicon detectors measurements of charged particle trajectories. The tracking detectors are enclosed by a calorimeter system consisting of a Liquid Argon electromagnetic $\boxed{4}$ and hadronic $\boxed{5}$ calorimeter, covering the forward and central regions, complemented by an electromagnetic and hadronic Spagetti Calorimeter $\boxed{12}$, in the backward region. The liquid argon calorimeter is surrounded by a superconducting coil $\boxed{6}$ providing a magnetic field of ~ 1.15 T parallel to the beam axis. Charged particles will follow a helical path in the magnetic field and from the measured trajectories the particle momenta can be determined. To eliminate the influence on beam particles from small

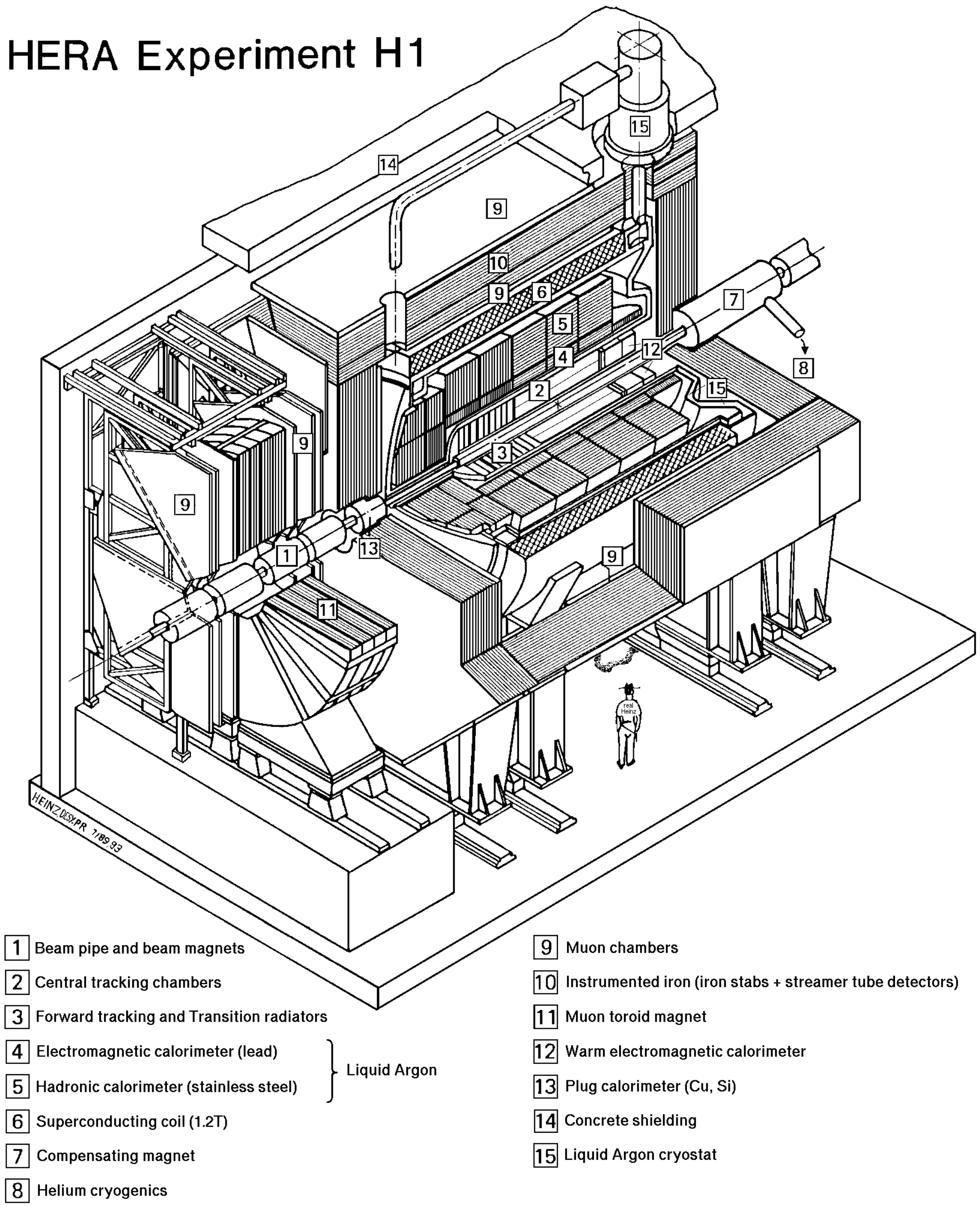


Figure 2.2: *Schematic cross-section of the H1 detector.*

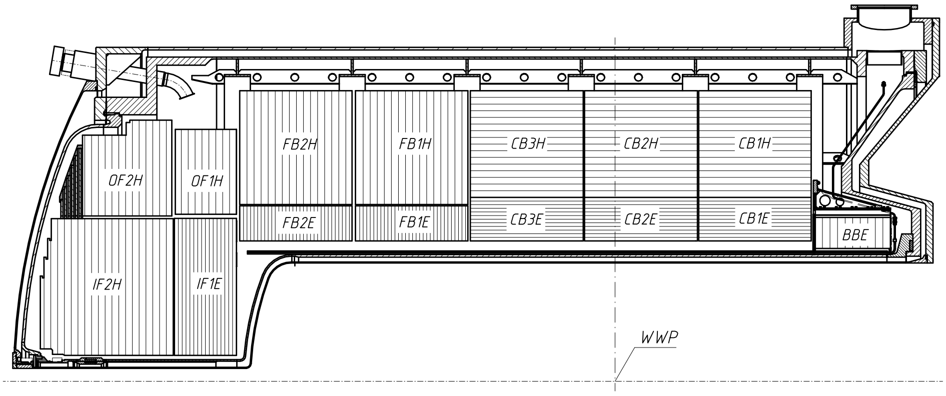


Figure 2.3: *A side view of the electromagnetic and the hadronic part of the LAr calorimeter, showing the segmentation into calorimeter wheels.*

transverse magnetic field components in the solenoidal field, an additional magnet close to the beam $\boxed{7}$ is used for compensation. The magnetic field from the superconducting coil is closed by the iron return yoke, which consists of thick iron slabs interleaved with streamer chambers. The instrumented iron $\boxed{10}$ is used to measure energy not deposited in the calorimeter system. The outside of the return yoke is covered by muon chambers $\boxed{9}$. Muons going in the forward region are detected and their momenta are measured by a muon spectrometer, consisting of drift chamber planes in front of and after a toroidal magnet $\boxed{11}$.

The main components of the H1 detector relevant for this analysis are the electromagnetic and hadronic calorimeters, the tracking system, the Time-of-Flight counters (section 2.2.3) and the trigger system. These are described in some detail below. A complete description of the H1 detector can be found in [81, 82].

2.2.1 Calorimeters

There are four calorimeters in the H1 detector which provide energy measurments of particles produced in the collisions: the Liquid Argon (LAr) calorimeter, the SPACAL calorimeter, the PLUG and the Tail Catcher (TC). The LAr calorimeter has the widest coverage for energy measurement of electromagnetically interacting particles and hadrons. The SPACAL is mainly used to measure the energy of electrons scattered at small angles for the determination of the event kinematics, but it also has a hadronic section. The TC complements the LAr calorimeter by providing a rough measurement of hadronic energy, which is not completely deposited in the LAr calorimeter. The TC is only important for jets with very high energies (above 120 GeV). The PLUG is used to measure particles in the extreme forward region $0.6^o < \theta < 3^o$. The TC and the PLUG are not used in this analysis.

The LAr calorimeter is described in some detail since on the one hand it is especially important for the analysis presented in this thesis and on the other hand the development of a new calibration method for jet measurements is part of the thesis work.

The Liquid Argon Calorimeter (LAr)

The LAr calorimeter [81, 82] covers the polar angle range $4^\circ < \theta < 154^\circ$ and consists of two sections: an inner electromagnetic section and an outer hadronic section. It is also segmented along the z-direction into eight wheels (see Figure 2.3), called the Forward Barrel calorimeters (FB1, FB2), the Central Barrel calorimeters (CB1, CB2, CB3), the Inner Forward calorimeters (IF), the Outer Forward (OF) hadronic calorimeters and the Backward Barrel Electromagnetic calorimeter (BBE). The wheels consist of eight identical octants in ϕ (see Figure 2.4). The LAr calorimeter is of sandwich type, which means that it contains absorber plates and LAr as the active material. The purpose of the absobers is to make the incoming particles interact so that a shower of energy degraded particles are produced. These may ionize the active material and the charge produced is a measure of the energy contained in the shower. The LAr calorimeter is segmented further into 45 000 calorimeter channels (cells), and the cells are combined into 256 trigger towers (Big Towers). The energy collected in the Big Towers is used by the H1 trigger system to select jet events. The electronic noise fluctuations correspond to an energy variation of between 10 and 30 MeV per cell on the average. Various noise suppression algorithms have been applied in order to minimize the influence of the electronic noise. However, from studies of MC simulated DIS events, it has been shown that the energy loss due to electronic noise suppression can be as large as 17% of the total signal [82] at low Q^2. After application of electronic noise suppression and some additional corrections, the charge collected in the calorimeter cells is converted into energy by a reconstruction algorithm [80], as will be described later

Electromagnetic calorimeters are optimized to measure the energies of photons and electrons (and positrons) through the production of electromagnetic showers, from interactions of the particles with the calorimeter material. The showers are produced by repeated bremsstrahlung and pair production processes until the shower particles have reached a critical energy, below which the photons produce electrons via the photoelectric effect or by Compton scattering and the electrons will lose their energy in ionization processes. The longitudinal extension of the electromagnetic shower depends strongly on the energy of the incident particle ($\propto ln(E)$), and the Z-value of the material used in the absorber plates. The transverse shape of an electromagnetic shower is given by the opening angles of the e^+e^--pairs and the emission angle of the bremsstrahlung photons. In the low energy part of the shower the multiple scattering of the electrons and positrons dominates, giving a transverse structure containing a high energy core and a low energy halo. The depth of an electromagnetic calorimeter is usually specified in radiation lengths (X_0), with 1 X_0 being the average length over which the energy of a particle is reduced by a factor 1/e. The absorbers in the LAr electromagnetic calorimeter of H1 are made of 2.4 mm thick lead plates, which for the full depth of the electromagnetic part of the calorimeter corresponds to a total of 20-30 radiation lengths depending on the orientation of the absorber plates and the impact angle of the particle. The orientation of the absorber plates in the different wheels is indicated by the vertical and horizontal lines in Figure 2.3. The reason for the different orientations is that the impact angle for particles, produced in the nominal collision point of HERA, should always be larger than 45^o. This is to minimize the variation in absorber material seen by particles and thereby a reasonably uniform energy resolution is obtained.

As mentioned above the experimentally determined energy is derived from the charge produced by the ionization of the LAr. The electromagnetic calibration constant, which converts the collected charge to visible energy, has to be determined from measurements in test beams of well defined energies. The electromagnetic scale E_0^i for cell i is defined as

$$E_0^i = C_{exp} \cdot Q^i \tag{2.1}$$

where C_{exp} is the experimentally determined calibration constant and Q^i is the charge deposited in cell i. More details can be found in [84]. The calibration factors for the electromagnetic stacks were obtained from the electron test-beams at CERN and MC simulations. For the hadonic stacks, additional MC simulations of electrons detected only by the hadonic stacks were used to scale the factors obtained from the electromagnetic stacks [80].

Hadronic showers are produced from strong interaction processes of the hadrons with nuclei in the calorimeter absorber plates. The transverse shape of a hadronic shower is much broader than that of an electromagnetic shower, and this can be used to distinguish between the two types of shower. The depth of the calorimeter, which is needed to fully contain a hadronic shower is much larger than what is needed to have a fully contained electromagnetic shower. For hadronic calorimeters the depth is given in terms of interaction lengths, λ, where 1 λ is the average length before a particle undergoes a nuclear interaction. The hadronic stacks in the LAr hadron calorimeter are made of 19 mm thick stainless steel plates. The total longitudinal thickness of the hadronic calorimeter sections varies between 5-8 interaction lengths depending on the orientation of the absorber plates and the polar angle of the incoming hadron.

Hadron showers contain not only charged hadrons but also π^0 and η, which decay mainly into two photons, and leptons produced in the hadronic interactions. Therefore, a hadronic shower can be divided into an electromagnetic component and a hadronic component. A significant fraction of the hadronic component is not measurable like the recoil energy of the atoms and soft neutrons, which get lost in the absorber plates. Muons produced from meson decays are not measured in the calorimeter and neutrino particles escape detection completely. Also the energy required to break up atomic nuclei will not be measured by the calorimeter. The energy of the electromagnetic component of the hadronic shower increases with the incident particle energy and may get larger than the hadronic component. Since a much larger fraction of an electromagnetic component is measured in the detector than is the case for the hadronic one, the difference in visible energy has to be corrected for. In the LAr calorimter 30% of the energy deposited by the hadronic shower is not measured.

Equalization of the energy response for electromagnetic and hadronic components can be obtained either by hardware compensation or, if the granularity of the calorimeter is high enough, by software compensation. In hardware compensation the method is to decrease the sensitivity to the electromagnetic component and simultaneously increase the sensitivity to the hadronic component. By choosing a high Z material like Uranium a larger fraction of the electromagnetic component is absorbed in the Uranium and consequently the visible energy will decrease. On the other hand neutrons in the hadronic component will produce spallation of the Uranium nuclei, giving products that will deposit their energy in the active material. This obviously increases the sensitivity to the hadronic component of the shower. Another possibility of hardware compensation is to use thin absorber plates, so that a significant fraction of the low energy neutrons escape into the active material. The interaction of the neutron in the active material will produce recoil protons which contribute to the calorimeter signal. By carefully tuning the thickness of the absorber and the active material full compensation can be achieved. A high Z material has to be chosen in order to keep the total depth of the calorimeter within limits. Although the energy resolution for hadrons will improve using hardware compensation, the energy resolution for electromagnetically interacting particles is deteriorated.

A different approach is to apply software compensation which can be done for calorimeters of high granularity, where the signals of the detector cells can be weighted in such a way that the response from electrons and hadrons depositing the same energy is equalized. During the event reconstruction, all cells from the electromagnetic and hadronic stacks are calibrated, using equation (2.1), to the electromagnetic scale. The method exploits the fact that local energy deposits of high density are mainly of electromagnetic origin. Thus, in a calorimeter with high granularity the amount of energy deposited in the cells can be used for statistical separation of the two components. Cells from the hadronic and the electromagnetic calorimeters are combined into groups, clusters, by an Electron/Pion separation method implemented in a reconstruction algorithm [83]. The purpose of the clustering algorithm is to collect cells into groups related to particle showers. Clusters not classified as electromagnetic are used for the clustering of the hadronic shower [80,82]. Showers from hadrons consist of many clusters due to the properties of the hadronic shower. Correction for energy losses due to the dead material in front of the calorimeter, calorimeter stacks and electronic noise suppression are done by the reconstruction algorithm. As mentioned above, in order to find the proper hadronic scale, which equalizes the response to the electromagnetic and the pure hadronic components of a hadron shower, a software weighting technique is applied to the cells which have been classified by the Electron/Pion separation method as belonging to hadronic clusters. In the reconstruction the weighted energy in a cell i, E^i_{rec}, belonging to a hadronic cluster is calculated from the cell energy corrected to the electromagnetic energy scale, E^i_0 , according to:

$$E^i_{rec} = [a_0 + a_1 \cdot exp(-\alpha E^i_0/V^i)]E^i_0 \tag{2.2}$$

where a_0, a_1 and α are the parameters of the weighting function and V^i is the volume of the cell. Note that the parameters are different for the electromagnetic and hadronic calorimeter parts. After the compensation the hadronic clusters are on the correct scale, i.e. the hadronic scale. With equal response the influence of large fluctuations in the hadronic shower composition on the energy reconstruction is suppressed.

The fine granularity for e/π separation, homogeneity of the signal response and the stability are some of the main reasons for choosing a sampling calorimeter with a liquid argon technology. The energy resolution of $\sigma_{em}(E)/E \approx$11% / $\sqrt{(}E/\mathrm{GeV}) \oplus$ 1% for the electromagnetic section and $\sigma_{had}(E)/E \approx$50% / $\sqrt{(}E/\mathrm{GeV}) \oplus$ 2% for the hadronic section was obtained from the test beam measurements at CERN.

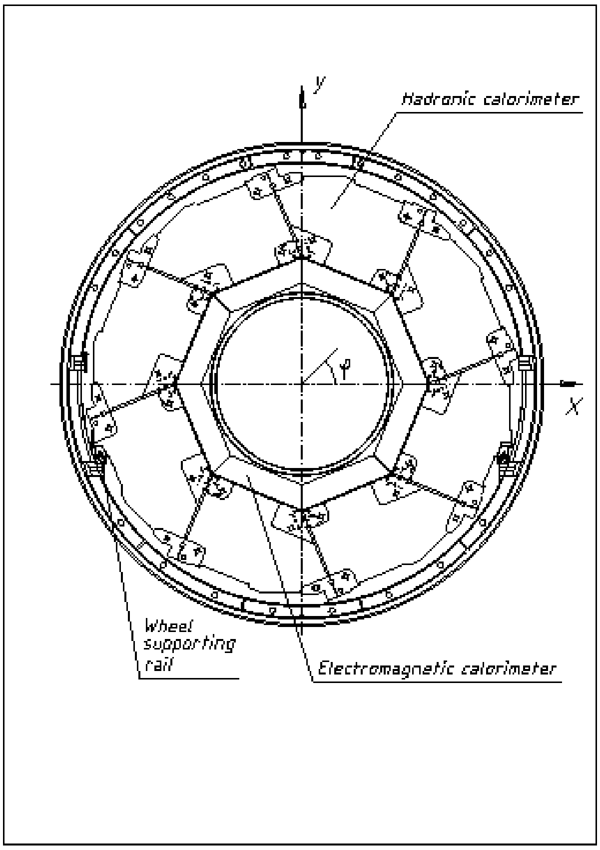


Figure 2.4: *A transverse view of a LAr calorimeter wheel.*

The SPACAL

During the shutdown in 1994/1995 the Backward Electromagnetic Calorimeter, BEMC, was replaced by the Spaghetti Calorimeter (SPACAL) [85–89]. It was designed to have a large geometric acceptance in the backward region, high electromagnetic energy and position resolution, a good pion-electron separation and good time resolution. The main purpose of the SPACAL is to accurately measure the energy and polar angle of the scattered electron from which the kinematics of DIS processes in the region of $0.4 < Q^2 < 150$ GeV^2 can be calculated. In this Q^2 region x_{Bj} values as low as 10^{-5} can by reached.

The SPACAL calorimeter consists of an electromagnetic section and a hadronic section as shown in Figure 2.5. The SPACAL is a sampling calorimeter with lead as absorber and scintillating plastic fibers as active material. Particles interacting with the absorber produce particle showers which cause scintillations in the plastic fibers. The light is collected and conducted to the photomultiplier tubes.

The scintillating fibres, with 0.5 mm diameter in the electromagnetic section and 1 mm diameter in the hadronic section, are embedded in lead. The electromagnetic section of the SPACAL contains 1192 cells, where 2 cells constitute one construction unit, as indicated by the thin lines in Figure 2.6. Eight two-cell units are grouped to form a supermodule, confined by the thicker lines in the figure. A VETO layer close to the beam pipe should only contain a limited energy deposition to guarantee a good reconstruction of the scattered electron energy (see also Chapter 3.3.3). The pupose of the hadronic SPACAL is to measure the energy leakage from the electromagnetic section and, in combination with it, determine the hadronic energy flow in the backward region. The hadronic section has a depth of approximately 1 λ and contains 128 modules in total. The coverage in polar angle of the SPACAL is $153^o < \theta < 177.5^o$.

A resolution of $\sigma_{em}(E)/E \approx 7\% / \sqrt(E/\mathrm{GeV}) \oplus 1\%$ for the electromagnetic sections and $\sigma_{had}(E)/E \approx 56\% / \sqrt(E/\mathrm{GeV}) \oplus 7\%$ for the hadronic section was measured in test beams at CERN and DESY.

2.2.2 The Tracking System

Figure 2.7 gives a schematic side view of the H1 tracking system, which can be subdivided into three parts: the central tracker, the forward tracker and the backward drift chamber.

The Forward Tracker

The Forward tracker [82] consists of three sections. Each section contains three planar drift chambers, a multi-wire proportional chamber, a passive transition radiation detector and a radial drift chamber. The forward tracker covers the angular region $5^o < \theta < 30^o$. The planar chambers consist of parallel drift cells, each having four wires. The three chambers are rotated by 60^o in ϕ with respect to each other. The radial chamber consists of 48 drift cells, with each cell containing 12 radially strung wires separated by 10 mm in the z direction. The multi-wire proportional chamber consists of tree chambers, which are used to provide a fast trigger for the forward tracks. The combined information from the

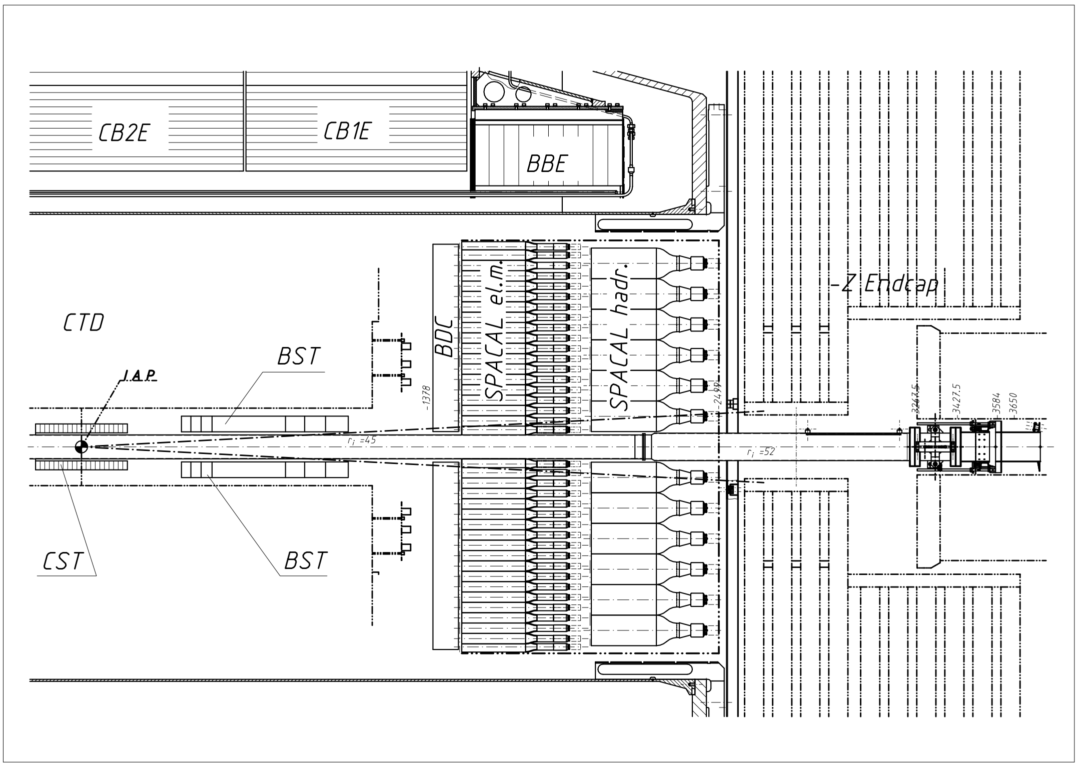


Figure 2.5: *A schematic side view showing the backward part of the inner detector with the BDC immediately in front of the electromagnetic and hadronic SPACAL.*

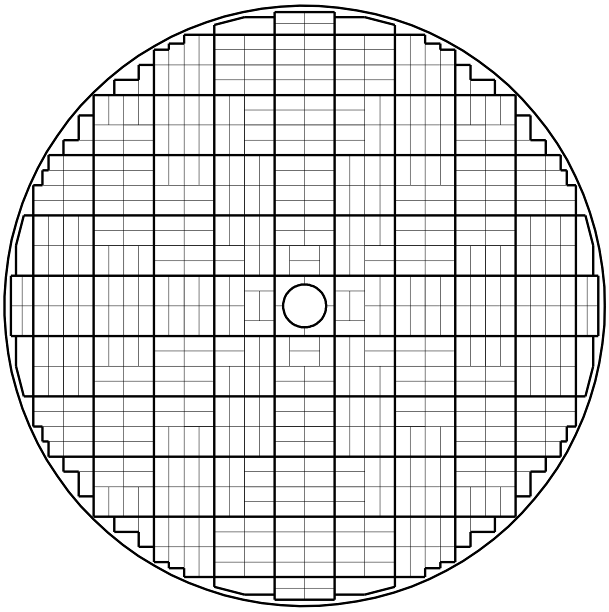


Figure 2.6: *The transverse (x,y) view of the electromagnetic SPACAL, where the thin lines indicate construction units consisting of two calorimeter cells, whereas the thicker lines mark the supermodules, containing 8 construction units.*

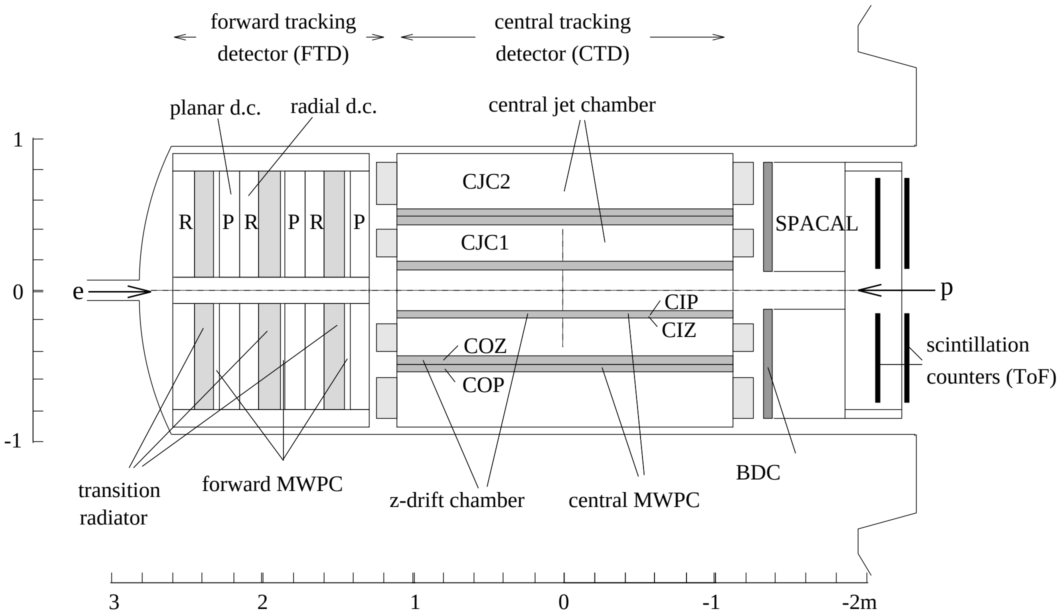


Figure 2.7: *A side view of the H1 tracking system with the different tracking chambers indicated.*

planar chambers, the radial chamber and the multi-wire proportional chamber provides the curved trajectories of forward going particles in the solenoidal field, from which the particle momenta can be calculated. It was designed to achieve a momentum resolution $\sigma_p/p^2 < 0.003\ \mathrm{GeV}^{-1}$ and an angular resolution $\sigma_{\phi,\theta} < 1$ mrad.

The Central Tracker

The central tracking system consists of the Central Silicon Tracker (CST), shown in Figure 2.5, the inner (CJC1) and outer (CJC2) Central Jet Chambers (CJC), the Central Silicon Tracker (CST), the Central Inner and Outer Z-chambers (CIZ/COZ) and the Central Inner and Outer Proportional chambers (CIP/COP) which are shown in Figure 2.7.

The CST [91, 92] is located nearest to the interaction point. It provides vertex information from precision measurements of charged particle tracks close to the interaction point. Including the precise CST $r\phi$-hits in the track reconstruction of the wire chambers improves the transverse momentum resolution.

CJC1 covers the angular range $11^o < \theta < 169^o$ and contains 30 drift cells with 24 sense wires each, whereas CJC2 covers $26^o < \theta < 154^o$ and contains 60 drift cells with 32 sense wires each (see Figure 2.8). Electrons released through ionization of the CJC gas by charged particles will drift towards the anode wires along the electric field between the anode- and the cathode-wires. From the known drift velocity of electrons in the gas and a measurement of their arrival time at the anode wire, given by the appearance of a signal, the position of the ionization point can be reconstructed in the $r\phi$-plane. A space resolution of $\sigma_{r\phi} = 130\ \mu$m per measuring point has been achieved. A large number of such measurements are made in order to reconstruct the particle trajectory. The z-position of a hit can be reconstructed using the charge division method, which makes use of the

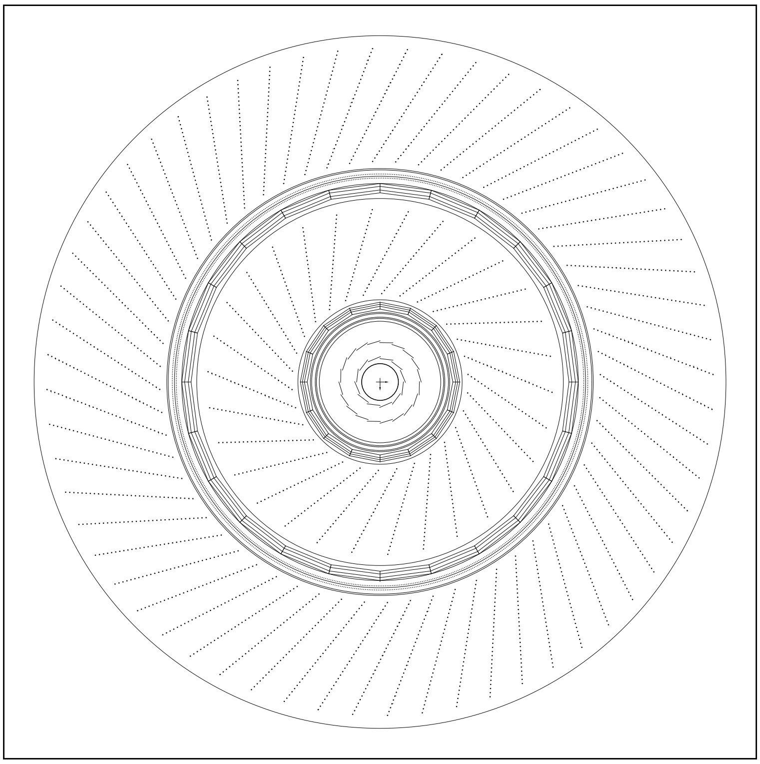


Figure 2.8: *A radial view of the central tracking system. The 32 and 60 sense wires in CJC1 and CJC2, respectively, is shown.*

fact that the amplitude of the signal is damped as it travels through the anode wire. By comparing the amplitudes of the signals at both ends of the signal wire the z-coordinate can be determined with a resolution of $\sigma_z = 22$ mm.

The CIZ and COZ drift chambers have wires strung perpendicular to the beam directions and are used to provide precise z-coordinates of the tracks. The CIZ chamber is located inside the CJC1, and the COZ chamber in between the CJC1 and CJC2. The polar angles covered by the CIZ and COZ chambers are $16^o < \theta < 169^o$ and $25^o < \theta < 156^o$, respectively. The information from these chambers improves the z-resolution by two orders of magnitude compared to a measurement with the CJC alone.

The CIP and COP Multi-Wire Proportional Chambers (MWPC), which are located close to the CIZ and COZ, respectively, have a time resolution better than the 96 ns separation between two HERA bunch crossings. By combining hits in these chambers a fast decision can be made on whether the tracks are produced in the vertex region or not. This information is used to select collision events on the first trigger level, which is discussed in more detail in section 2.2.4.

The Backward Drift Chamber

The Backward Drift Chamber (BDC) [82] consists of four double layers of drift cells, covering an angular range of $153^o < \theta < 176^o$. The layers are subdivided into octants, with the drift cells oriented so as to provide an accurate measurement of r i.e. the distance to

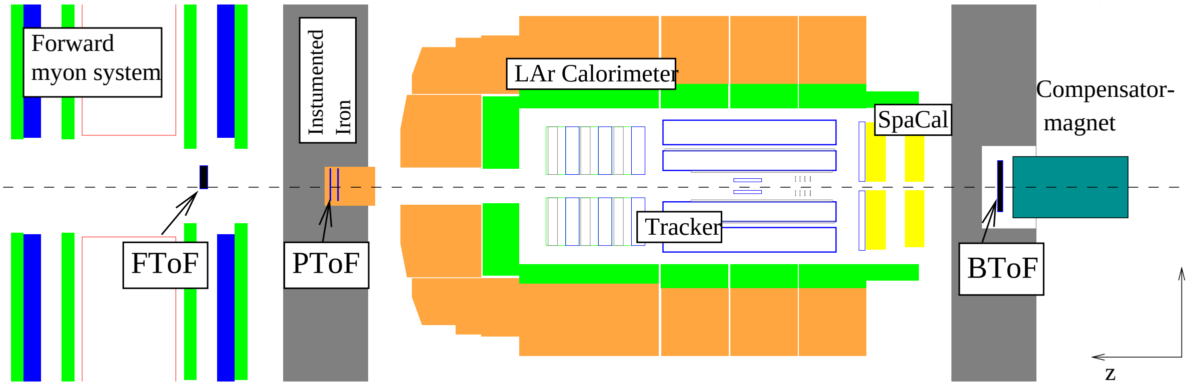


Figure 2.9: *A side view of the H1 detector with the TOF system indicated.*

the beam axis. Each double layer is rotated by 11.25^o with respect to the neighbouring ones, allowing for an extraction of the ϕ coordinate of a track. The BDC was designed for both triggering purposes and to deliver short track segments of particles entering the SPACAL. In this analysis the BDC and SPACAL information is combined to identify and reconstruct the scattered electron and also to measure hadrons in the backward region.

2.2.3 The Time-of-Flight System

The Time-of-Flight (ToF) system [82] consists of scintillator layers placed at various distances from the interaction point. The positions of the different scintillator detectors composing the ToF system are shown in Figure 2.9. They are the BToF (backward ToF) at z=-275 cm, the PToF (Plug ToF) at z=+540 cm and the FToF (forward ToF) at z=+790 cm. In addition there are two layers of Veto scintillator walls, not shown in the figure. The big Veto wall at z=-810 cm covers essentially the entire surface of the H1 detector, whereas the small Veto wall is positioned close to the beam pipe. The system is calibrated such that the ToF measured in detectors upstream and downstream should give a difference that is close to zero for particles produced in the interaction region, whereas larger time differences would come from particles produced at some distance from the nominal collision point by interactions of the beams with the walls of the beam pipe or the residual gas in the vacuum tube, called beam-wall and beam gas events, respectively. The ToF resolution of 1 ns has resulted in a definition of the vertex region given by $-35 < z_{vtx} < +35$ cm.

2.2.4 The Trigger System

The purpose of the trigger system in H1 [81] is to select physically interesting events out of a large sample in a very short time. Only a small fraction of the total events is selected as *ep* collision events. The main backgrounds in the H1 detector originate from beam-gas and beam-wall events, synchrotron radiation and cosmic rays. Events considered to be uninteresting are also rejected by the H1 trigger system. The trigger system consists of four on-line trigger levels (L1-4) and one off-line (L5) trigger level. The beams cross every 96 ns which corresponds to a beam crossing frequency of 10 MHz. The event rate presented to the first trigger level (L1) is 100 kHz. This has to be decreased to 1-10 Hz after the fourth trigger level (L4) in order to limit the data stored for later analysis to a reasonable level.

At the first trigger level, the information from each part of the detector is stored in buffers (pipelines) to avoid dead time until the L1 trigger has taken a decision, which occurs within 2.3 μs, corresponding to 25 bunch crossings. An event is selected by the L1 trigger if the event fulfills at least one of the 128 subtrigger requirements. The subtriggers built out of logical combinations of 256 trigger elements obtain their information from different components of the H1 detector. If a subtrigger condition is fullfilled then the "raw" bit is set (true). Subtriggers with high counting rates are down-scaled (prescaled) with respect to those with lower counting rates. Thus, the actual number of events passing the subtrigger will be smaller than the true number of events fulfilling the trigger conditions. The "actual" bit will be set for the subtrigger if the event is selected. If the "actual" bit of at least one of the subtriggers is set then the event is transferred to the second level trigger, L2.

The input rate of the L2 trigger is about 1 kHz. The L2 trigger consists of Neural Network (L2NN) and Topological (L2TT) triggers. The L2TT uses stored topology information of the background mapped into a grid in ϕ and θ. L2 provides a decision in about 20 μs. If the event is accepted by L2 it is sent to the fourth level trigger. During the 1990-2000 run period the L3 trigger was not implemented.

At the fourth trigger level an online event reconstruction is performed and the events are classified into different physical classes. All events assigned to physics classes are kept.

2.2.5 The Luminosity System

The number of events, N, from a specific process with the cross section σ can be extracted from the relation $L = N/\sigma$, where L is the integrated luminosity measured in a time interval over which the analysis is performed. The so called integrated luminosity is obtained from $L = \int \mathcal{L}\ dt$, where $\mathcal{L}$ is the instantaneous luminosity, measured in units of cm^{-2}sec^{-1}. At HERA the instantaneous luminosity is determined from the measured rate of Bethe-Heitler events, $ep \rightarrow e^{'}p\gamma$, the cross section, of which has been calculated to a high degree of accuracy [92]. Thus, by measuring the number of events from a specific process and knowing L for the measuring period, the cross section of that process can be calculated. The luminosity system [81] consists of the electron tagger located 33.4 m from the nominal interaction point in the direction of the electron beam and the photon detector, which is placed at z = -102.8 m. A Bethe-Heitler process is recognized from a coincidence between signals in the photon and electron luminosity detectors.

Chapter 3

Event Reconstruction and DIS Selection

The analyses presented in this thesis has used data collected by the H1 detector in the run periods of the years 1999-2000 and 2006-2007. The kinematics of DIS events are determined from measuring the polar angle and the energy of the scattered positron (the electron method). In this chapter the reconstruction of the electron will be described as well as the event selection based on various phase space cuts. The resolution of the reconstructed event variables is discussed. Furthermore, the construction of combined objects of calorimeter clusters and particle tracks, which are used in the jet finding, are explained. Finally, the jet finding algorithm is described.

3.1 Run Selection

A run is defined as a collection of events measured during over a period of time during which the experimental conditions remained approximently the same. Runs with fully operational detector components relevant to the studies presented in this thesis are selected. A fully operational detector means that the detector has been operated at its nominal high voltage. The detectors important for this analyis are the SPACAL and LAr calorimeters, the CJC, forward tracker, CIP, COP and BDC tracking chambers, the TOF counters, the trigger system and the luminosity system. Runs with low luminosity and poor beam conditions have been excluded. The total number of events which fulfil the selection criteria during the 1999-2000 run period corresponds to a luminosity of 64 pb^{-1}.

3.2 The Event Vertex

To reduce the contribution from beam-gas and beam-wall events the reconstructed z-position of the measured interaction point, is required to be close to the nominal collision point. The nominal z-vertex defines the origin of the H1 coordinate system, at which the center of gravity of the electron and proton bunches are tuned to collide. The proton beam

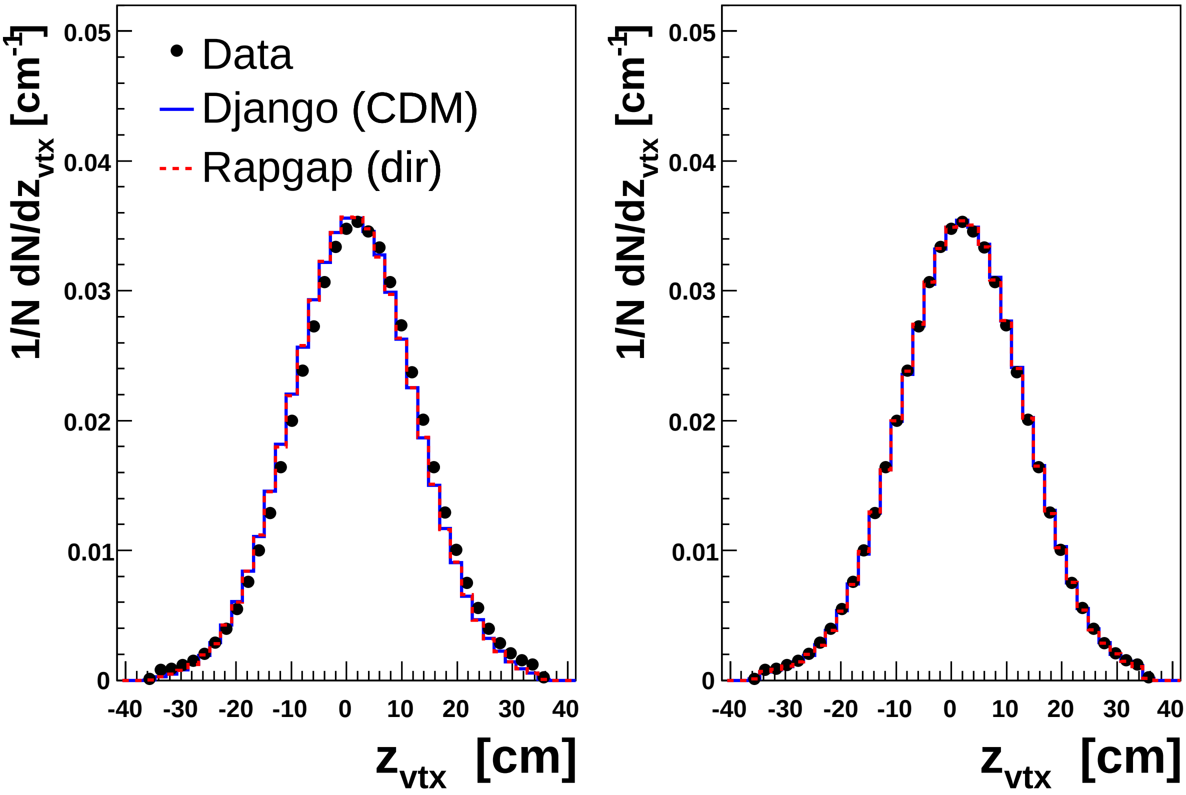


Figure 3.1: *The measured z-vertex distribution before (left) and after (right) the z-vertex weights have been applied.*

may contain so called 'satellite bunches', which are created during the injection from PETRA to HERA. These satellite bunches are shifted by a few nanoseconds with respect to the main proton bunch and consequently collisions between the positrons and such satellite bunches will have z-vertices shifted compared to the nominal collision point. The z-profile of a bunch has a Gaussian shape and the typical length of a proton bunch in terms of σ is 11 cm whereas the typical length of an electron bunch is 1 cm. The z-vertex is required to lie within ±35 cm from the nominal interaction point. In the left hand plot of Figure 3.1 the experimentally obtained vertex distribution for a specific run is compared to the one generated by the H1 detector simulation program. It is observed that the simulated distribution does not agree with the data. This is due to the fact that the position of the actual collision point is normally not exactly at the nominal vertex position and it is also not the same from run to run. It actually depends on how the beam conditions from individual beam fills are optimized and such variations have not been accounted for in the detector simulation. In order to take a shift of the measured vertex distribution into account the Monte Carlo events are reweighted to describe the z-profile. After the reweighting the agreement is excellent as can be seen in the right hand plot of Figure 3.1b.

3.3 The Electron Identification

The scattered electron is identified by requiring a shower in the electromagnetic part of the SPACAL, connecting to a charged particle track. The energy measurement of the scattered electron is provided by the electromagnetic SPACAL calorimeter in the kinematic region $5 < Q^2 < 100$ GeV^2. A track is defined by hits in the BDC and its direction is given by combining these hits with a well reconstructed interaction vertex. The information from the hadronic part of the SPACAL is used to improve the correct recognition of the scattered electron.

3.3.1 Trigger Selection and Prescale Weights

The first step in the procedure to identify and measure the scattered electron is to find events with a shower in the electromagnetic part of the SPACAL, by demanding the SPACAL subtriggers for DIS events to be set. Data from the 1999-2000 run period used here must fulfill the total subtrigger requirement of (S0∨S3∨S61)[1]. Each subtrigger consists logically of different trigger elements. The logical structure of the subtriggers is given in Appendix A. Below only the main trigger elements of the subtriggers are considered.

The so called trigger towers in the SPACAL contain four (2×2) contigous SPACAL cells. A simple inclusive electron trigger (IET) requires at least one trigger tower to have a deposited energy above some threshold. The size of the trigger tower has been chosen such as to ensure the transverse containment of the electromagnetic shower at any scattering angle within the acceptance of the SPACAL. Using a sliding window method, the trigger towers are required to partly overlap in such a way that also when the impact point is at one of the borders of one trigger tower the shower will be fully contained in one of the overlapping trigger towers. The subtrigger S0 is fulfilled if at least one of the outer trigger towers contains an energy larger than 6 GeV. This is denoted SPCLe_IET > 2. The subtrigger S3 has the additional requirement compared to the S0 trigger that the total energy deposit in the SPACAL should be at least 12 GeV within a certain time window. The combined requirement is according to the H1 notation written as SPCLe_IET > 2∨SPCLe_ToF_E_2. Due to the inclusive nature of the S0 and S3 subtriggers, they have high trigger efficiencies but also high rates. Therefore, S0 and S3 are prescaled during the data taking in order to suppress the number of events being read out. The average prescales of the S0 and S3 triggers are 3.28 and 1.49, respectively, for the data taking period 1999-2000. Consequently, a large fraction of the events will be rejected due to the prescale and the events accepted have to be given a large prescale weight.

The total requirement of the subtrigger S61 is given by:

$$(SPCLe\_IET\ >\ 2 \vee SPCLe\_IET\_Cen\_3) \wedge DCRPh\_THig \wedge zVtx\_sig,$$

where SPCLe_IET_Cen_3, similar to SPCLe_IET > 2, requires an energy deposition of more than 6 GeV but for one of the central trigger towers. The trigger element DCRPh_THig is fullfilled if a track candidate with transverse momentum of 400-800 MeV is found in the CJC and zVtx_sig indicates a signature for the vertex. S61 is thus a more exclusive trigger

---

[1] The symbols ∨ and ∧ stands for logical OR and AND, respectively

and has accordingly a lower trigger rate compared to the S0 and S3 triggers. On the other hand it has a lower efficiency and is assigned a lower prescale factor (1.14).

Using the trigger combination (S0$\vee$S3$\vee$S61), the weight calculated from the prescale may suffer from large variations from event to event. Also events with the same logical pattern (e.g. S0 and S3 is true but s61 is false) belonging to different runs could be assigned different prescale weights due to the fact hat the luminosity of different runs varied significantly over the 1999-2000 run period. Therefore the prescale weights were smothed by averaging the event weights over the run luminosities [93].

3.3.2 Electron Kinematics

As well as the fulfilled trigger condition (S0$\vee$S3$\vee$S61), which flags an electromagnetic shower in the SPACAL, a correct identification of the scattered electron also needs verification that the shower is not produced by a photon. This is done by requiring hits in the BDC matching the position of the electromagnetic shower. The scattering angle, θ_e, and the azimuthal angle, ϕ_e, can be calculated to high precision by combining the information from the BDC with a well reconstructed interaction vertex, as obtained from charge tracks in the central tracker. The energy of the scattered electron, E_e, is thus provided by the cluster energy measured in the SPACAL. To ensure a well confined shower and to suppress contributions from photoproduction events, where the scattered electron disappears down the beam pipe and a hadron from the hadronic final state imitates a scattered electron, it is required that:

$$E_e > 9 \text{ GeV}. \tag{3.1}$$

Further cuts to reduce background events from photoproduction are given in section 3.3.3. The angular requirement

$$156 < \theta_e < 175^\circ \tag{3.2}$$

restricts the electron to fall inside the coverage of the SPACAL and the BDC, avoiding the θ_e region with low hit efficiency.

3.3.3 Technical cuts

In the following, additional cuts to improve the reconstruction of the scattered electron in the electromagnetic SPACAL calorimeter are presented. These cuts are applied both to the data and the detector simulated MC events.

Rejection of "Fake" Electron Candidates

The transverse size of a cluster is a property that can be used to efficiently distinguish between hadrons and electrons detected by the SPACAL calorimeter, since showers from hadrons are expected to be broader than those induced by electrons. The transverse size can be estimated by the cluster radius which is defined as

$$R_{cl} = \frac{\Sigma_i E_i \sqrt{(x_{cl} - x_i)^2 + (y_{cl} - y_i)^2}}{E_{cl}},$$

where x_{cl} and y_{cl} define the center of gravity of the cluster, whereas E_{cl} is the cluster energy. The central impact position of each cell i contributing to the cluster is represented by the coordinates x_i, y_i, and its energy content by E_i. The energy of a cluster is obtained by summing over all contributing cells, $E_{cl} = \Sigma_i E_i$, and the center-of-gravity of the cluster is given by:

$$x_{cl} = \frac{\Sigma_i E_i x_i}{\Sigma_i E_i} \quad \text{and} \quad y_{cl} = \frac{\Sigma_i E_i y_i}{\Sigma_i E_i}$$

.

For the SPACAL it has been found that a cut on the cluster radius of:

$$R_{cl} < 3.5 \ \ \text{cm},$$

optimizes the separation between showers produced by electrons and hadrons [94]. A further rejection of hadrons, potentially being identified as electrons, is obtained using the information from the hadronic part of the SPACAL. Hadrons hitting the SPACAL will normally start showering in the electromagnetic part of the SPACAL but will deposit a large fraction of their energy in the hadronic part. Electrons, however, will lose almost all their energy in the electromagnetic part of the SPACAL. Adding up the energy measurements of the cells in the hadronic SPACAL, which belong to the electron candidate, will allow to define an energy cut, E_{had}, to reject hadrons. This cut has been set to:

$$E_{had} < 0.5 \text{ GeV}.$$

As mentioned above the signature of a scattered electron is a shower in the electromagnetic part of the SPACAL correlated with a particle track. The electron will produce hits in the BDC, which is positioned immediately in front of the SPACAL. The space points extracted from the BDC information provide a short track segment that can be extrapolated to the entrance surface of the SPACAL to give an impact position. This should agree with the center of gravity of the cluster within certain limits. The cut applied is

$$\Delta R_{BDC} < 3.0 \text{ cm},$$

where ΔR_{BDC} is the distance between the cluster center in the SPACAL and the calculated impact position.

In the initial state, the scalar $E - P_z$ is obtained by summing the four-momenta of the electron and the proton, giving:

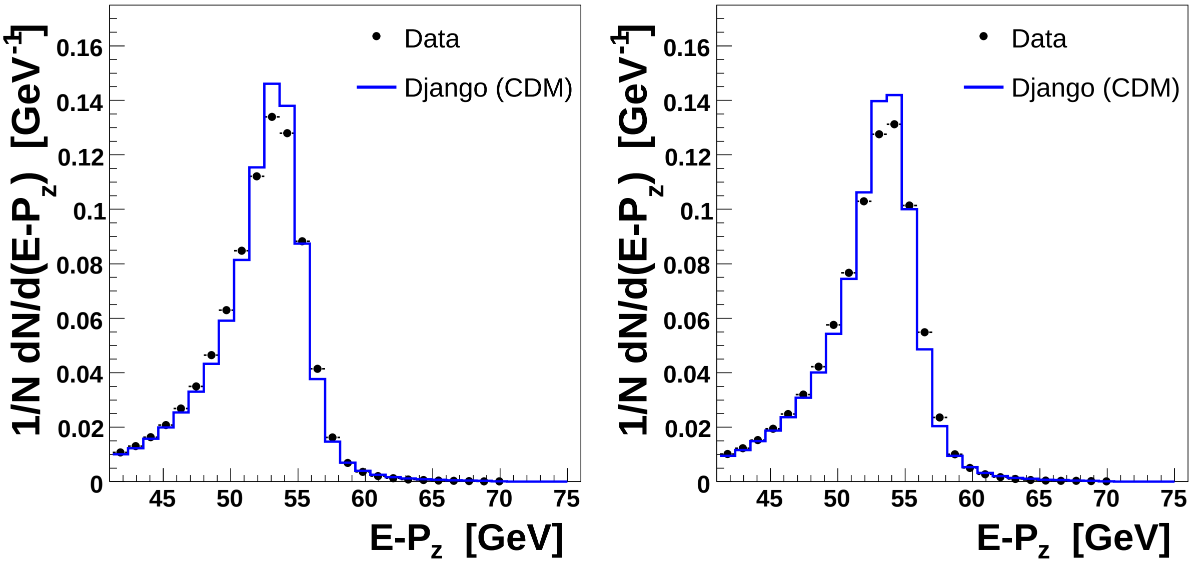


Figure 3.2: *The reconstructed $E - P_z$ before (left) and after (right) the cluster calibration has been performed.*

$$E - P_z = [920 - 920 + 27.5 - (-27.5)] \text{ GeV} = 55 \text{ GeV} = 2E_e^0$$

where E_e^0 is the energy of the incoming electron beam. Due to the momentum and energy conservation the total sum of the scalar $E - P_z$ in the final state is equal to the initial one. $E - P_z$ in the final state is

$$\sum_j E_j - P_{z,j} = 2E_e^0 = E - P_z,$$

where the summation j extends over all measured particles (also the scattered electron in the SPACAL) in the H1 detector. Due to the restricted acceptance of the H1 detector (leading to particle leakage) and the limited accuracy in the energy measurement of the final state particles, $E - P_z$ is smeared around $2E_e^0$.

In photoproduction the scattered electron is not detected ($\theta_e \rightarrow 180^o$) and a charged hadron can then be identified as the scattered electron candidate. In these cases the $E - P_z$ is much smaller than $2E_e^0$, due to the unmeasured scattered electron.

This is also the case in events where the incoming electron radiates a highly energetic photon which escapes detection. To reduce these effects the cut

$$35 \text{ GeV} \ < \ E - P_z \ < \ 70 \text{ GeV} \qquad (3.3)$$

is applied. In Figure 3.2 the $E - P_z$ distributions, after applying the selections in Equation(3.3), is shown as calculated from combined objects of clusters and tracks. The left hand plot shows the distribution before the calibration of the energy measurement has been applied to the cluster energies and it is observed that the distribution is not peaked at $2E_e^0$. This is however the case after the corrections from the calibration have been applied as seen from the right hand side plot of Figure 3.2, although the shape of the distribution is not perfectly reproduced. More about combined objects can be found in Section 3.5 and the calibration of the calorimeter energy measurements is presented in Chapter 4.

SPACAL Cell Selections

Due to the steeply falling Q^2 spectrum, the scattered electrons are strongly peaked in the backward region ($\theta_e = 180^o$). A large fraction of the electrons detected in the SPACAL will thus hit the detector close to its inner edge. Part of the shower from such electrons will not be measured due to transverse energy leakage. In order to avoid such problems so called VETO regions close to the beam pipe have been defined in the SPACAL. Requiring the energy in these regions, E_{veto}, to be less than 1 GeV ensures that the electron shower is essentially fully contained in the SPACAL.

Problems with wrongly measured shower energies also appear in the case of dead cells or cells with high counting rates. Scattered electrons at the outer regions of the SPACAL will of course also suffer from transverse energy leakage in the same way as electrons close to the inner edge. To avoid the problematic regions a number of fiducial cuts have been introduced. These decrease the acceptance of the SPACAL but increase the efficiency of the S0 and S3 triggers. The following regions of the SPACAL were excluded for the 1999-2000 year data:

$$\begin{gathered}
(-49.6 < x_{cl} < -36.9) \text{ and } (-34.2 < y_{cl} < -20.8) \\
(-13.0 < x_{cl} < \quad 0.2) \text{ and } (-55.5 < y_{cl} < -41.0) \\
(44.8 < x_{cl} < 57.5) \text{ and } (35.4 < y_{cl} < 49.9) \\
(-17 < x_{cl} < 9) \text{ and } (-9 < y_{cl} < 17) \\
y_{cl}^2 + x_{cl}^2 < 15^2 \\
y_{cl}^2 + x_{cl}^2 > 70^2,
\end{gathered}$$

where x_{cl} and y_{cl} are given in cm with respect to the z-axis. Figure 3.3 shows the hit positions of the scattered electrons in the SPACAL before (left plot) and after (right plot) the fiducial cuts have been applied. The fiducial cuts are also applied to the detector simulated Monte Carlo samples.

3.4 DIS Sample

Different methods, based on different quantities measured in the detector, are available for calculating the kinematics of a DIS event. The kinematic variables, which have to be reconstructed are any two of the variables Q^2, y and x_{Bj} since they are related through $Q^2 = s \cdot x \cdot y$, where s is the center-of-mass energy. Below, three of the methods for calculating the event kinematics will be described. The resolutions of the kinematic variables have been investigated by comparing the values generated by the Django(CDM) event generator with the values reconstructed from detector simulated MC data. The method giving the best resolution will be used for comparisons with the DIS data in various control plots.

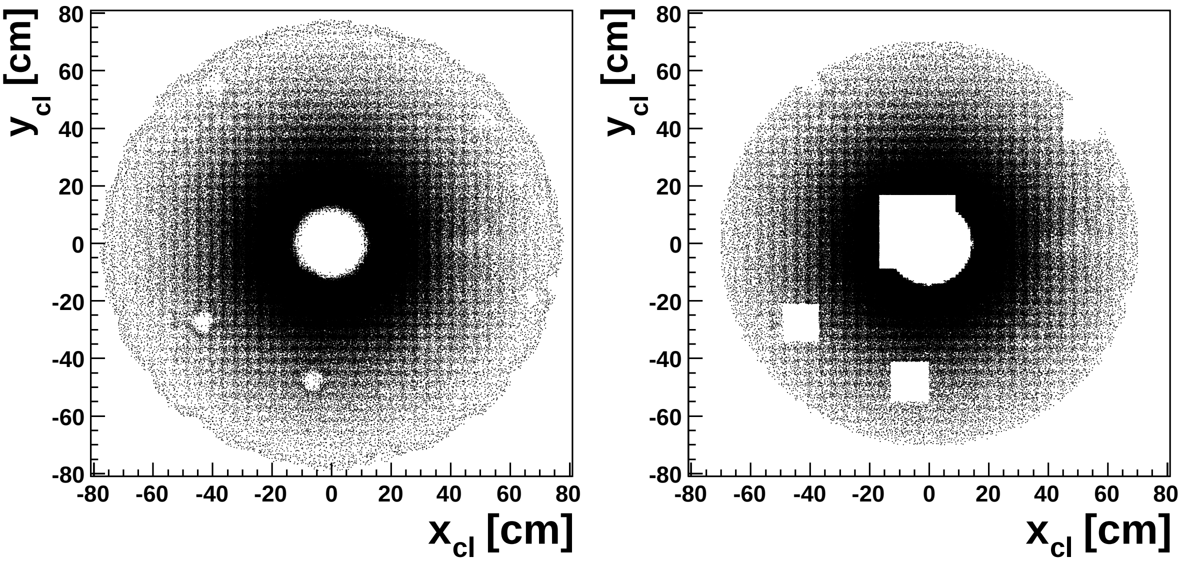


Figure 3.3: *The position of clusters* (x_{cl}, y_{cl}) *produced by the scattered electrons in the SPACAL before (left plot) and after (right plot) the fiducial cuts have been applied.*

3.4.1 Reconstruction of the Event Kinematics

The three methods presented here are the Electron (e), Sigma (Σ), and the Electron Sigma (eΣ) methods.

The electron method: The most commonly used method to reconstruct the kinematics of events in the low Q^2 region is the electron method. In this method the event kinematics are obtained from measuring the scattered electron alone. The variables Q^2, y and x_{Bj} are obtained from the relations:

$$\begin{aligned} Q^2 &= Q_e^2 &&= 4E_e^0 E_e cos^2(\theta_e/2) \\ y &= y_e &&= 1 - \frac{E_e}{E_e^0} sin^2(\theta_e/2) \\ x_{Bj} &= x_{Bj,e} &&= \frac{Q_e^2}{y_e \cdot s}. \end{aligned} \tag{3.4}$$

where E_e^o and E_e are the energies of the incoming and scattered electron, respectively, and θ_e is the polar angle of the scattered electron.

At high Q^2 and low y the reconstruction of the event kinematics by the electron method is sensitive to QED radiation. The emission of a hard photon from the initial state electron will not be noticed since it will proceed undetected in the beam tube. The collision energy for such events will be lower than the nominal one, which will lead to a wrong determination of the event kinematics by the electron method.

The Σ- and the $e\Sigma$-methods: To reduce the effect of QED radiation, the hadronic final state particles can be used to calculate the event kinematics [95, 96]. Although the energy of the hadronic final state is measured with much lower precision than that of the scattered electron, it is directly related to the true collision energy. In these two methods the energy and momentum of all measured final state hadrons are used to calculate the total hadronic energy, E_h, and longitudinal momentum, $P_{z,h}$.

$$E_h - P_{z,h} = \sum_i E_i - P_{z,i} \tag{3.5}$$

The summation i extends over all measured particles in the angular range ($3^o < \theta_i < 175^o$) except for the scattered electron, which is excluded. The quantity $E_h - P_{z,h}$ is mostly sensitive to the energy calibration of the LAr calorimeter in the central region, whereas the effect from particles in the forward and backward regions is smaller. Thus particles lost through the holes in the detector for the beam pipe have little influence on the quantity $E_h - P_{z,h}$. The DIS variables Q^2, y and x_{Bj} are reconstructed in the Σ-method according to:

$$\begin{aligned} Q^2 &= Q^2_\Sigma &= \frac{E_e^2 sin^2\theta_e}{1 - y_\Sigma} \\ y &= y_\Sigma &= \frac{E_h - P_{z,h}}{E - P_z} \\ x_{Bj} &= x_{Bj,\Sigma} &= \frac{Q^2_\Sigma}{y_\Sigma 4 E_e E_p} \end{aligned} \tag{3.6}$$

Since the term $E_h - P_{z,h}$ is calculated using hadrons measured by the LAr calorimeter, the Σ-method is suffers from the poor resolution of the calorimeter. In the $e\Sigma$-method [95, 96] a mix of the electron and Σ-methods is used:

$$\begin{aligned} Q^2 &= Q^2_{e\Sigma} &= Q^2_e \\ y &= y_{e\Sigma} &= \frac{2E_e(E_h - P_{z,h})}{(E - P_z)^2} \\ x_{Bj} &= x_{bj,e\Sigma} &= x_{Bj,\Sigma} \end{aligned} \tag{3.7}$$

The $e\Sigma$-method provides a better reconstruction of the event variables than the Σ-method.

3.4.2 Resolution of the Event Kinematics

The precision in the determination of the event kinematics can be investigated by comparing Monte Carlo events on generator level and after detector simulation. The distributions of the ratio y_{rec}/y_{gen} have been used to determine the mean value $\langle y_{rec}/y_{gen} \rangle$ and the resolution, in terms of standard deviations, σ, as a function of y_{gen}. From the measurement of the mean value, shown in Figure 3.4, it is observed that the kinematics reconstructed by the electron method gives better agreement with the hadron level MC data than the Σ- and $e\Sigma$-methods, although especially the $e\Sigma$ method exhibits improved performance as the y-value increases. The electron method gives also a better resolution (smaller σ-values) than the Σ- and $e\Sigma$-methods, except for the lowest y_{gen} bin, as can be seen in Figure 3.5. The resolution increases with higher values of y, but most dramatically for the electron-

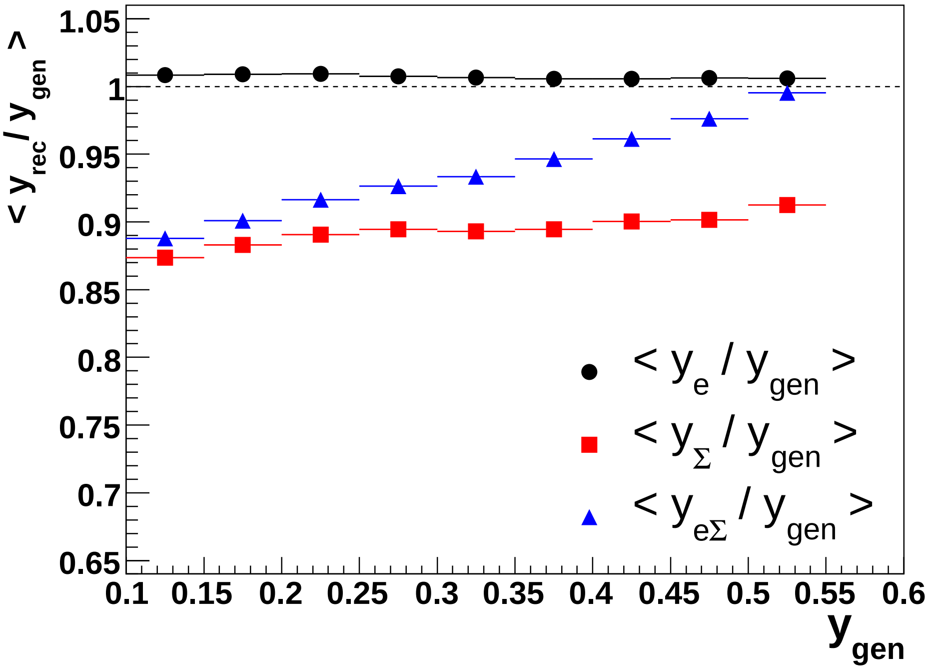


Figure 3.4: *Distributions of* $\langle y_{rec}/y_{gen}\rangle$ *(mean values) for the e-,* Σ*- and e*Σ*-methods obtained from Django(CDM) Monte Carlo data, shown as a function of* y_{gen}.

and $e\Sigma$-methods. The problems of the Σ- and $e\Sigma$-methods are related to the rather poor energy measurement of the hadronic final state which suffers from the limited detector acceptance in θ, particle leakage, mismeasurements of the particle energy and poor energy calibration of the LAr calorimeter (only in second order). A behaviour similar to that observed for the reconstruction of the y-variable can be seen in the reconstruction of the x_{Bj} variable, shown in the Figures 3.6-3.7, resulting in the same conclusions. In the electron- and $e\Sigma$-methods, Q^2 is reconstructed in the same way and therefore they have identical perfomancies as shown in the Figures 3.8-3.9. The Σ-method fails to give a reasonably good reconstruction of Q^2 and has the worst resolution.

3.4.3 The Phase Space and DIS Event Selection

From the study described in section 3.4.2 it is obvious that the electron method offers by far the most accurate reconstruction of the event kinematic. This is a reflection of the fact that energy measurements of electromagnetically interacting particles are much more accurate than those of hadrons. The resolutions of E_e and θ_e are shown in Figure 3.10. The good resolution of the electron momentum is also important for the calibration of the LAr calorimeter, which is presented later in this thesis.

The following phase-space cuts are applied in order to obtain a well defined DIS sample:

$$\begin{aligned} 5 &< Q_e^2 < 100 \text{ GeV}^2 \\ 0.1 &< y_e < 0.7 \end{aligned}$$

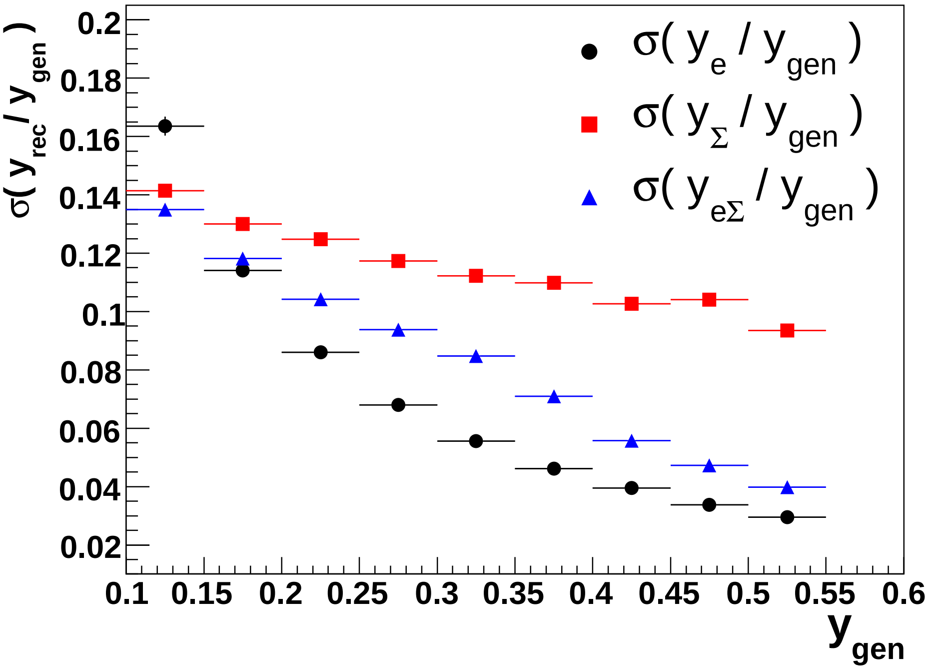


Figure 3.5: *The* σ *values of the* y_{rec}/y_{gen} *distributions for the e-,* Σ*- and e*Σ*-methods obtained from Django(CDM) Monte Carlo data, shown as a function of* y_{gen}.

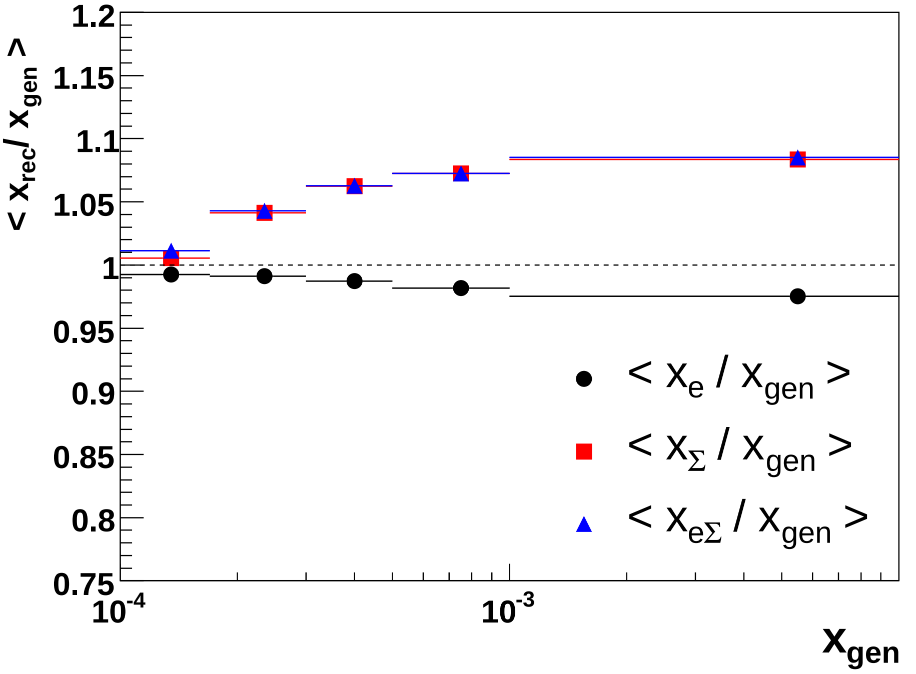


Figure 3.6: *Distributions of* $\langle x_{rec}/x_{gen}\rangle$ *(mean values) for the e-* Σ*-, and e*Σ*-methods obtained from Django(CDM) Monte Carlo data, shown as a function of* x_{gen}.

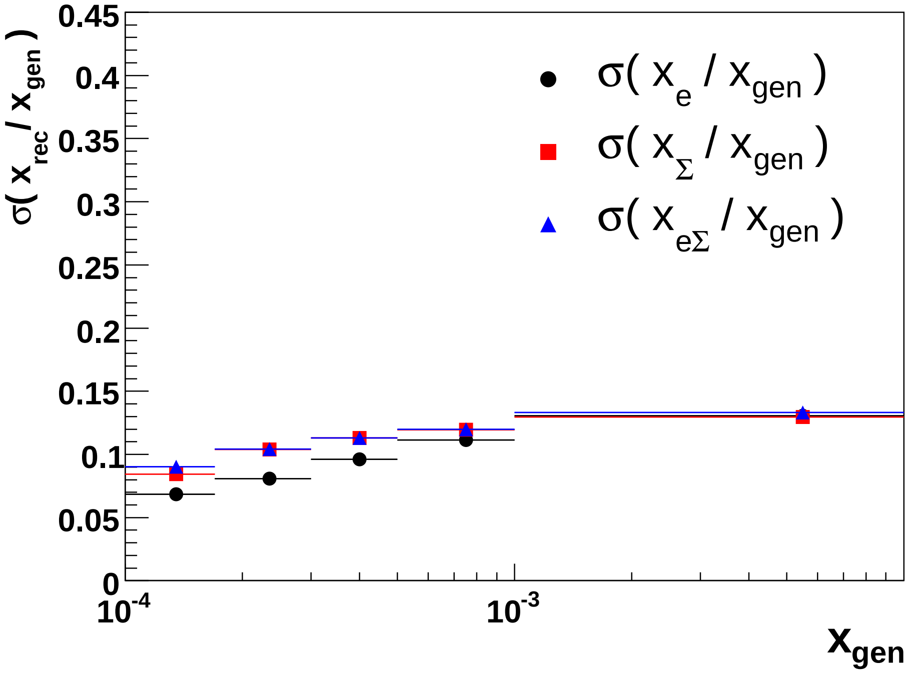


Figure 3.7: *The* σ *values of the* x_{rec}/x_{gen} *distributions for the e-,* Σ*- and e*Σ*-methods obtained from Django(CDM) Monte Carlo data, shown as a function of* x_{gen}.

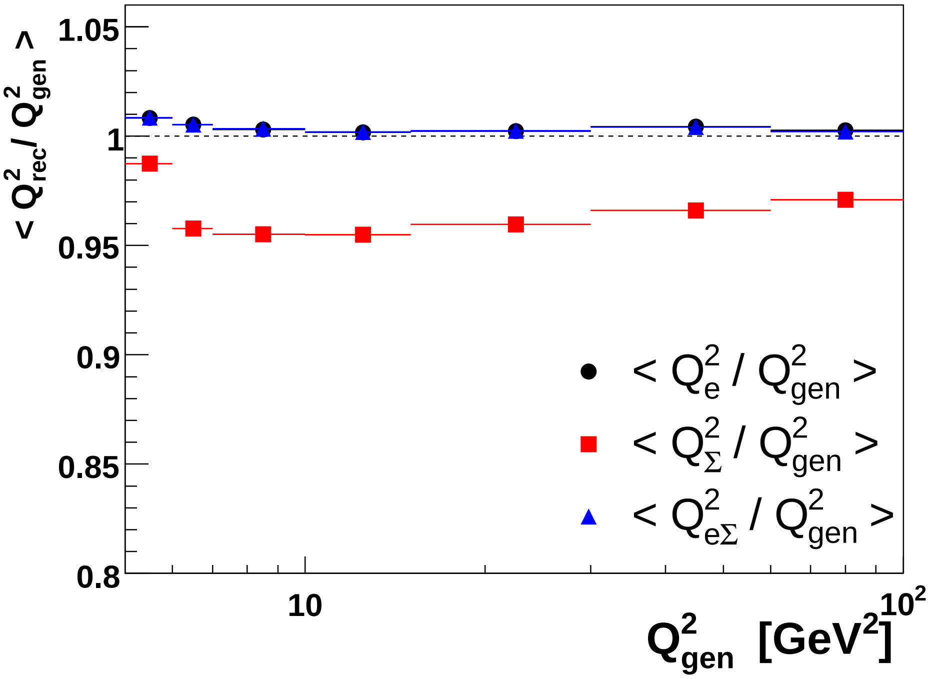


Figure 3.8: *Distributions of* $\langle Q^2_{rec}/Q^2_{gen}\rangle$ *for the e-,* Σ*- and e*Σ*-methods obtained from Django(CDM) Monte Carlo data, shown as a function of* Q^2_{gen}*. Note that* $Q^2_{e\Sigma} = Q^2_e$ *follows from the definition of e*Σ *method.*

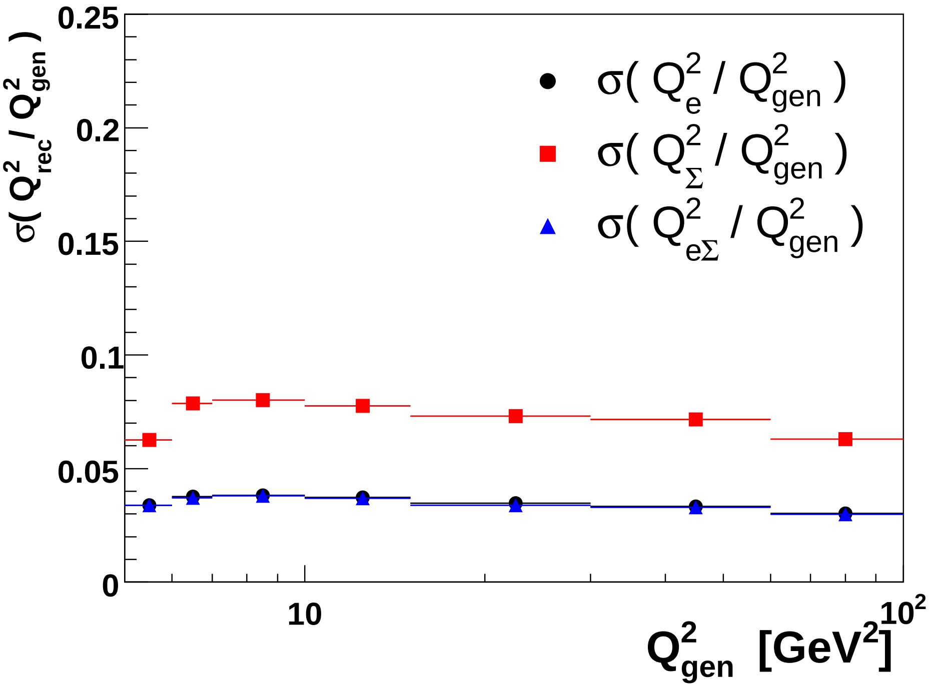


Figure 3.9: *The σ values of the Q^2_{rec}/Q^2_{gen} distributions for the e-, Σ- and eΣ-methods obtained from Django(CDM) Monte Carlo data, shown as a function of x_{gen}. Note that $Q^2_{e\Sigma} = Q^2_e$ follows from the definition of eΣ method.*

In addition to the geometrical cut $156^o < \theta_e < 175^o$, a cut in Q^2 is applied, which is consistent with the scattered electron being restricted to the SPACAL acceptance region. However, this cut also helps to reduce contributions from photoproduction.

The cuts on detector level which define the selection of DIS events are summarized in Table 3.1.

3.4.4 Quality of the DIS Cuts and Control Plots

The DIS events, selected by applying the cuts in Table 3.1, have been used to produce various control plots in order to investigate how well predictions from MC event generators compare to the experimental data. In Figure 3.11 distributions of Q^2, y, x_{Bj}, θ_e, ϕ_e and E_e are shown and compared with the predictions of the Rapgap(dir) and Django(CDM) programs at the detector level. The distributions are normalised to the integrated luminosity. The two MC programs are able to predict the DIS cross section quite well, except for a small deviation at large x_{Bj}. It should however be noticed that the vertical scale is logarithmic.

Some of the cuts in the DIS selection are are especially introduced in order to suppress the photoproduction background. The background remaining after these cuts have been estimated to be less than 1%.

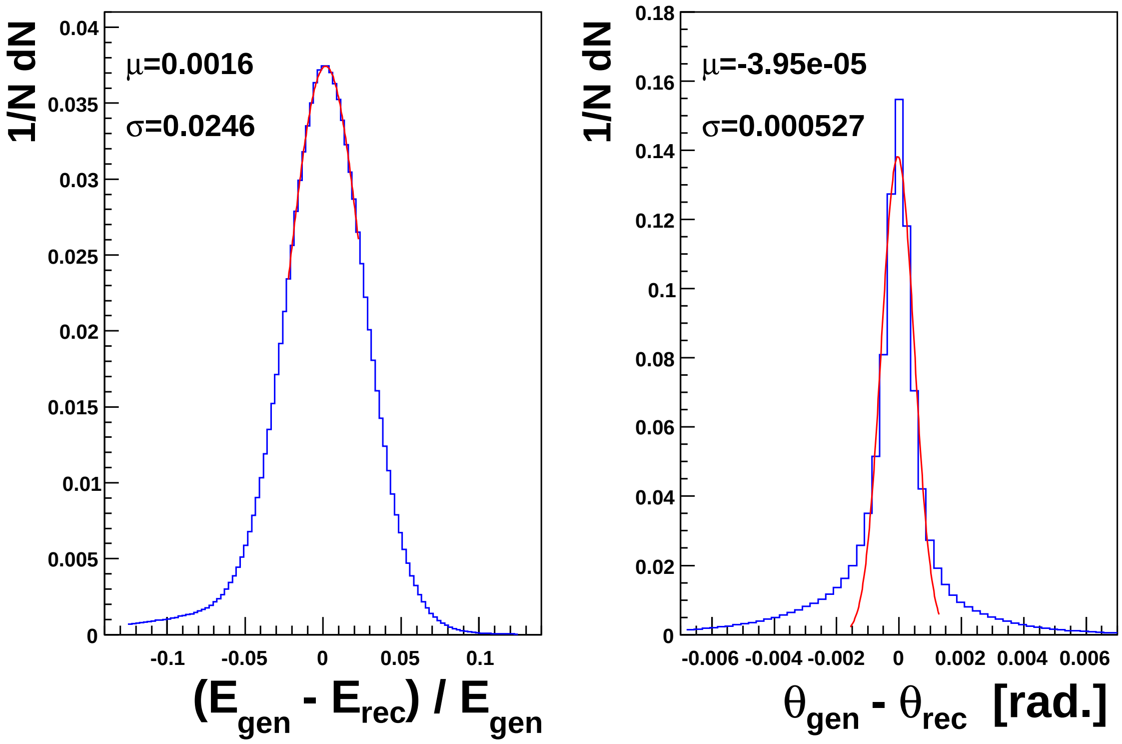


Figure 3.10: *The resolutions of the energy and polar angle of the scattered electron obtained using Django(CDM) Monte Carlo data. The mean value (μ) and the σ are obtained from a Gaussian fit around the peak.*

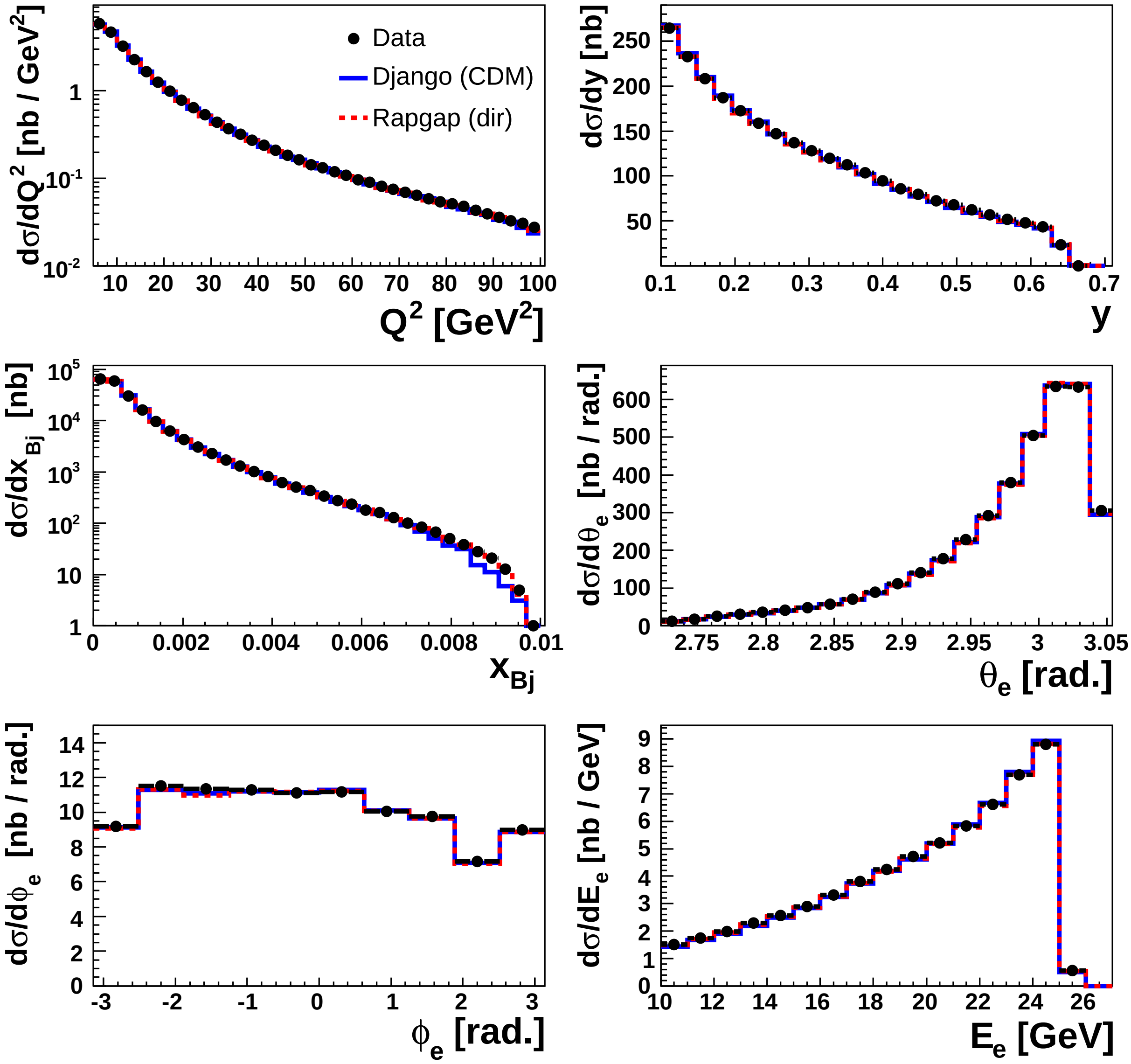


Figure 3.11: *Control distributions for the DIS sample, showing* Q^2*,* y*,* x_{Bj}*,* θ_e*,* ϕ_e *and* E_e*. The cross sections are compared to detector simulated Monte Carlo data from Rapgap(dir) and Django(CDM), respectively.*

	DIS Cuts	
5 GeV^2 <	Q^2	< 100 GeV^2
9 GeV <	E_e	
0.1 <	y	< 0.7
156° <	θ_e	< 175°
35 GeV <	$E - P_z$	< 70 GeV
	$\|z_{vtx}\|$	< 35 cm
	R_{cl}	< 3.5 cm
	E_{had}	< 0.5 GeV
	R_{BDC}	< 3. GeV
	E_{veto}	< 1.0 GeV
	fiducial cuts	
	S0∨S3∨S61 (data only)	
	Run selections (data only)	

Table 3.1: *A summary of cuts used for the DIS event selection.*

3.5 The Hadroo2 Algorithm

The quarks and gluons emitted in the scattering process will produce hadrons through the fragmentation process, giving rise to the hadronic final state. Decays of short lived particles may produce photons and leptons, which have to be distinguished from the scattered electron. Photons and neutral hadrons will only be measured in the calorimeters, whereas charged particles may be measured both in the tracking devices and the calorimeters. However, some of the charged particles will have too low momenta to reach the calorimeter and some might have so high momenta that their curvature is too small to be measured in the tracking device. In order to obtain the best possible measurement of the hadronic final state particles it has to be determined in which momentum ranges the tracking information and the calorimeter information should be used. Thus, a reconstructed track of a charged hadron has to be associated with a shower in the calorimeters in order to perform a comparison of the energy measurements. In this process tracks from isolated electrons and muons are not considered, since they most probably originate from decays of short lived hadrons before the decay products enter the detector volume. The requirements for defining particles as isolated are given in Appendix B.

The purpose of the Hadroo2 reconstruction algorithm [97], is thus to associate a selected track (see Appendix B for definition) of a charged hadron to a cluster and determine whether the momentum measurement provided by the tracker or the energy measurement by the calorimeter is the most accurate to reconstruct the kinematics of the particle candidate. By assuming all tracks to originate from pions the resolution in the energy measurement from the track reconstruction can be determined through:

$$\frac{\sigma_{E_{track}}}{E_{track}} = \frac{1}{E_{track}}\sqrt{P_{track}^2 \cdot \cot^2\theta \cdot \sigma_\theta^2 + \frac{\sigma_{P_T}^2}{sin^2\theta}} \tag{3.8}$$

where σ_{P_T} and σ_θ are the errors in P_T and θ, respectively, obtained from the track fitting error information. The energy of the track is given by:

$$E^2_{track} = P^2_{track} + m^2_\pi \tag{3.9}$$

The error which can be expected from measuring this track in the LAr can be estimated according to:

$$\left(\frac{\sigma_E}{E}\right)_{LAr,expect.} = \frac{\sigma_{E_{LAr,expect.}}}{E_{track}} = \frac{0.5}{\sqrt{E_{track}}} \tag{3.10}$$

where $\sigma_{E_{LAr,expect.}}$ is the expected uncertainty from the energy measurement in the LAr calorimeter, of a track with an energy E_{track} as measured by the tracking system. From measurements in test beams the energy resolution of the hadronic section of the LAr calorimeter has been determined to be $50\%/\sqrt{E}$.

If $\sigma_{E_{track}}/E_{track} < \sigma_{E_{LAr,expect.}}/E_{track}$ then the track energy measurement is preferred. For a charged particle originating from the interaction point and measured by both the tracker and the calorimeter, the tracker will always provide the best measurement unless the particle momentum is above 12 GeV in the forward region and 25 GeV in the central region. In such cases the track related cluster in the calorimeter has to be identified and removed in order to avoid double counting. Due to the fact that the visible energy, generated by a particle with energy E_{track} in the LAr calorimeter, is always smaller than the true energy of the particle, a direct energy comparison can not be used to identify the corresponding cluster. The starting point for identifying the cluster is to extrapolate the track along a helix up to the surface of the calorimeter and from there connect it to a straight line tangential to the helix. The energy in the calorimeter, which is associated to the particle track, $E_{assoc.}$, is given by the sum of all clusters which are confined by the volume of a cone centered around the extrapolated track direction and with an opening angle of 67.5^o, overlapping with one cylinder of radius 25 cm in the electromagnetic section and one cylinder of radius 50 cm in the hadronic part.

The next step is to compare $E_{assoc.}$ with E_{track} including fluctuations in the cluster and track measurements. In the cases where one finds:

$$E_{assoc.} < E_{track}\left[1 + 1.96\sqrt{\left(\frac{\sigma_{E_{track}}}{E_{track}}\right)^2 + \left(\frac{\sigma_E}{E}\right)^2_{LAr\ expected}}\right] \tag{3.11}$$

the total energy of $E_{assoc.}$ has to be removed and the track information is used. In other cases an energy E_{track} is removed from $E_{assoc.}$ and the energy difference $E_{assoc.} - E_{track}$ is assumed to originate from neutral particles or other charged particles.

If $\sigma_{E_{track}}/E_{track} > \sigma_{E\ LAr\ expected}/E_{track}$ is true then three cases are considered:

1. $|E_{track} - E_{assoc.}| < 1.96\sigma_{E_{assoc.}}$, where $\sigma_{E_{assoc.}} = 0.5\sqrt{E_{assoc.}}$
   The energy measurement from the track agrees with that of the clusters within the fluctuations.

2. $E_{track} > E_{assoc} + 1.96\sigma_{E_{assoc.}}$
   The energy measurement from the track exceeds that of the clusters including the upward fluctuation.

3. $E_{track} < E_{assoc.} - 1.96\sigma_{E_{assoc.}}$
   The energy measurement from the track is smaller than that of the clusters including the downward fluctuation.

In the cases 1 and 2 the calorimeter measurement is used and the track is removed. In case 3 the track measurement is chosen and the cluster energy is removed as explained above.

3.6 Jet Reconstruction

In high energy particle collisions the interesting physics is related to the properties of the partons and leptons involved in the primary interaction. In contrast to leptons the partons can not be measured directly in the detector, since colour confinement prevents them from appearing as free particles. Instead hadrons are produced through the fragmentation and hadronisation processes. Provided the parton momentum is high enough collimated flows of hadrons are created, called particle jets. As the parton momentum increases the jets tend to be more pencil-like, which simplifies their identification. Although the fragmentation and hadronization processes are not understood from first principles, they can be well described by phenomenological models. In all fragmentation models the kinematic properties of a jet are strongly correlated with those of the parton from which it evolved. One should keep in mind that jets at the detector level (data and detector simulated MC data) and at the hadron level (after fragmentation and hadronization of partons) are colourless objects which are used to describe the underlying properties of quarks and gluons, which carry colour charges. Despite that, jets will reflect the properties of the partons involved in the short distance interactions, provided the jet is sufficiently well defined.

Different types of algorithms have been developed in order to reconstruct jets. However, the definition of a jet is not unambiguous. The picture of a jet as a collimated flow of particles implies that a minimum energy has to be available for the jet formation. Due to the fact that the definition of a jet is not unique, the existing jet-finding algorithms are based on different observables in the jet reconstruction. The success of an algorithm is given by its ability to reproduce the four vector of the original parton i.e. its momentum and direction of flight.

The procedure to group hadrons together into jets is by no means obvious. The algorithm used in this analysis is the longitudinally invariant $k_\perp$ clustering algorithm [98–100]. The advantage of the $k_\perp$ algorithm is that the clustering procedure is safe from infrared divergencies because soft partons are first combined with harder partons and it is also safe from collinear divergencies since two collinear partons are first combined with the original parton. Particles which are found to belong to the same jet, are merged together into a pseudoparticle using the $k_\perp$ recombination scheme [101]. Eventually, the pseudoparticles, represented by momentum vectors, will correspond to the jet momentum vectors.

3.6.1 The Longitudinally Invariant $k_\perp$ Algorithm

In the longitudinally invariant $k_\perp$ algorithm two resolution variables $d_{i,B}$ and $d_{i,j}$ are defined. They measure the distance of particle $\{i\}$ with respect to the direction of the beam, B, and the separation $\{i,j\}$ of the particle $\{i\}$ to all other particles $\{j\}$. Particles are merged into a pseudoparticle if their momenta are nearly parallel. The algorithm starts with an original list of particles and proceeds as follows:

1. Calculate the distance between the particle $\{i\}$ and the beam direction

   $$d_{i,B} = P_{T,i}^2 R^2.$$

   where R represents the radius of a cone in the η-ϕ plane which is given the value 1 in this analysis.

2. Define for each pair of particles $\{i,j\}$, the longitudinally invariant distance

   $$d_{i,j} = min(P_{T,i}^2, P_{T,j}^2) \cdot \left((\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2\right)$$

3. Find the minimum of $d_{i,j}$ and $d_{i,B}$.

4. If the minimum is given by $d_{i,j}$, then the particle i and j are merged together into a massless pseudo particle k according to the $k_\perp$ recombination scheme:

   $$\begin{aligned} P_{T,k} &= P_{T,i} + P_{T,j} \\ \eta_k &= \frac{P_{T,i}\eta_i + P_{T,j}\eta_j}{P_{T,k}} \\ \phi_k &= \frac{P_{T,i}\phi_i + P_{T,j}\phi_j}{P_{T,k}}. \end{aligned}$$

   Replace particles $\{i,j\}$ with the pseudo particle k.

5. If the minimum is given by $d_{i,B}$, then the particle $\{i\}$ is removed from the list and defined as an jet.

6. Move back to step one until no particles are left in the list.

Chapter 4

Calibration of Jet Energy Measurements

As already discussed in section 2.2.1 the energy measurement of hadronic showers suffers from the limited energy fraction of the shower that is measurable (the visible energy) which leads to a systematic shift and poor energy resolution. Thus, the measured energy is significantly smaller than the energy carried by the hadron that initiated the shower. The losses in the energy measurement have to be corrected for so as to reconstruct, as closely as possible, the energy of the initial hadron. The correction factors are applied in order to bring the measured data to the correct scale i.e. the absolute energy scale of the measurement. The absolute scale of track momentum measurements and its uncertainty, has been determined to better than 1 %, whereas the absolute energy scale for cluster measurements in the hadronic calorimeter is known to substantially worse precision (several percent). The uncertainty in the determination of the absolute energy scale will have a direct impact on how well the energy of hadronic final state particles (HFS particles) can be measured. The precision in the determination of the energy scale also depends on how well the detector simulation is done. Not only has the performance of the calorimeter itself to be accurately described but also the material budget in front of the calorimeter has to be known in detail in order to correctly reproduce the energy measurement of the calorimeter.

Different calibration methods have been used by the H1 experiment, with the aim of reducing the uncertainty and improving the absolute scale in the measurement of cluster energies. The Iterative method [102] and the High Pt Jet Calibration method [97, 103] are the most frequently used ones. The Iterative method provides a global calibration of hadronic energy measurements i.e. it includes all calorimeter clusters in the event, while in the High Pt Jet Calibration method only clusters belonging to the reconstructed jets in the event are considered. The High Pt Jet Calibration method has been developed for calibrating energy of jets with tranverse momenta larger than 10 GeV and for high Q^2 events ($Q^2 > 100$ GeV^2).

In this chapter, a method of calibrating measurements of jets with transverse momenta below 10 GeV, the Low Pt Jet Calibration method, is described. This method is inspired by the method presented in [97, 103]. Data collected in the years 1999-2000 and 2006-2007 have been used to test the performance of the Low Pt Jet Calibration method. Since

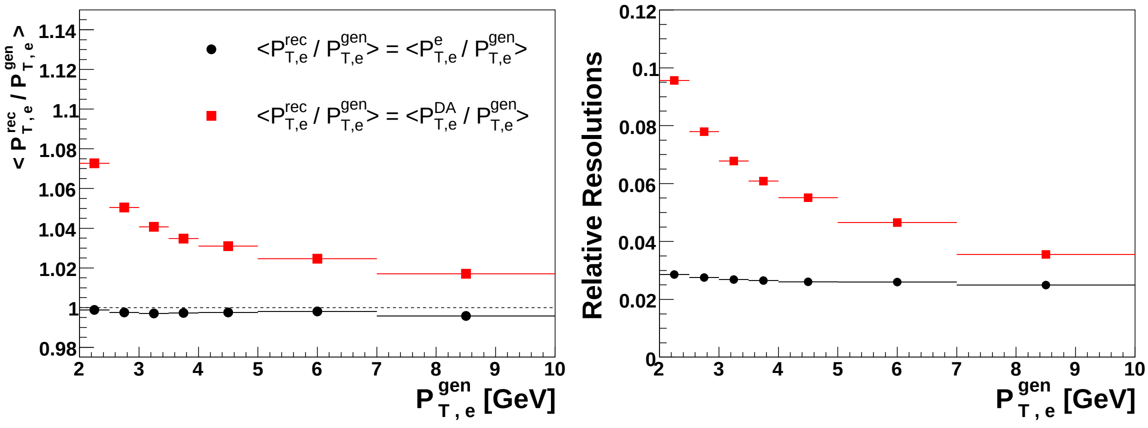


Figure 4.1: *The mean value of the ratio between the reconstructed and generated* $P_{T,e}$ *distributions (left) and the relative resolution, defined as the standard deviation,* σ*, of the* $P^{rec}_{T,e}/P^{gen}_{T,e}$ *distribution (right) , both plotted as a function of* $P^{gen}_{T,e}$*. A comparison is made between the cases where* $P_{T,e}$ *has been reconstructed using the double angle method and electron method, respectively.*

clusters are considered to be massless objects, energy is equivalent to momentum and in the following these concepts are used without distinction.

In section 4.1 general considerations about the jet energy calibration are given. The selection criteria used to define the calibration sample are presented and motivated in section 4.2. In section 4.3 the calculation of calibration constants is discussed, and the performance of the calibration method is examined in section 4.4. A summary is given in section 4.5 together with a discussion of possible future improvements.

4.1 General Considerations

At HERA, the electron and proton beams collide head-on, which means that the total transverse momentum of the initial state is zero. Thus, the total transverse momentum of the final state also has to be zero due to momentum conservation. This means that the transverse momentum of the scattered electron has to be balanced by the transvers momentum of the hadronic final state, in the laboratory system. However, the experimental measurement suffers from limitations given by the detector, such as the acceptance, the resolutions of energy and momentum measurements and particles that escape detection (like neutrinos). Provided the scattered electron can be measured with high precision, the influence of the detector effects can be studied by using the measured transverse momentum of the scattered electron as reference in a comparison with the measured transverse momentum of the hadronic final state. It should be noted that the Hadroo2 algorithm was used previoulsy (section 3.5) to reconstruct clusters in the LAr calorimeter, related to an incoming hadron, and thus isolated electrons and muons were excluded. If, however, the scattered electron is used as reference all particles in the hadronic final state have to be included i.e. also isolated electrons and muons.

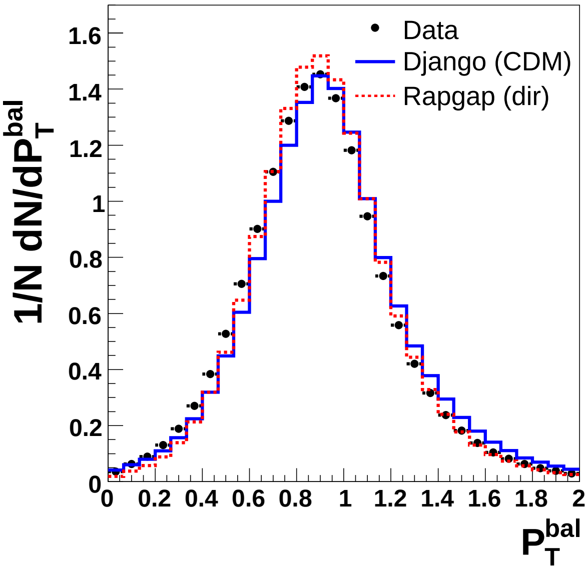


Figure 4.2: *The P_T balance distribution, $P_T^{bal} = P_{T,h}/P_{T,e}$, shown for uncalibrated data from DIS events, and for the Monte Carlo predictions of Rapgap(dir) and Django(CDM).*

At high Q^2 the scattered electron is detected in the LAr calorimeter, in which case the best reconstruction of its transverse momentum is offered by the double angle method, where P_T^{da} is given by:

$$P_T^{da} = \frac{2E_e^0}{tan\frac{\theta_e}{2} + tan\frac{\theta_h}{2}} \tag{4.1}$$

with

$$tan\frac{\theta_h}{2} = \frac{E_h - P_{z,h}}{P_{T,h}}. \tag{4.2}$$

The polar angle, θ_h, extracted from Equation (4.2), is called the angle of the hadronic system in the event. $E_h - P_{z,h}$ is given by the Equation (3.5) and $P_{T,h}$ is defined as:

$$P_{T,h} = \sqrt{(\sum_i P_{x,i})^2 + (\sum_i P_{y,i})^2}. \tag{4.3}$$

Due to its good performance at high Q^2, the double angle method has been used to calibrate the energy measurements of high P_T jets [97,103].

At low Q^2, the scattered electron is detected in the SPACAL for which the electron method provides the best determination of its transverse momentum, $P_{T,e}$. The performance of the $P_{T,e}$-reconstruction can be estimated by comparing the reconstructed and generated values from Monte Carlo generated events. Figure 4.1 shows the mean value of the ratio between the reconstructed and generated $P_{T,e}$ distributions, $\langle P_{T,e}^{rec}/P_{T,e}^{gen}\rangle$, together with the relative resolution, defined as the standard deviation, σ, of the $P_{T,e}^{rec}/P_{T,e}^{gen}$-distribution, both plotted as a function of $P_{T,e}^{gen}$. A comparison between the electron method and double angle method, in the low Q^2-region, proves that the generated $P_{T,e}$-values are very well reconstructed by the electron method, whereas the double angle method gives a much worse reconstruction, especially at low $P_{T,e}$. The resolution in the reconstruction is also much better for the electron method compared to the double angle method. It should, however, be kept in mind that the $P_{T,e}$-reconstruction from the electron method also depends on the uncertainty in the absolute energy scale of the SPACAL, whereas the error in the absolute energy scale of the LAr calorimeter cancels to first order if the double angle method is applied. Still, the electron method provides by far the most accurate determination, due to the high precision in the energy calibration of the electromagnetic SPACAL. Thus, the electron method will be used for the investigations of the balance between the transverse momenta of the scattered electron and the HFS particles. The transverse momentum balance is defined by:

$$P_T^{bal} = \frac{P_{T,h}}{P_{T,e}} \tag{4.4}$$

where $P_{T,h}$ is the total transverse momentum of the HFS particles. The limited resolution of the experimental measurement will lead to Gaussian like P_T-distributions, whereas the losses due to acceptance and non-measurable particles will give a systematic shift of the distributions compared to the P_T of the scattered electron.

The P_T^{bal} distribution extracted from experimental data, is compared with those obtained from detector simulated Monte Carlo data using the Rapgap(dir) and Django(CDM) programs, in Figure 4.2, for low Q^2 events. As can be seen the Rapgap(dir) distribution reproduces the data somewhat better than that of Django(CDM). It can also be noticed that the peak value of the distributions is around 0.9, indicating that the amount of energy lost in the event is about 10%, on the average. The calculated mean value of the P_T^{bal} distributions is used as an indicator of the absolute energy scale i.e. how well the total P_T of the hadronic system is reconstructed. In order to gain some knowledge of how well the response of the detector is described by the detector simulation and how well the events are reconstructed, the double ratio of the mean P_T^{bal}-values for data and Monte Carlo events can be studied. The double ratio is defined as:

$$DR(P_T^{bal}) = \langle\frac{P_{T,h}}{P_{T,e}}\rangle_{data}/\langle\frac{P_{T,h}}{P_{T,e}}\rangle_{MC} \tag{4.5}$$

where $\langle P_{T,h}/P_{T,e}\rangle$ is the mean value of the P_T^{bal} distributions. If $DR(P_T^{bal})$ is not equal to unity, different amounts of transverse momentum are reconstructed from the data and MC events, suggesting that our knowledge of the detector and/or the event reconstruction is not good enough. Thus, the double ratio, $DR(P_T^{bal})$, measures the systematic uncertainty of the HFS transverse momentum. Since $P_{T,h} \approx E_h sin(\theta_h)$ and the uncertainty in the θ_h reconstruction is expected to be small compared to the $P_{T,h}$ measurement, $DR(P_T^{bal})$ is essentially equivalent to the systematic uncertainty in the absolute energy scale. From the methods described in [97,102,103] it has been demonstrated that the absolute energy scale can be determined from the double ratio method to better than 2% with an uncertainty less than 2%.

The method described above can be used to find calibration constants for measurements of the total hadronic final state but in the case of jets the principle is to re-weight the energy measurement of clusters belonging to jets only. The weights are extracted separately for the different wheels of the LAr calorimeter, see Figure 2.3. The position of the wheels, given in polar angle, and their coverage, in terms of rapidity, is specified in Table 4.1. The calibration weights for the jets in the low Q^2 region are extracted from a sample containing exactly one reconstructed jet, above some momentum threshold, and the scattered electron detected by the SPACAL (Jet calibration sample), i.e. QPM like events. In the ideal case the total transverse momentum of the hadronic final state, $P_{T,h}$, would be equal to the tranverse momentum of the jet, provided the jet algorithm clusters 'all' HFS particles into the jet, and consequently the mean value of the P_T^{bal} distribution would reflect the absolute energy scale of the jet measurement. However, according to the string fragmentation model it is not unambigous which particles belong to the jet in a QPM event and which particles belong to the proton remnant. Problems might also be caused by two-jet events where one of the jets is too soft to fulfill the jet transverse momentum cut and therefore the events will be treated as QPM events. Different selections have been studied with the aim of reducing the fraction of problematic events in the calibration sample. Two MC programs, Rapgap(dir) and Django(CDM), are used in order to study the model dependence of the calibration method.

	IF1	IF2	OF	FB	CB3	CB2	CB1	BBE
θ [*deg.*]	$7-10$	$10-15$	$15-30$	$30-55$	$55-80$	$80-110$	$110-135$	$135-155$
$\mid \Delta\eta \mid$	0.36	0.41	0.71	0.68	0.43	0.54	0.53	0.62

Table 4.1: *The different regions in polar angle of the LAr wheels to be calibrated and their coverage in* η.

4.2 The Calibration Sample

The DIS calibration sample was selected by applying the standard cuts for the DIS event selection as presented in Table 3.1, except for the following changes:

- $E_e > 15$ GeV has been imposed to improve the electron energy reconstruction, but also to reduce the contamination from photoproduction processes.
- In order to increase the acceptance of QPM jet events at low θ angles (the forward region) the lower cut on y has been removed.
- In order to further reduce the contamination from photoproduction processes and the effects from QED radiation, a narrower window of the scalar $E - P_z$ has been defined, 52 GeV $< E - P_z <$ 70 GeV.
- Exactly one isolated electron in the SPACAL is required.

The jets are found and reconstructed in the laboratory frame by applying the inclusive $k_\perp$ algorithm. Events in the DIS calibration sample, with a jet of $P_{T,jet} > 3$ GeV within the angular range $7^o < \theta_{jet} < 155^o$, are selected. In order to calculate calibration weights for the individual calorimeter wheels, the full jet has to be contained within the wheel i.e. the transverse size of the jet has to be smaller than the rapidity coverage of the wheel.

The transverse size of a jet is represented by an energy weighted radius, R_w, defined by:

$$R_w \;=\; \sum_i E_i\sqrt{(\phi_{jet}-\phi_i)^2+(\eta_{jet}-\eta_i)^2}\;/\;E_{jet},$$

where E_{jet}, ϕ_{jet} and η_{jet} are the energy of the jet, and the azimuthal angle and rapidity of the jet axis, respectively, whereas E_i, ϕ_i and η_i are the energy, azimuthal angle and rapidity of particle i in the jet. The sum runs over all particles in the jet. In Figure 4.3 the R_w-distributions are shown in different bins of $P_{T,jet}$, obtained from data and Monte Carlo events generated by Rapgap(dir) and Django(CDM). Again it is seen that Rapgap(dir) gives a better description of the data compared to Django(CDM). Further the average transverse jet size is seen to increase from $R_w \approx 0.32$, for jets with P_T in the range 6 - 7 GeV, to $R_w \approx 0.35$ for jets with P_T 3.5 - 4 GeV. Thus, the transverse size of the jets is always smaller than or comparable to the coverage, in η, of the separate LAr wheels (see Table 4.1), even at the lowest $P_{T,jet}$ of the calibration sample.

The uncertainty in the absolute scale of the jet transverse momentum reconstruction is given by $DR(P_T^{bal})$, as defined by Equation (4.5). The mean value of the P_T^{bal} distributions is obtained by fitting a Gaussian around the peak of the distribution. In order to obtain a distribution around the peak value, it is obvious that $P_{T,h}$ must be allowed to become smaller than $P_{T,e}$. This requirement is met by introducing a cut on $P_{T,e}$ that is larger than the P_T cut used in the jet reconstruction. For this reason the transverse momentum of the scattered electron has to fulfill:

$$P_{T,e} \;>\; 4 \text{ GeV}. \tag{4.6}$$

For a pure one-jet event the polar angle of the hadronic system θ_h essentially represents the scattering angle of the struck quark, θ_{jet}, except for small deviations due to the string fragmentation. For events where all particles, except the remnant particles, are clustered into one jet by the jet algorithm, the polar angle of the jet, θ_{jet}, and the angle of the hadronic system, θ_h, will coincide by default. The angle θ_h is determined from a measurement of the hadronic final state particles according to Equation (4.2) and is thus directly dependent on the LAr calibration, which is undesireable. By using instead the y-value determined from either the electron method, Equation (3.4), or the $e\Sigma$- method, Equation (3.7), θ_h is given by:

$$tan(\theta_{h,e\Sigma}/2) \;=\; \frac{y_{e\Sigma}}{1-y_{e\Sigma}}tan(\theta_e/2). \tag{4.7}$$

However, as was noted in section 3.4.1 the reconstruction of the event kinematics by the electron method gives bad resolution in the low y-region partly due to the influence of QED radiation. Figure 4.4 contains plots of the angular distributions of the jet axis, θ_{jet}, and of the angle of the hadronic system as determined by the double angle method (θ_h), by the $e\Sigma$-method ($\theta_{h,e\Sigma}$) and by the electron method ($\theta_{h,e}$), for data and the Monte Carlo samples of Rapgap(dir) and Django(CDM). As can be seen, the agreement between data and the Monte Carlo distributions is good except in the forward region, where the influence of the beam remnant is expected to be large. The distributions of θ_h and $\theta_{h,e\Sigma}$ have similar shapes although they do not agree well with the θ_{jet} distribution. The $\theta_{h,e}$ distribution is

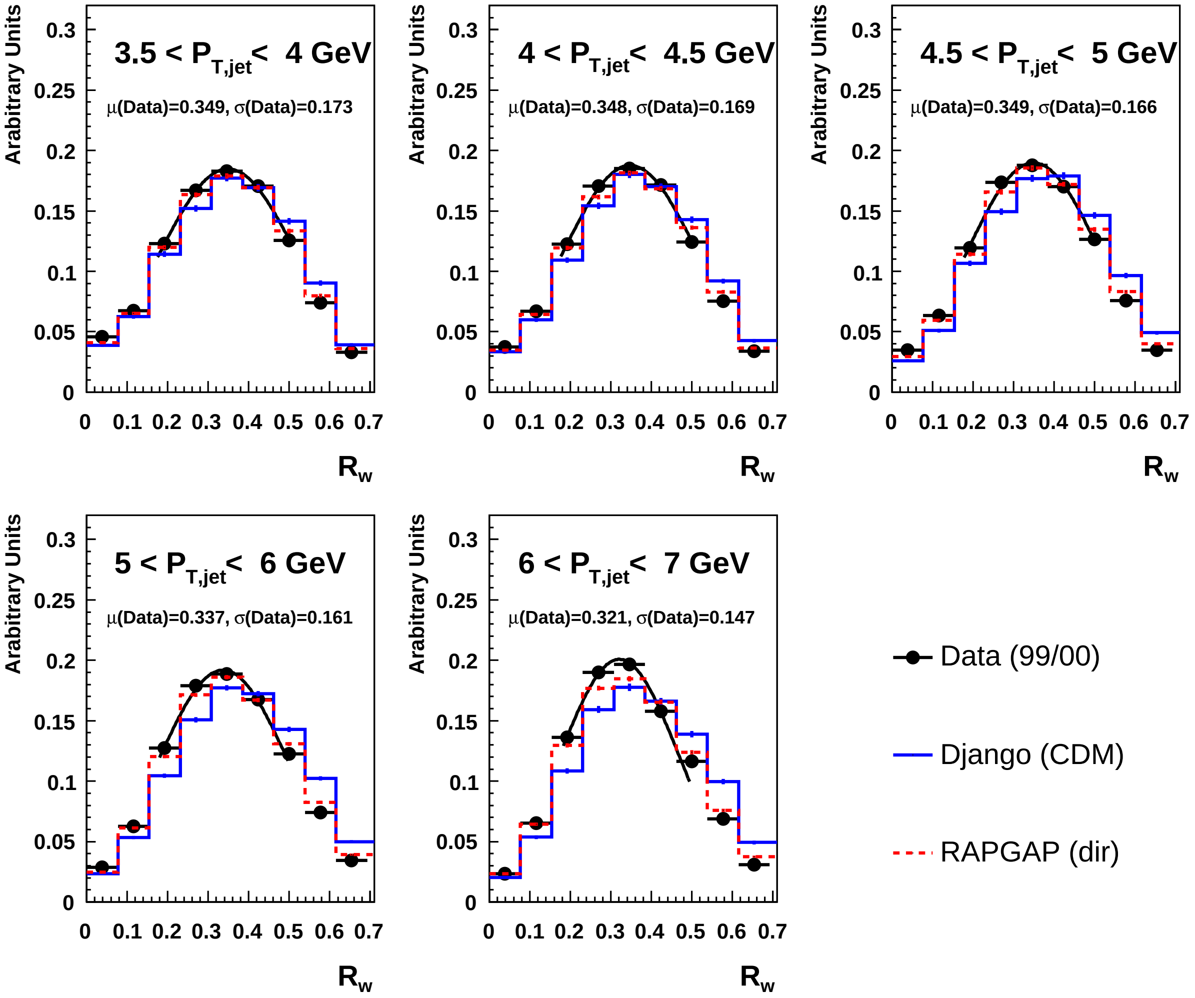


Figure 4.3: *Distributions of the transverse jet size* R_w*, where* $R_w = \sum_i E_i \sqrt{(\phi_{jet} - \phi_i)^2 + (\eta_{jet} - \eta_i)^2} / E_{jet}$*. Data are compared to the predictions of Rapgap(dir) and Django(CDM) in different regions of* $P_{T,jet}$*. The mean value,* μ*, and the standard deviation,* σ*, are obtained by fitting a Gaussian to the central part of the data distributions.*

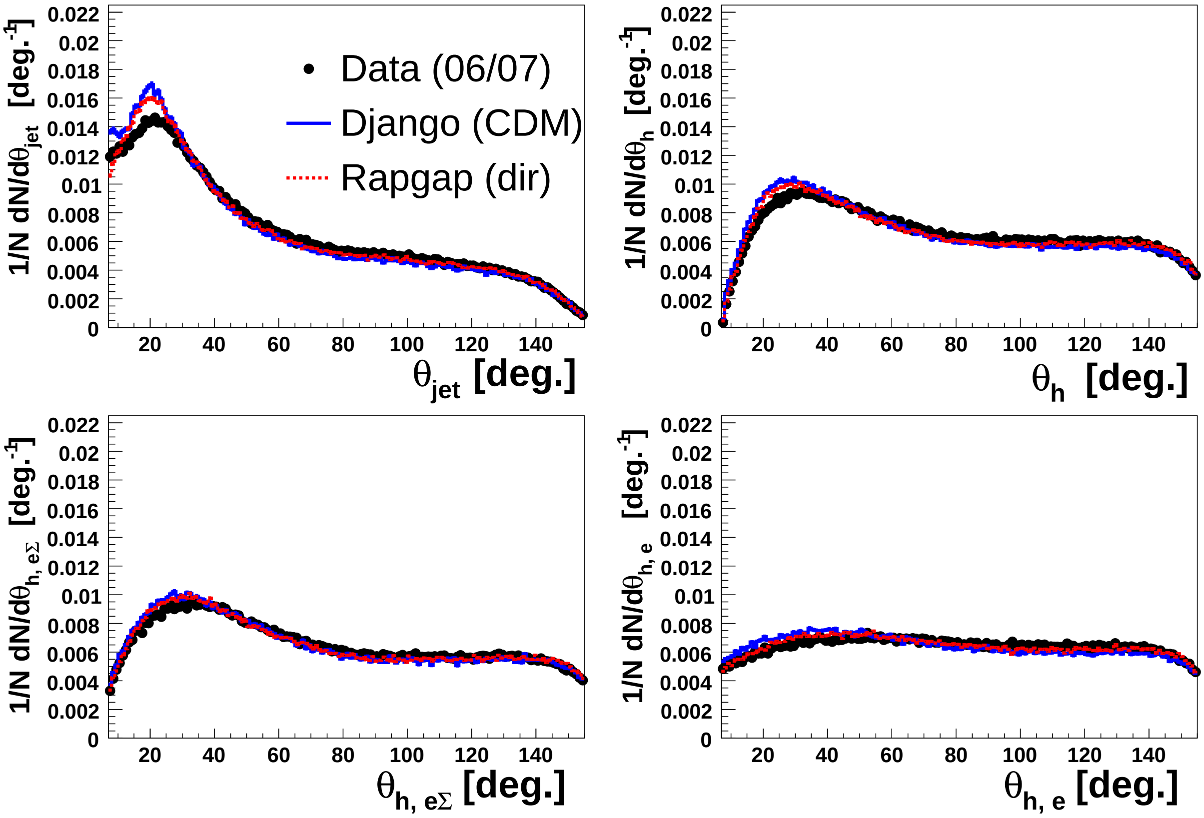


Figure 4.4: *The distributions of polar ang* θ_{jet}, θ_h, $\theta_{h,e\Sigma}$ *and* $\theta_{h,e}$. *Where* θ_{jet} *is the polar angle of the jet,* θ_h *is the polar angle of the hadronic system,* $\theta_{h,e\Sigma}$ *and* $\theta_{h,e}$ *are the polar angle of the hadronic system calculated with the* $e\Sigma$ *and* e *methods, respectively. Data are compared to the predictions of Rapgap(dir) and Django(CDM).*

much flatter than the others. Based on the considerations above the $e\Sigma$ method has been used to reconstruct the angle of the hadronic system.

In principle the calibration of the jet energy measurement should account for the loss of jet particles that fall outside the acceptance of the LAr calorimeter but should avoid including particles that belong to the proton remnant. Since the influence of the beam remnant gets problematic in the very forward direction it is necessary to require that the angle of the hadronic system is reconstructed well inside the solid angle of the calorimeter. Therefore the following cut is imposed:

$$10^o \quad < \quad \theta_{h,e\Sigma} \quad < \quad 150^o. \tag{4.8}$$

To further constrain the selection of QPM events, the difference between the polar angle of the jet and of the total hadronic system is required to be small,

$$|\theta_{h,e\Sigma} - \theta_{jet}| \quad < \quad 20^o. \tag{4.9}$$

For the purpose of extracting the calibration constants of jet energy measurements the definition of $P_{T,h}$ and P_T^{bal} have been modified in the following way:

$$P_{T,h}^{'} = \sqrt{(\sum_i \overline{P}_T^i - \sum_{i\in jet} \overline{P}_T^i + \overline{P}_{T,jet})^2}. \tag{4.10}$$

The summation i extends over all measured particles, represented by clusters and tracks, in the angular region $3^o < \theta_i < 175^o$, except for the scattered electron. Here, the jet is considered as one object, given by the $k_\perp$ algorithm, which carries the struck parton momentum and therefore, to avoid double counting, the transverse momenta of the particles belonging to the jet, $\sum_{i\in jet} \overline{P}_T^i$, are subtracted from the transverse momenta of all particles in the event, $\sum_i \overline{P}_T^i$, in Equation (4.10). From MC studies it was found that this definition of $P_{T,h}^{'}$ provides a better description of the generated momentum of the hadronic system.

The ratio between $P_{T,jet}$ and $P_{T,h}^{'}$ is demanded to be close to one, in order to reject events where a large fraction of the transverse momentum is carried by particles, which are not allocated to the jet.

$$0.8 < \frac{P_{T,jet}}{P_{T,h}^{'}} < 1.4 \tag{4.11}$$

After these additional cuts have been applied the DIS calibration sample has been reduced to a Jet calibration sample.

The imbalance of the transverse momentum of the event can in the general case be written as $(1 - P_T^{bal})$. In the case of QPM events the imbalance, corrected for clusters not included in the jet, is given by $(1 - P_T^{bal}) \cdot (P_{T,JetClus}/P_{T,Clus})$, where $P_{T,JetClus}$ is the transverse momentum of the clusters in the jet and $P_{T,Clus}$ is the transverse momentum of all particles in the event. It should be remembered that it is the uncorrected cluster energies that enter into $P_{T,JetClus}$. Thus, a revised definition of the transverse mometum balance for QPM events is introduced:

$$P_T^{bal\,'} = 1 - (1 - P_T^{bal}) \cdot (\frac{P_{T,JetClus}}{P_{T,Clus}}). \tag{4.12}$$

This modification minimizes the contribution to the momentum imbalance from particles not allocated to the jet. In Figure (4.5-4.8) the control plots of the Jet calibration sample are shown. In Figure 4.5 and Figure 4.6 the distributions of $P_{T,jet}/P_{T,h}^{'}$ and E_{jet}/E_h, are shown for different regions of θ_{jet}, respectively. From Figure 4.6 it is seen that events with jets reconstructed in the central region have a significant fraction of the total energy outside the jet ($E_{jet}/E_h \ll 1$), whereas Figure 4.5 shows that most of the transverse momentum of the event is carried by the jet ($P_{T,jet}/P_{T,h}^{'} \approx 1$). The data are well described by Django(CDM), while Rapgap (dir) does not reproduce the energy distribution of QPM events.

Distributions of the DIS kinematic variables are shown for the Jet calibration sample in Figure 4.7. As observed Rapgap(dir) provides an almost perfect description of the data, whereas the predictions of Django(CDM) exhibit some deviations.

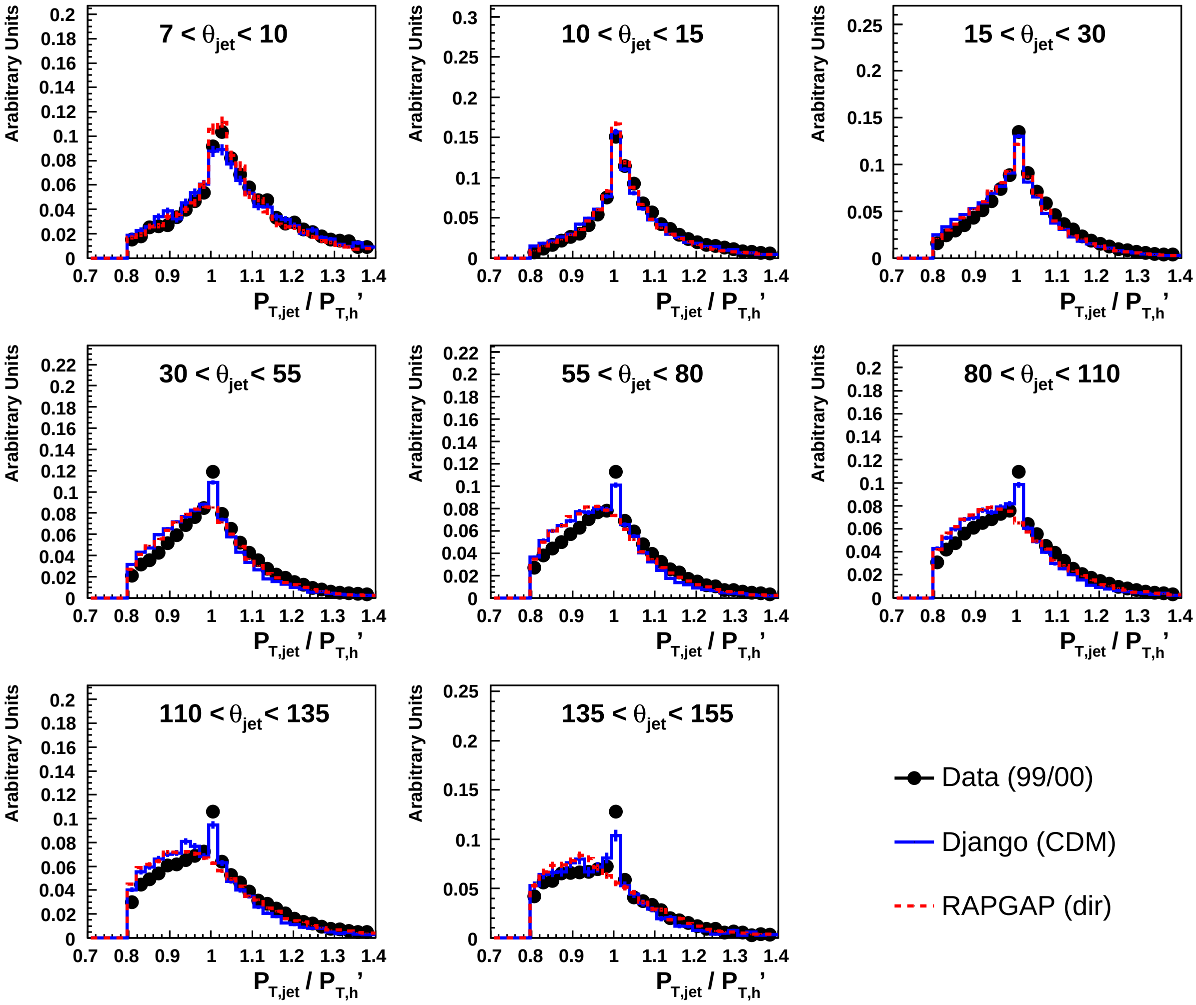


Figure 4.5: *Distributions of the $P_{T,jet}/P'_{T,h}$ ratio presented for different regions in θ_{jet} for the Jet calibration sample.*

In Figure 4.8 the distributions of θ_{jet} and θ_h are compared to the predictions of Rapgap(dir) and Django(CDM) for the Jet calibration sample. It can be observed that the two distributions have shapes that are more similar than those observed for the corresponding distributions of the DIS calibration sample (see Figure 4.4).

The absolute scale of the measurements, is given, as mentioned already, by the mean values of the $P_T^{bal'}$ distributions, obtained from a fit of a Gaussian distribution around the peak of the $P_T^{bal'}$ distributions. The fit covers one standard deviation from the mean value. In Figure 4.9 the $P_T^{bal'}$ distributions are shown together with the fits in the angular range $30^o < \theta_{jet} < 55^o$ for the Jet calibration sample. In Figure 4.10 the mean values of the $P_T^{bal\ '}$ distributions, $\langle P_T^{bal\ '}\rangle$, are plotted as a function of $P_{T,e}$ in different bins of θ_{jet} for the Jet calibration sample, using data from the 1999-2000 run periods, together with the distributions obtained from the Monte Carlo samples of Rapgap(dir) and Django(CDM). The $P_{T,e}$ dependence of $\langle P_T^{bal'}\rangle$ is extracted from fits using the function:

$$F^{bal}(\theta_{jet}, P_{T,e}) \quad = \quad A_{\theta_{jet}}(1 + exp(-B_{\theta_{jet}} - C_{\theta_{jet}}P_{T,e})) \qquad (4.13)$$

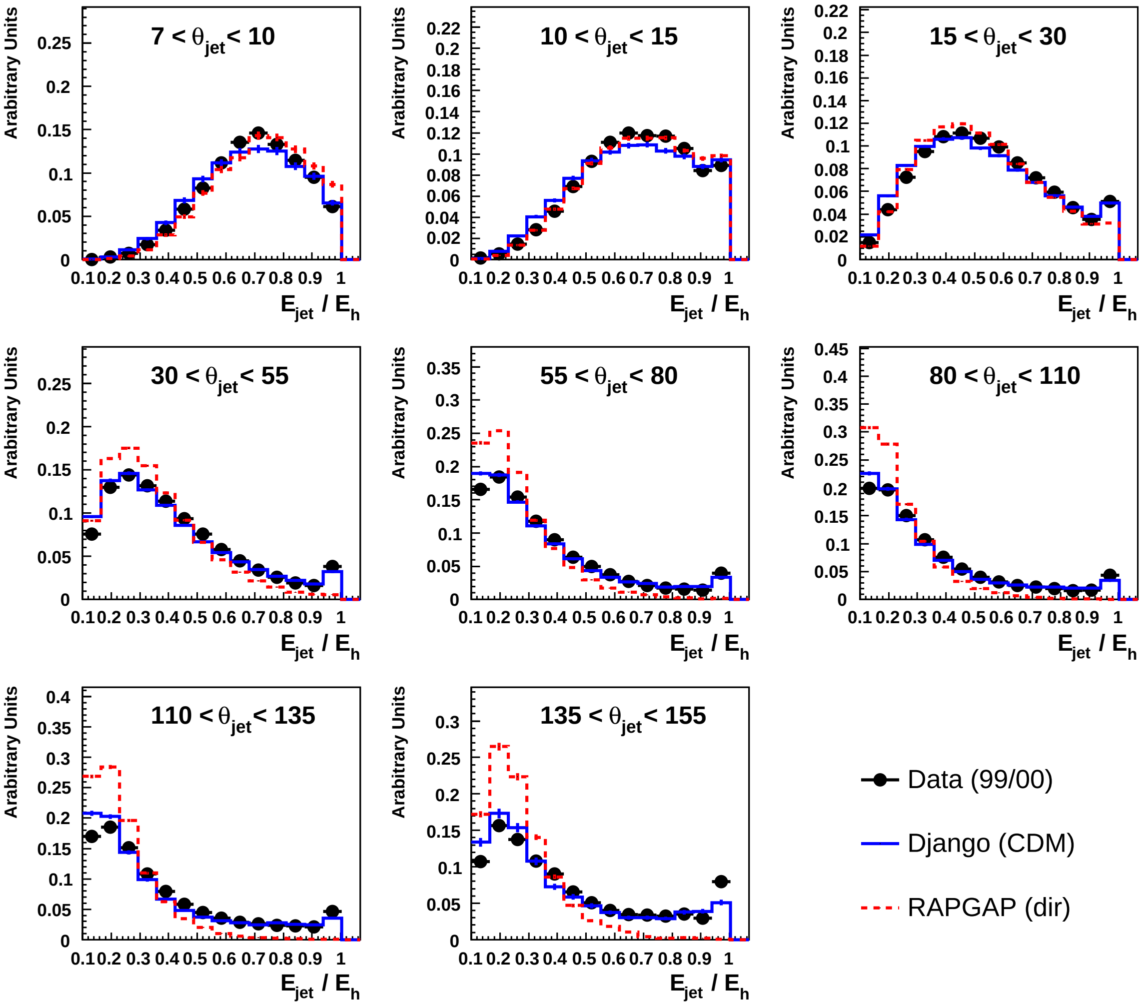


Figure 4.6: *Distribution of the E_{jet}/E_h ratio presented for different regions in θ_{jet} for the Jet calibration sample.*

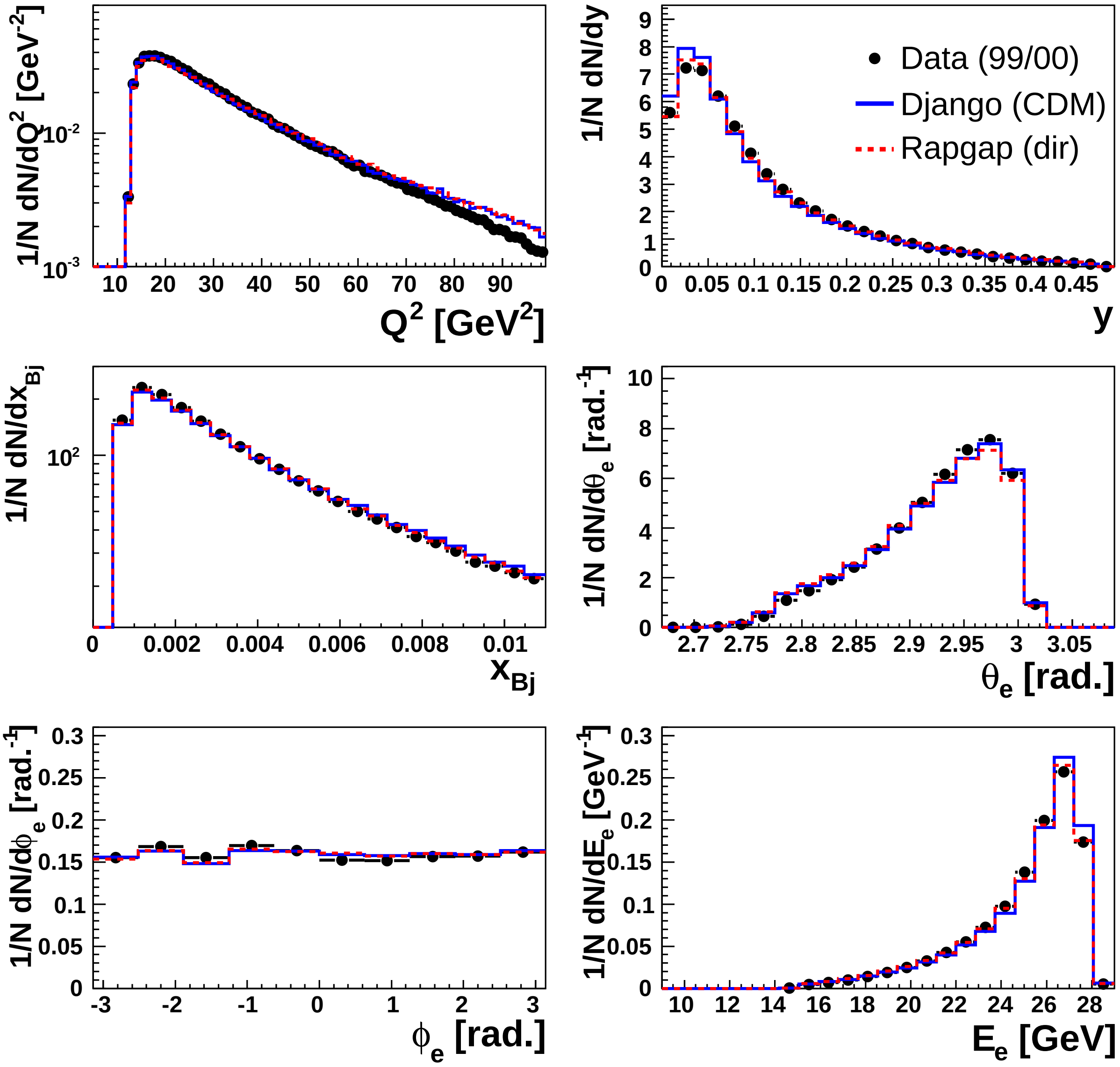


Figure 4.7: *Control distributions of the variables Q^2, y, x_{Bj}, θ_e, ϕ_e and E_e for the Jet calibration sample. The experimental data are compared to detector simulated Monte Carlo data from Rapgap(dir) and Django(CDM). The distributions have been normalised to the total number of events.*

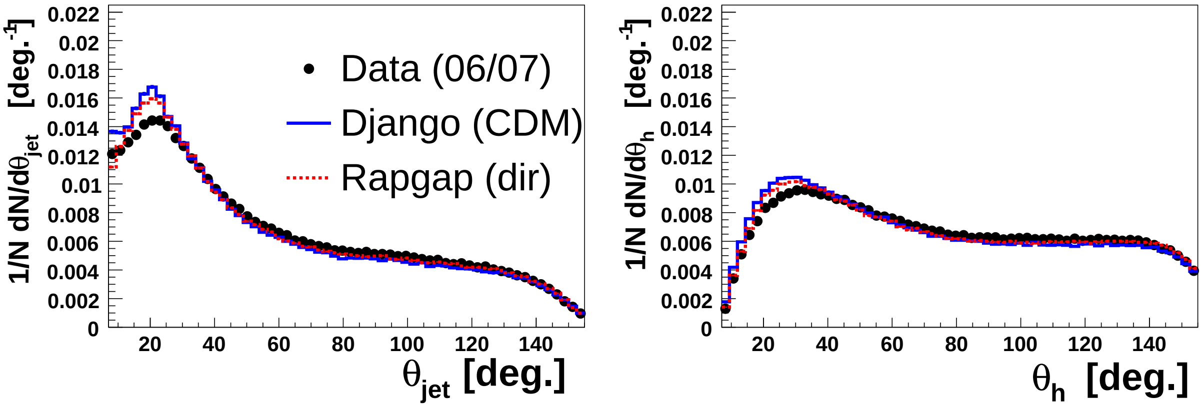


Figure 4.8: *The distributions of θ_{jet} and θ_h for the Jet calibration sample, normalised to the total number of events.*

where $A_{\theta_{jet}}$, $B_{\theta_{jet}}$ and $C_{\theta_{jet}}$ are the calibration coefficients to be determined. Less than 1% difference between the Monte Carlo predictions of Rapgap(dir) and Django(CDM) is observed in all θ_{jet} bins.

In Figure 4.11 the corresponding plots are shown for data taken in the 2006-2007 run periods. It should be mentioned that the understanding of the detector and the performance of the event reconstruction algorithm in the years 2006-2007, were not completely satisfactory as the results, presented in this thesis, were extracted. Thus, the predictions of $P_T^{bal\,\prime}$ from the Monte Carlo programs exhibit large deviations with respect to the data in the forward region.

4.3 Correction Factors

The absolute scale of the jet measurement as a function of the reference momentum $P_{T,e}$ is determined using the function $F^{bal}(\theta_{jet}, P_{T,e})$, separately for the different wheels of the LAr calorimeter (θ bins). A correction factor, f, is applied to clusters belonging to jets of transverse momenta $3 < P_{T,jet} < 9$ GeV, which have been reconstructed in the laboratory frame within the angular region of $7^o < \theta_{jet} < 155^o$. The correction factor is given by:

$$f \quad = \quad \frac{1 - F^{bal}(\theta_{jet}, P_{T,jet}) \times (1 - C_{cls})}{F^{bal}(\theta_{jet}, P_{T,jet}) \times C_{cls}} \tag{4.14}$$

where:

$$C_{cls} \quad = \quad \frac{P_T^{JetClus}}{P_T^{JetTracks} + P_T^{JetClus}} \tag{4.15}$$

is the transverse momentum fraction of the jet taken from the cluster information compared to the total transverse momentum of the jet given by the sum of cluster and track measurements. $F^{bal}(\theta_{jet}, P_{T,jet})$ is the fitted function in Equation (4.13) using $P_{T,jet}$ as reference. It is clear that since $P_{T,e}$ gives the transverse momentum balance to the total

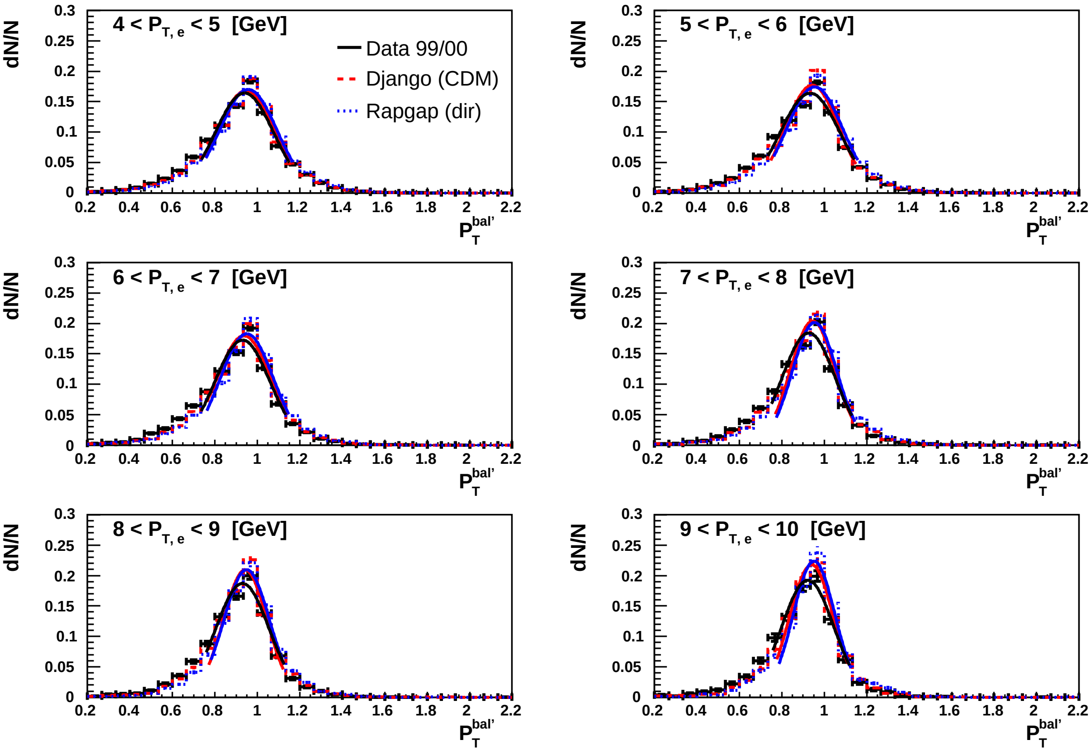


Figure 4.9: $P_T^{bal'}$ *distributions for uncalibrated experimental data (99/00), and for Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs, in the angular range* $30^o < \theta_{jet} < 55^o$ *for the calibration sample. Fits around the mean values are also shown.*

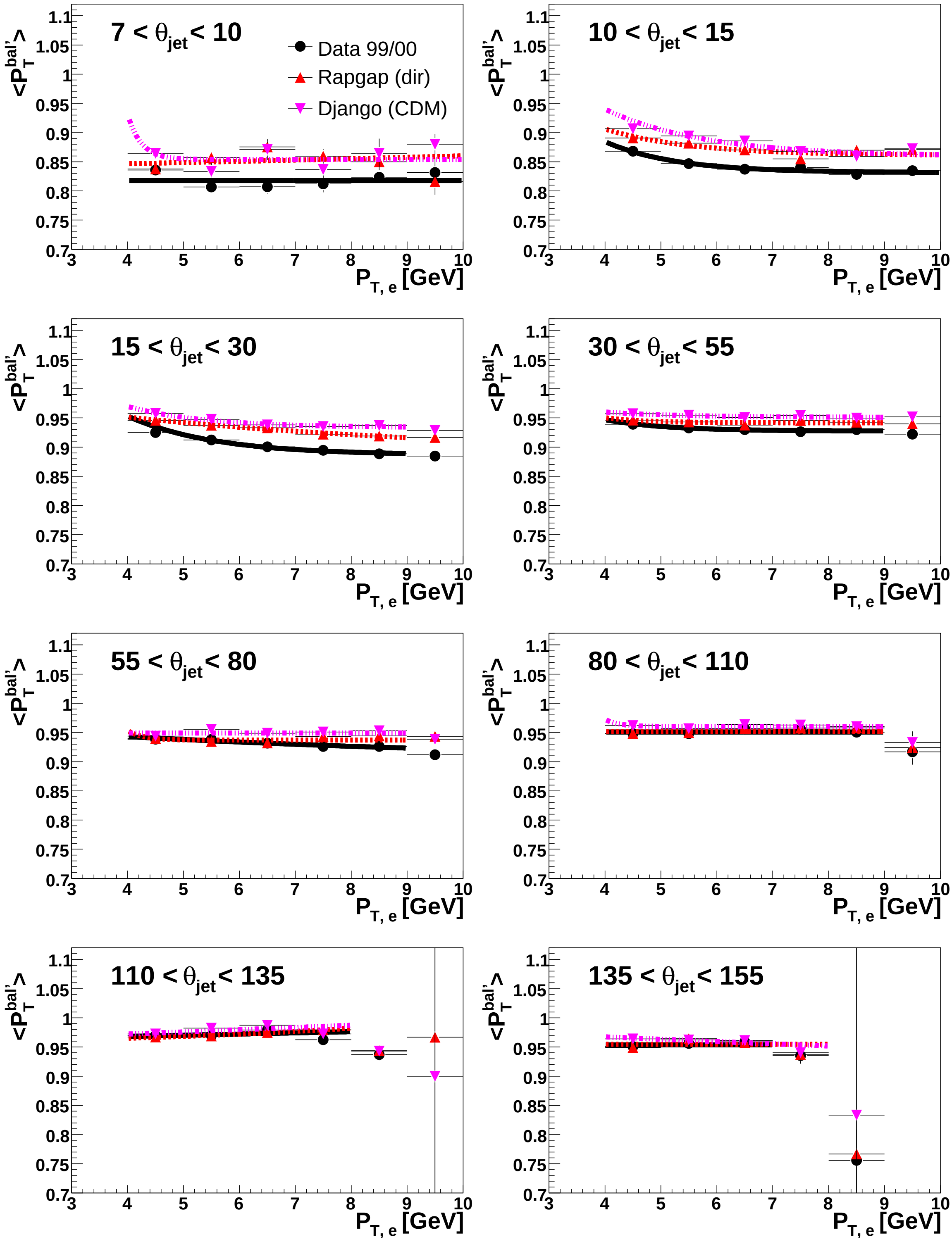


Figure 4.10: *$\langle P_T^{bal'} \rangle$-values as a function of $P_{T,e}$ in bins of θ_{jet} for experimental data collected in the 99/00 run period and Monte Carlo data generated by the Rapgap(dir)and Django(CDM) programs. The function $F^{bal}(\theta, P_{T,e}) = A_{\theta_{jet}}(1 + exp(-B_{\theta_{jet}} - C_{\theta_{jet}}\ P_{T,e}))$ have been fitted to data, Rapgap(dir) and Django(CDM) points and are represented by the full, dashed and dash-dotted lines, respectively.*

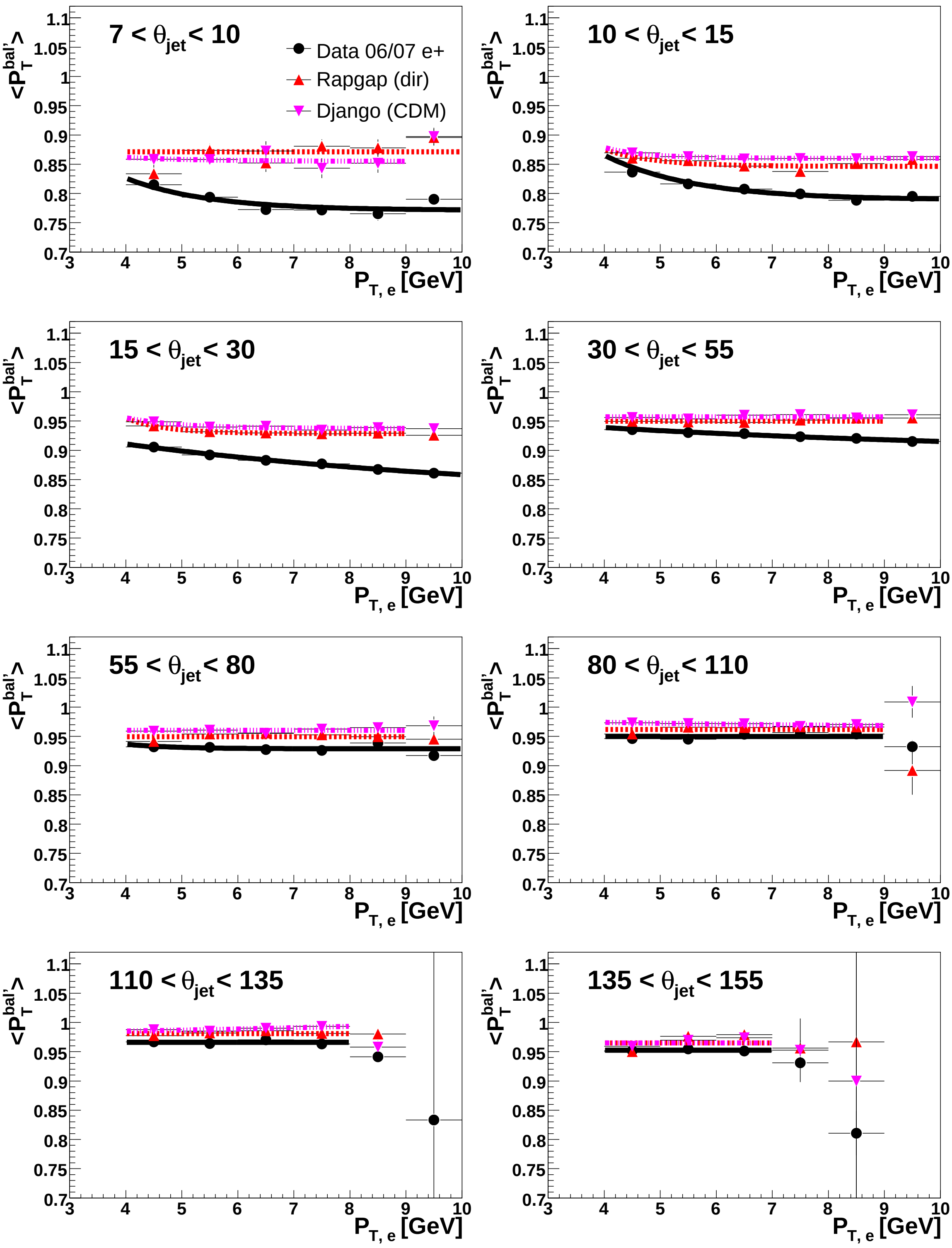


Figure 4.11: *$\langle P_T^{bal'} \rangle$-values as a function of $P_{T,e}$ in bins of θ_{jet} for experimental data collected in the 06/07 run period and Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. The function $F^{bal}(\theta, P_{T,e}) = A_{\theta_{jet}}(1 + exp(-B_{\theta_{jet}} - C_{\theta_{jet}}\ P_{T,e}))$ have been fitted to data, Rapgap(dir) and Django(CDM) points and are represented by the full, dashed and dash-dotted lines, respectively.*

hadronic final state system, it can not be used as reference for the jet measurements. However, the absolute scale of the jet momentum measurement is not known, but has to be determined in the following iterative process. First, the uncalibrated cluster information is inserted in Equation (4.15) and the obtained value of C_{cls} is used to extract a correction factor f_1 from Equation (4.14). In the second step the correction factor f_1 is applied to the cluster energies and a new value of C_{cls} is calculated. This can now be inserted into Equation (4.14) to obtain an improved correction factor f_2. This procedure can be repeated but it has been found that two iterations are enough to get a sufficiently good correction factor [97].

The transverse momentum measurement of jets with $P_{T,jet}$ larger than 9 GeV is calibrated using the High Pt Jet Calibration method, where a slightly different function is used for F^{bal}:

$$F^{bal}(\theta_{jet}, P_{T,e}) \quad = \quad A_{\theta_{jet}}(1 - exp(-B_{\theta_{jet}} - C_{\theta_{jet}} P_{T,e})). \tag{4.16}$$

Clusters which have not been allocated to the jet have not been included in the calibration procedure described above but their energy measurement may be calibrated using the Iterative calibration method. Here the correction factors are different for electromagnetic and hadronic clusters, and for different regions in the LAr calorimeter.

4.4 The Calibration Performance

In order to test the performance of the Low Pt Jet calibration the DIS calibration sample was used with the additional requirement that the events contained exactly one jet with $P_{T,jet} > 3.5$ GeV. Two calibration methods have been considered:

- The transverse momentum measurement of jet clusters is calibrated using the Low and High Pt Jet Calibration methods. In the following this procedure is denoted M3. After the calibration has been completed the jet algorithm is applied once again to the full event i.e. including the energy calibrated jet clusters, the energy uncalibrated clusters not allocated to the jet and the tracks. Hereby information is gained on how frequently a cluster, which was not allocated to the jet in the initial jet reconstruction step, will be included into the jet after energy calibration.
- In order also to include clusters not allocated to the jet in the calibration procedure, a combination of the Low and High Pt Jet Calibration methods with the Iterative method have to be used. Thus, in this method, called M5, all clusters in the event are calibrated.

After the calibration constants have been obtained, all variables that are related to the hadronic final state are recalculated, e.g. $P_{T,h}$, and θ_h, since the performance of the calibration method is tested using the total hadronic final state.

The model dependence of the calibration method is obtained by first extracting the calibration constants using Rapgap(dir) and Django(CDM), respectively, and thereafter applying each constant to both Rapgap(dir) and Django(CDM). This was done using the M3 calibration procedure.

An additional control sample with at least two reconstructed jets of $P_{T,jet} > 3.5$ GeV, the dijet sample, is defined.

4.4.1 The One-Jet Test Sample

Various plots have been produced to illustrate the performance of the calibration procedure. The Figures 4.12, 4.14, 4.15 and 4.17 have been organized in the same way. The plots a) show distributions before calibration. The distributions in the plots b) and d) have been obtained by applying correction factors calculated using the Django(CDM) data, whereas in the distributions c) and e) corrections factors calculated from Rapgap(dir) data have been used. In the plots b) and c) the M3 calibration procedure was followed, while in the plots d) and e) the M5 scheme was used.

In Figure 4.12 $\langle P_T^{bal} \rangle$ is shown as a function of θ_{jet} for experimental data (99/00), and for the Monte Carlo samples of Rapgap(dir) and Django(CDM). Before calibration (Figure 4.12a) the scale is systematically below unity, decreasing to 0.75-0.8 in the forward region for both data and the Monte Carlo samples. With the M3 calibration procedure, the deviations of the absolute scale from unity are smaller than or about 6%, except for the lowest θ_{jet}-bin, where the deviation is almost 10%. Using the M5 calibration scheme improves the situation such that the deviations are approaching 2%, again with the exception of the extreme forward region, where it remains 10%. This deviation originates mainly from particles escaping detection in the forward region, as can be seen from plotting the transverse momentum balance on generator level, in a plot of $P_T^{bal,gen} = P_{T,h}^{gen}/P_{T,e}^{gen}$, shown in Figure 4.13. Although consistent with unity over most of the angular range, the deviations increases up to 6% when approaching the most forward going region. It should be mentioned that an attempt to correct for the missing transverse momentum in the forward region led to a deterioration in the resolution of the jet four-momentum reconstruction due to an overcalibration of the jet energy.

The uncertainty of the absolute energy scale, $DR(P_T^{bal})$, as a function of θ_{jet}, is presented in Figure 4.14. Before calibration the uncertainty is outside the 2% window in the forward direction, whereas after calibration the uncertainty is well inside 2%. In this representation some dependence on the correction factors calculated from the Rapgap(dir) and Django(CDM) samples can be observed which is of the order of 1% . A difference of about 1.5% in $DR(P_T^{bal})$ obtained from Rapgap(dir) and Django(CDM) is also observed. The M3 calibration scheme gives somewhat smaller uncertainties than the M5 scheme.

Figure 4.15 and Figure 4.17 show how the $\langle P_T^{bal} \rangle$ and $DR(P_T^{bal})$ distributions depend on $P_{T,e}$. An imbalance is observed as $P_{T,e}$ increases and this remains after calibration, although it gets smaller. The difference between data and the MC predictions, seen before calibration, disappears once the calibration has been performed. Events of the one-jet test sample predominantly contain forward jets as $P_{T,e}$ gets high. This is demonstrated by Figure 4.16, where the correlation between θ_{jet} and $P_{T,e}$ is shown. The fraction of one-jet events where the jet is in the forward region and $P_{T,e}$ is in the range up to 6.5 GeV, constitutes around 25%, and should be compared to the fraction of forward jet events where $P_{T,e}$ is above 6.5 GeV, which is as high as 45%. Thus, the increase of the imbalance at high $P_{T,e}$ is due to particle leakage in the forward region i.e. the imbalance observed in Figure 4.12 for small θ_{jet} and in Figure 4.15 for high $P_{T,e}$ have the same origin. From Figure 4.17 it is established that the uncertainty of the absolute scale, $DR(P_T^{bal})$, is smaller than 2% over the full $P_{T,e}$ range for both the M3 and M5 schemes, although M3 exhibits a somewhat better performance.

The mean value of the ratio between the reconstructed and generated $P_{T,jet}$ distributions, $\langle P_{T,jet}^{rec}/P_{T,jet}^{gen}\rangle$, together with the relative resolution, defined as the standard deviation, σ, of the $P_{T,jet}^{rec}/P_{T,jet}^{gen}$ distribution, are both plotted as a function of $P_{T,jet}^{gen}$ in Figure 4.18 and as a function of θ_{jet}^{gen} in Figure 4.19. An improvement in the absolute energy scale of the jet measurements in going from uncalibrated to calibrated data can be observed over the full $P_{T,jet}^{gen}$ range. The relative jet energy resolution is somewhat worse for calibrated data compared to uncalibrated at lower $P_{T,jet}^{gen}$ but becomes better at higher $P_{T,jet}^{gen}$. This is due to the fact that some jets with low transverse momenta in the MC data, which originally failed to fulfill the $P_{T,jet} > 3.5$ GeV requirement on detector level might do so after the calibration and thus deteriorates the energy resolution of calibrated data compared to uncalibrated. The relative resolution is somewhat worse for energy calibrated data compared to uncalibrated data over the full angular range. In Figures 4.20 and 4.21 the mean value of the ratio between the reconstructed and generated $P_{T,h}$ distributions, $\langle P_{T,h}^{rec}/P_{T,h}^{gen}\rangle$, together with the relative resolution, defined as the standard deviation, σ, of the $P_{T,h}^{rec}/P_{T,h}^{gen}$ distribution, are shown as a function of $P_{T,jet}^{gen}$ and θ_{jet}^{gen}. The $P_{T,jet}^{gen}$ dependence of uncalibrated data almost disappears once the data have been calibrated. The relative resolution shows little dependence of $P_{T,jet}^{gen}$. The θ_{jet}^{gen} dependence of the absolute scale and the relative resolution is similar to what was observed for jets.

Plots similar to the above have been produced to study the performance of the calibration using data from the 2006-2007 run periods. The data, compared to Rapgap(dir) and Django(CDM), are presented in Figures 4.22 and 4.25, where only calibration factors calculated from the Django(CDM) sample using the calibration scheme M3. The overall performance is similar to that found in the 1999-2000 data sample.

The performance of the Method M5 for the 2006-2007 data sample is not presented, since it is similar to that found in the 1999-2000 data sample i.e. the performance of the method M5 is close to that of M3.

4.4.2 The Two-Jet Test Sample

The absolute scale and the uncertainty in the absolute scale obtained from the two-jet sample collected in the years 1999-2000 are shown as a function of $P_{T,e}$ in Figures 4.26 and 4.27, and as a function of θ_h in Figures 4.28 and 4.29. The data are compared to Rapgap(dir) and Django(CDM). Only correction factors extracted from the Django(CDM) sample were used together with the calibration procedure M3. Again the general behaviour of the distributions are similar to those of the one-jet sample, with the uncertainty of the absolute energy scale within 2% for the energy calibrated data.

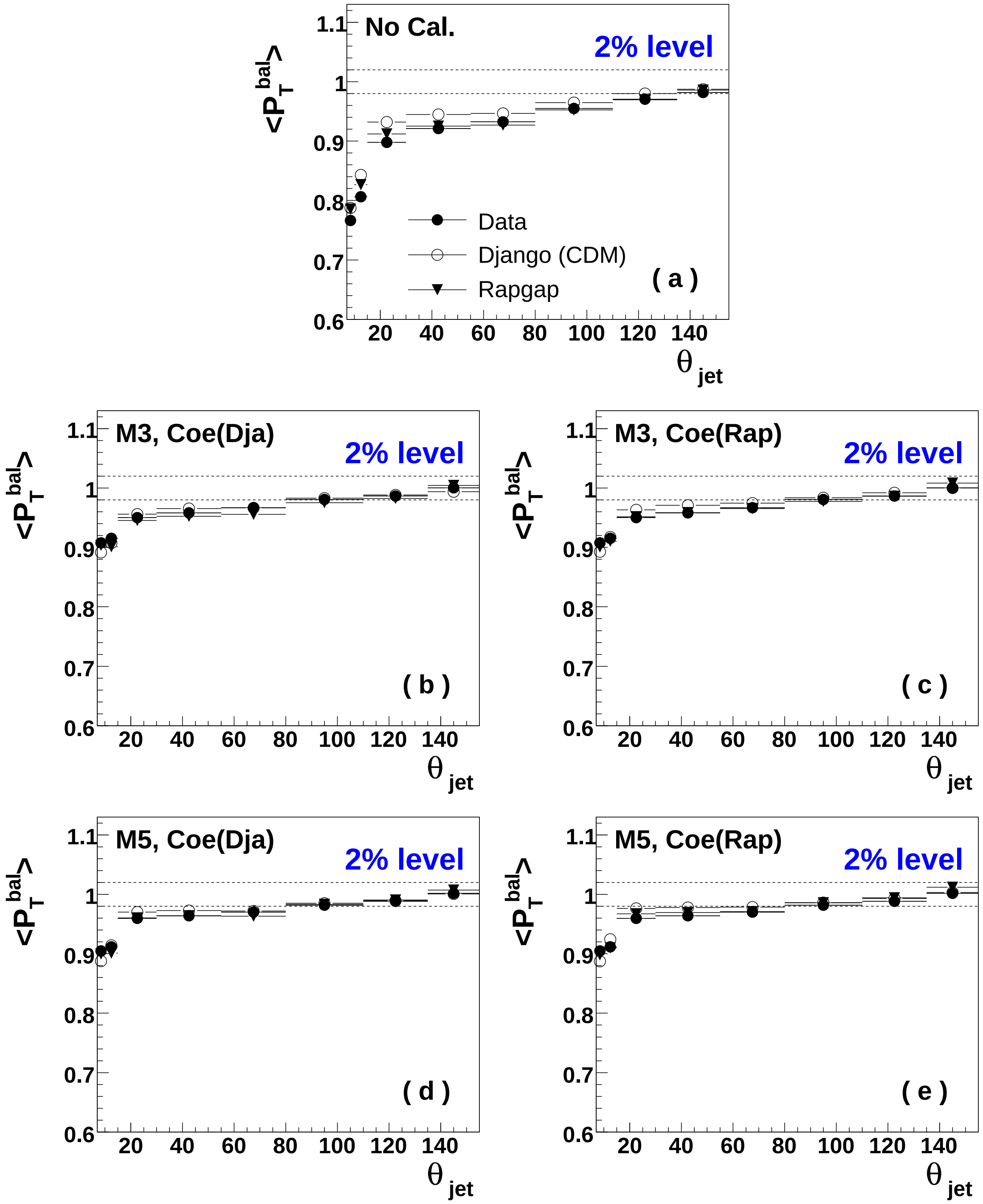


Figure 4.12: $\langle P_T^{bal} \rangle$ *as a function of* θ_{jet} *for the one-jet test sample, extracted from experimental data (99/00) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data, in plots (b) and (c) distributions for data calibrated according to the M3 scheme and in plots (d) and (e) distributions calibrated according to the M5 procedure. In the plots (b) and (d) Django(CDM) was used to calculate the calibration factors, whereas in plots (d) and (e) the calibration factors were obtained using Rapgap(dir).*

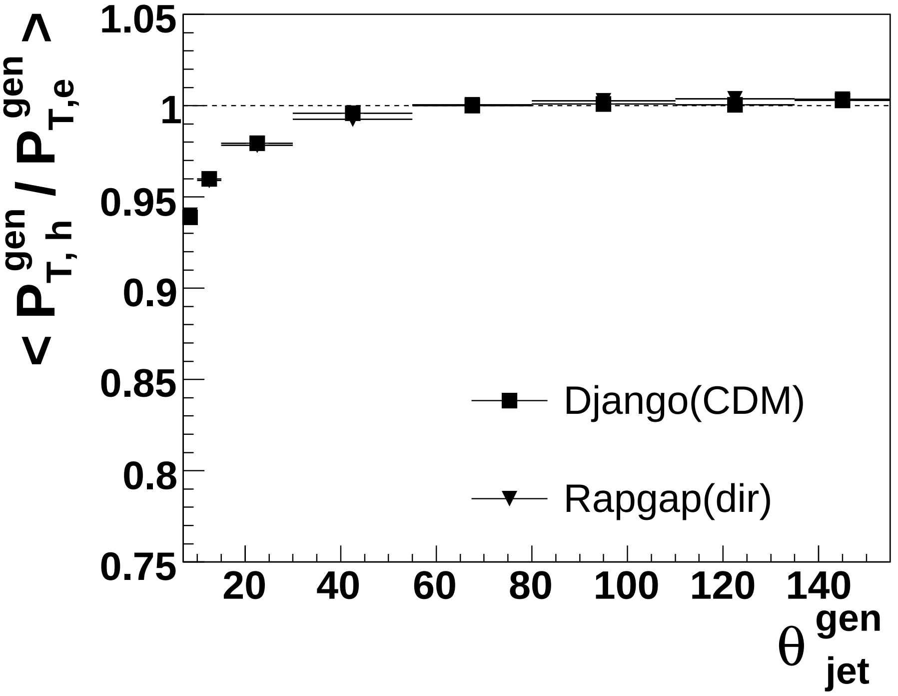


Figure 4.13: *The mean value of the P_T balance as a function of θ_{jet} on generator level over the full range of the detector coverage, $3^o < \theta < 176^o$.*

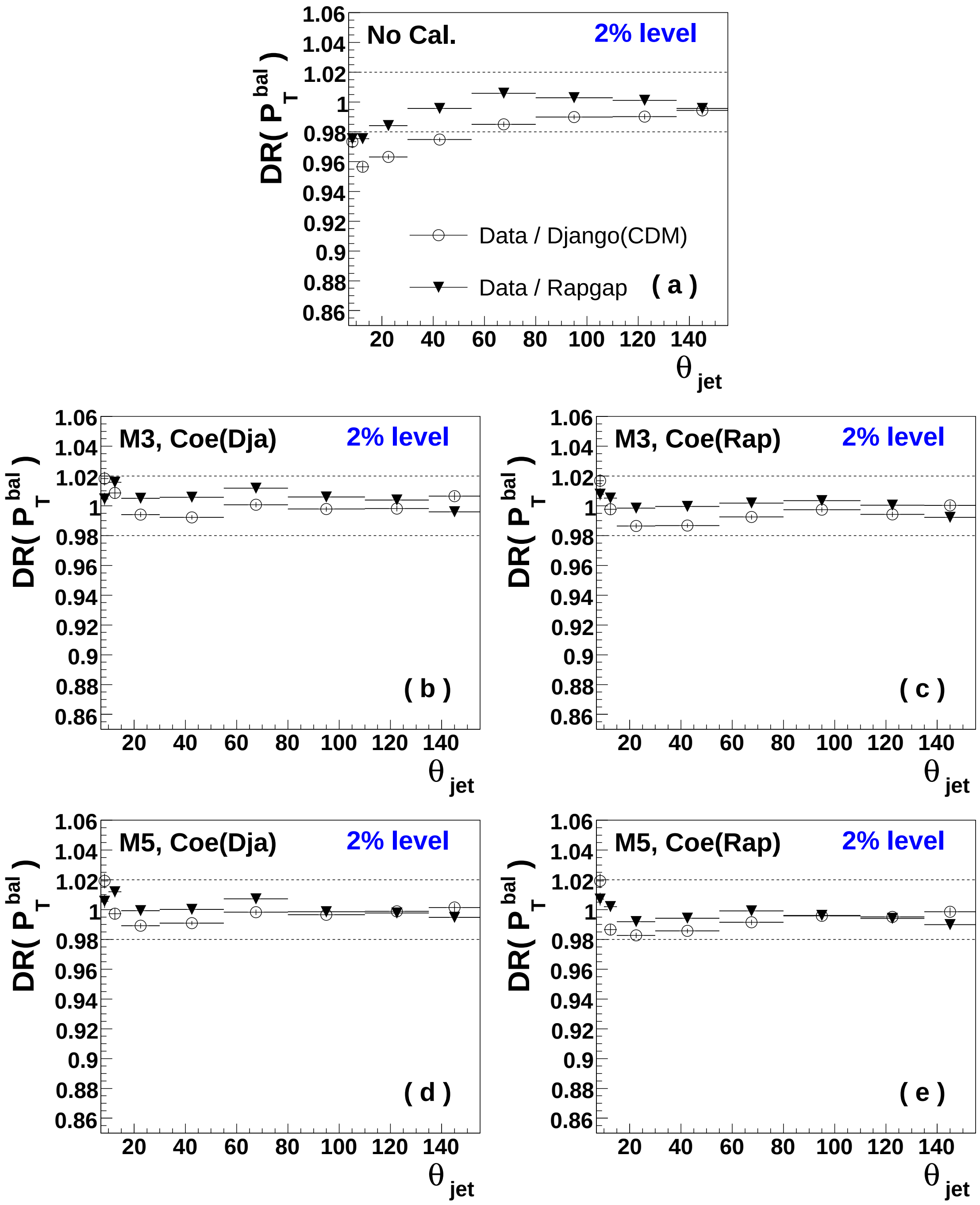


Figure 4.14: *$DR(P_T^{bal})$ distributions as a function of θ_{jet} for the ratio between $\langle P_T^{bal}\rangle$ of experimental data (99/00), using the one-jet test sample, and of Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data, in plots (b) and (c) distributions for data calibrated according to the M3 scheme and in plots (d) and (e) distributions calibrated according to the M5 procedure. In the plots (b) and (d) Django(CDM) was used to calculate the calibration factors, whereas in plots (d) and (e) the calibration factors were obtained using Rapgap(dir).*

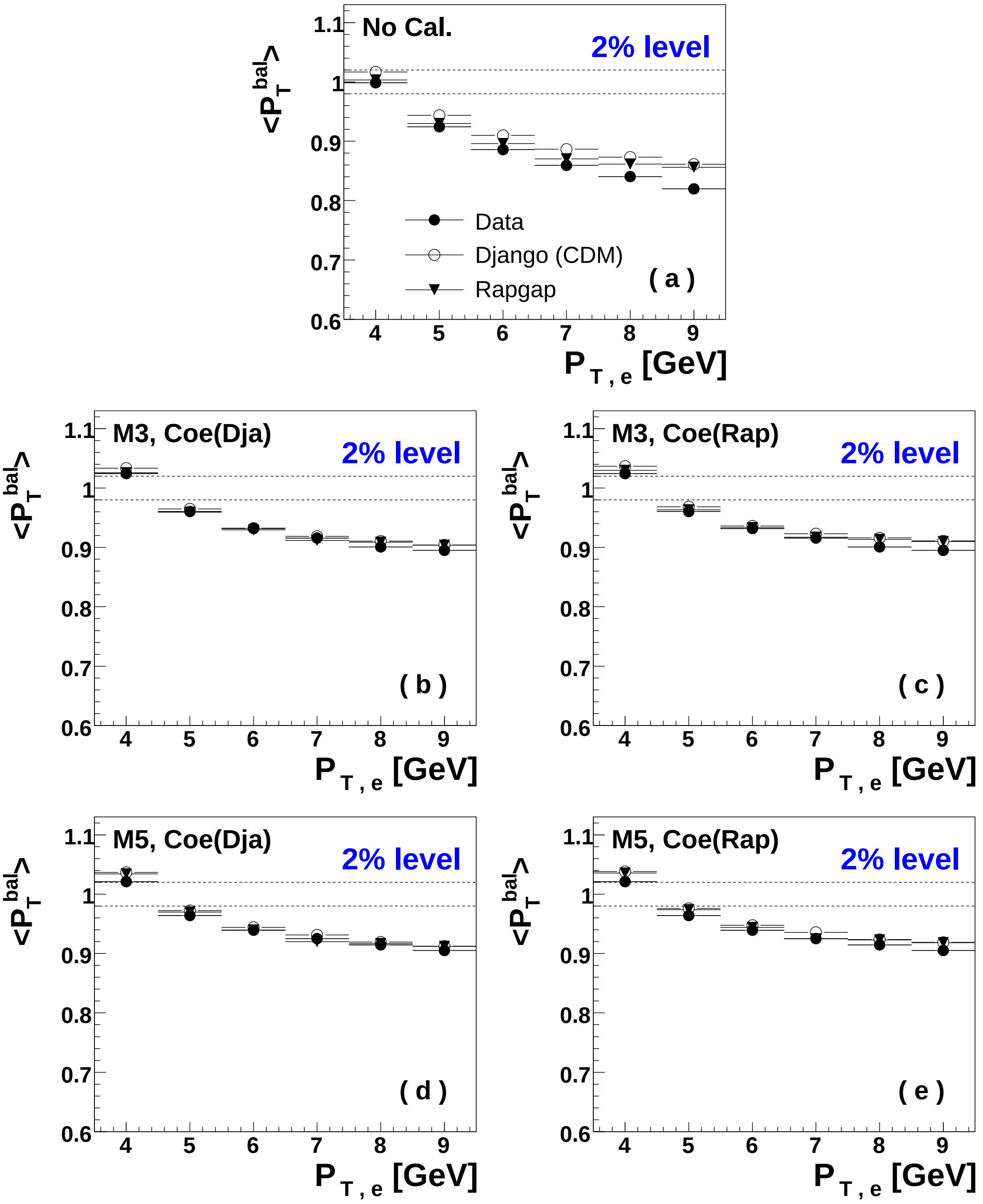


Figure 4.15: $\langle P_T^{bal}\rangle$ *as a function of* $P_{T,e}$ *for the one-jet test sample, extracted from experimental data (99/00) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data, in plots (b) and (c) distributions for data calibrated according to the M3 scheme and in plots (d) and (e) distributions calibrated according to the M5 procedure. In the plots (b) and (d) Django(CDM) was used to calculate the calibration factors, whereas in plots (d) and (e) the calibration factors were obtained using Rapgap(dir).*

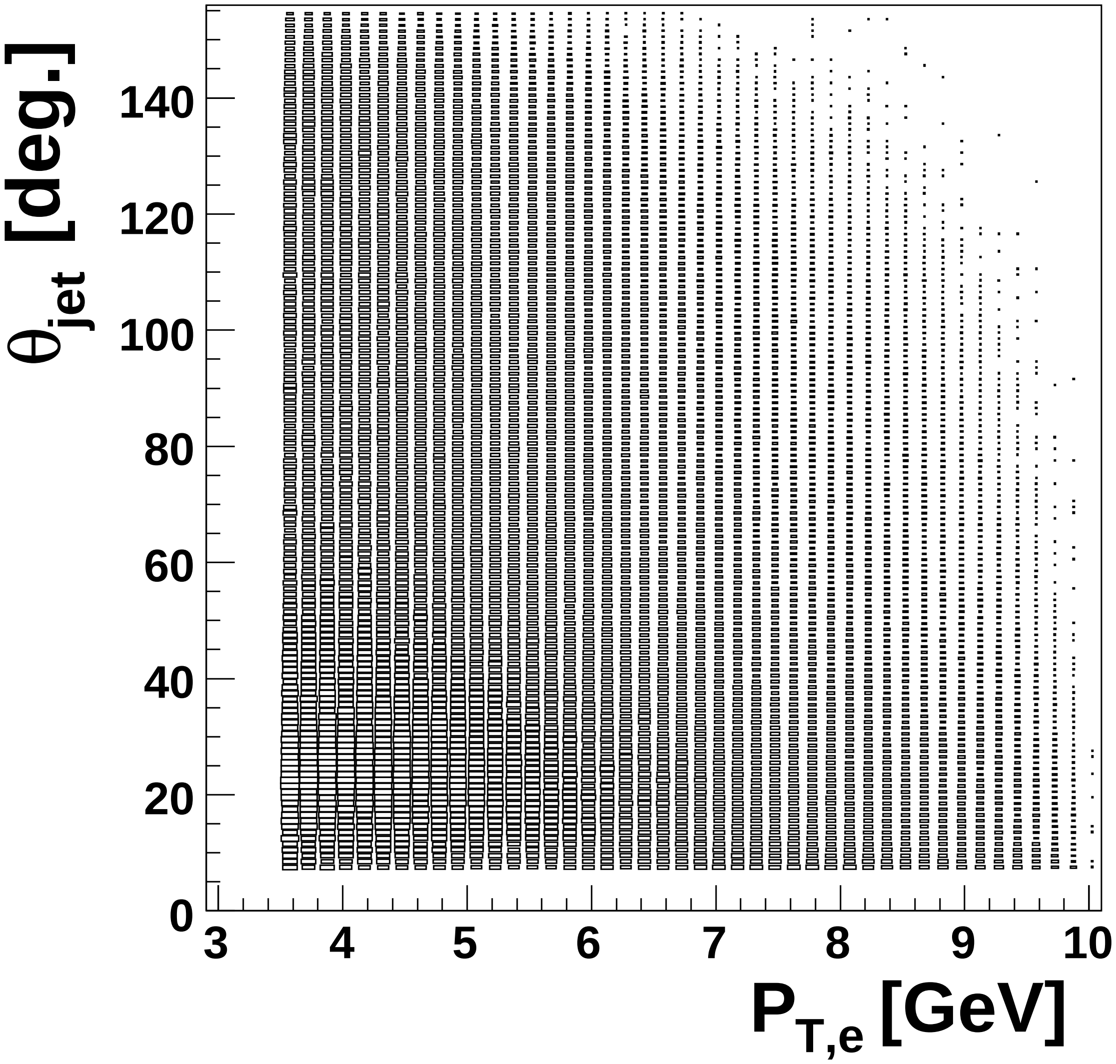


Figure 4.16: *The correlation between the polar angle of the QPM jet, θ_{jet}, and the transverse momentum of the scattered electron, $P_{T,e}$, for the one-jet test sample.*

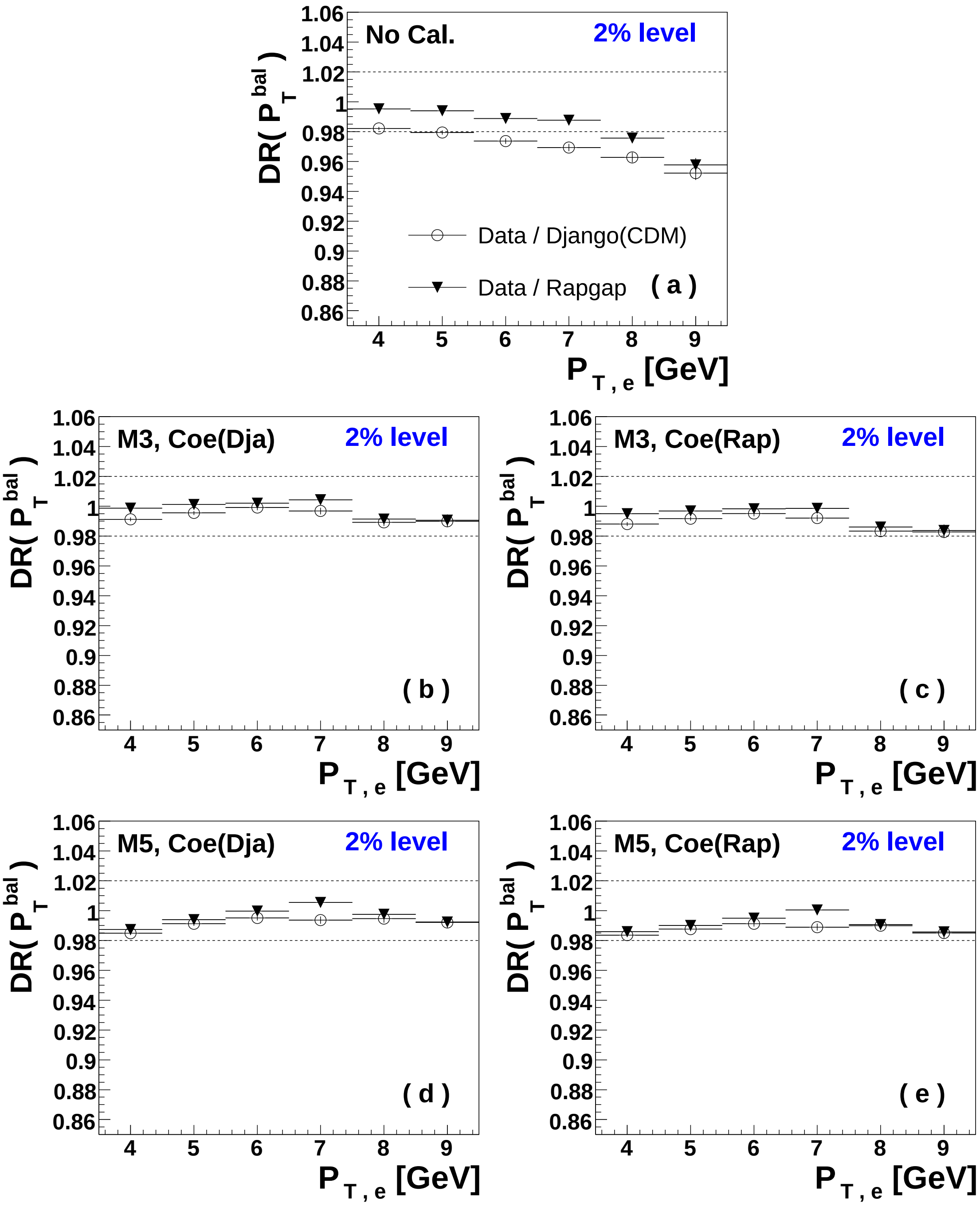


Figure 4.17: *$DR(P_T^{bal})$ distributions as a function of $P_{T,e}$ for the ratio between $\langle P_T^{bal}\rangle$ of experimental data (99/00), using the one-jet test sample, and of Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data, in plots (b) and (c) distributions for data calibrated according to the M3 scheme and in plots (d) and (e) distributions calibrated according to the M5 procedure. In the plots (b) and (d) Django(CDM) was used to calculate the calibration factors, whereas in plots (d) and (e) the calibration factors were obtained using Rapgap(dir).*

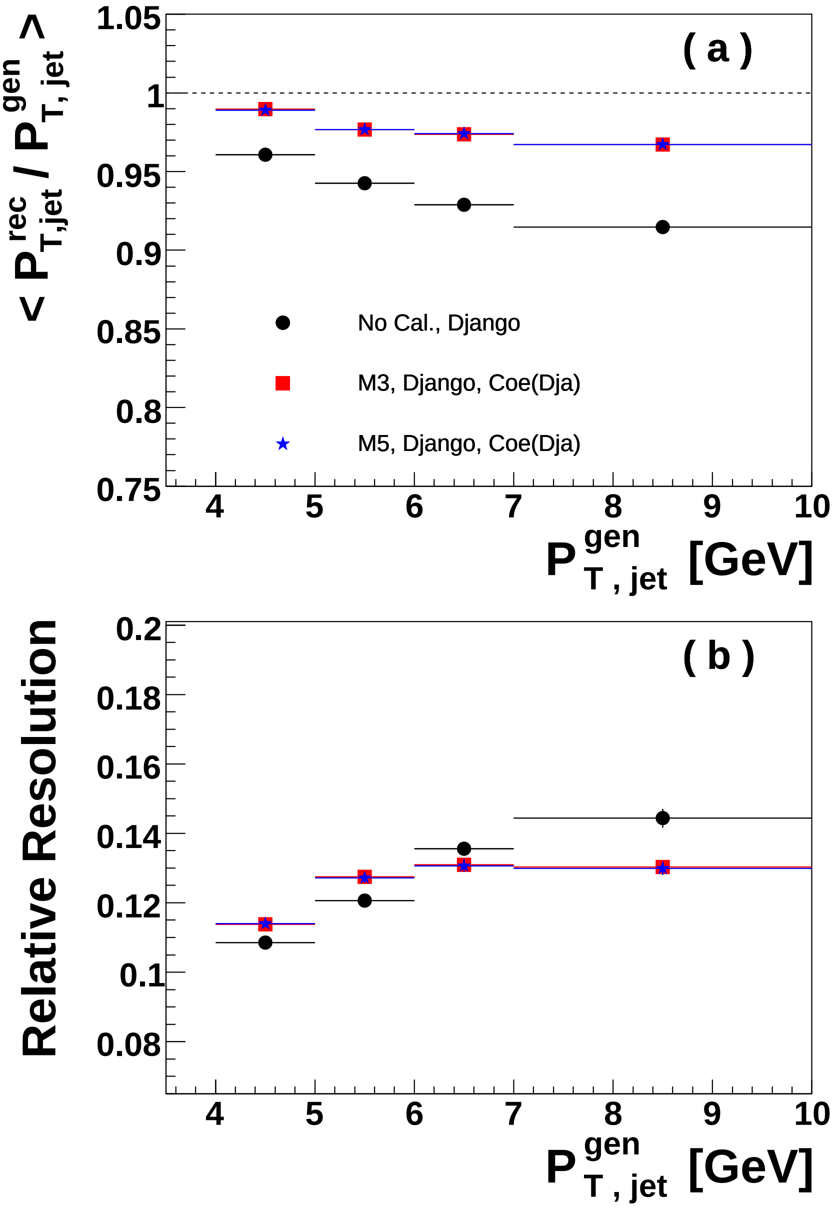


Figure 4.18: *The mean value (a) and the relative resolution (b) of the* $P^{rec}_{T,jet}/P^{gen}_{T,jet}$ *distributions as function of* $P^{gen}_{T,jet}$ *for the one-jet test sample generated by the Django(CDM) program. Compared are the distributions of uncalibrated data (●), and those where the calibration was performed according to the M3 (■) and the M5 (⋆) recipe, using calibration coefficients obtained from Django(CDM). The definition of the relative resolution is given in the text.*

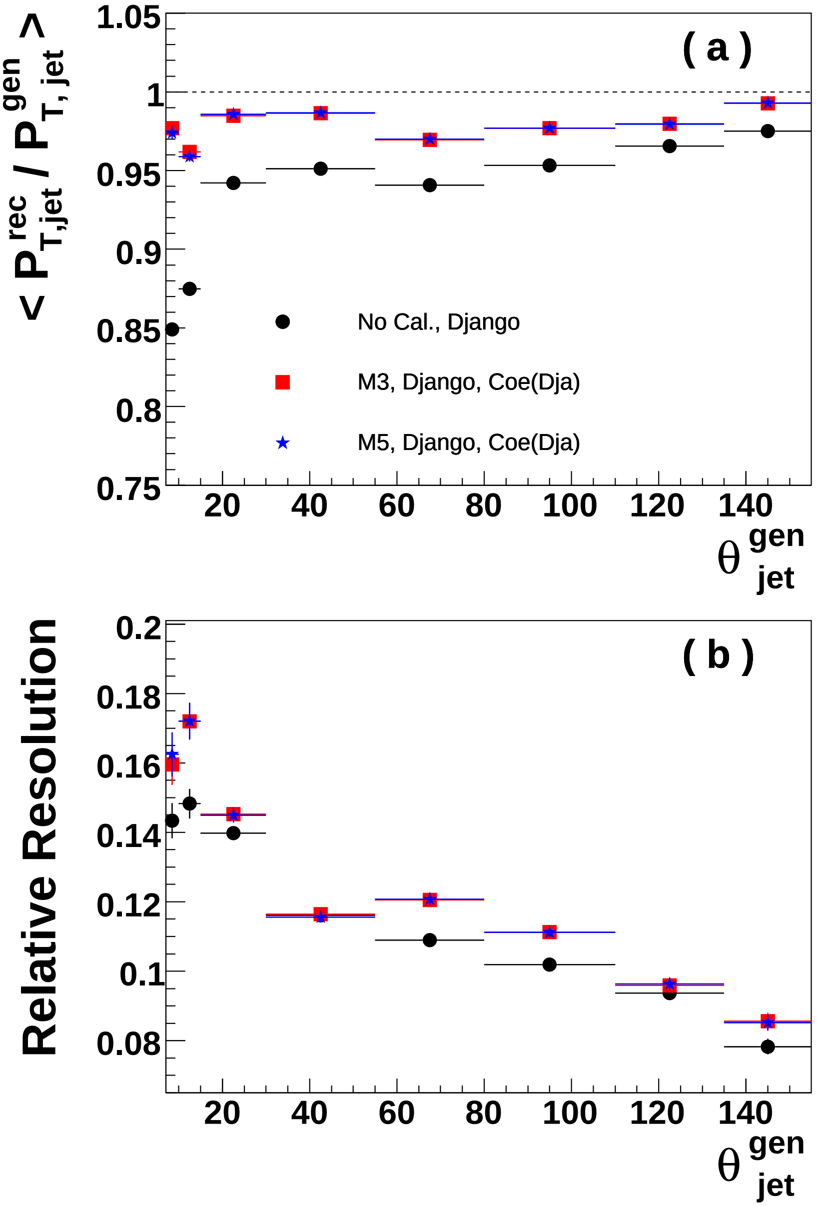


Figure 4.19: *The mean value (a) and the relative resolution (b) of the* $P^{rec}_{T,jet}/P^{gen}_{T,jet}$ *distributions as function of* θ^{gen}_{jet} *for the one-jet test sample generated by the Django(CDM) program. Compared are the distributions of uncalibrated data (●), and those where the calibration was performed according to the M3 (🟥) and the M5 (⋆) recipe, using calibration coefficients obtained from Django(CDM). The definition of the relative resolution is given in the text.*

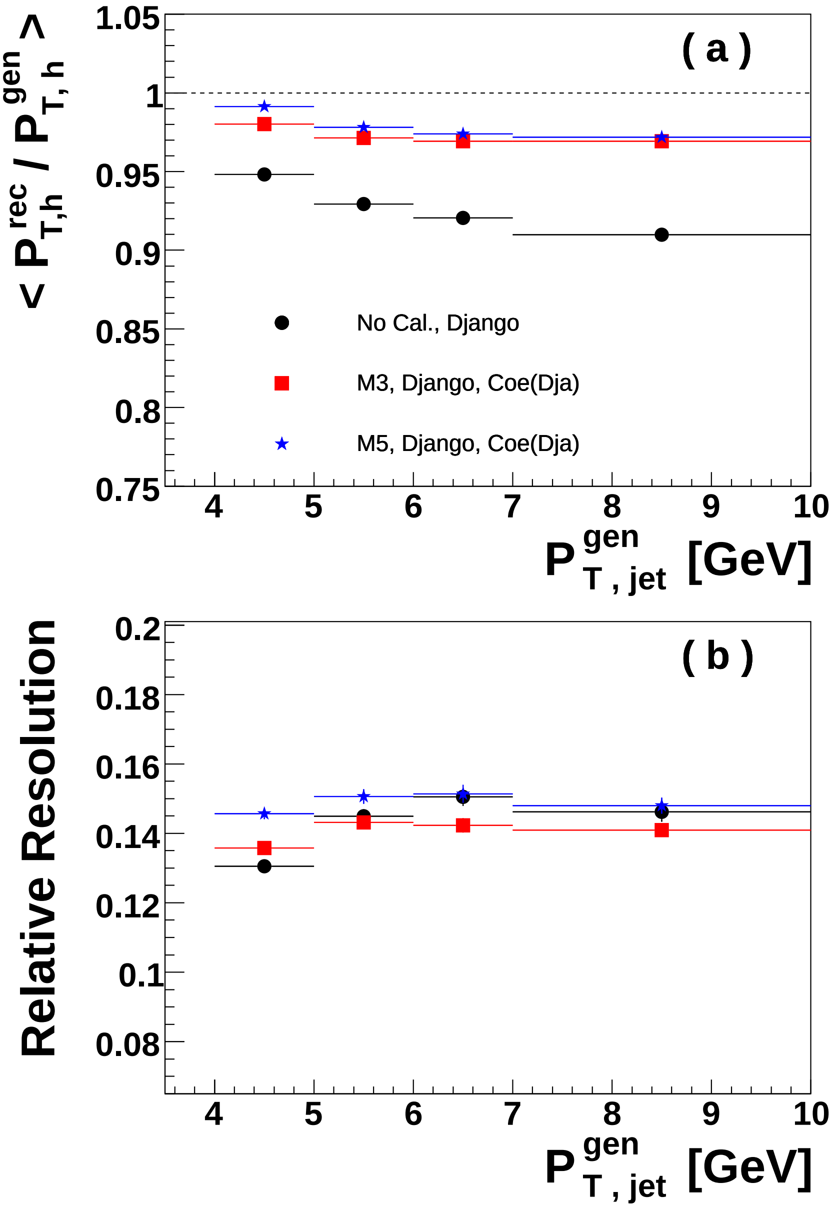


Figure 4.20: *The mean value (a) and the relative resolution (b) of the* $P_{T,h}^{rec}/P_{T,jet}^{gen}$ *distribution as function of* $P_{T,jet}^{gen}$ *for the one-jet test sample generated by the Django(CDM) program. Compared are the distributions of uncalibrated data (●), and those where the calibration was performed according to the M3 (■) and the M5 (⋆) recipe, using calibration coefficients obtained from Django(CDM). The definition of the relative resolution is given in the text.*

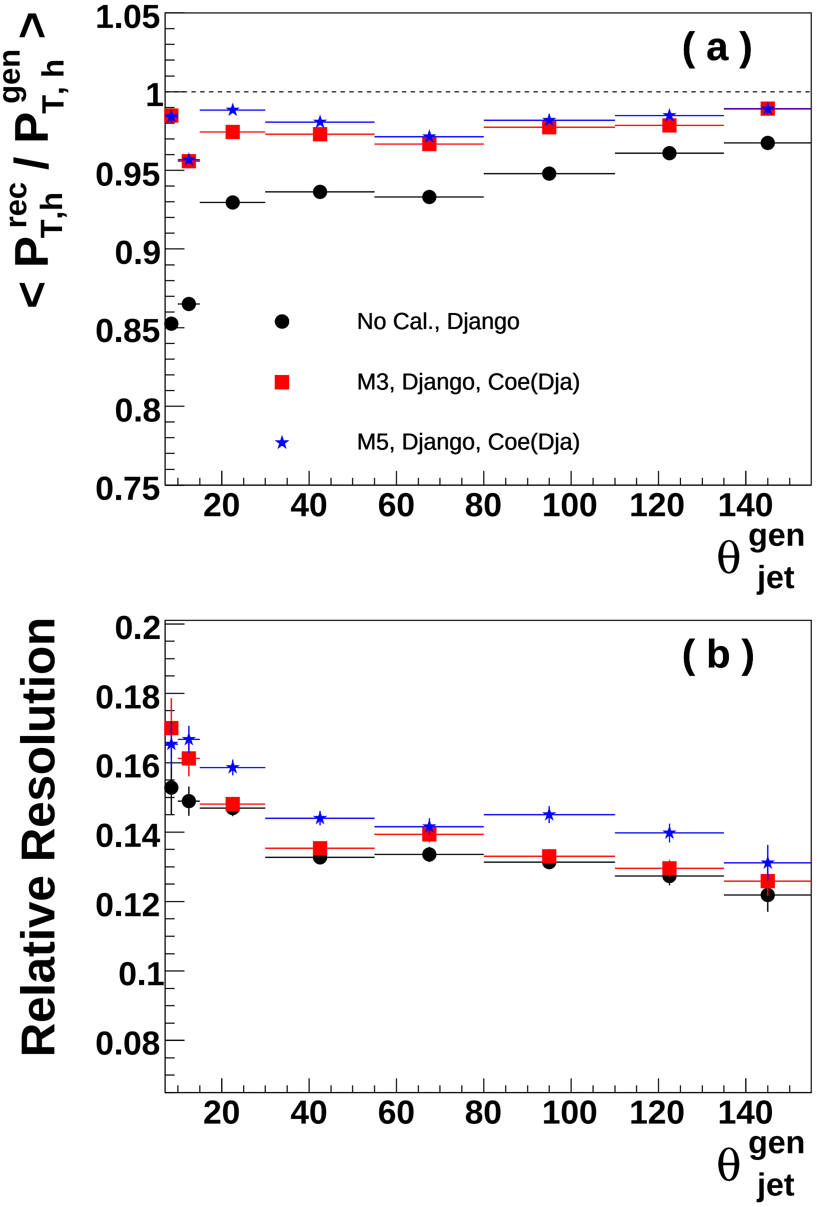


Figure 4.21: *The mean value (a) and the relative resolution (b) of the* $P_{T,h}^{rec}/P_{T,h}^{gen}$ *distribution as function of* θ_{jet}^{gen} *for the one-jet test sample generated by the Django(CDM) program. Compared are the distributions of uncalibrated data (●), and those where the calibration was performed according to the M3 (■) and the M5 (⋆) recipe, using calibration coefficients obtained from Django(CDM). The definition of the relative resolution is given in the text.*

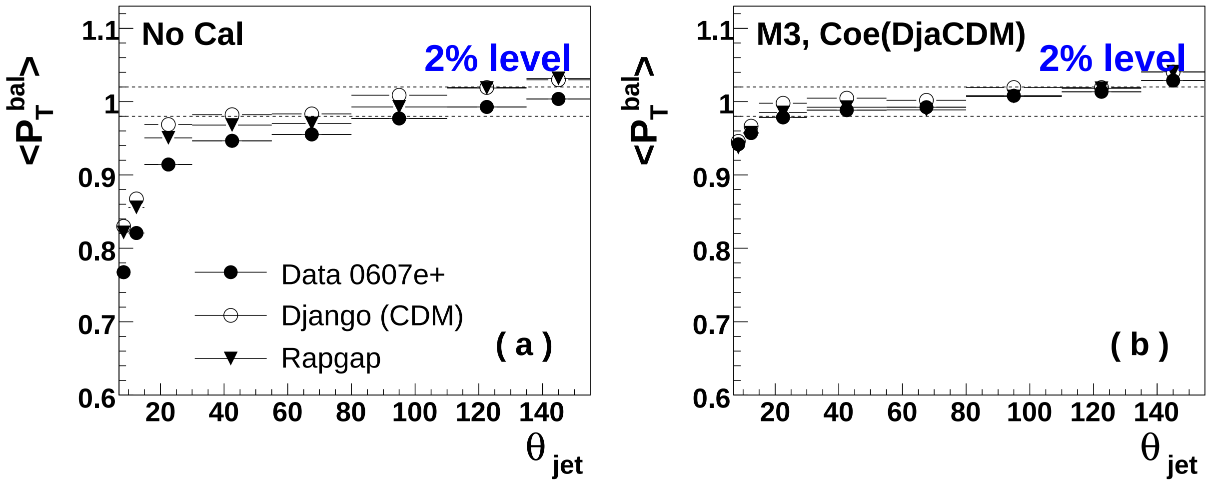


Figure 4.22: $\langle P_T^{bal} \rangle$ *distributions as a function of* θ_{jet} *for the one-jet test sample, extracted from experimental data (06/07) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3 scheme, where Django(CDM) was used to calculate the calibration factors.*

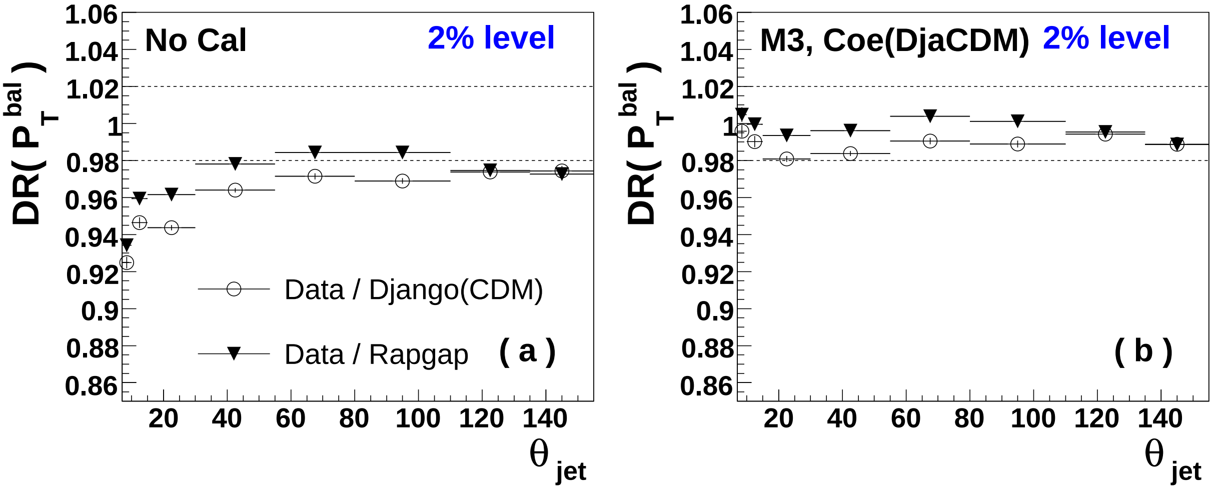


Figure 4.23: $DR(P_T^{bal})$ *distributions as a function of* θ_{jet} *for the ratio between* $\langle P_T^{bal} \rangle$ *of experimental data (06/07), using the one-jet test sample, and of Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3, where Django(CDM) was used to calculate the calibration factors.*

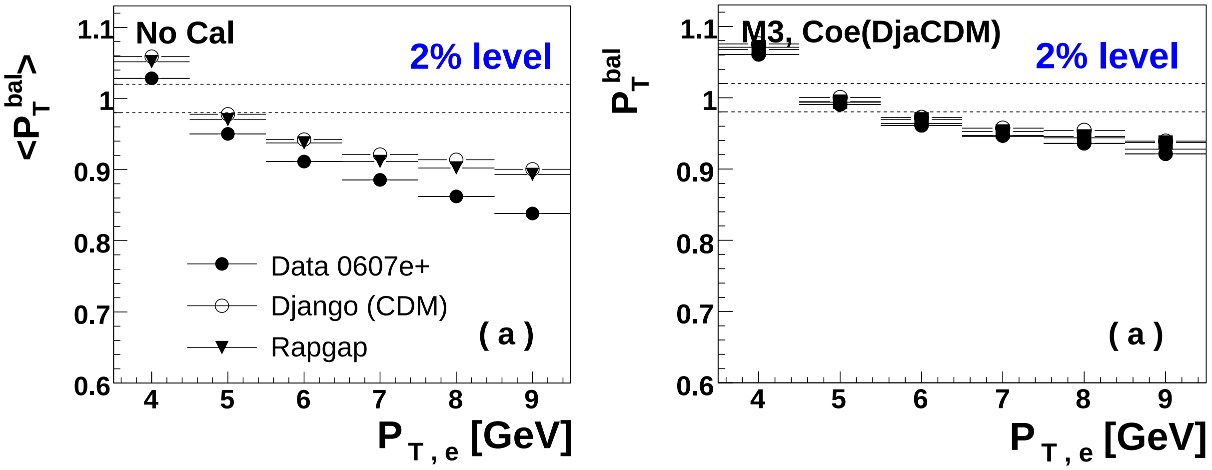


Figure 4.24: $\langle P_T^{bal} \rangle$ *distributions as a function of* $P_{T,e}$ *for the one-jet test sample, extracted from experimental data (06/07) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3 scheme, where Django(CDM) was used to calculate the calibration factors.*

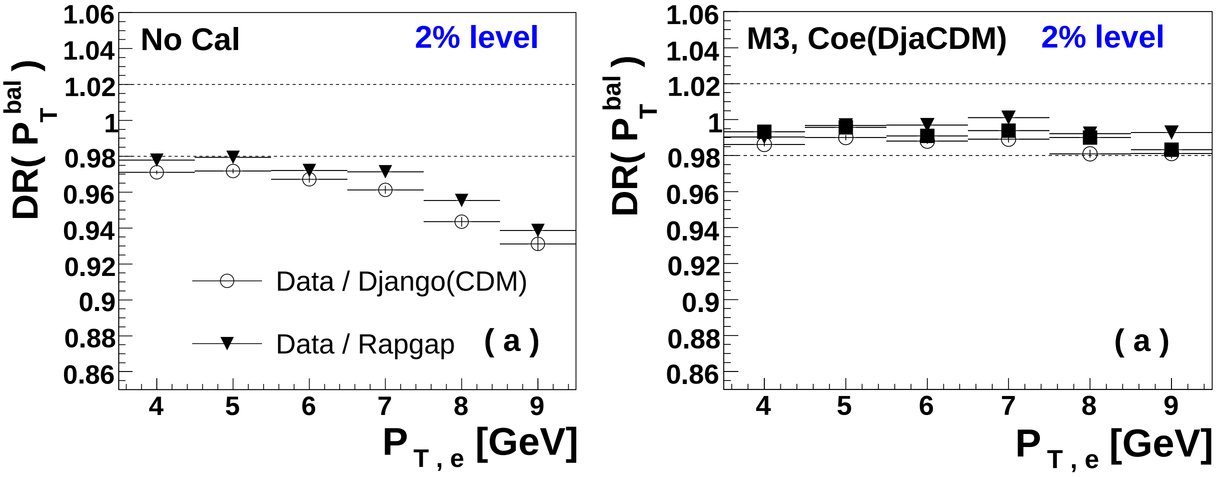


Figure 4.25: $DR(P_T^{bal})$ *distributions as a function of* $P_{T,e}$ *for the ratio between* $\langle P_T^{bal} \rangle$ *of experimental data (06/07), using the one-jet test sample, and of Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3, where Django(CDM) was used to calculate the calibration factors.*

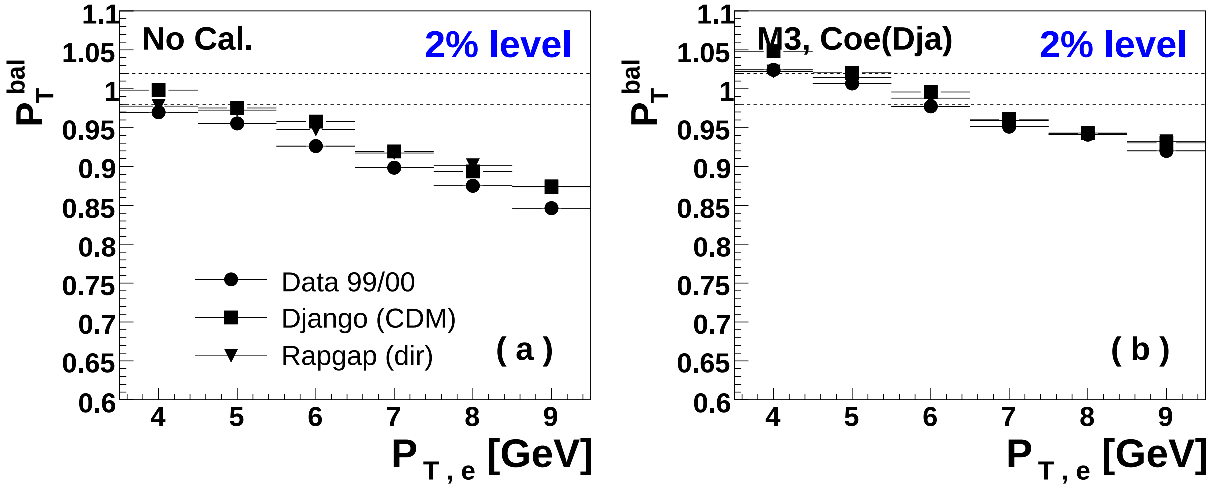


Figure 4.26: $\langle P_T^{bal} \rangle$ *distributions as a function of* θ_{jet} *for the two-jet test sample, extracted from experimental data (99/00) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3 scheme, where Django(CDM) was used to calculate the calibration factors.*

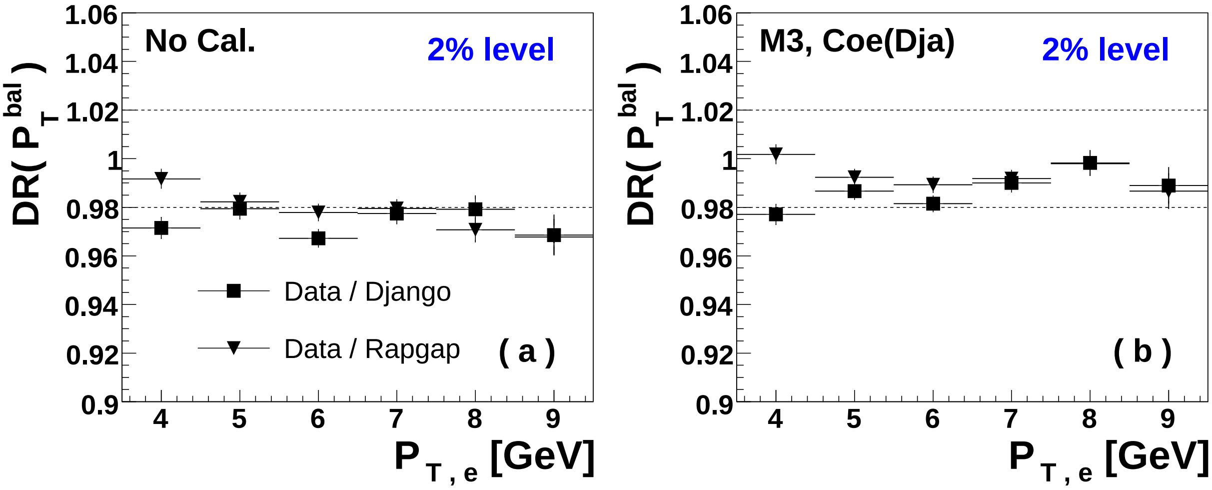


Figure 4.27: $DR(P_T^{bal})$ *distributions as a function of* $P_{T,e}$ *for the ratio between experimental data (99/00), using the two-jet test sample, and Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3, where Django(CDM) was used to calculate the calibration factors.*

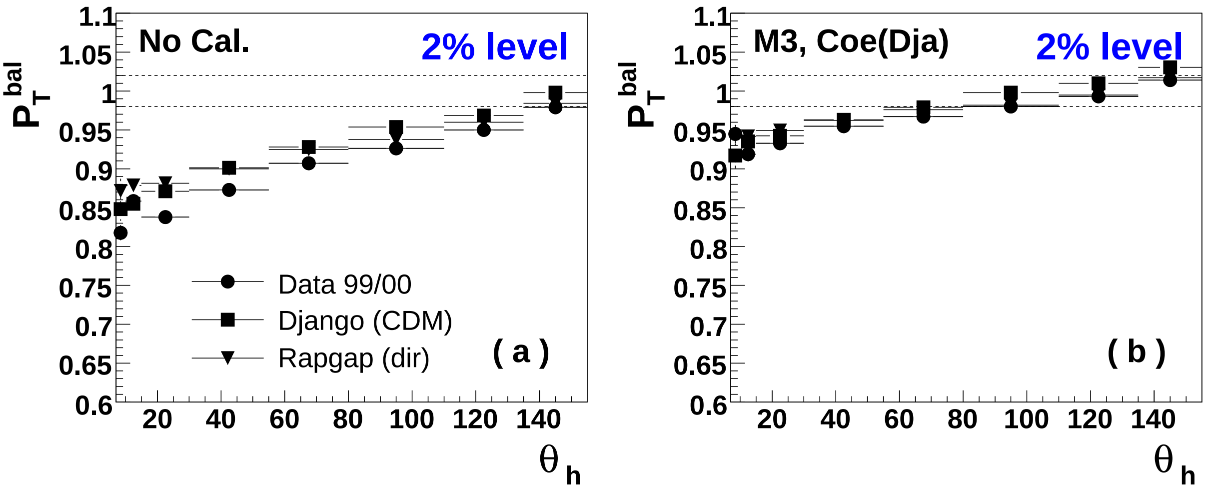


Figure 4.28: $\langle P_T^{bal} \rangle$ *distributions as a function of* θ_h *for the two-jet test sample, extracted from experimental data (99/00) and from Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3 scheme, where Django(CDM) was used to calculate the calibration factors.*

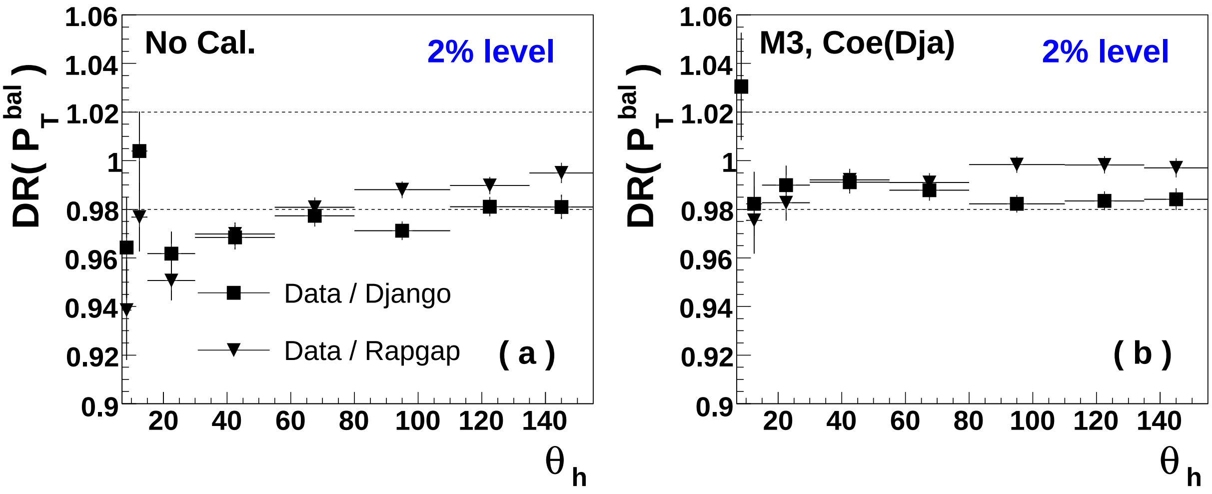


Figure 4.29: $DR(P_T^{bal})$ *distributions as a function of* θ_h *for the ratio between experimental data (99/00), using the two-jet test sample, and Monte Carlo data generated by the Rapgap(dir) and Django(CDM) programs. Shown are in plot (a) the distributions for uncalibrated data and in plot (b) distributions for data calibrated according to the M3, where Django(CDM) was used to calculate the calibration factors.*

4.5 Summary, Comments and Outlook

Measurements of jet energies suffer from statistical fluctuations due to the evolution of the hadronic showers but also from systematic errors due to the fact that a large fraction of the jet energy will not be measured in the detector. In physics analyses involving measurements of jet energies in H1, the uncertainty in the absolute scale of the energy measurement has so far provided the biggest contribution to the total systematic error. Thus, it is strongly motivated to try to improve the jet calibration procedure. In H1, a calibration method to reduce the uncertainty in the absolute scale of jet energy measurements at $P_T > 10$ GeV is available. Here, an improved method to calibrate energy measurements of low P_T jets, $P_T < 9$ GeV, has been introduced and tested for data taken during the years 1999-2000 and 2006-2007.

The SPACAL is used to measure the scattered electron of DIS events at low Q^2. The absolute scale of the electron energy measurement in the SPACAL is known to better than 1% and thus the measured energy of the scattered electron can be used as reference to compare with the energy measurement of hadronic energy in the LAr calorimeter. In order to extract the absolute scale of jet energy measurements, DIS events with exactly one reconstructed jet were selected. In the ideal case the transverse momentum of the scattered electron should be balanced by the transverse momenta of the hadrons in the jet, if one disregards the problem of associating some soft particles to either the jet or the proton remnant. Thus, the ratio $P_T^{bal} \equiv P_{T,h}/P_{T,e}$ is a measure of the absolute energy scale and the double ratio $DR(P_T^{bal}) = \langle P_T^h/P_{T,e}\rangle_{data}/\langle P_T^h/P_{T,e}\rangle_{MC}$ gives the uncertainty in the absolute scale. The electron method is used to reconstruct the transverse momentum of the scattered electron. Different selection criteria have been tried in order to improve the purity of the one-jet calibration sample.

In the Low Pt Jet Calibration method only those clusters which have been measured in the LAr calorimeter and allocated to the jet are included in the calibration i.e. energy contributions to the jet from track measurements are not considered. The calibration factors are calculated for different θ regions, corresponding to the polar angle coverage of the LAr calorimeter wheels.

Using the Jet calibration sample, selected from the 99/00 data, an absolute energy scale could be determined to better than 2% over the full central region ($15^o < \theta < 155^o$) and to about 10% in the forward region ($\theta < 15^o$). It was shown that (4-6)% of the total imbalance in the forward region is caused by the limited θ acceptance of the calorimeter. In the central region it should be possible to determine the absolute calibration to better than 2% but this needs further investigation. The limited acceptance of the LAr calorimeter in the forward direction is also responsible for a large fraction of the transverse momentum imbalance at high $P_{T,e}$.

The uncertainty in the absolute energy scale for jet measurements is about 2% after calibration. The relative resolution of the jet transverse momentum measurement is deteriorated after calibration for $P_{T,jet} <$ 6 GeV. One explanation for this could be that a global event quantity, $P_T^{bal} = P_{T,h}/P_{T,e}$, based on the final state hadrons, is used to calibrate jets, which are regarded as single objects. An improved calibration might be possible using a different definition of P_T^{bal} and/or a different selection of the one-jet sample, which might reduce the influence of clusters outside the jet.

Two Monte Carlo event generators, Rapgap(dir) and Django(CDM), are used to estimate the dependence of the calibration on the choice of model. The model dependence is observed to be on the level of 1.5%. The difference in performance of the M3 and M5 calibration schemes is marginal. M3 is better in some respects and M5 in other. This indicates that clusters not allocated to a jet have significant impact on the calculation of the correction factors.

The calibration of data from the 2006-2007 run periods, show the same overall performance as for data taken in 1999-2000. Due to the large data sample collected in 2006-2007, it might be possible to use events from diffractive scattering to derive the calibration coefficients. Since there is no colour string between the scattered parton and the proton in diffractive events, it is less ambigous to reconstruct the jet and there should be no contributions from clusters not belonging to the jet. Diffractive events are selected by requiring a rapidity gap in the forward region, with no hadronic activites.

Chapter 5

Multiple Interactions in DIS

Following the discussions in section 1.5.2 an analysis dedicated to study a possible signal of MPI in DIS at HERA is presented in this chapter. The analysis has been inspired by the MPI study at the Tevatron [62, 63], where the multiplicity and energy flow of charged particles have been measured in the regions transverse to the hardest jet of the event. The current analysis follows the same principle by investigating the jet production rate and the amount of transverse momentum carried by the jets in azimuthal areas transverse to the direction of the jets of the hard scattering. Since interactions in addition to the primary interaction involve lower momenta by default, the transverse momentum required to define jets has been lowered from 5 GeV for the hardest jets to 3.5 GeV for so called *mini-jets*. Two different event samples have been studied, where one is based on the selection of inclusive 1-jet events and the other inclusive 2-jet events, where the latter is a subsample of the first.

In the following, the main ideas of the analysis method are given and the effects of different event topolgies are discussed after which the event selection is described. The results are presented and discussed with respect to comparisons with different model predictions. In order to draw the correct conclusions the predictions of the models have been investigated in terms of scale dependence, choice of PDF, mini-jet production and reproduction of basic event observables.

5.1 Measuring Philosophy

The challenge in finding a significant signal from MPI is essentially to separate the final state products of the hard scattering from those originating from additional interactions. As an example Figure 5.1 shows schematic views of a single parton exchange process and an MPI process, both delivering four jets. Per definition additional interactions are always softer than the hard scattering and in general they are much softer i.e. the transverse momenta of the particles produced are small. Although secondary interactions obviously are kinematically dependent on the momenta involved in all preceding interactions, one may to a first approximation assume that the different interactions occur independently of each other. This means that the products of additional interactions can occupy any region

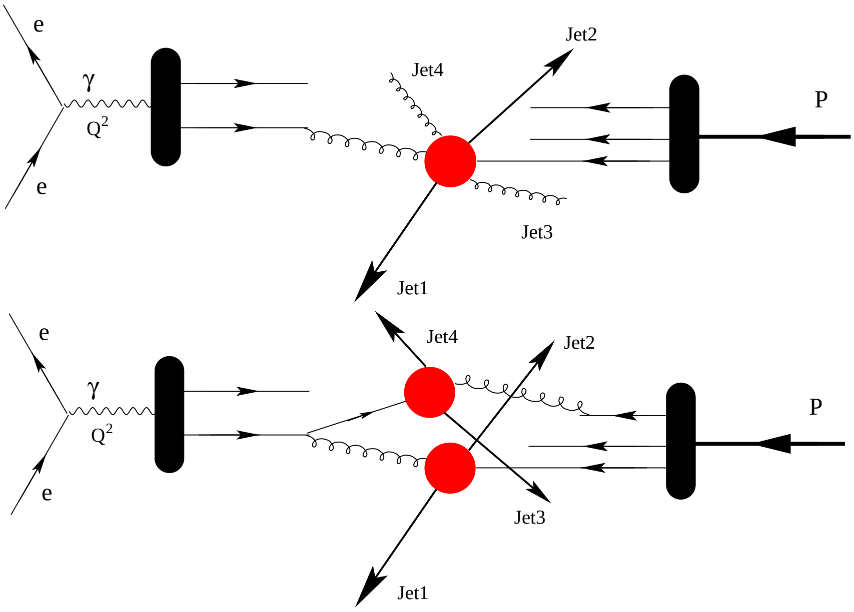


Figure 5.1: *An example of a four jet final state produced in ep-scattering from single parton exchange and multiple parton interaction, respectively.*

of the available angular phase space, independent of where the hard scattering products go.

The basic philosophy of the measurement is to study jet production in the Hadronic Centre-of-Mass (HCM) frame, where the virtual photon collides head-on with the proton.

$$\bar{P}_\gamma + \bar{P}_p = \bar{0} \tag{5.1}$$

In this system the requirement of a hard jet must always be accompanied by jets which balance its transverse momentum. The dominating topology contains two jets back-to-back and all final states with more jets will be suppressed by factors of α_s. In order to explain the basic idea of the measurement, the simplest topology is discussed to start with. In this case regions in azimuthal angle, which are oriented back-to-back, can be defined in such a way that they essentially contain all the transverse momentum carried by two oppositely directed jets. Consequently, angular regions which are transverse to the directions of the hard jets should be the ones where the sensitivity to a signal from additional scatterings is the highest. Consistent with the cone radius defined in the longitudinally invariant $k_\perp$-algorithm four regions in $\Delta\phi^*$ [1] have been chosen as given below:

$$\begin{array}{rcccl} & & |\Delta\phi^*| < 60^o & & \text{called the toward region} \\ & & |\Delta\phi^*| > 120^o & & \text{called the away region} \\ 60^o & < & \Delta\phi^* & < 120^o & \text{and} \\ -120^o & < & \Delta\phi^* & < -60^o & \text{are called the transverse regions} \end{array} \tag{5.2}$$

---

[1] Variables labeled with a * are given in HCM frame.

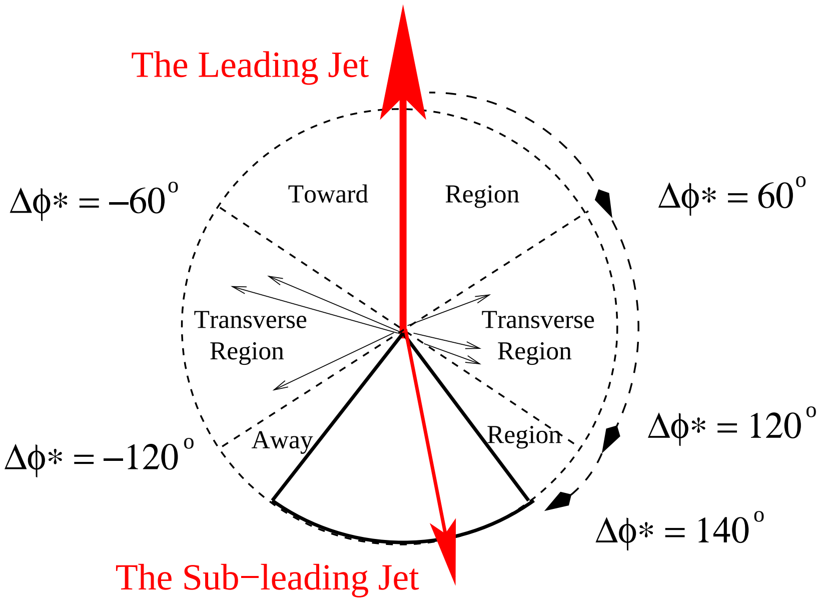


Figure 5.2: *An illustration of an MPI event shown in the azimuthal plane, with the toward, away and transverse regions indicated.* $\Delta\phi^*$ *is defined with respect to the leading jet in the HCM system. The sub-leading jet is restricted to the region* $|\Delta\phi^*| > 140°$, *marked by the bold line.*

Here, $\Delta\phi^*$ is the azimuthal angle relative to the jet with the highest transverse momentum reconstructed in the HCM frame, called the *leading jet*, in the event. The angle ϕ^* is defined relative to the transverse momentum of the scattered electron in HCM. Figure 5.2 illustrates the orientation of the four $\Delta\phi^*$ regions with respect to the two hard jets and possible products from additional interactions. Following the measuring concept outlined above a selection of events with two oppositely directed jets would obviously give the cleanest measuring conditions. On the other hand, an experimental requirement of two jets being exactly back-to-back would lead to an unacceptably low statistical sample. Therefore the angular requirements have to be somewhat relaxed and in this analysis inclusive 2-jet events have been extracted by requiring the second hardest jet, called the *sub-leading jet*, to fall inside the angular region $|\Delta\phi^*| < 140^o$, as indicated by the area surrounded by the thick solid line in Figure 5.2. This specifically means that the jet axis has to fall within this region. Although essentially all the transverse momentum carried by the leading jet is restricted to the *toward region*, the sub-leading jet may spill over some of its transverse momentum into one of the transverse regions depending on its direction. In order to limit the amount of transverse momentum entering into the transverse regions from the sub-leading jet, the angular acceptance region of the jet has been somewhat restricted compared to that of the *away region*. In any case the fact that the jets are not exactly back-to-back means that the phase space is opened up for production of additional jets, which in some cases may populate the transverse regions. This is of course even more true in the case that only hard jet is required in the event selection, i.e. in the inclusive 1-jet sample. The latter sample has however the advantage of containing more events. Thus, as always, there is a trade off between large statistical samples and clean measuring conditions. In this analysis both the inclusive 1-jet sample and the inclusive 2-jet sample have been analysed.

The problem of hard scattering jets, from single parton interactions, delivering momentum into the transverse regions will be discussed in more detail in the next section.

Jets produced in the primary scattering process will carry higher momenta than those coming from additional interactions. Regarding this the leading jet in the inclusive 1-jet sample, and both the leading and sub-leading jets in the inclusive 2-jet sample are required to have momenta higher than 5 GeV. The additional jets have to carry at least 3.5 GeV in order to be accepted as mini-jets. For each event the two transverse regions defined in Equation (5.2) have been separated into a *high activity region* and a *low activity region*, depending on which region contains the highest scalar sum of the transverse momenta of the mini-jets.

As discussed already in section 1.6, DIS may give rise to MPI in both direct and resolved photon processes, whereas the Pythia and Herwig event generators have only implemented MPI and SUE, respectively, for resolved photon processes. Experimentally, the contributions to mini-jet production from direct and resolved photon processes may be separated by subdividing the available phase space into different regions. Thus, the inclusive 1-jet sample is measured in three different bins of Q^2, where in the lowest bin the resolved photon process is enhanced. The inclusive 2-jet sample is split up with respect to x_γ, which gives the fraction of the photon four-momentum carried by the parton involved in the hard scattering. A value of x_γ equal to unity means direct photon interaction, since the photon itself carries all the momentum into the hard scattering. In resolved processes, on the other hand, the photon fluctuates into a quark pair and thus only a fraction of the photon momentum enters the hard scattering via one of the quarks and consequently $x_\gamma < 1$. However, x_γ can not be determined directly but has to be reconstructed from the jets produced in the hard interaction.

Since the parton density dramatically increases with decreasing x_{Bj}, the probability for having MPI increases correspondingly. Small x_{Bj} corresponds to large W^2 for a given Q^2, since $x_{Bj} = Q^2/(Q^2 + W^2)$, and therefore the mini-jet production has been investigated in two regions of W. The requirement of a high momentum jet in the forward region i.e. small polar angles with respect to the proton, is a good selection criterion for low x events in DIS, as has been shown in studies of forward jet production [104, 105]. This motivates sub-divided the samples into two regions of pseudorapidity in which the leading jet is either central or forward.

The results of the experimental measurements on mini-jets are compared to the predictions of various QCD based models. The strategy of the analysis is as follows:

- Measure the mini-jet activities in regions expected to be sensitive to MPI (transverse regions).
- Compare results on mini-jet production with the predictions of Monte Carlo programs based on different QCD models in DIS (e.g. DGLAP, CCFM and CDM) in order to study how well the modelling of higher order processes agrees with the measurements.
- Compare the experimental results with the predictions of Monte Carlo generators, which include the generation of additional activities like MPI (Pythia) and SUE (Herwig).

5.2 Jet Toplogy in the HCM Frame

In this section the jet topologies of three and four jet production via single parton interaction are discussed. Jets from such events constitute the main background to the jets from MPI.

In the laboratory frame the transverse momentum of the hadronic final state produced in an *ep*-collision is balanced by the transverse momentum of the scattered electron. For high collision energies, it can too a good approximation be assumed that the transverse momentum of the jets is responsible for most of the hadronic transverse momentum. Boosted to the HCM the total transverse momentum of the jets becomes zero. Jets from the hard scattering might appear in the regions transverse to the leading jet if at least three jets are produced. Examples of such a jet topolgies are shown in Figure 5.3. The definition of the leading jet is that it is the jet of highest momentum (> 5 GeV), and it is used to define $\Delta\phi^* = 0$. The jets in the figure have been ordered in transverse momenta according to $P_{T,Jet1} > P_{T,Jet2} > P_{T,Jet3} > P_{T,Jet4}$. The most common scenario for a three jet event is when the sub-leading $Jet2$ is accompanied by a softer jet, $Jet3$, as illustrated in Figure 5.3a. The most probable topology is that the third jet also goes into the away region. Although the angular restrictions of $Jet2$ together with the transverse momentum requirement of $Jet3$ somewhat limit the possible topologies in the ϕ^* plane, it is still possible that the third jet enters into the transverse region with enough momentum to fulfill the requirements of a mini-jet. The kinematic restrictions of a Mercedes star event, as shown in Figure 5.3b, is such that it might just be able to produce three jets, where none is going into the away region. However, the most probable case when no jet is observed in the away region is that the jets do not fall inside the acceptance in rapidity. The four jet scenario might give two jets which populate the transverse regions, as shown in Figure 5.3c. In a four jet event the jets may be produced either by a single parton interaction or by a two parton exchange. In the first scenario the transverse momenta of the two pairs of jets $Jet1, Jet2$ and $Jet3, Jet4$ do not necessarily balance. However, this must be true for the second case. By defining the high and low activity transverse regions, on an event by event basis, most of the contributions from hard scattering are expected to fall in the high acitivity region. Contributions from MPI are to first approximation assumed to be uncorrelated with the primary interaction and should not prefer one or other of the transverse regions. Thus, higher order effects might preferably be studied in the high activity region, whereas the effects of MPI should be most visible in the low activity region.

The interpretation of the data suffers from two difficulties. Experimentally, the contributions from higher order processes can not be identified on an event by event basis, and such jets entering into the transverse regions can thus not be rejected. Theoretical calculations of 2- and 3-jet cross sections can not be used to estimate the effect statistically, since the leading and sub-leading jets in most of the cases are essentially back-to-back, which leads to divergencies in the calculations and thus unphysical results are obtained (see section 1.4.5). For inclusive jet production, where only one reconstructed jet is required, a NNNLO calculation (next-to-next-to-next-to-leading order, corresponding to 1+3 jets) would be needed in order to estimate jet emissions into all four $\Delta\phi^*$ regions. Such calculations do not yet exist. Instead phenomenological Monte Carlo models have to be applied, where higher order contributions are accounted for by initial and final state radiation. A necessary requirement for reliable conclusions concerning MPI to be drawn

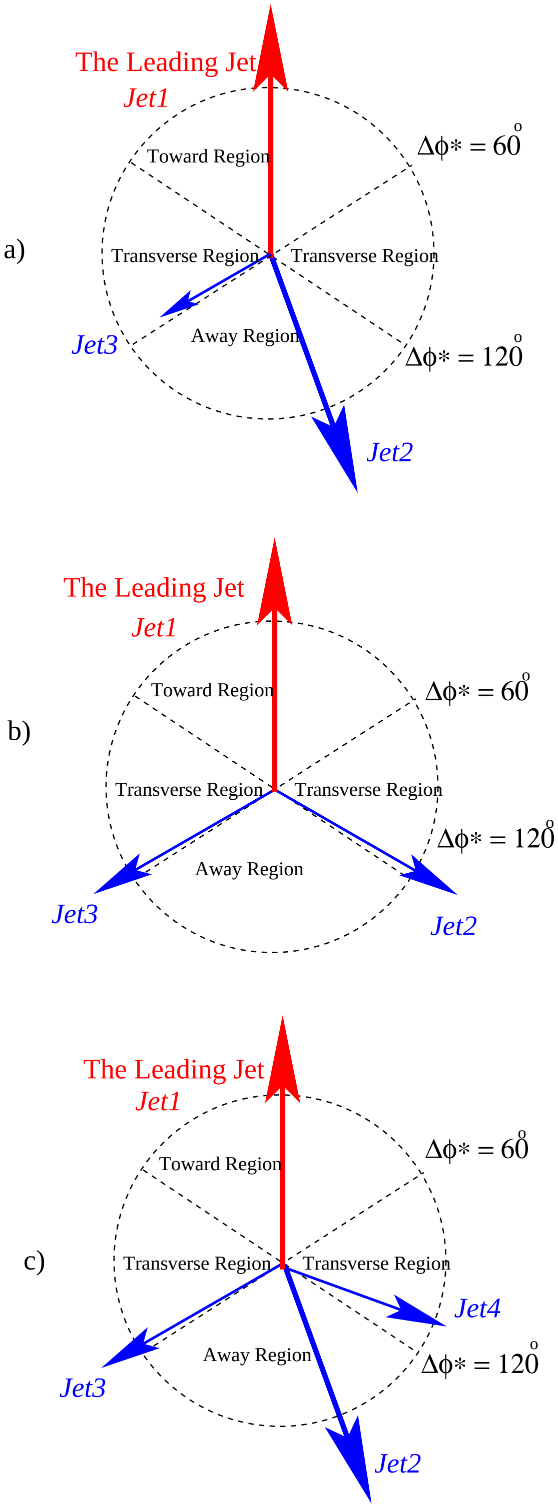


Figure 5.3: *Examples of possible jet topologies for three and four jet events from a hard scattering process, leading to jet activities in the transverse regions.* $\Delta\phi^*$ *is defined with respect to the leading jet in the HCM system. The jets are ordered by decreasing transverse momenta* $P_{T,Jet1} > P_{T,Jet2} > P_{T,Jet3} > P_{T,Jet4}$.

from comparisons of the model predictions with data, is that the models should be able to reproduce the properties of the hard scattering, i.e. the jet activities in the toward and away regions should be reasonably well described.

Secondly, deviations from the predictions of these models in the transverse regions would indicate activities in addition to the modelled hard interactions but would not necessarily constitute evidence for MPI. Only one model exists which describes MPI in DIS. Unfortunately this model in not developed to be valid for high Q^2 and comparisons to it can therefore not be expected to be very conclusive.

5.3 Event Selections

Data and detector simulated Monte Carlo events passing the selection criteria for DIS events, presented in Table 3.1, have been used to extract the jet samples in this analysis.

As already mentioned, the analysis is carried out in the HCM frame. All hadronic final state particles measured in the event have been energy corrected using the Low and High Pt Jet Calibration method and the Iterative method (denoted M5 in Chapter 3.6) in the laboratory frame and are boosted to the HCM frame. The photon four-momentum, P_γ, used to define the boost transformation to the HCM frame is reconstructed from the four-momentum P_e^0 and P_e of the incoming and the scattered electron, respectively:

$$P_\gamma = (E_\gamma, \bar{P}_\gamma) = P_e - P_e^0.$$

The longitudinally invariant k_T algorithm, using the k_T recombination scheme, is used to find and reconstruct jets in the HCM frame with $P^*_{T,jets} > 3.5$ GeV, from all measured particles, except the scattered electron.

To ensure well defined mini-jets, the requirement that the jet transverse momentum has to be larger than 3.5 GeV, must be fulfilled in the laboratory frame as well. The jet axis in the laboratory frame has to fall inside the pseudo-rapidity range $-1.5 < \eta_{jets} < 2.79$, which is well covered by the LAr calorimeter. In order to apply jet cuts in the laboratory system, the same particles as were assigned to each jet in the HCM frame, were used to calculate the jet kinematics in the laboratory system.

The inclusive 1-jet sample is defined by demanding that the transverse momentum of the leading jet (lj) in the laboratory frame, $P_{T,lj}$, and the HCM frame, $P^*_{T,lj}$, are both larger than 5 GeV.

In the inclusive 2-jet sample the transverse momenta of both the leading and sub-leading jets (sl) are required to be larger than 5 GeV both in the laboratory frame and the HCM system. As explained already, the jet axis of the sub-leading jet is restricted to the angular region $|\phi^*_{lj} - \phi^*_{sl}| > 140°$, which is well inside the away region.

Since x_γ can not be measured directly it has to be reconstructed from the jets produced in the hard interaction. Two methods for the x_γ reconstruction are considered:

$$x_\gamma = \frac{\sum_{i=1}^{2} P^*_{T,i}\, e^{-\eta^*{}_i}}{2E^*_\gamma} \tag{5.3}$$

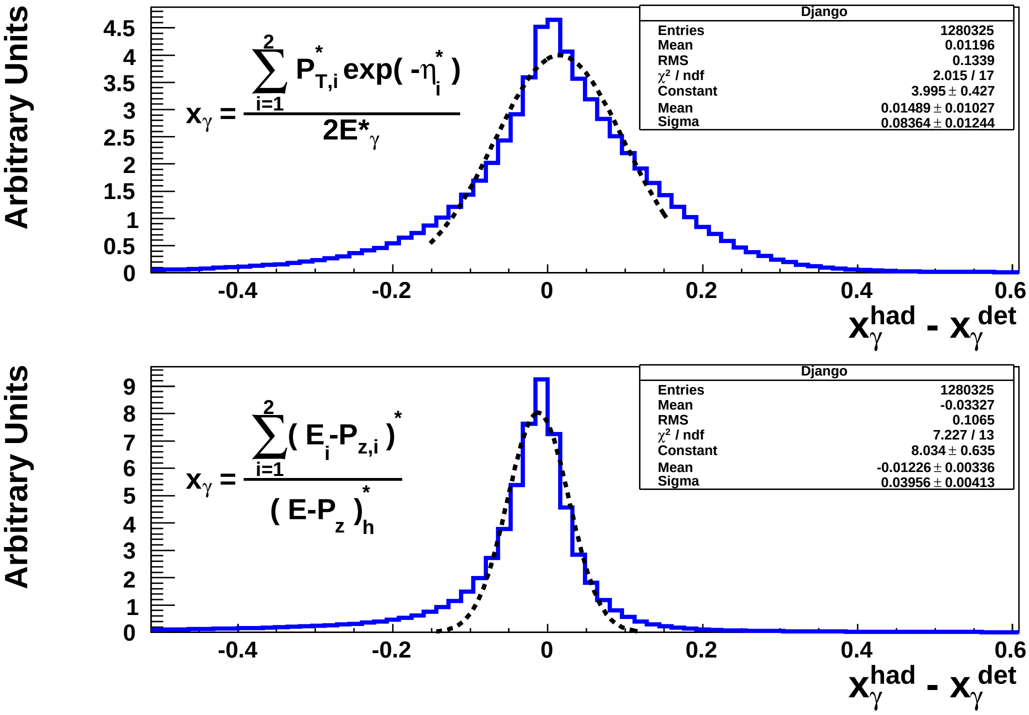


Figure 5.4: *The resolution in x_γ calculated from Django(CDM) Monte Carlo events, using two different methods. A fit of a gaussian distribution around the peak value of the distribution is used to extract the mean value and the standard deviation (σ), as given in the figure.*

and

$$x_\gamma \;=\; \frac{\sum_{i=1}^{2} E_i^* - P_{z,i}^*}{(E^* - P_z^*)_h}. \tag{5.4}$$

All quantities are given in the HCM frame, where E_i^*, $P_{z,i}^*$, $P_{T,i}^*$ and $\eta^*{}_i$ correspond to the energy, the longitudinal momentum, the transverse momentum and the pseudorapidity, respectively, of the leading jet (i=1) and the sub-leading jet (i=2). E_γ^* is the photon energy and $(E^* - P_z^*)_h$ is defined by Equation (3.5). The inclusive 2-jet sample has been used to study the resolution of x_γ for each of the two reconstruction methods. Figure 5.4 shows the distribution of the difference between x_γ on detector, x_γ^{det}, and hadron level, x_γ^{had}. For a quantitative comparison, the central part of the distribution has been fitted to a Gaussian distribution. The results of the fits are given in the figure. As can be seen the method 5.4 provides a better resolution than Equation (5.3). This is mainly due to the resolution of the cluster energy measurement. In method 5.3, the numerator depends on the cluster measurement while the denominator does not and thus x_γ depends strongly on the resolution of the LAr measurement. In method 5.4, both the denominator and numerator depend on the LAr measurement resulting in a x_γ definition less sensitive to the resolution of the cluster energy measurement. In this thesis, method 5.4 is used.

The multiplicity and the transverse momentum of the mini-jets in each of the four $\Delta\phi^*$ regions, are calculated for each event. The final results are presented in terms of the average multiplicity, $\langle N \rangle$, and the average transverse momentum, $\langle PtSum^* \rangle$, defined as:

$$\langle N \rangle = \frac{\sum_{i=1}^{N_{ev}} N_i}{N_{ev}} \tag{5.5}$$

$$\langle PtSum^* \rangle = \frac{\sum_{i=1}^{N_{ev}} PtSum_i^*}{N_{ev}}. \tag{5.6}$$

Here, N_i is the multiplicity of mini-jets in event i and $PtSum_i^*$ is the scalar sum of their transverse momenta, both observables calculated separately for the four $\Delta\phi^*$ regions. N_{ev} is the number of events included in the calculation. The results are extracted for the following bins of the available phase space.

- The inclusive 1-jet sample is subdivided in 2 samples in η_{lj}: the central inclusive 1-jet sample $-1.5 < \eta_{lj} < 1.5$ and the forward inclusive 1-jet sample $1 < \eta_{lj} < 2.79$. Each η_{lj} sample is subdivided further into 3 bins in Q^2 and 2 bins in W:

  $$3Q^2 \times 2W \text{ bins} = [5-10, 10-25, 25-100] \text{ GeV}^2 \times [100-200, 200-300] \text{ GeV}.$$

- The inclusive 2-jet sample is subdivided into four bins:

  $$2x_\gamma \times 2W \text{ bins} = [0-0.7, 0.7-1] \times [100-200, 200-300] \text{ GeV}.$$

In table 5.1 the jet cuts applied to the forward and central inclusive 1-jet samples and to the inclusive 2-jet sample are summarized. The central η_{lj} region is defined by the coverage of the central tracker, whereas the forward region has been defined in such a way that the statistics obtained in the two regions is approximately the same. This means that the forward region has been allowed to overlap with the central region somewhat. Thus, the central and forward regions are not strictly separated. It should be noted that this subdivision only concerns the leading jet, whereas the mini-jets are measured in the full range of $-1.5 < \eta < 2.79$ for all three event samples.

5.4 Measurement of the Mini-jet Production

The measured data has to be corrected for detector effects in order to perform a comparison to model predictions, which are given on hadron level. Corrections for the influence of the detector are done by simulating its perfomance in a Monte Carlo program. The correctness of the detector simulation has to be verified through comparisons to data in various control plots. The Monte Carlo generators, Rapgap(dir) and Django(CDM), have been used together with the detector simulation program H1SIM to generate data on detector level. The detector geometry, measuring resolutions and efficiencies of subdetectors are the main information used as input to the detector simulation program. Two different methods to correct the measured data for detector effects have been applied, the *bin-by-bin method* and the *unfolding method*. These methods are described and the results are compared and discussed. The most important systematic uncertainties in the measurement have been identified and their magnitudes have been calculated.

Central Inclusive 1-Jet Sample

Leading Jet	Sub-leading Jet	Mini-jets
$-1.5 < \eta_{lj} < 1.5$ $P_{T,lj} > 5$ GeV $P^*_{T,lj} > 5$ GeV		$-1.5 < \eta_{jets} < 2.79$ $P_{T,jets} > 3.5$ GeV $P^*_{T,jets} > 3.5$ GeV

Forward Inclusive 1-Jet Sample

Leading Jet	Sub-leading Jet	Mini-jets
$1 < \eta_{lj} < 2.79$ $P_{T,lj} > 5$ GeV $P^*_{T,lj} > 5$ GeV		$-1.5 < \eta_{jets} < 2.79$ $P_{T,jets} > 3.5$ GeV $P^*_{T,jets} > 3.5$ GeV

Inclusive 2-Jet Sample

Leading Jet	Sub-leading Jet	Mini-jets
$-1.5 < \eta_{lj} < 2.79$ $P_{T,lj} > 5$ GeV $P^*_{T,lj} > 5$ GeV	$-1.5 < \eta_{sj} < 2.79$ $P_{T,sj} > 5$ GeV $P^*_{T,sj} > 5$ GeV $\lvert\phi^*_{lj} - \phi^*_{sl}\rvert > 140°$	$-1.5 < \eta_{jets} < 2.79$ $P_{T,jets} > 3.5$ GeV $P^*_{T,jets} > 3.5$ GeV

Table 5.1: *A summary of the jet cuts for the various samples: the central inclusive 1-jet sample, the forward inclusive 1-jet sample and the inclusive 2-jet sample.*

5.4.1 Trigger Efficiency

It is important to have a high trigger effeciency to avoid that events of interest are rejected. To study the efficiency of the trigger used in the analysis its acceptance has been compared to a set of so called monitor triggers. A trigger is composed of a set of trigger elements, which are defined in Appendix A. It is important that the trigger used in the analysis, here called the physics trigger, and the monitor trigger have no trigger elements in common. The physics trigger, $PyTr$, used in this analysis, is composed of the following trigger elements (S0∨S3∨S61), which relies upon information from the CJC, the SPACAL. The combination of trigger elements (S39∨S64∨S66∨S67∨S77) is used for the monitor trigger, $MoTr$, which only depends on the LAr calorimeter information. The trigger efficiency, ϵ, is defined by:

$$\epsilon = \frac{N(PyTr \ \wedge \ MoTr)}{N(MoTr)},$$

where $N(PyTr \wedge MoTr)$ is the number of events accepted by both the physical trigger and the monitor trigger, whereas $N(MoTr)$ is the number of events triggered by the monitor trigger. The trigger efficiency for the inclusive 1-jet sample is shown as a function of η_{lj}, ϕ_{lj} and $P_{T,lj}$ in Figure 5.5. For the inclusive 2-jet sample the efficiency is shown in Figure 5.6

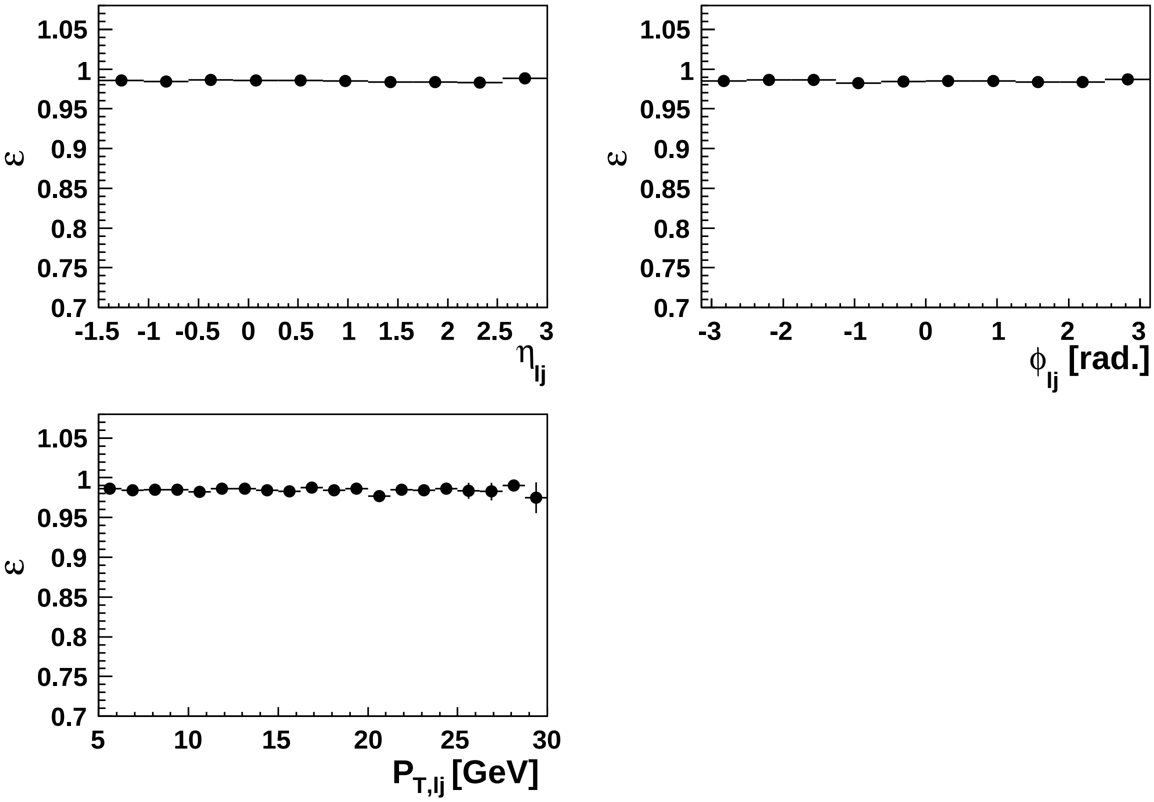


Figure 5.5: *The trigger efficiency, ϵ (see text for definition), as a function of η_{lj}, ϕ_{lj} and $P_{T,lj}$ for the inclusive 1-jet sample.*

as a function of the same variables, for both the leading and sub-leading jets. As can be seen the efficiency is close to 100% for both samples.

5.4.2 Control Plots

In this section data are compared to the predictions of Rapgap(dir) and Django(CDM), where the Monte Carlo data include corrections from QED radiation, and have been subject to a full detector simulation. Control plots have been produced to check how tracks and clusters are reproduced in terms of various kinematic variables. The measurement of the scattered electron and the reconstruction of kinematic variables, using the electron method, have been checked. Further, the description of the leading and sub-leading jet kinematics in different parts of the available phase space have been investigated. Finally, the kinematics of mini-jets has been studied.

Tracks and Clusters

In Figure 5.7 the distributions of pseudo-rapidity, transverse momentum and azimuthal angle of HFS particles (cluster and tracks) are shown for data, Django(CDM) and Rapgap(dir) for the inclusive 1-jet sample. The plots on the left hand side show the distributions in

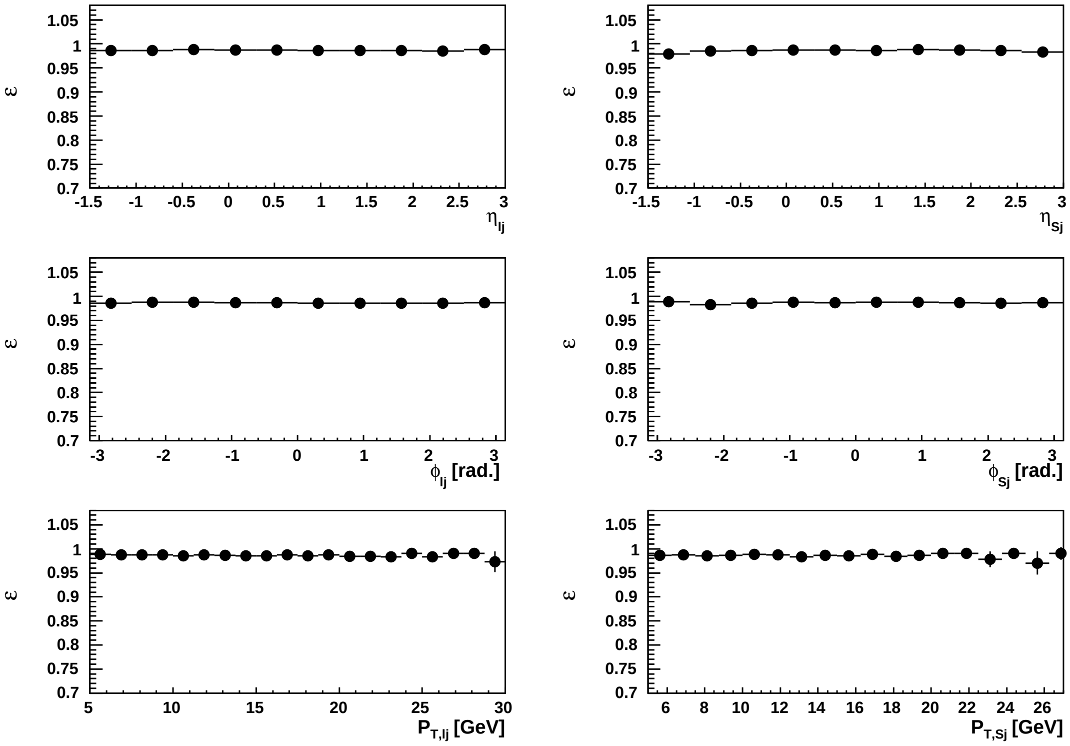


Figure 5.6: *The trigger efficiency, ϵ (see text for definition), as a function of psuedo-rapidity, azimuthal angle and transverse momentum of the leading jet (plots to the left) and the sub-leading jet (plots to the right) for the inclusive 2-jet smaple.*

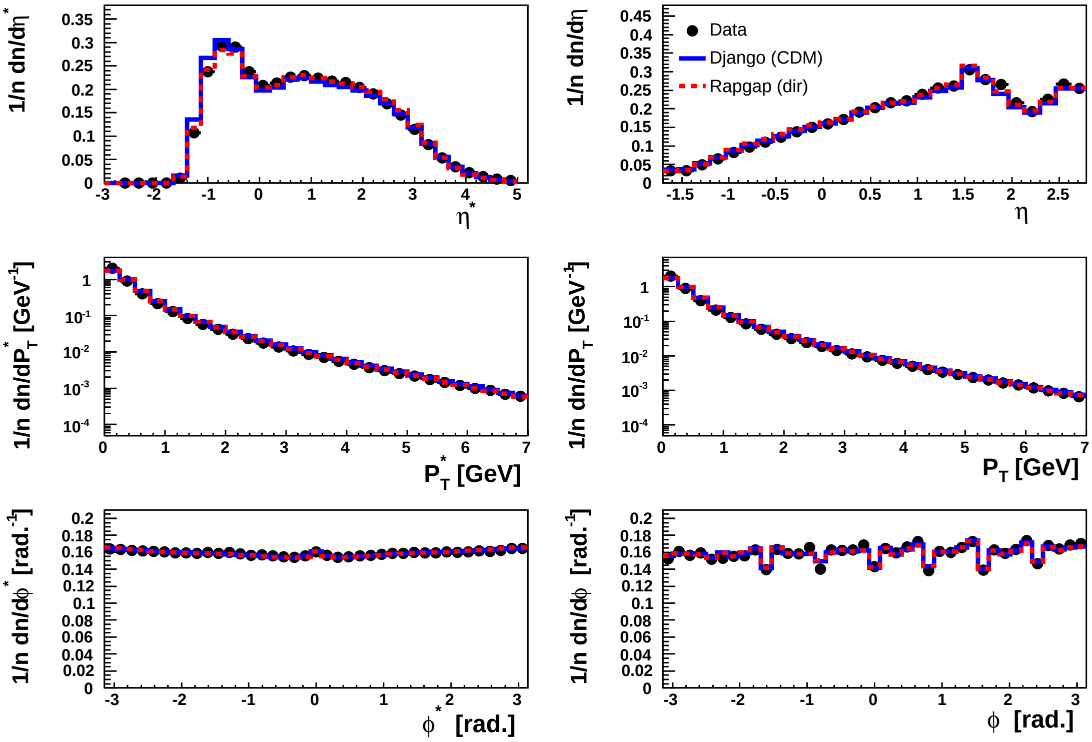


Figure 5.7: *The distributions of pseudo-rapidity, transverse momentum and azimuthal angle for HFS particles in the inclusive 1-jet sample, presented in both the HCM (plots to the left) and the laboratory frame (plots to the right). Data are compared to the predictions of Django(CDM) and Rapgap(dir). The distributions are normalised to the total number of entries.*

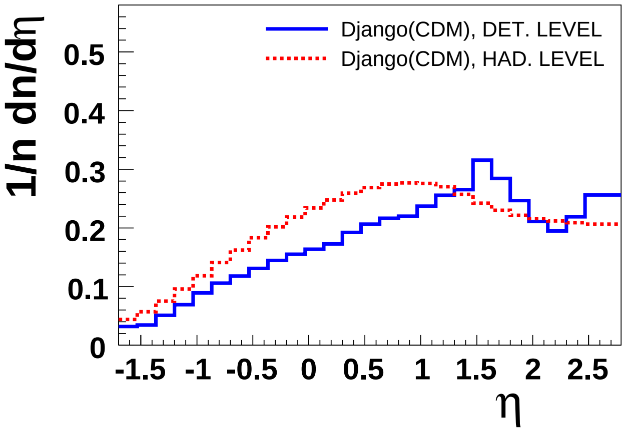


Figure 5.8: *The pseudo-rapidity distribution for final state particles of the inclusive 1-jet sample, on detector level (solid line) and hadron level (dotted line). The distributions have been calculated using the Django(CDM) event generator*

the HCM frame, whereas those on the right hand side are in the laboratory frame. The distributions are normalised to the total numer of entries. As observed, the MC predictions are in good agreement with data, although Django(CDM) overshoots the data somewhat at negative values of η^*, which is in the proton fragmentation region. The dips seen in the ϕ spectrum is caused by the edges of the calorimeter octants, that build up the LAr calorimeter, but this behaviour is well reproduced by both MC's. In order to find the explanation for the two peaks at $\eta \approx 1.7$ and $\eta \approx 2.6$, the pseudo-rapidity distributions on detector and hadron level have been compared using events generated by Django(CDM). From Figure 5.8 it can be seen that the peaks are not present on hadron level. A detailed study of this phenomenon in [106], shows that the peak at $\eta \approx 1.7$ originates from nuclear scattering of protons against a beam collimator, used for shielding against synchrotron radiation, while the peak at $\eta \approx 2.6$ is mainly produced by electronics noise in the LAr calorimeter. Also these effects are successfully described by the detector simulation. In any case the influence of these effects on the production of high momentum jets is small.

In Figure 5.9 the distributions of pseudo-rapidity, transverse momentum and azimutal angle are shown for individual particles in the inclusive 1-jet sample, both in the HCM frame and the laboratory system. Both MC programs give an excellent description of the data. The distributions of the distance of closest approach from the extrapolated track to the measured vertex, dca', and the track length, R_{length}, projected onto the (x, y) plane at the z_{vertex}, are shown in Figure 5.10 together with the number of hits per track, N_{hits}. The overall description by the MC's is good, although Rapgap(dir) is somewhat below data in the dca' distribution and the number of hits is slightly less than what data shows.

The same level of agreement is obtained in the corresponding distributions for particles in the inclusive 2-jet sample.

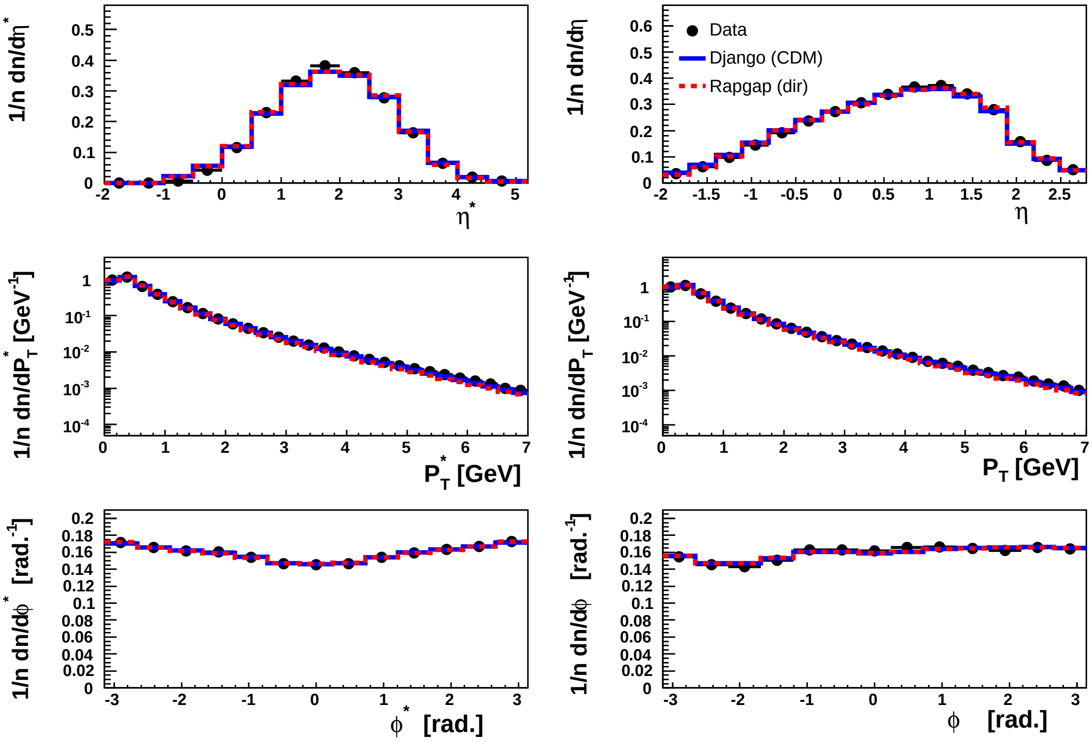


Figure 5.9: *The distributions of pseudo-rapidity, transverse momentum and azimuthal angle for tracks in the inclusive 1-jet sample, presented for the HCM frame (plots to the left) and the laboratory frame (plots to the right). Data are compared to the predictions of Django(CDM) and Rapgap(dir). The distributions are normalised to the total number of entries.*

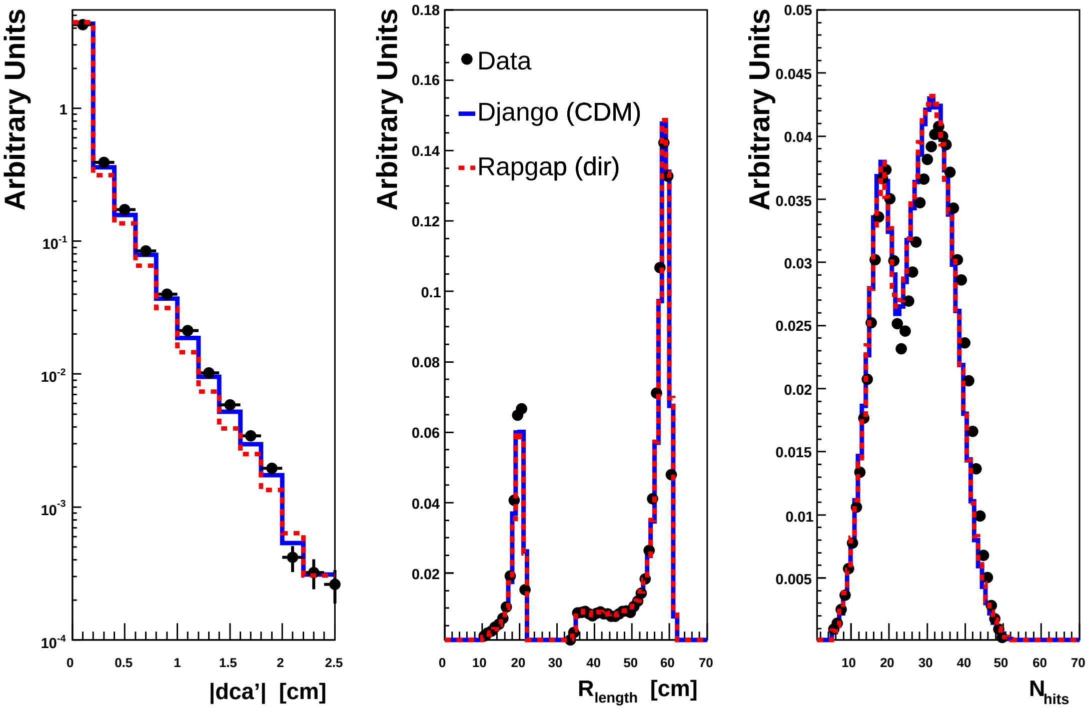


Figure 5.10: *The distributions of the closest distance between the extrapolated track and the reconstructed vertex, dca′, the track length,* R_{length}*, and the number of hits on a track,* N_{hits}*, for the inclusive 1-jet sample. Data are compare to the predictions of Django(CDM) and Rapgap(dir).*

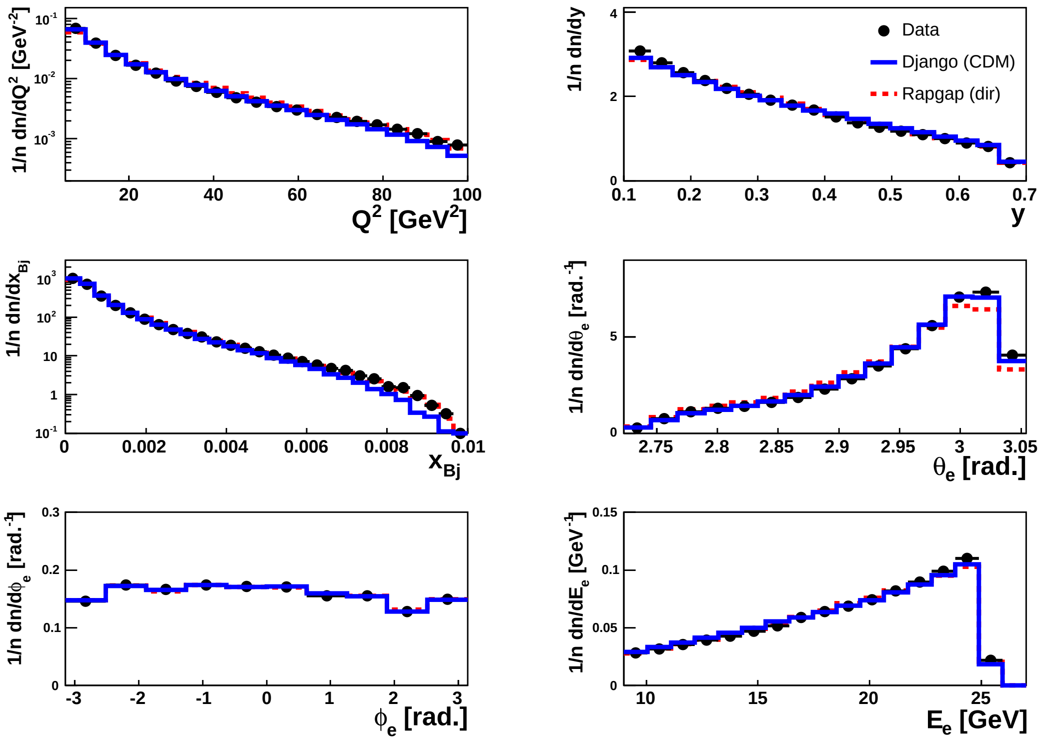


Figure 5.11: *The distributions of Q^2, y, x_{Bj}, θ_e , ϕ_e and E_e for the inclusive 1-jet sample. Data are compared to the prediction of Rapgap(dir) and Django(CDM). The distributions are normalised to the total number of events.*

The Scattered Electron and DIS Quantities

In Figures 5.11 and 5.12 the distributions of the DIS variables Q^2, y, x_{Bj} and the scattered electron variables θ_e, ϕ_e and E_e are compared to the predictions of Rapgap(dir) and Django(CDM) for the inclusive 1-jet and 2-jet samples, respectively. The distributions are normalised to the total number of events. Only small deviations are observed in the MC description of data. For example a somewhat better description of the high Q^2 and x_{Bj} regions is given by Rapgap(dir) compared to Django(CDM) for the inclusive 1-jet sample. On the other hand the scattering of electrons by small angles (large θ_e) is slightly better reproduced by Django(CDM) for both the inclusive 1- and 2-jet samples. The W-distributions, normalised to the total number of events, are shown for the DIS sample, where no jets are required, and for the inclusive 1- and 2-jet samples in Figure 5.13. It is observed that the distribution peaks towards higher W as the number of jets required in the sample increases. This is expected since more energy is needed in order to produce more jets. Both MC's provide good agreement with data for all three samples. A good description of these distributions is essential, since the data are subdivided into bins of W.

Leading and Sub-leading Jet Kinematics

The description of the jet kinematics is the most important demand on the Monte Carlo generators used to correct data for detector effects, but it is also the most difficult task since

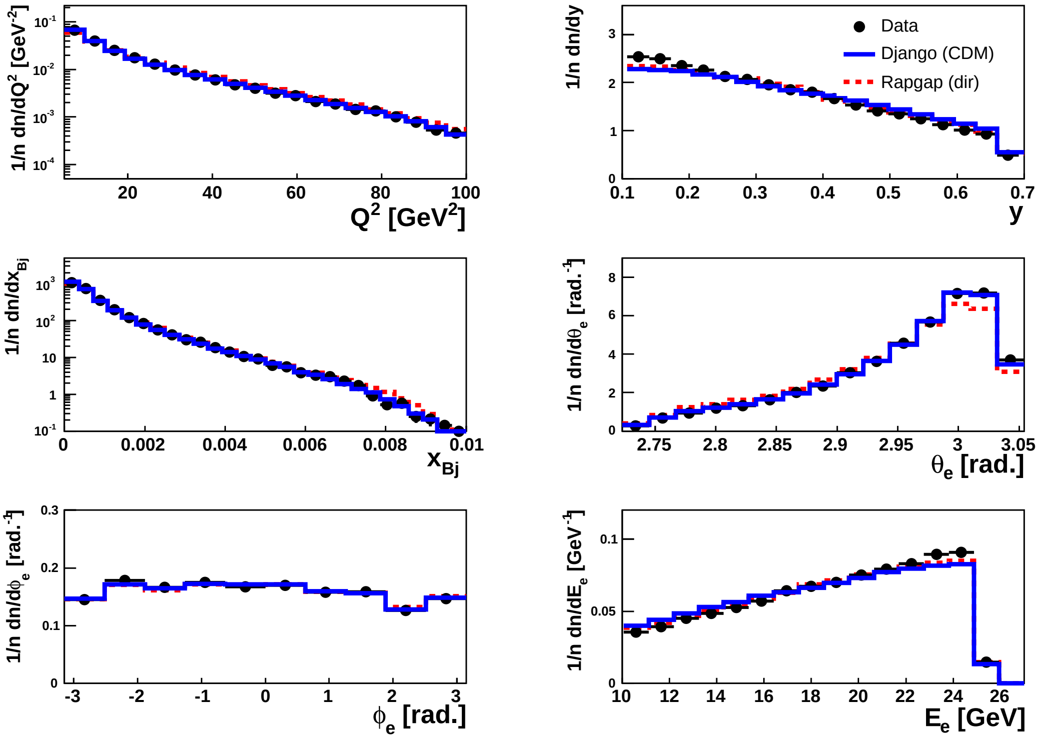


Figure 5.12: *The distributions of* Q^2, y, x_{Bj}, θ_e , ϕ_e *and* E_e *for the inclusive 2-jet sample. Data are compared to the prediction of Rapgap(dir) and Django(CDM). The distributions are normalised to the total number of events.*

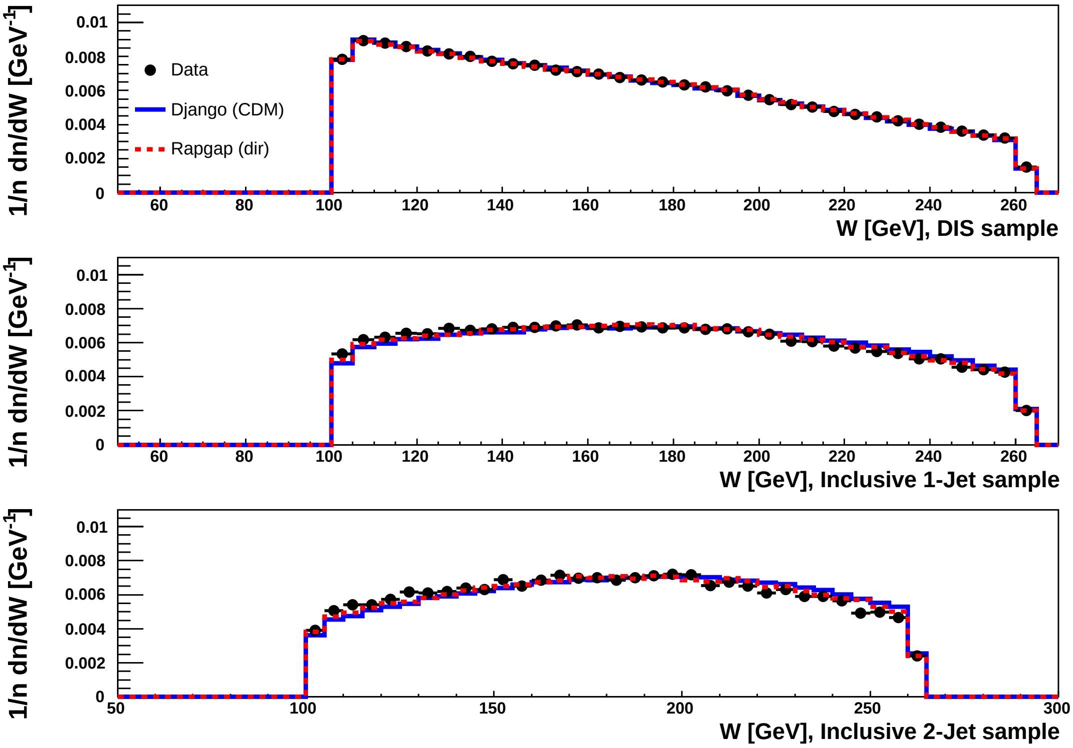


Figure 5.13: *The W-distributions for DIS processes (top), the inclusive 1-jet sample (middle) and the inclusive 2-jet sample (bottom). Data are compared to the prediction of Rapgap(dir) and Django(CDM). The distributions are normalised to the total number of events.*

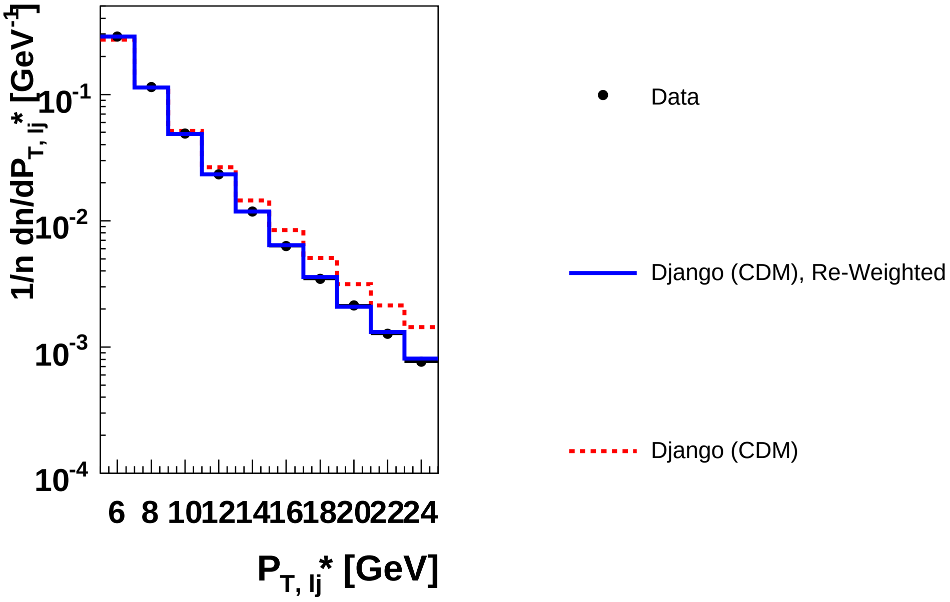


Figure 5.14: *The $P^*_{T,lj}$-distribution for data and MC events from Django(CDM) before (dashed lines) and after (solid line) re-weighting.*

it relies upon a good understanding of QCD. Especially it is not easy to correctly describe distributions of jet variables for events where several jets are required, since this needs a good modelling of higher order QCD processes. In these cases the Monte Carlo events might have to be re-weighted such that the generated distributions agree with data. In Figure 5.14 the distribution of $P^*_{T,lj}$ for data is compared to the prediction of Django(CDM) for the inclusive 1-jet sample. As can be seen, the $P^*_{T,lj}$ distribution is not well reproduced by the CDM model, whereas after re-weighting the agreement is excellent. It should be mentioned that the Monte Carlo events are re-weighted on both detector level and hadron level, but only in cases where the required jets are found on both levels. Different methods of re-weighting have been tested in order to check their influence on the systematic errors and it was found to be small. In this analysis the $P^*_{T,lj}$ and η^*_{lj} in the inclusive 1-jet sample have been adjusted to agree with data. For the inclusive 2-jet sample, the $P^*_{T,lj}$, η^*_{lj}, x_γ and $|\phi^*_{lj} - \phi^*_{sl}|$ have been corrected.

The distributions of $P^*_{T,lj}$, ϕ^*_{lj} and η^*_{lj} for the forward inclusive 1-jet sample are shown in Figure 5.15 for $100 < W < 200$ GeV, and for $200 < W < 300$ GeV in Figure 5.16. In Figures 5.17 and 5.18 the same distributions are shown for central inclusive 1-jet sample in the low and high W regions, respectively. The distributions are normalised to the total number of events. The data are compared to the predictions of Django(CDM) and Rapgap(dir) after the re-weighting procedure has been performed. All distributions are well described by both Monte Carlo event generators. Kinematic cuts, introduced in the laboratory system to ensure that jets are reconstructed well inside the LAr acceptance, are responsible for the non-flat ϕ-distribution in the HCM.

The same distributions are shown for the inclusive 2-jet sample, separately for the

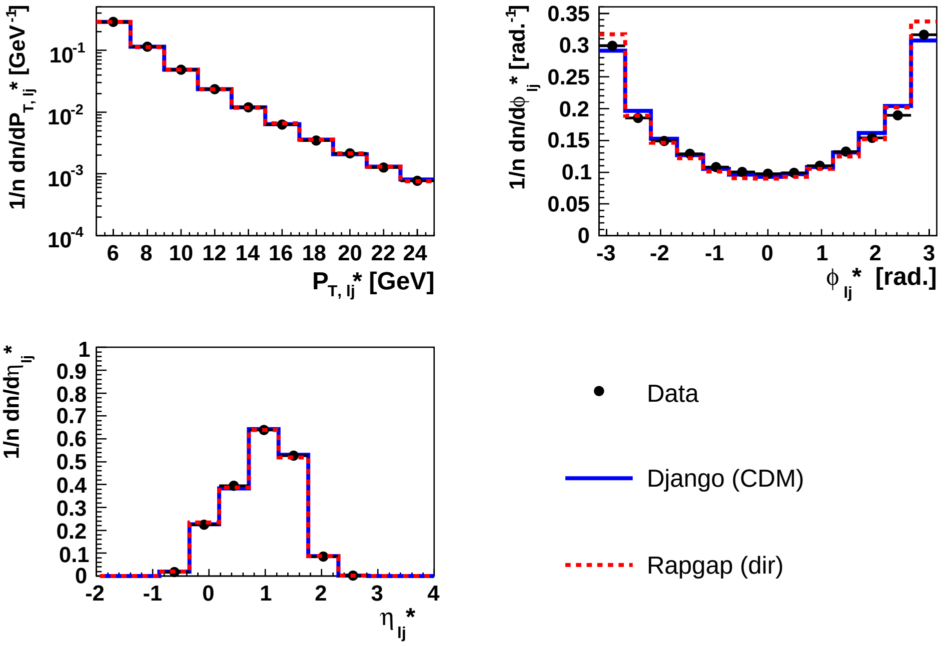


Figure 5.15: *The distributions of* $P^*_{T,lj}$, ϕ^*_{lj} *and* η^*_{lj} *of data compared to the predictions of Django(CDM) and Rapgap(dir) for the forward inclusive 1-jet sample at low* W*. The distributions are normalised to the total number of events.*

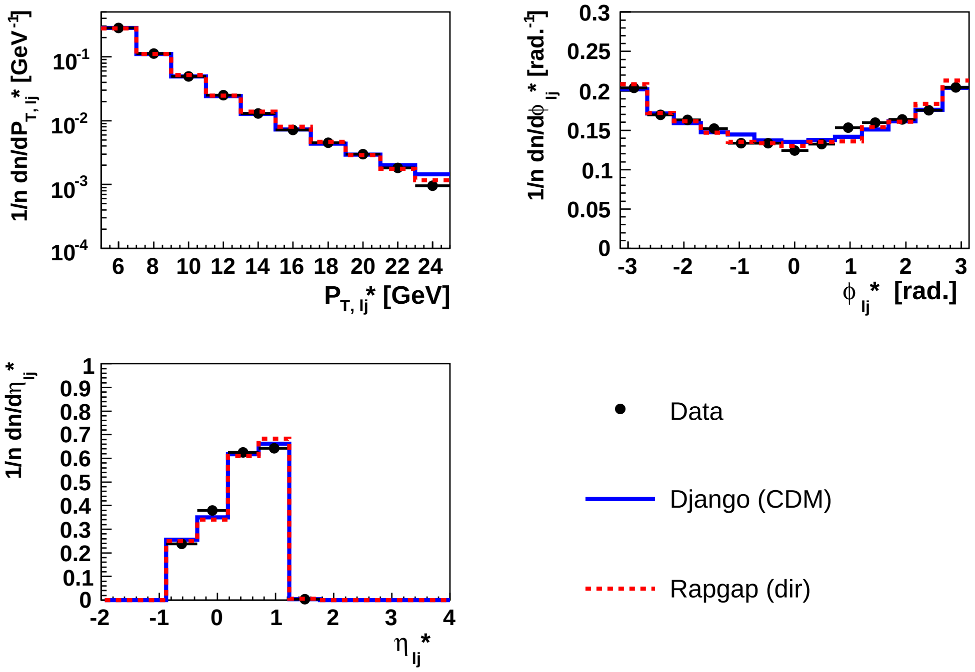


Figure 5.16: *The distributions of* $P^*_{T,lj}$, ϕ^*_{lj} *and* η^*_{lj} *of data compared to the predictions of Django(CDM) and Rapgap(dir) for the forward inclusive 1-jet sample at high* W*. The distributions are normalised to the total number of events.*

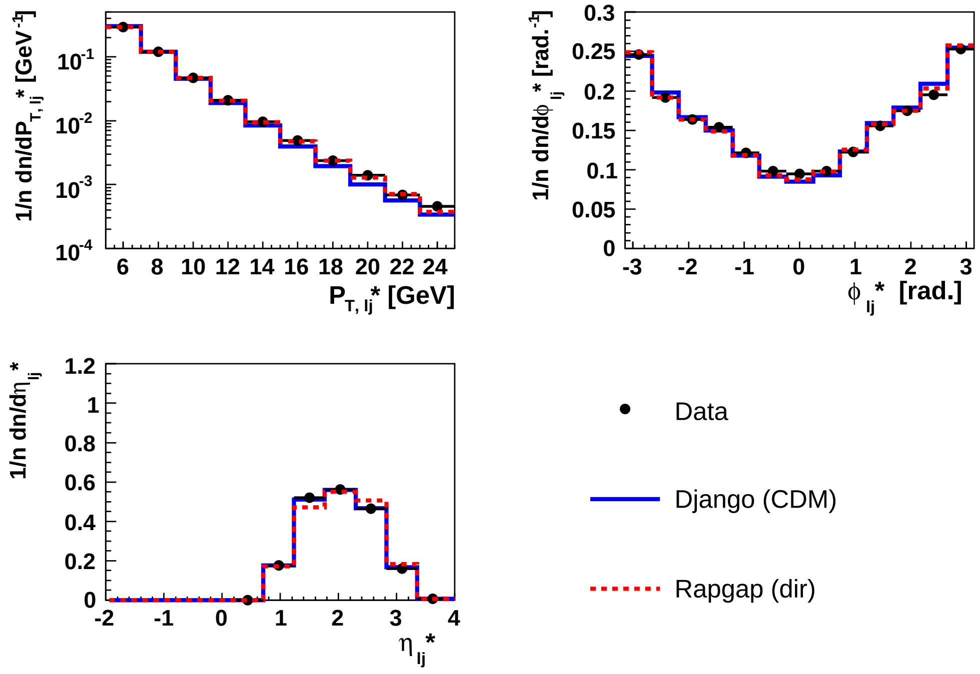


Figure 5.17: *The distributions of* $P^*_{T,lj}$*,* ϕ^*_{lj} *and* η^*_{lj} *of data compared to the predictions of Django(CDM) and Rapgap(dir) for the central inclusive 1-jet sample at low* W*. The distributions are normalised to the total number of events.*

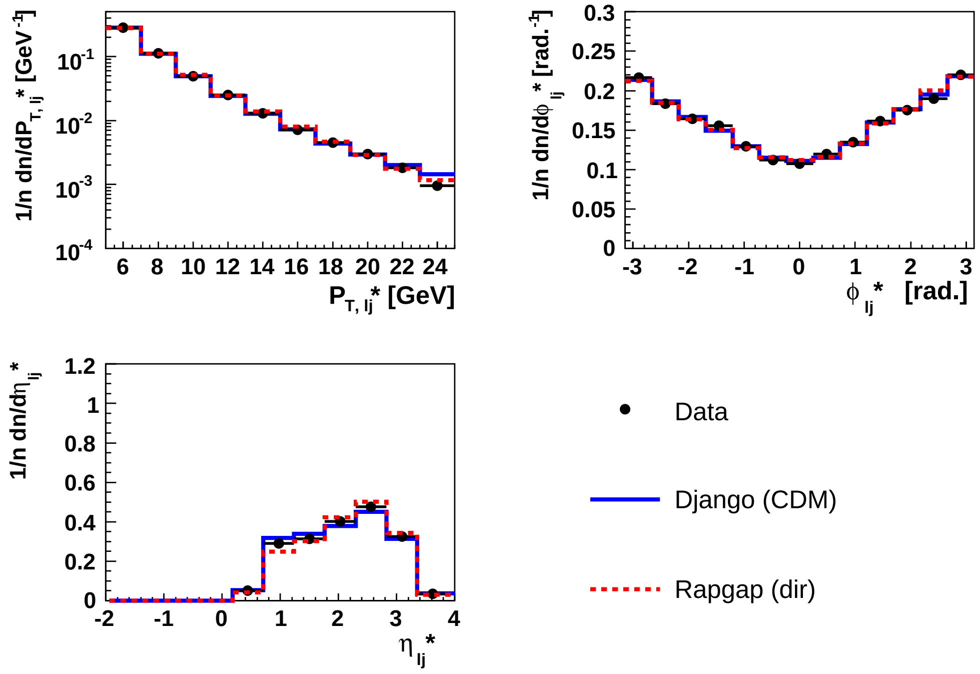


Figure 5.18: *The distributions of* $P^*_{T,lj}$*,* ϕ^*_{lj} *and* η^*_{lj} *of data compared to the predictions of Django(CDM) and Rapgap(dir) for the central inclusive 1-jet sample at high* W*. The distributions are normalised to the total number of events.*

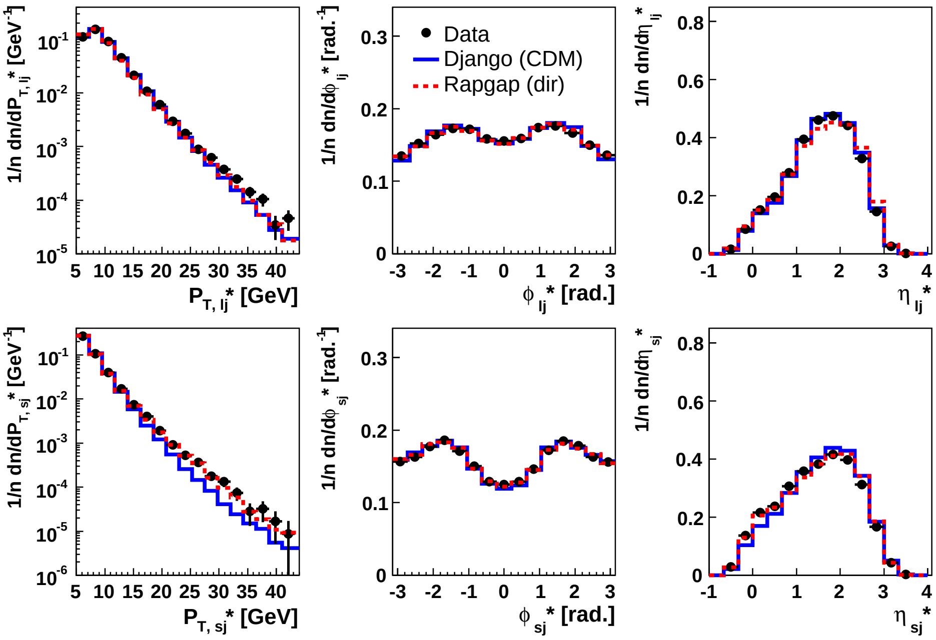


Figure 5.19: *The distributions of the transverse momentum, the pseudo-rapidity and the azimuthal angle of the leading (upper plots) and sub-leading jet (lower plots) of data compared to the predictions of Django(CDM) and Rapgap(dir) for the inclusive 2-jet sample at low W. The distributions are normalised to the total number of events.*

leading and sub-leading jets in Figure 5.19 for the low W region and in Figure 5.20 for the high W region. Although the agreement between MC predictions and data in general is good, some deviations can be observed for Django(CDM), especially in the $P^*_{T,sl}$ spectra.

The distributions of x_γ and $|\phi^*_{lj} - \phi^*_{sl}|$, normalised to the total number of events, are shown for the inclusive 2-jet sample in Figures 5.21 and 5.22, respectively. The agreement between the MC predictions and data is good except for large values of x_γ. As observed the x_γ-distribution has a significant tail towards small values, indicating a substantial contribution of resolved photon events in the inclusive 2-jet sample. The $|\phi^*_{lj} - \phi^*_{sl}|$-distribution is peaked at 180° as expected, since the inclusive 2-jet sample is dominated by events with exactly two reconstructed jets, resulting in a back-to-back topology.

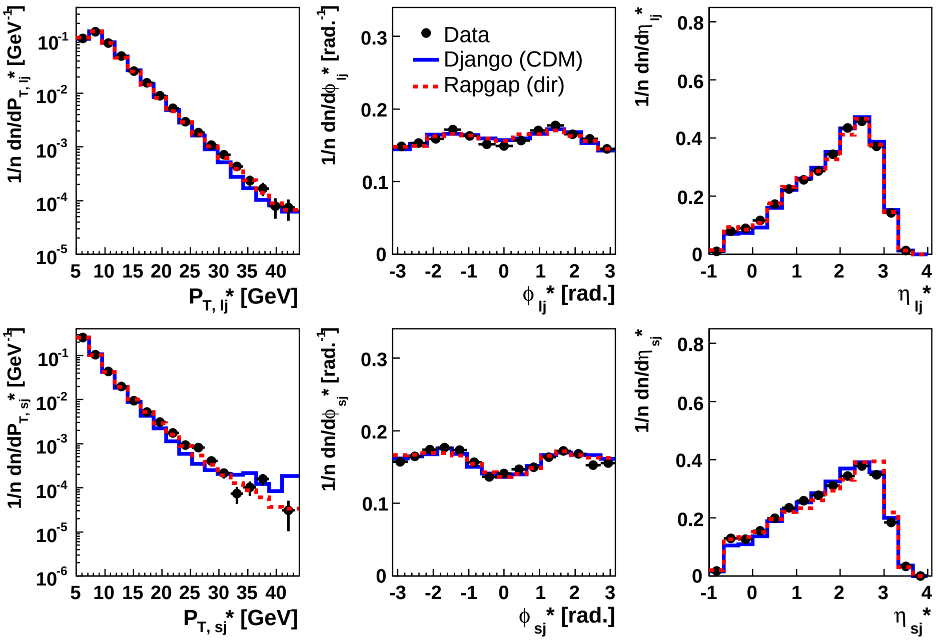


Figure 5.20: *The distributions of the transverse momentum, the pseudo-rapidity and the azimuthal angle of the leading (upper plots) and sub-leading jet (lower plots) of data compared to the predictions of Django(CDM) and Rapgap(dir) for the inclusive 2-jet sample at high W. The distributions are normalised to the total number of events.*

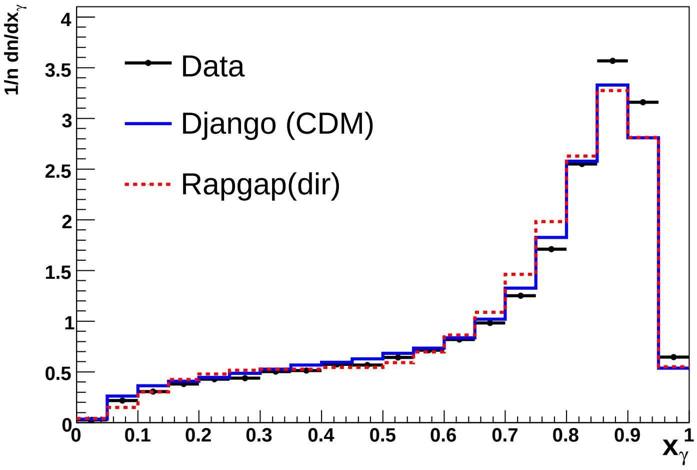


Figure 5.21: *The x_γ-distribution of data compared to the predictions of Django(CDM) and Rapgap(dir) for the inclusive 2-jet sample. The distributions are normalised to the total number of events.*

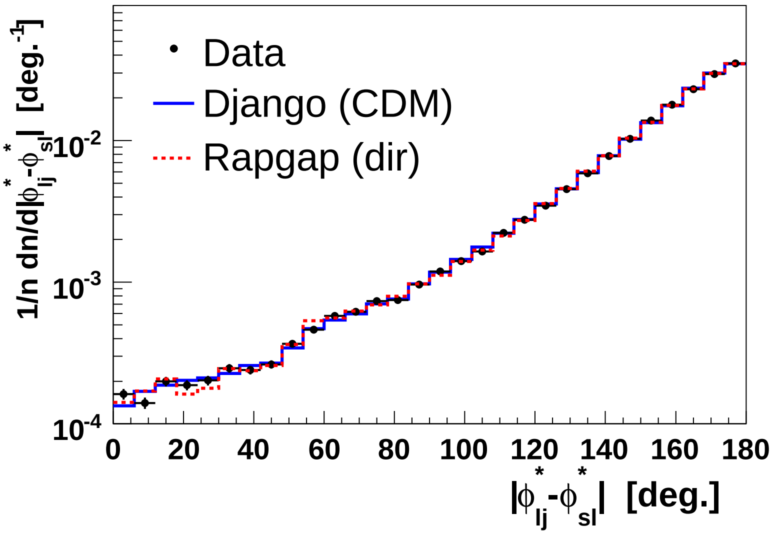


Figure 5.22: *The distributions of* $|\phi^*_{lj} - \phi^*_{sl}|$ *of data compared to the predictions of Django(CDM) and Rapgap(dir) for the inclusive 2-jet sample before the cut* $|\phi^*_{lj} - \phi^*_{sl}| > 140°$ *have been applied. The distributions are normalised to the total number of events.*

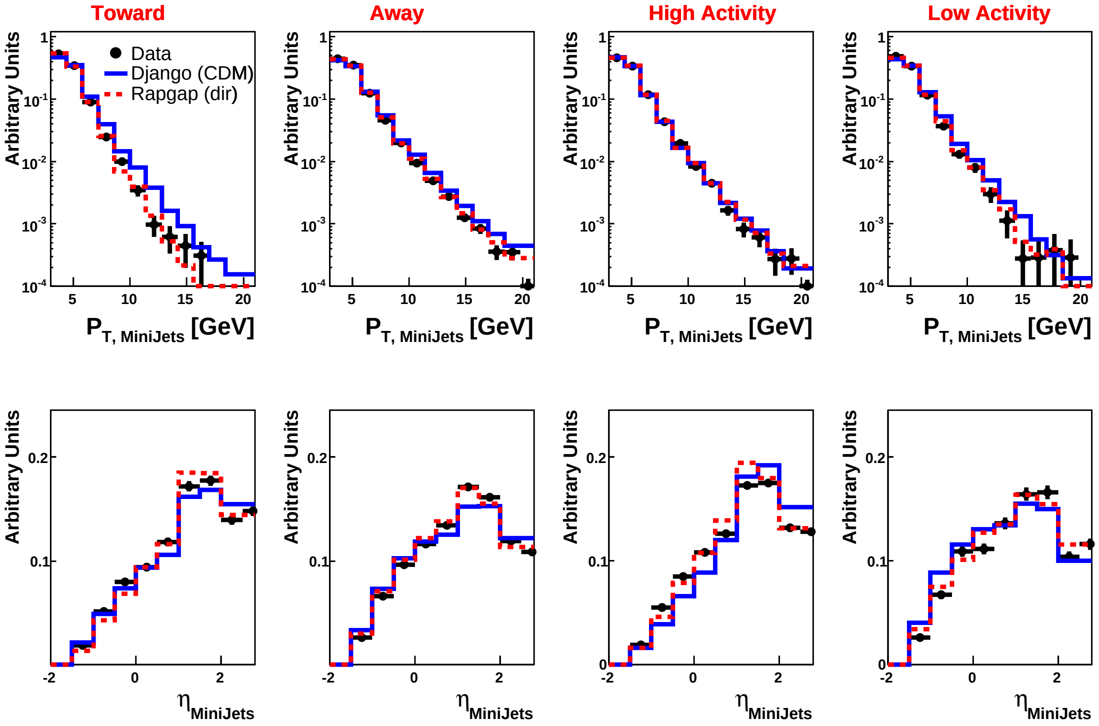


Figure 5.23: *The distributions of* $P_{T,\ MiniJets}$ *(upper plots) and* $\eta_{MiniJets}$ *(lower plots) for the inclusive 1-jet sample, with the leading jet going forward, shown separately for the toward, away, high- and low activity regions, for data compared to the predictions of Django(CDM) and Rapgap(dir) at low* W*.*

Kinematics of Mini-jets

Figure 5.23 shows the distributions of transverse momentum, $P_{T,MiniJets}$, and pseudorapidity, $\eta_{MiniJets}$, for mini-jets from the inclusive 1-jet sample, where the leading jet is in the forward region. The distributions are shown separately for the toward, away, high- and low activity regions at low W. The leading jet in the toward region has been excluded from the plots. It is observed that Rapgap(dir) gives a better agreement with data compared to Django(CDM). The same level of agreement is found for the central inclusive 1-jet sample and for the inclusive 2-jet sample.

5.4.3 Purity and Stability

The limitations of the detector perfomance results in errors on reconstructed quantities, which leads to detector level distributions that are smeared and shifted with respect to the true values. Thus, an event reconstructed in bin i on the hadron level might be reconstructed in a different bin on the detector level. This effect is called bin migration and can be studied using Monte Carlo simulations. In order to study migration between bins, the two variables purity, (P), and stability, (S), are used.

$$P_i = \frac{N_i(HAD \ \wedge \ DET)}{N_i(DET)} \tag{5.7}$$

$$S_i = \frac{N_i(HAD \ \wedge \ DET)}{N_i(HAD)}. \tag{5.8}$$

Here, $N_i(DET)$ and $N_i(HAD)$ are the number of events reconstructed in bin i on the detector and hadron levels, respectively, and $N_i(HAD \ \wedge \ DET)$ is the number of events reconstructed in the same bin i on hadron and detector level. For a perfect detector, purity and stability are equal to unity. The purity, P_i, measures the fraction of events which have been correctly reconstructed in bin i. The stability, S_i, gives the fraction of events in bin i which will not migrate out of that bin at detector level. A large fraction of these migrations are from/to neighboring bins. One obvious way to decrease the migrations is to increase the bin size.

Figure 5.24 shows the stability and purity in bins of Q^2 as calculated by Django(CDM) and Rapgap(dir) for the inclusive 1-jet sample, where the leading jet has been reconstructed in the forward region, both for low and high values of W. In all cases the stability and purity are better than 50%, and the prediction of Django(CDM) and Rapgap(dir) differ at most by 10%. For the sample where the leading jet is reconstructed in the central region, the stability and purity exceeds 60%, as seen from Figure 5.25, which means that there are fewer migrations. This is due to the good momentum resolution of the central tracker, which is reflected in the momentum resolution of jets. In the central region the calculations of the two Monte Carlo programs give almost idential values.

Figure 5.26 shows the Django(CDM) and Rapgap(dir) calculations of the stability and purity as a function of x_γ for the inclusive 2-jet sample, separately for the low and high W-regions. The purity is on the level of 60% or above. The stability is on the level of 60% for high values of x_γ and between 40% and 50% in the low x_γ-region.

The main reason for migrations in the jet sample is the fairly poor energy resolution of jet clusters.

5.4.4 Correction Factors

The experimental data are corrected for detector and QED effects by taking the ratio between MC generated data on hadron level without QED radiation and on detector level including QED radiation. Due to the limited resolution of the detector it may happen that events generated in one kinematic bin migrate into another, as discussed in section 5.4.3. In analyses were the migrations are small the *bin-by-bin method* can be used to correct the data. However, if the migrations are substantial the *unfolding method*, which takes the effects of migrations into account, has to be used. In this analysis the Monte Carlo event generators Rapgap(dir) and Django (CDM) have been used to calculate the correction factors for both methods and the results have been compared.

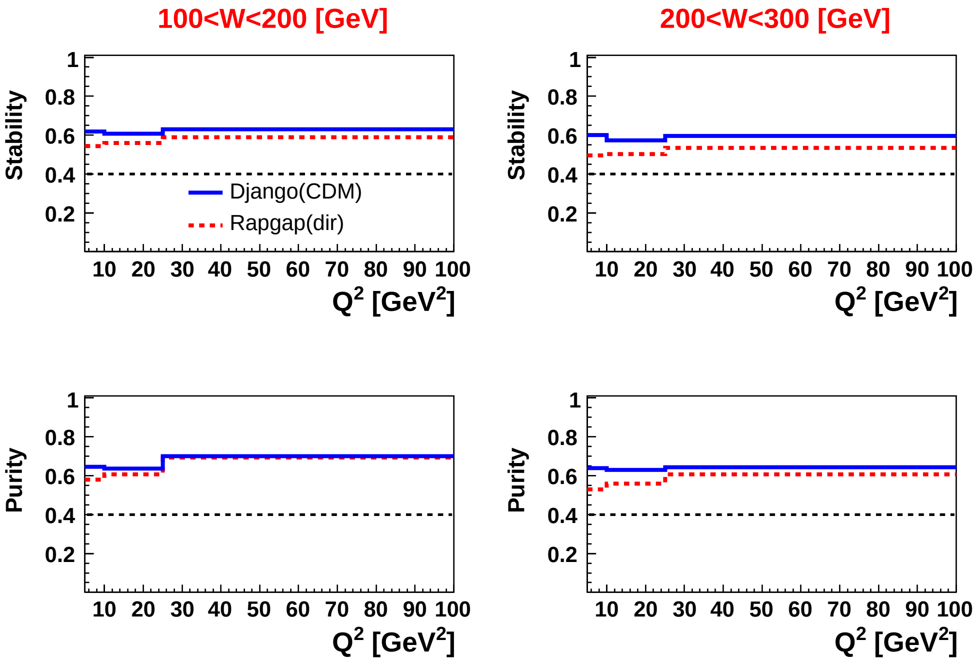


Figure 5.24: *The Stability (S) and the Purity (P) as a function of Q^2 for the forward inclusive 1-jet sample in two bins of W, calculated by Django(CDM) and Rapgap(dir). The thin broken line at 0.4 represents the lowest acceptable value of the stability and purity.*

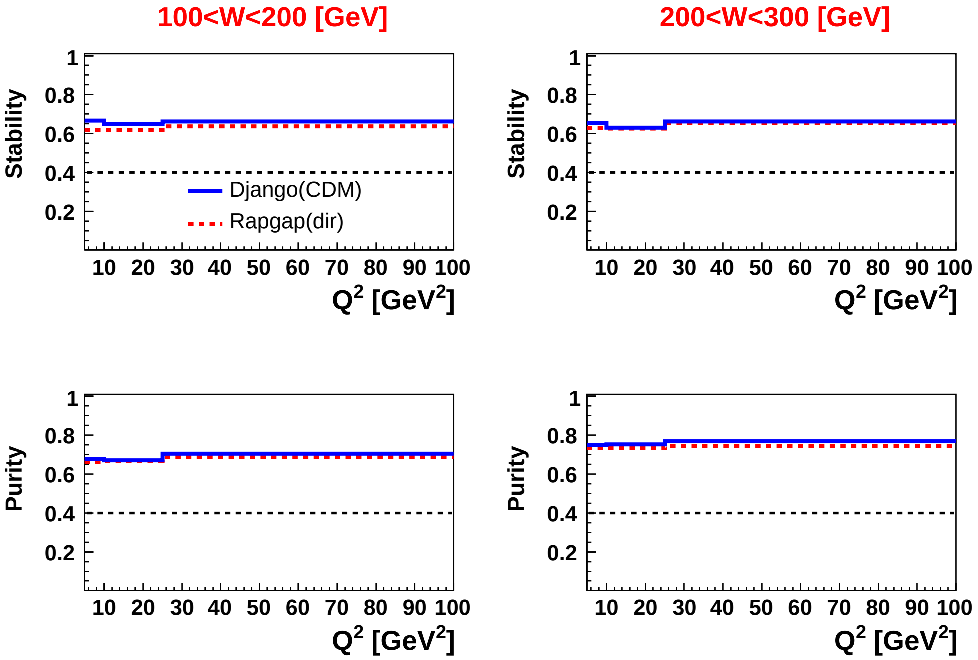


Figure 5.25: *The Stability (S) and the Purity (P) as a function of Q^2 for the central inclusive 1-jet sample in two bins of W, calculated by Django(CDM) and Rapgap(dir). The thin broken line at 0.4 represents the lowest acceptable value of the stability and purity.*

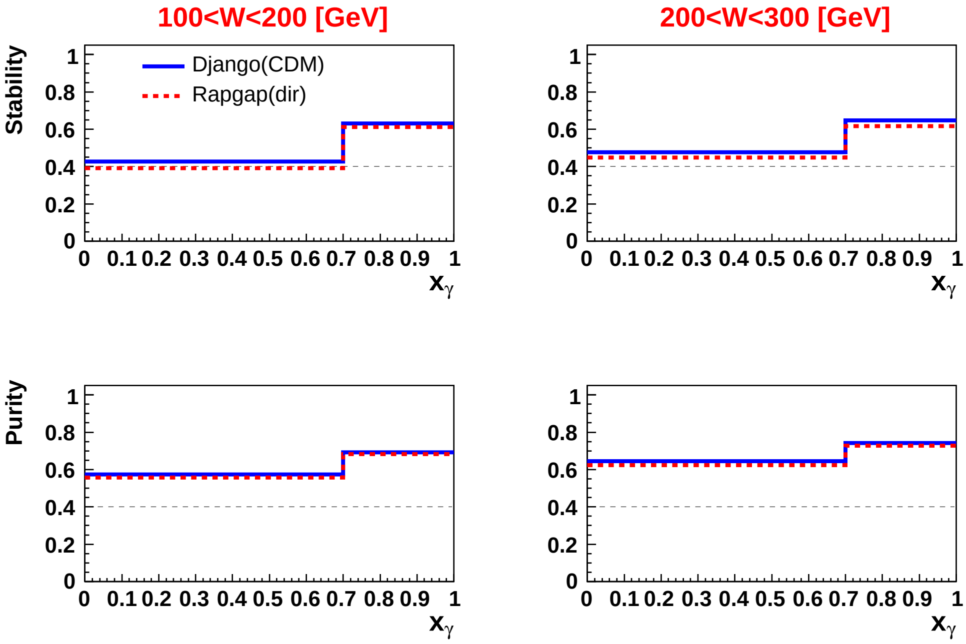


Figure 5.26: *The stability (S) and purity (P) as a function of x_γ for the inclusive 1-jet sample in two bins of W, calculated by Django(CDM) and Rapgap(dir). The thin broken line at 0.4 represents the lowest acceptable value of the stability and purity.*

5.4.5 The Bin-by-Bin Correction Method

In the bin-by-bin method the correction factor, C_i, with respect to an observable A in bin i is defines as:

$$C_i(\langle A\rangle) = \frac{\langle A\rangle_{i,Had,NonRad}}{\langle A\rangle_{i,Det,Rad}}, \qquad \langle A\rangle = \langle N\rangle, \langle PtSum^*\rangle. \tag{5.9}$$

Here, $\langle A\rangle_{i,Had,NonRad}$ and $\langle A\rangle_{i,Det,Rad}$ are the observable A on the hadron level without QED radiation and detector level with QED radiation, respectively, in bin i. Below, a selection of correction factors of $\langle N\rangle$ and $\langle PtSum^*\rangle$ for the inclusive 1- and 2-jet samples for various $\Delta\phi^*$ and phase space regions are presented.

In the Figures 5.27-5.30 the corrections factors $C(\langle N\rangle)$ calculated using Django(CDM) and Rapgap(dir), are plotted as a function of Q^2 for the inclusive 1-jet sample, subdivided into the central and forward η-regions of the leading jet, and two W ranges. This set of plots are shown separately for the four regions in azimuthal angle, the toward (Figure 5.27), the away (Figure 5.28), the high (Figure 5.29) and low (Figure 5.30) activity regions. As observed the correction factors are essentially equal to unity in the whole phase space coverage of the toward region. In the away region the correction factors are in general larger being highest in the case of the leading jet going forward and the invariant mass of the hadronic system being low. A possible explanation is that in the forward η-region of the LAr calorimeter the energy resolution of the leading jet is poor. The corrections of $\langle N\rangle$ in the high and low activity regions are everywhere below 30%.

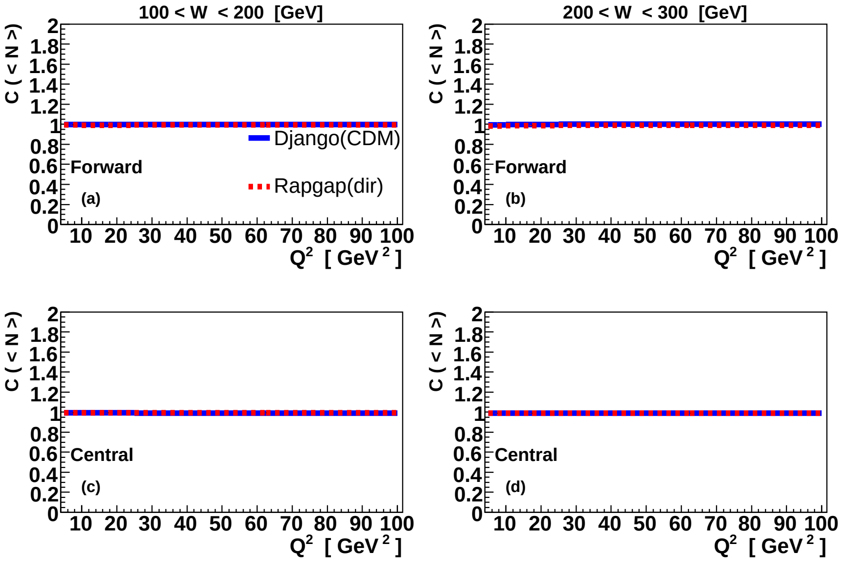


Figure 5.27: *The correction factors for the average mini-jet multiplicity,* $C(\langle N\rangle)$*, as a function of* Q^2 *for the inclusive 1-jet sample, shown for the toward region in two bins of* W*.*

For the inclusive 2-jet sample $C(\langle N\rangle)$ is shown as a function of x_γ for the two W regions in Figures 5.31 and 5.32. The results on $\langle N\rangle$ for the toward, away, high and low activity regions, respectively, are shown in the Figures 5.31 and 5.32. The corrections factors are close to unity in the toward and away regions, whereas in the transverse regions the average corrections from Django(CDM) and Rapgap(dir) are below 20%.

The correction factor $C(\langle PtSum^*\rangle)$ for the 1- and 2-jet samples are shown in Figures (5.33-5.34) and in Figure 5.35, respectively. As can be seen, the corrections factors for $\langle PtSum^*\rangle$ are similar to those of $\langle N\rangle$.

The fact that the corrections factors calculated from Django(CDM) and Rapgap(dir) are different, is a reflection of the different models used for parton production. Therefore, the average correction factors obtained from the two MC generators are used. Also, half of the difference between the correction factors is taken as an estimate of the systematic uncertainty (see also section 5.4.7).

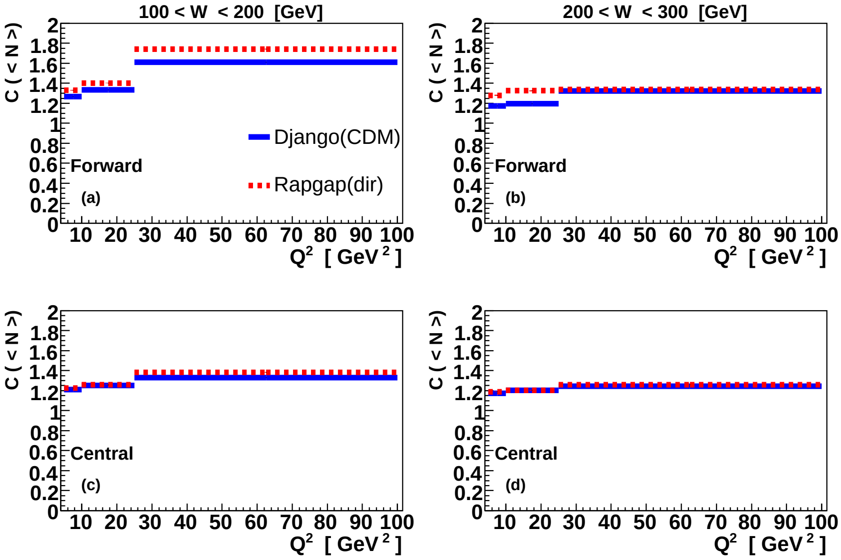


Figure 5.28: *The correction factors for the average mini-jet multiplicity,* $C(\langle N \rangle)$, *as a function of* Q^2 *for the inclusive 1-jet sample, shown for the away region in two bins of* W.

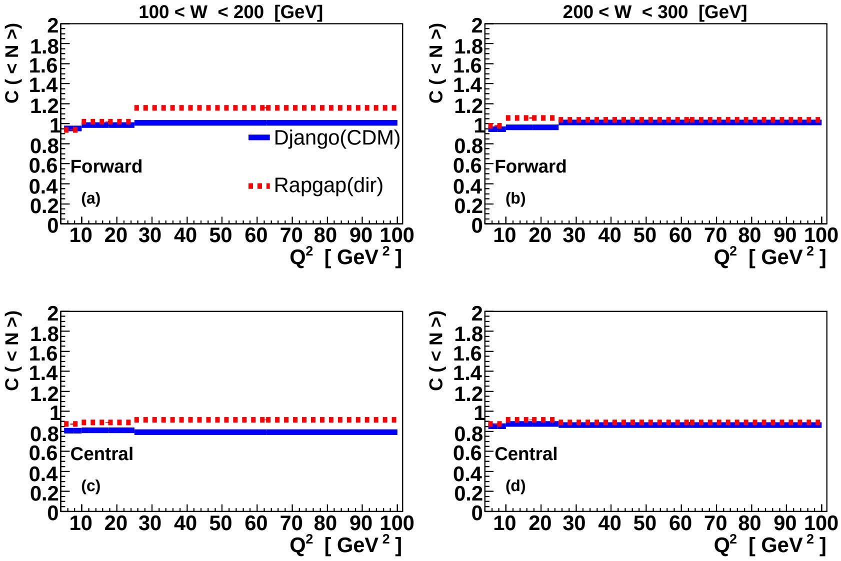


Figure 5.29: *The correction factors for the average mini-jet multiplicity,* $C(\langle N \rangle)$, *as a function of* Q^2 *for the inclusive 1-jet sample, shown for the high activity region in two bins of* W.

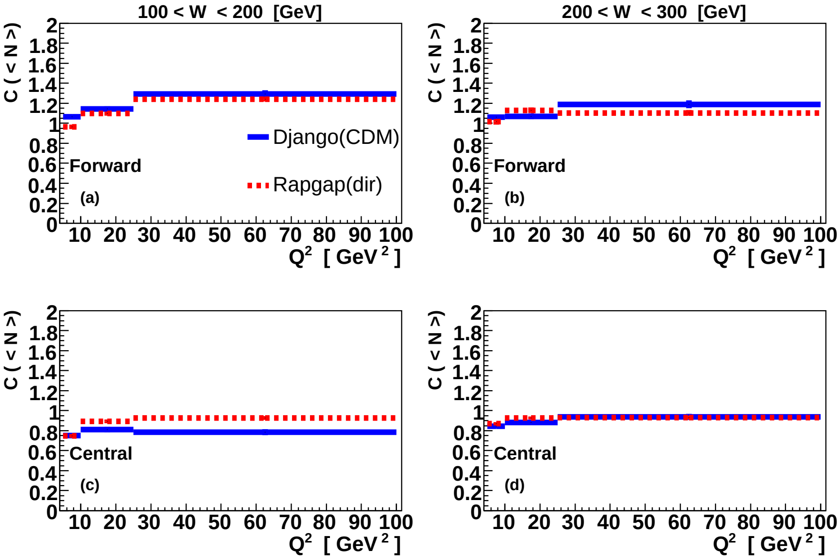


Figure 5.30: *The correction factors for the average mini-jet multiplicity, $C(\langle N \rangle)$, as a function of Q^2 for the inclusive 1-jet sample, shown for the low activity region in two bins of W.*

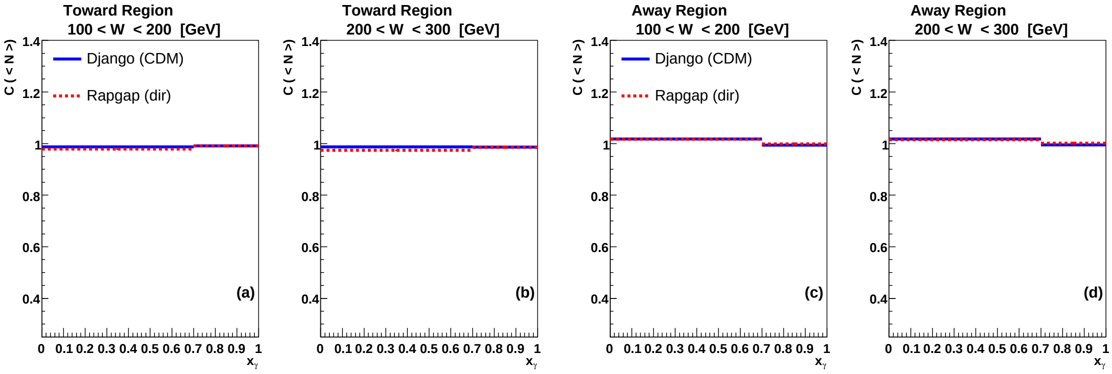


Figure 5.31: *The correction factors for the average mini-jet multiplicity, $C(\langle N \rangle)$, as a function of x_γ for the inclusive 2-jet sample, shown separately for the toward and away regions in two bins of W.*

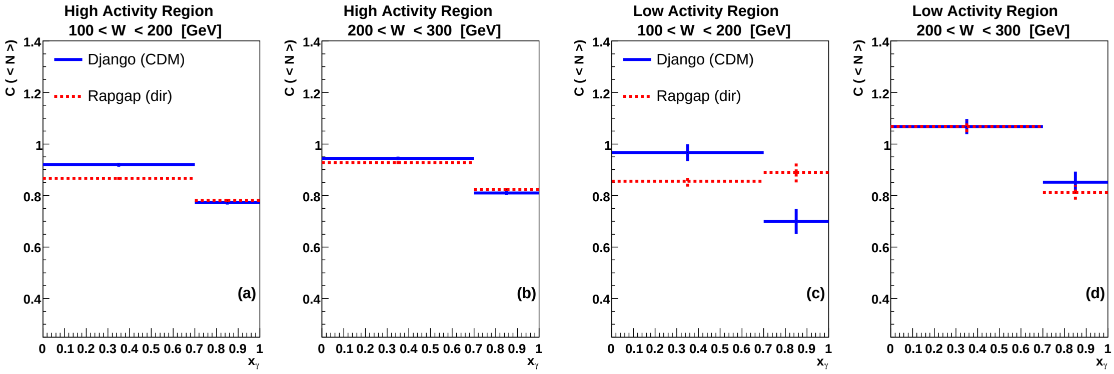


Figure 5.32: *The correction factors for the average mini-jet multiplicity, $C(\langle N \rangle)$, as a function of x_γ for the inclusive 2-jet sample, shown separately for the high and low activity regions in two bins of W.*

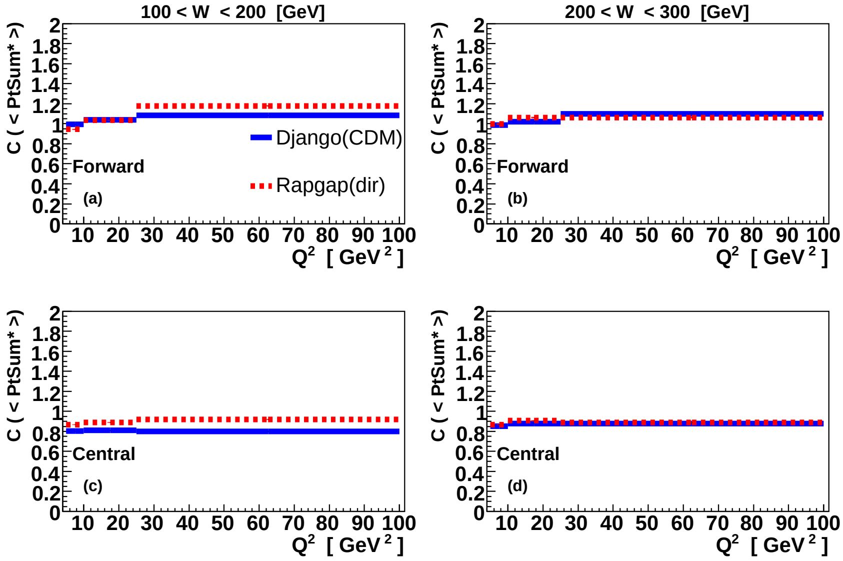


Figure 5.33: *The correction factors for $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, shown for the low activity region, with the leading jet going forward and central, respectively, and in two bins of W.*

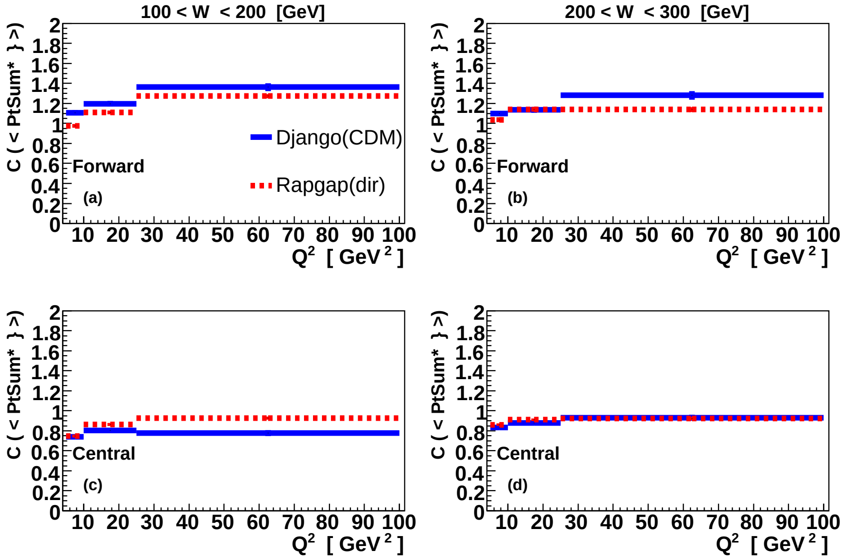


Figure 5.34: *The correction factors for* $\langle PtSum^* \rangle$ *as a function of* Q^2 *for the inclusive 1-jet sample, shown for the low activity region, with the leading jet going forward and central, respectively, and in two bins of* W.

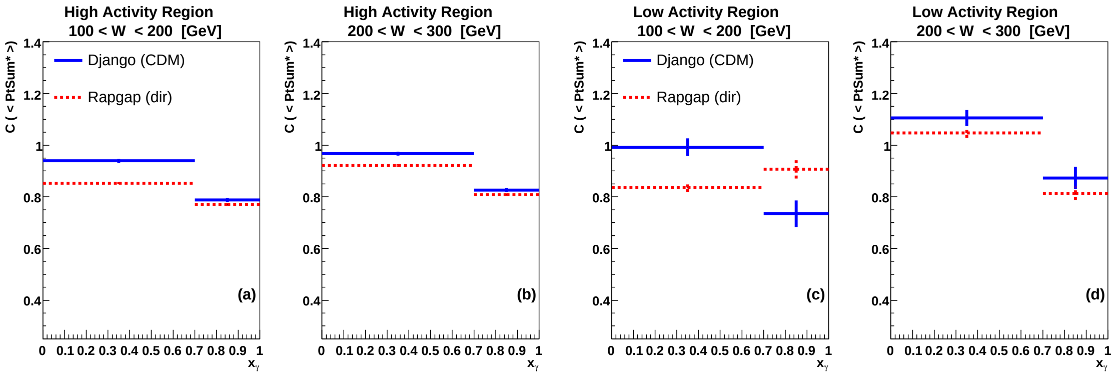
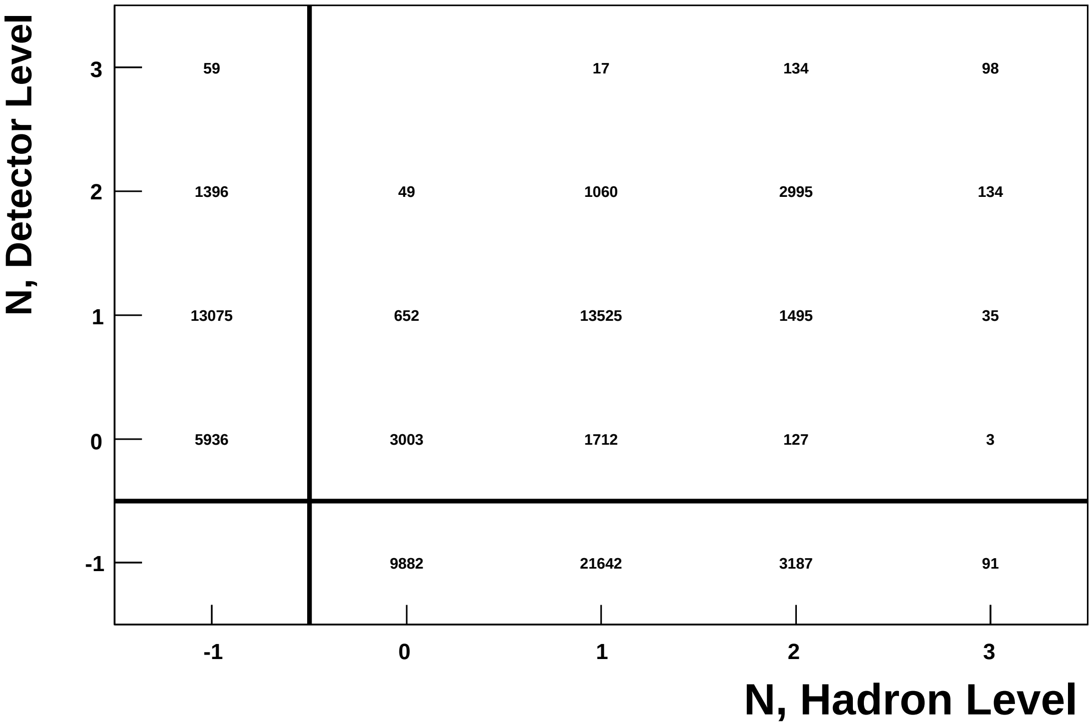


Figure 5.35: *The correction factors for* $\langle PtSum^* \rangle$ *as a function of* x_γ *for the inclusive 2-jet sample in the high- and low activity regions in two bins of* W.



	-1	0	1	2	3
3	59		17	134	98
2	1396	49	1060	2995	134
1	13075	652	13525	1495	35
0	5936	3003	1712	127	3
-1		9882	21642	3187	91

Figure 5.36: *The smearing matrix of the mini-jet multiplicity, N, for the central inclusive 1-jet sample in the away region, $5 < Q^2 < 10$ GeV^2 and $200 < W < 300$ GeV as generated by Rapgap(dir).*

5.4.6 The Unfolding Method

The migrations in some variable A, are given by a smearing matrix, S, which provides the number of events generated in bin i on hadron level and found in bin j on detector level, i.e. the smearing matrix reflects the influence of the detector on the recorded data. Figure 5.36 shows the smearing matrix of the mini-jet multiplicity in the away region for the inclusive 1-jet sample, where the leading jet is going in the forward direction, $5 < Q^2 < 10$ GeV^2 and $200 < W < 300$ GeV. The diagonal elements are as expected most populated, since they correspond to the number of events where equally many mini-jets are reconstructed as are produced. This essentially only means that the number of migrations out of one specific bin is equal to the number of migrations into that bin. The off-diagonal elements give the net number of migrations into neigbouring bins with respect to the bins where they were generated and they should be substantially smaller if the degree of migrations is reasonably low. The migrations between non-neighbouring bins especially should be quite small if the bin-by-bin method is to be used. The bins $(0, j)$ represent the case where a leading jet but no further jets has been generated on hadron level although j jets (excluding the leading jet) were found on detector level. Similarly, the situation where i jets (excluding the leading jet) have been generated but no jets (excluding the leading jet) were reconstructed on detector level, is given by the bins $(i, 0)$. The bins $(-1, j)$ and $(i, -1)$ contains event where not even the leading jet ($P_T > 5$ GeV) has been generated and reconstructed, respectively. The bin $(-1,-1)$ consequently corresponds to the case where a leading jet was neither generated nor reconstructed, which are events of no relevance for the unfolding procedure since it doesn't contain any information about migrations. The migration of the mini-jet multiplicity between the $\Delta\phi^*$ regions is not taken into account in the best way. This due to the fact that there is no trivial way of linking a jet on hadron level to the corresponding jet on detector level. This informations is needed to study for example migration of mini-jets from the away regions into the high activity regions. It should be mentioned here that the migrations of the event between the Q^2 and W is small. One smearing matrix is used for each bin of Q^2, W, η_{lj} and $\Delta\phi^*$.

By applying the smearing matrix to the generated data the measured distributions will be obtained. However, since we start out with the experimental measurements and want to extract the distributions on hadron level we need to use the inverted smearing matrix. Thus the relation between the hadron and detector levels can be written as:

$$h = S^{-1}d,$$

where h and d is the unfolded mini-jet multiplicity distribution on hadron and detector level, respectively. However, the inversion of the smearing matrix may suffer from statistical uncertainties or large statistical fluctuations in some of the bins and therefore more sophisticated methods might have to be used, like the regularized unfolding method [107]

The method used here, in order to avoid the problems with matrix inversion, is the iterative method based on Bayes' theorem [108–110]. The method has been used in many analyses by H1 [111–114]. The task is to determines the unfolding matrix, which delivers the hadron level data if applied to the measured data. This in principle requires knowledge about distributions on both levels, which is unfortunately not the case since the hadron level distribution is the one that has to be determined. However a reasonable ansatz concerning the 'true' hadron level distribution can be made and the distribution after unfolding can be assumed to be a better approximation to the 'true' distribution than the starting distribution. The unfolded distribution can then be used as input for the following step in an iterative procedure to get closer to the 'true' distribution. The number of iterations needed can be found by unfolding detector simulated MC events from one MC generator, using the smearing matrix and a starting distribution produced by a different MC program. The unfolded distribution of the first MC generator is then compared to the true hadron level distribution from the same MC. The deviations between the unfolded and true distributions is given by:

$$\chi = \sum_i \frac{|n_{i,unfold} - n_{i,had}|}{n_{i,had}}$$

which is summed over all bins i.

In Figure 5.37 the mini-jet multiplicity, N, of Rapgap(dir) and Django(CDM) is shown at detector level for the central inclusive 1-jet sample in the away region (left plot) and the low activity region (right plot) for $5 < Q^2 < 10$ GeV^2 and $200 < W < 300$ GeV.

Figure 5.38 shows χ as a function of the number of iterations obtained from the unfolding of $\langle N\rangle$ in the away region (left plot) and the low acivity region (right plot) for the inclusive 1-jet sample with the leading jet being central, $5 < Q^2 < 10$ GeV^2 and

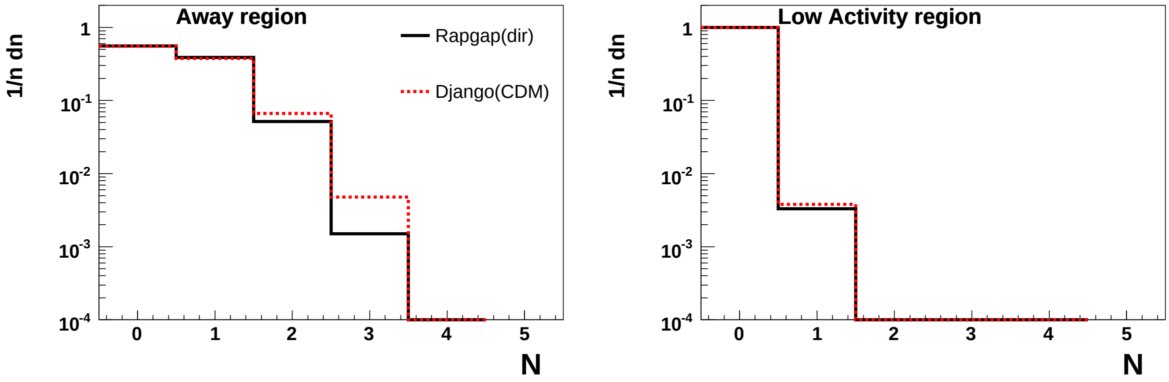


Figure 5.37: *The mini-jet multiplicity distribution for the inclusive 1-jet sample, with the leading jet being central, shown separately for the away and low activity regions,* $5 < Q^2 < 10$ *GeV*^{2} *and* $200 < W < 300$ *GeV, as calculated by Rapgap(dir) and Django(CDM).*

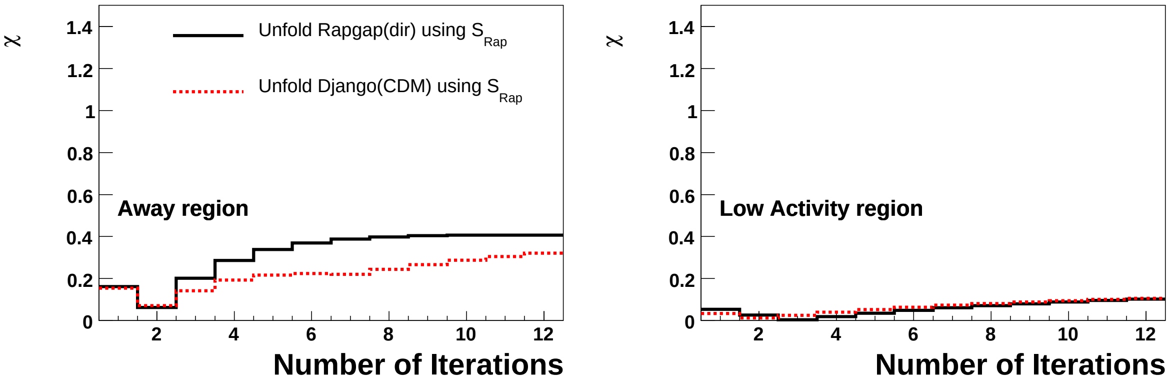


Figure 5.38: *The difference,* χ*, as a function of the number of iterations for the inclusive 1-jet sample, with the leading jet being central, shown separately for the away and low acitivity regions, 5<* $Q^2 <$ *10 GeV*^{2} *and 200* $< W <$ *300 GeV, as calculated by Rapgap(dir) and Django(CDM) using the smearing matrix provided by the Rapgap(dir) generator,* S_{Rap}*.*

$200 < W < 300$ GeV. Here, Rapgap(dir) and Django(CDM) are unfolded using the smearing matrix provided by Rapgap(dir), S_{Rap}. As can be seen the best agreement with the hadron level distribution is obtained after 2-3 iterations for both MC samples. A systematic investigation shows that 1-2 iterations are optimal for unfolding the inclusive 1-jet sample in the toward and away regions, whereas for the high and low activity regions usually a few more iterations are needed.

Figure 5.39 shows the smearing matrix of the mini-jet multiplicity for the inclusive 2-jet sample in the low activity region requiring $200 < W < 300$ GeV. The inclusive 2-jet sample suffers from larger migrations compared to the inclusive 1-jet sample, as was already concluded from the lower values of the stability and purity (see section 5.3.3). Again the optimal number of iteration needed in order to unfold the data is determined by MC simulations. Figure 5.40 shows χ versus the number of iterations for the inclusive 2-jet sample in the low activity region and at low W values, giving a minimum already at

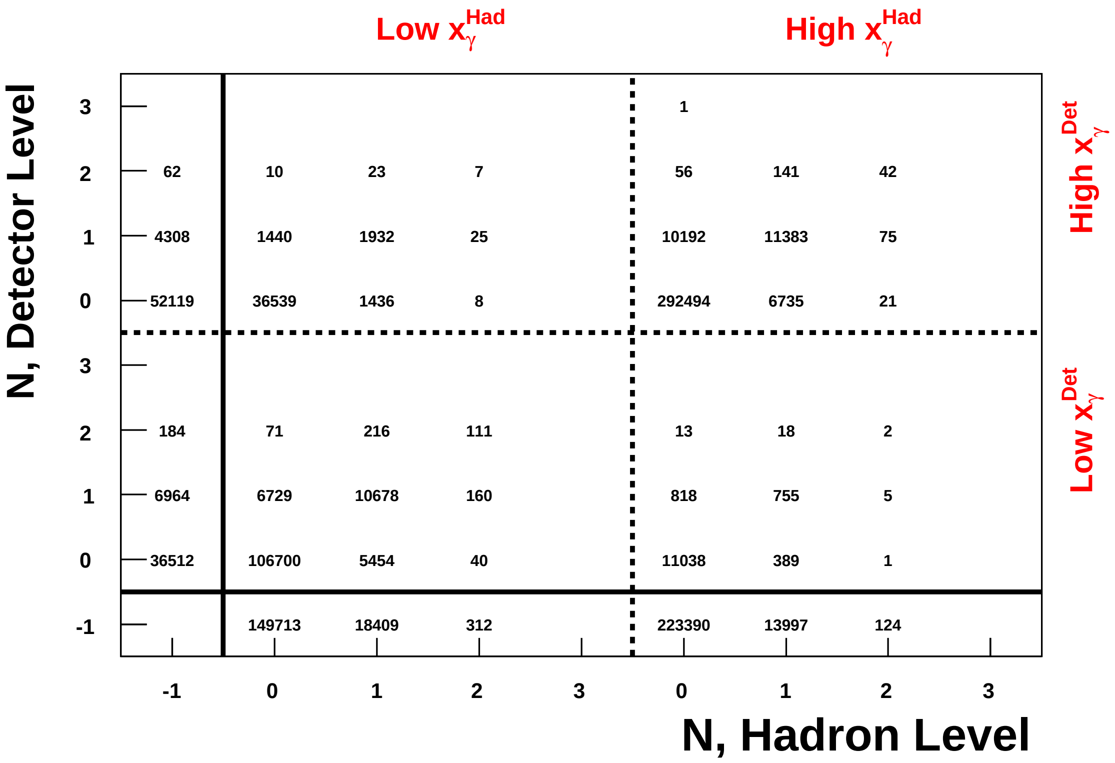


Figure 5.39: *The smearing matrix of the mini-jet multiplicity, N, for the inclusive 2-jet sample in the away region and for $200 < W < 300$ GeV as generated by Rapgap(dir).*

the first iteration.

The results of using the two correction methods on $\langle N \rangle$ are compared in Figures 5.41 and 5.42 for the inclusive 1-jet and 2-jet samples, respectively. It can be observed that the two methods agree within errors in all regions. Thus, the bin-by bin method can be used with confidence for this analysis.

5.4.7 Systematic Errors

The systematic uncertainties considered in this analysis are originating from the following sources:

- The uncertainty in the hadronic energy scale of the HFS energy was found to be ± 2% in Chapter 4. The influence of this on the mini-jet multiplicity is determined with Rapgap(dir) and Django(CDM), where the HFS energies have been increased and decreased by this amount.
- The absolute value of the electromagnetic energy scale in the SPACAL is determined to an accuracy of ± 1%.
- The scattering angle of the electron is measured to a precision of ±1 mrad.

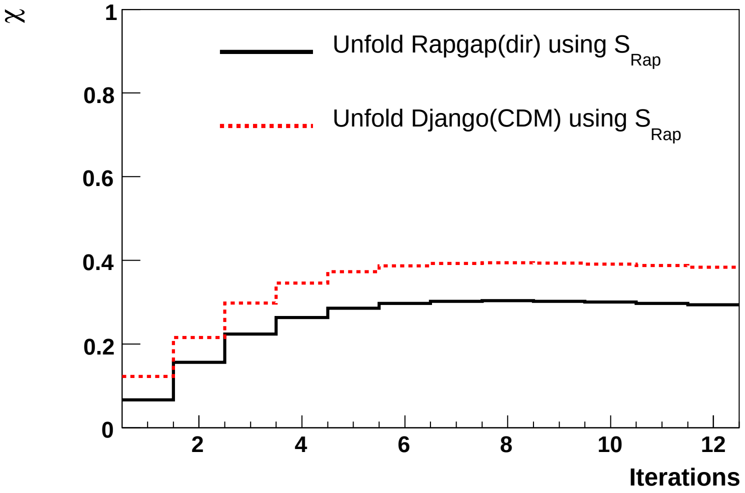


Figure 5.40: *The difference, χ, as a function of the number of iterations for the inclusive 2-jet sample, shown separately for the away and low acitivity regions and $200 < W < 300$ GeV, as calculated by Rapgap(dir) and Django(CDM) using the smearing matrix provided by the Rapgap(dir) generator, S_{Rap}.*

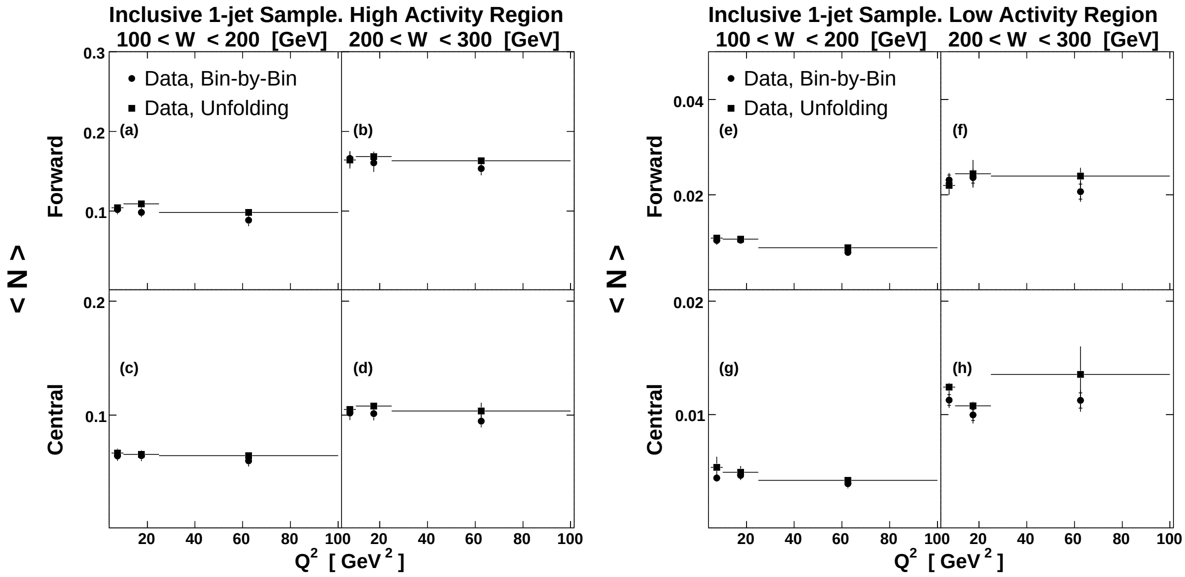


Figure 5.41: *The average mini-jets multiplicity, $\langle N \rangle$, on hadron level as a function of Q^2, obtained by the bin-by-bin method (circles) and the unfolding method (squares) for the inclusive 1-jet sample, with the leading jet going forward and central, respectively, shown separately for the high and low activity regions in two bins of W.*

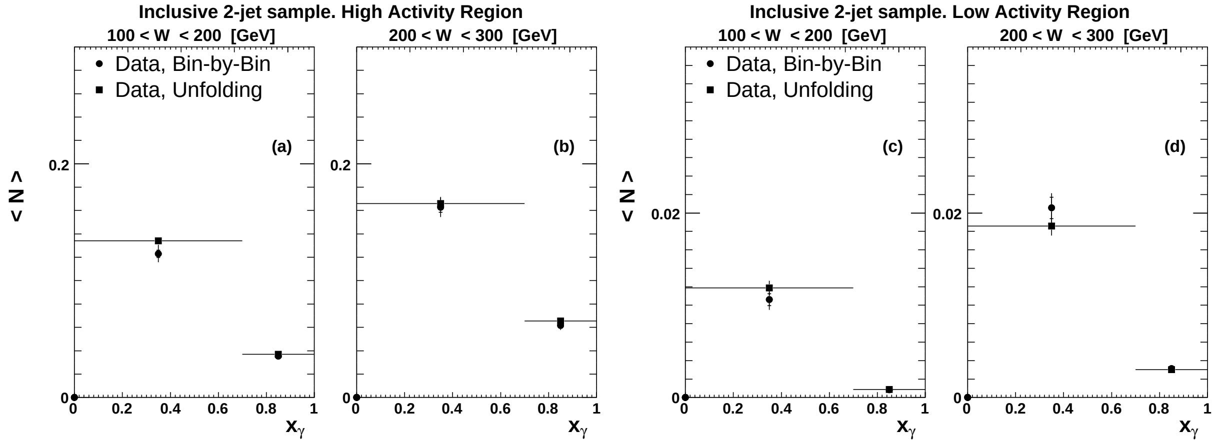


Figure 5.42: *The average multiplicity of mini-jets, $\langle N \rangle$, on hadron level as a function of x_γ, obtained by the bin-by-bin method (circles) and the unfolding method (squares) for the inclusive 2-jet sample, shown separately for the high and low activity regions in two bins of W.*

- The systematic uncertainties in the correction factors due to the model dependence is taken to be half of the difference between the correction factors obtained by Rapgap(dir) and Django(CDM).

A summary of the systematic errors of $\langle N \rangle$ introduced from the various sources mentioned above are given in Table(5.2). The largest systematic uncertainties are coming from the model dependence of the correction factors which can be as large as 10 %. All systematic uncertainties and the statistical error are added in quadrature to get the total error of the measurement.

Systematic sources	Inc. 1-jet Average Syst.	Inc. 2-jet Average Syst.
The HFS energy uncertainty (± 2%)	2 %	2 %
SPACAL EM energy scale (± 1 %)	< 1 %	< 1 %
Angle of the scattered electron(±1 mrad)	<1 %	<1 %
Model dependence ($\frac{C(Rap)-C(Dja)}{2}$)	8 %	8 %

Table 5.2: *Summary of the systematic errors of $\langle N \rangle$.*

5.5 Results

In this chapter the results on the average mini-jet multiplicity, $< N >$, and the total transverse momentum, $< PtSum^* >$, carried by the mini-jets are presented and compared to the predictions on hadron level by models not including QED effects. The statistical errors are calculated as the standard error of the mean value according to:

$$\sqrt{\frac{\langle A^2\rangle - \langle A\rangle^2}{N_{ev}}}, \qquad \langle A\rangle = \langle N\rangle, \langle PtSum^*\rangle. \tag{5.10}$$

where N_{ev} is the number of events. The statistical errors are given by the inner error bars while the systematic errors correspond to the outer error bars in the figures. In this section, Rapgap dir+res denotes Rapgap including direct and resolved photon procceses. The notation Lepto(CDM) indicates that the Lepto generator is using CDM for intial and final state radiations. The following Monte Carlo generators are used for data comparison: Rapgap dir+res, Lepto(CDM), Pythia with and without MPI, Herwig with and without SUE, and Cascade.

The presentation of the results is organised such that the inclusive 1-jet sample is discussed first and thereafter the inclusive 2-jet sample. For each sample a comparison to the various MC predictions is initially presented for the toward and away regions in order to see how well the different models are able to describe the hard scattering. This is followed by a description of the results obtained in the transverse regions, where a signal of MPI is most likely to be visible.

5.5.1 The Inclusive 1-jet Sample

The results on $\langle N\rangle$ and $\langle PtSum^*\rangle$ as a function of Q^2 are shown in the Figures 5.43 and 5.44, respectively, in the toward and away regions for the inclusive 1-jet sample, with the leading jet in the central and forward regions and for the two W bins. The general observation from data is that $\langle N\rangle$ in the toward regions is sligtly above unity as expected since the presence of the leading jet is required. In the away region $\langle N\rangle \approx 0.7$, which is also reasonable due to the fact that the jets balancing the leading jet in 3-jet events might spill over to the transverse regions or be outside the acceptance of mini-jets in psuedorapidity, $-1.5 < \eta_{jets} < 2.79$.

In Figures 5.43 and 5.44 comparisons are made with the Rapgap dir+res and Lepto(CDM) models. An inspection of the results on $\langle N\rangle$ shows that there is agreement with data by both MC models within the measuring errors in the toward region. Rapgap dir+res is also able to reproduce the data in the away region if the leading jet is going forward but exhibit some deviations if the leading jet is central. Lepto(CDM) does quite well in all regions except for the away region, in case where the leading jet is forward and W is high. The description of $\langle PtSum^*\rangle$ by Rapgap dir+res is satisfactory except for centrally going leading jets in the away region, where the observed deviation is related to the fact that the number of mini-jets in this region undershoots the data as seen in Figure 5.43. Lepto(CDM) provides a good description in all regions except in the toward region if the leading jet is going forward. This is a reflection of the well-known fact that CDM generates a higher energy flow in the forward region than is observed in the data, as already concluded from the $P^*_{T,lj}$ distribution shown in Figure 5.14. The $\langle PtSum^*\rangle$ and $\langle N\rangle$ dependence on Q^2 is essentially described by both models.

In Figures 5.45 and 5.46 the predictions on $\langle N\rangle$ and $\langle PtSum^*\rangle$ of Rapgap dir+res and Lepto(CDM) are compared to data in the high and low activity regions. The general observation for the transverse regions is that the predictions of both MC programs significantly undershoot the data for both observables in all regions. The high acitivity region has a mini-jet activity, which is roughly a factor ten higher than in the low activity region. This can be explained by contributions from jets produced in the hard scattering, which are not confined to the away region. It can also be observed that the jet activities are higher when the leading jet is going forward compared to when it is central. It should be noticed that Rapgap dir+res predicts a weak increase in $\langle N \rangle$ and $\langle PtSum^* \rangle$ with increasing Q^2, whereas Lepto(CDM) indicates a decrease with increasing Q^2. Data seem to support a weak decrease.

Figures 5.47-5.50 show the predictions of Pythia and Pythia MPI compared to data in the various regions. As expected the inclusion of MPI has very little influence on the results in the toward and away regions, since these are dominated by the hard scattering products. Although the mini-jet multiplicity is well described by Pythia in the toward region only the lowest Q^2-bin is reproduced in the away region. As Q^2 increases the deviations from data become increasingly large. The situation is even worse for $\langle PtSum^* \rangle$, where Pythia only provides agreement with data in the low Q^2 region for both the toward and away regions. Since Pythia fails to reproduce the high Q^2 behaviour in the regions of the hard scattering, it can obviously not be trusted for high Q^2 in the transverse regions either. Comparing the Phythia predictions with data in the transverse regions for the lowest Q^2 bin alone, it is noticed that the inclusion of MPI improves the agreement significantly and especially in the low activity region, where an MPI signal should appear most clearly, there is good agreement.

The predictions of Herwig and Herwig 10% SUE are shown in Figures 5.51-5.54 together with data. Again, the effects of SUE in the toward and away regions are almost negligible as expected. The agreement with data on $\langle N \rangle$ is excellent in the toward region but slightly high in the away region. The $\langle PtSum^* \rangle$ distributions are also well described in the toward region but significantly above data in the away region. This is consistent with the mini-jet multiplicity being somewhat high in this region. The predictions of Herwig in the transverse regions are substantially below data for both $\langle N \rangle$ and $\langle PtSum^* \rangle$ The inclusion of 10% SUE adds some jet activity in the low Q^2 bins of the transverse regions, but it is far from sufficient to reproduce the data. Increasing the amount of SUE would further improve the situation in the transverse regions but would at the same time worsen the situation in the regions of the hard scattering products.

Finally the predictions of Cascade and Cascade J2003 set2 are compared to data in Figures 5.55-5.58. Although not perfect, Cascade J2003 set2 provides by far the best agreement with data, both on $\langle N \rangle$ and $\langle PtSum^* \rangle$. This is the only model that is able to reproduce the data in the transverse regions.

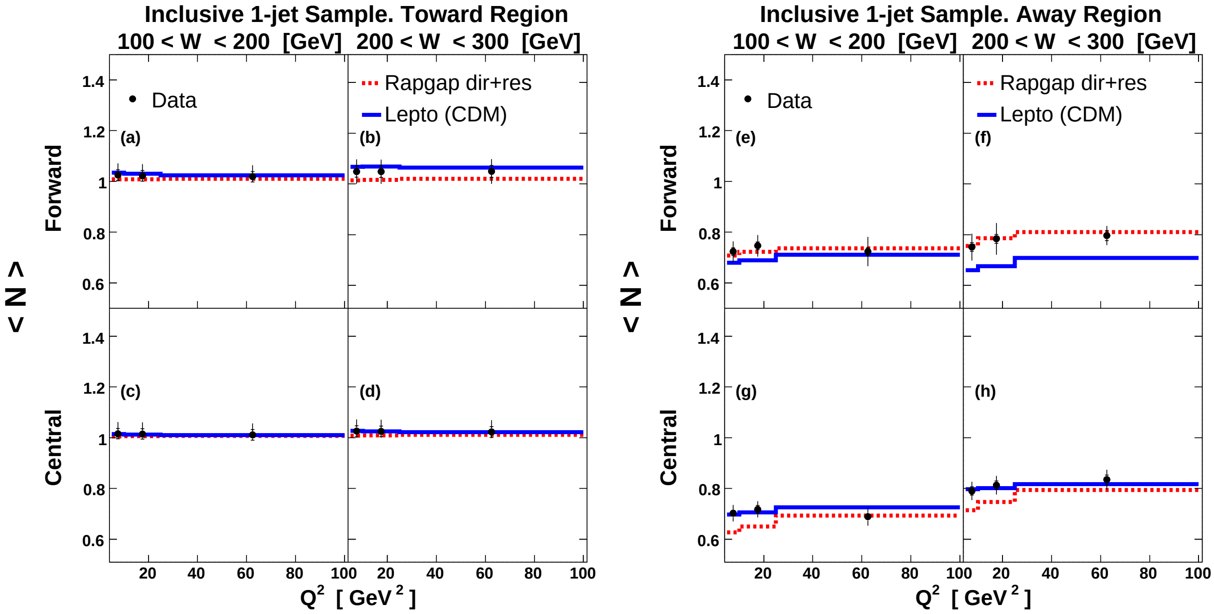


Figure 5.43: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Lepto(CDM) and Rapgap dir+res.*

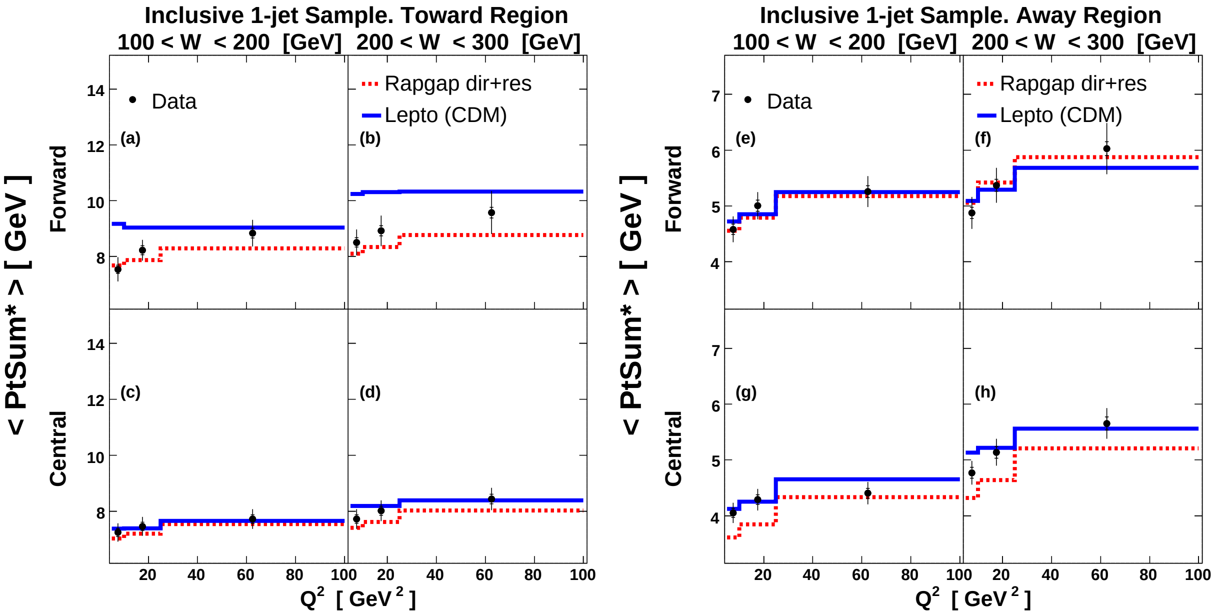


Figure 5.44: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Lepto(CDM) and Rapgap dir+res.*

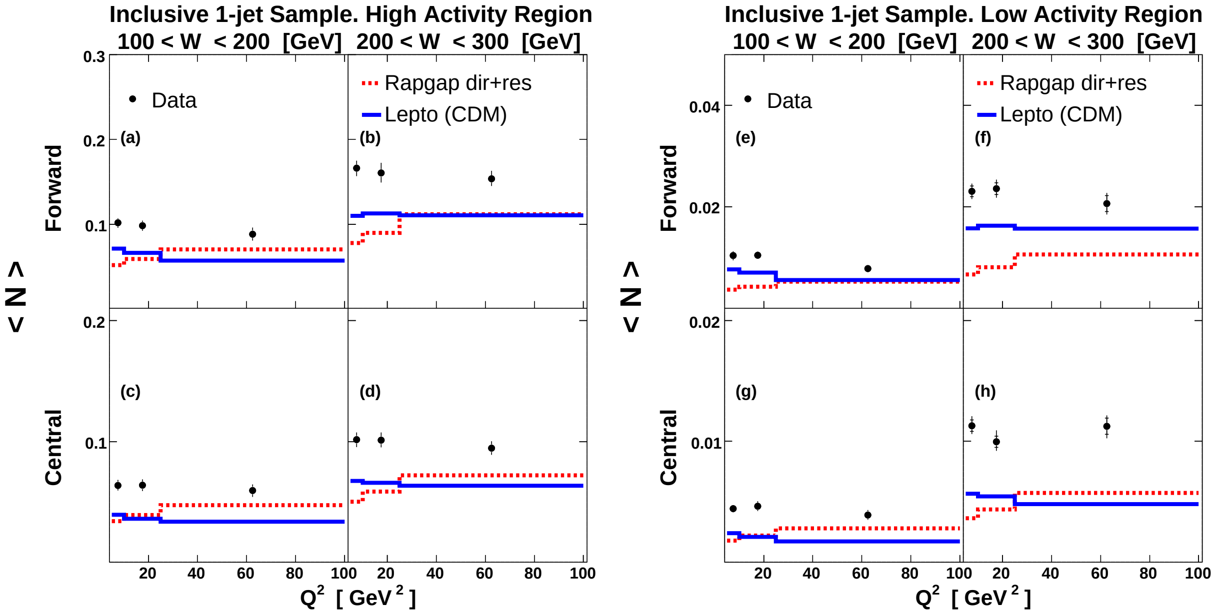


Figure 5.45: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the high and low activity regions in two bins of W. Data are compared to the predictions of Lepto(CDM) and Rapgap dir+res.*

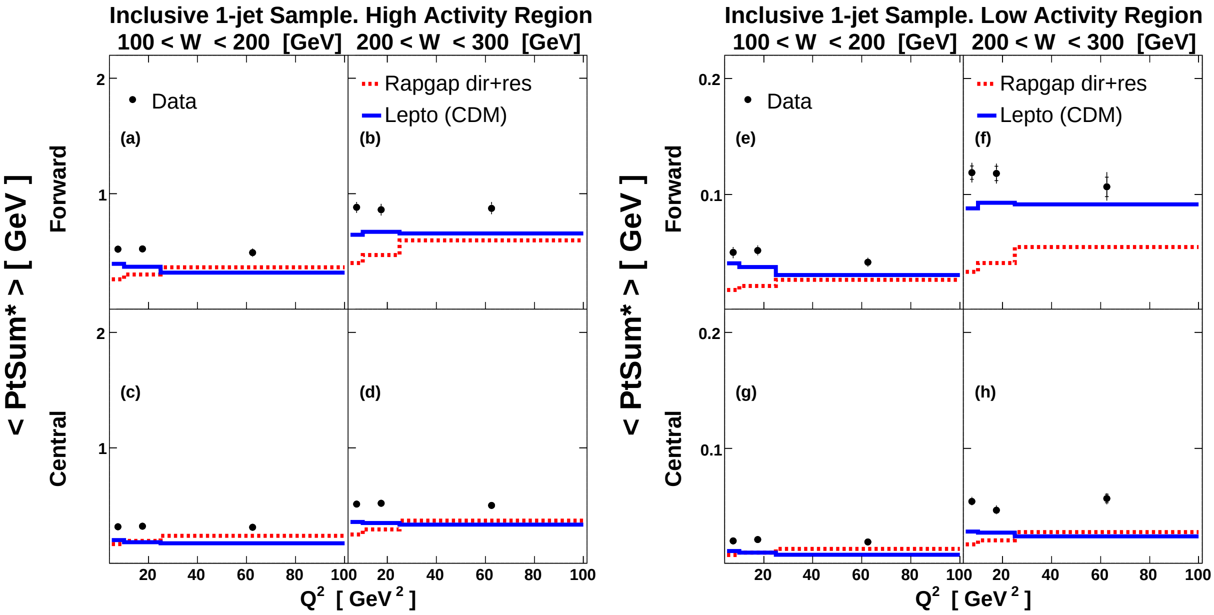


Figure 5.46: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the high and low activity regions in two bins of W. Data are compared to the predictions of Lepto(CDM) and Rapgap dir+res.*

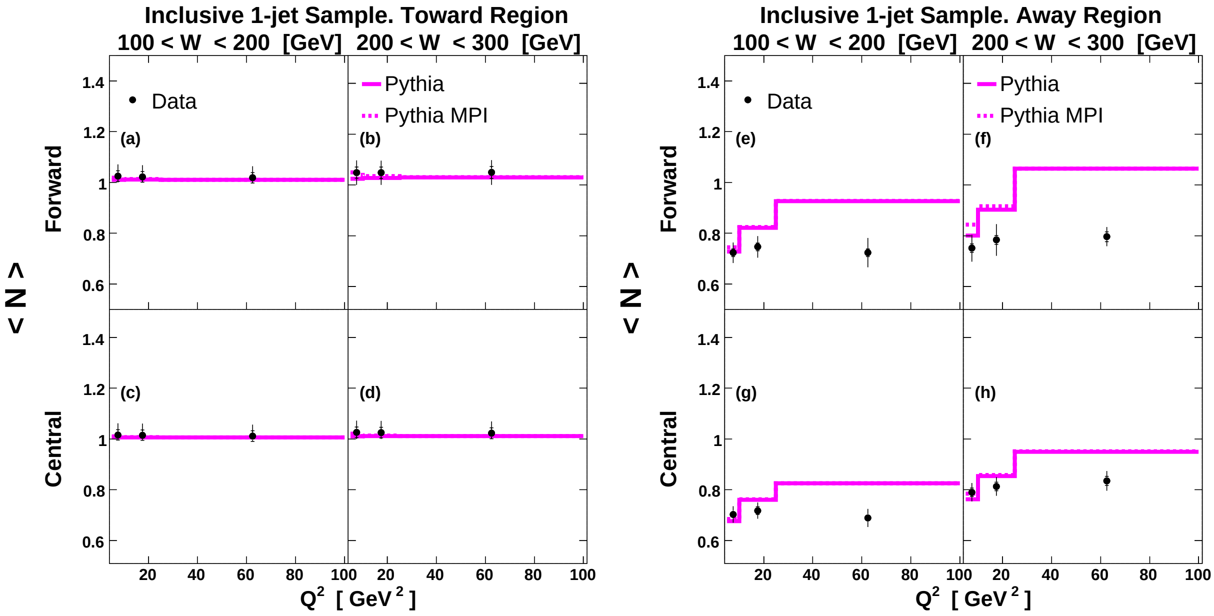


Figure 5.47: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Pythia and Pythia MPI.*

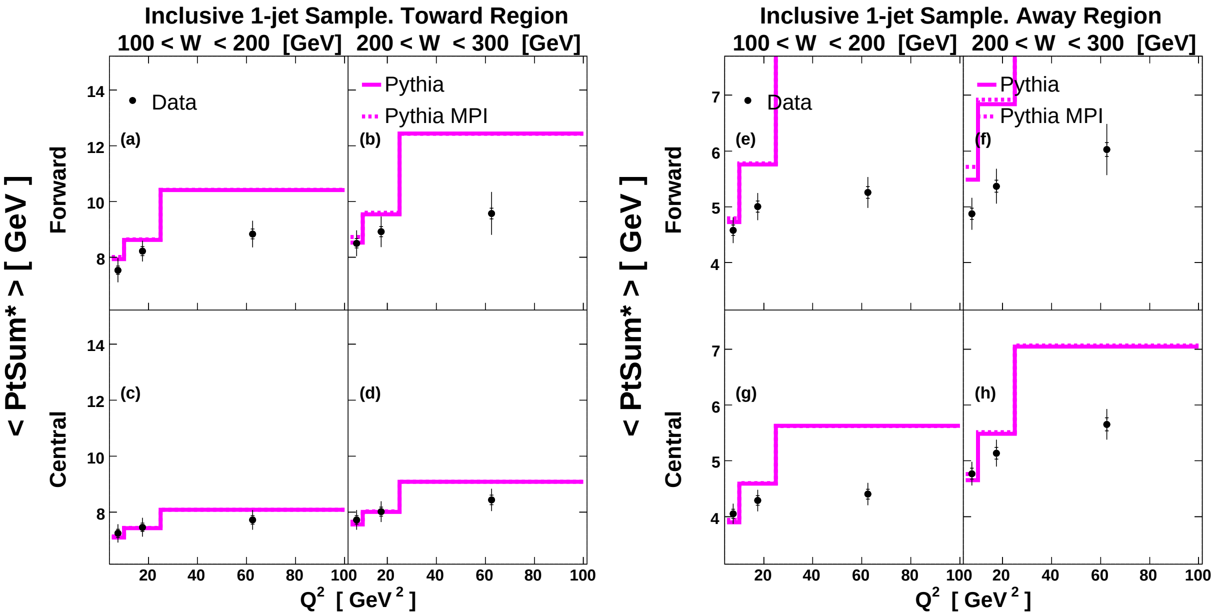


Figure 5.48: *The average of the total transverse momentum of mini-jets, $\langle PtSum^{*} \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. . Data are compared to the predictions of Pythia and Pythia MPI.*

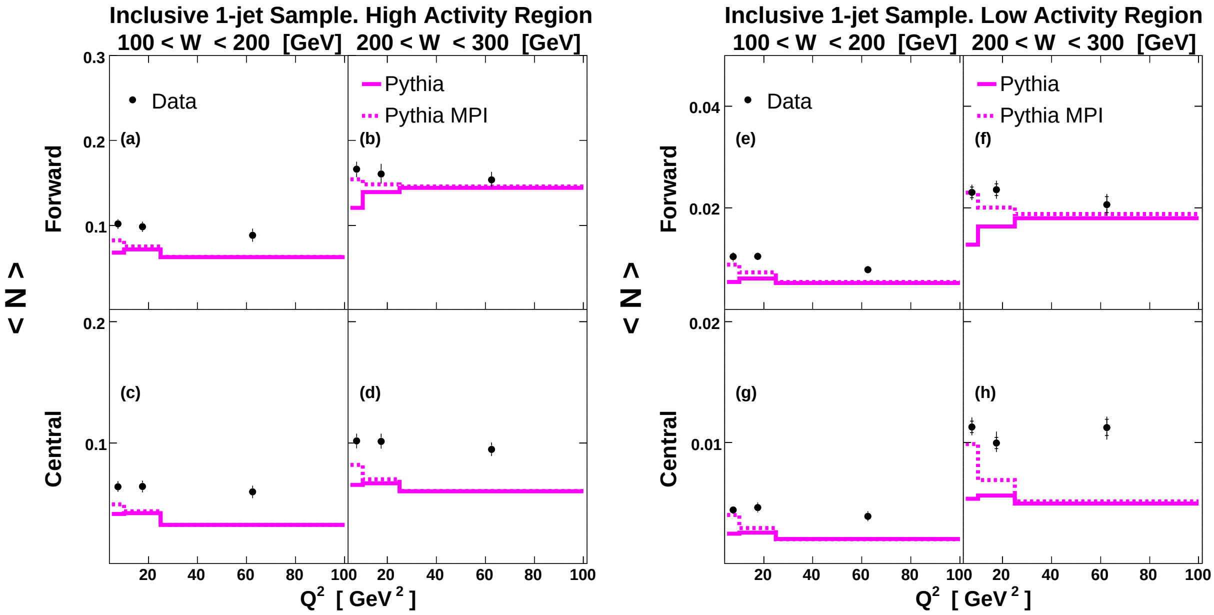


Figure 5.49: *The average multiplicity of mini-jets,* $\langle N \rangle$*, as a function of* Q^2 *for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of* W*. Data are compared to the predictions of Pythia and Pythia MPI.*

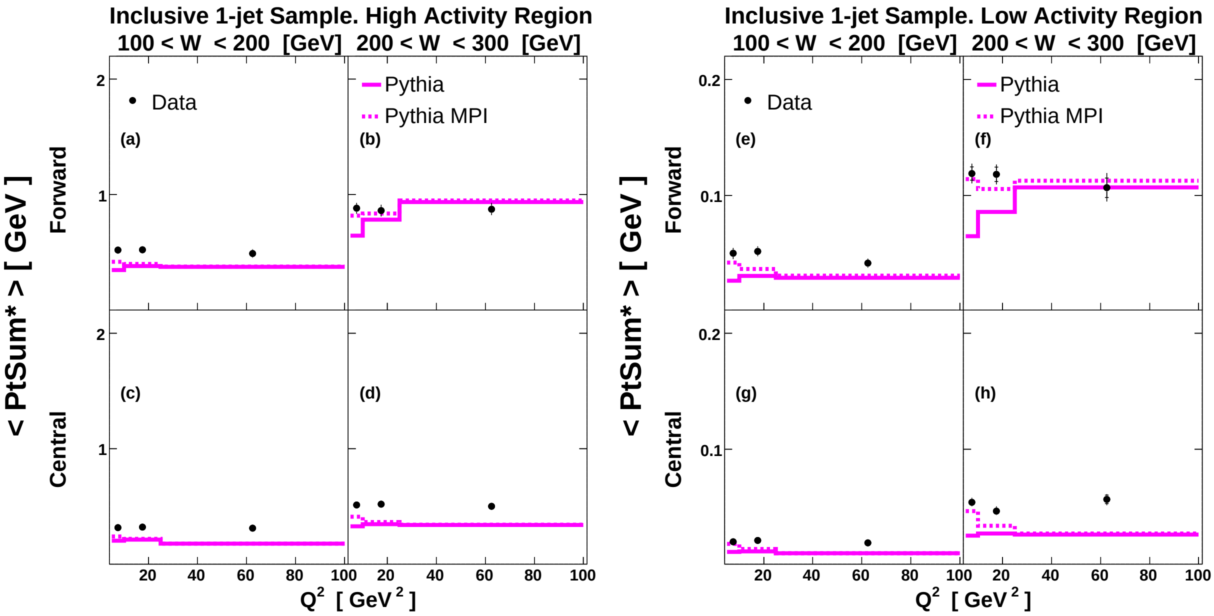


Figure 5.50: *The average of the total transverse momentum of mini-jets,* $\langle PtSum^* \rangle$ *as a function of* Q^2 *for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of* W*. Data are compared to the predictions of Pythia and Pythia MPI.*

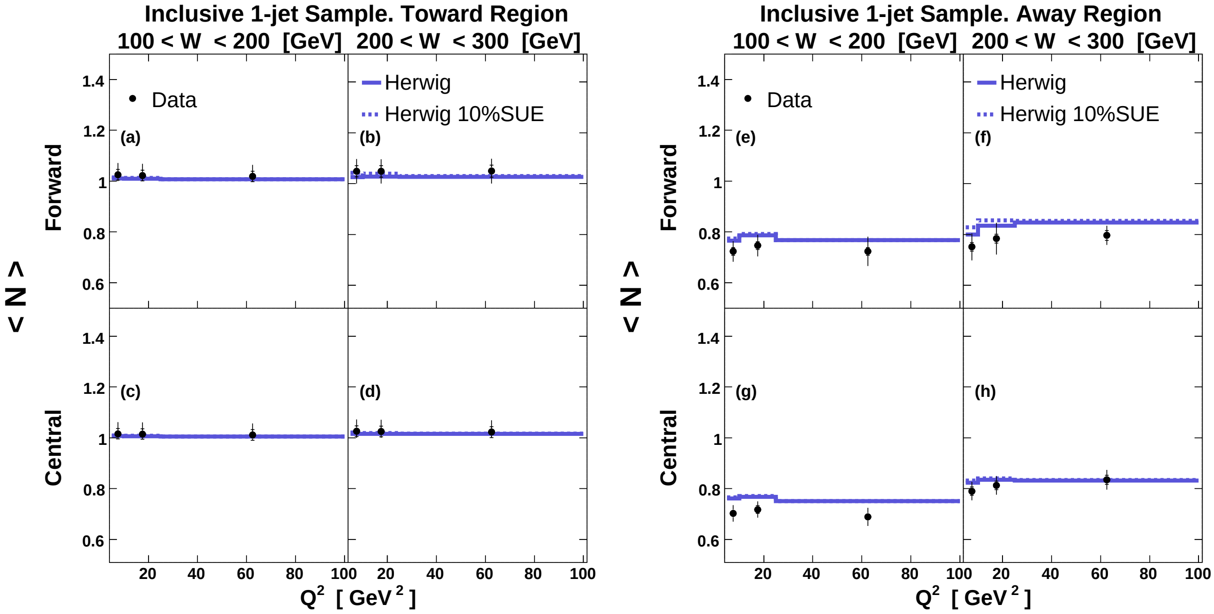


Figure 5.51: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Herwig and Herwig 10%SUE.*

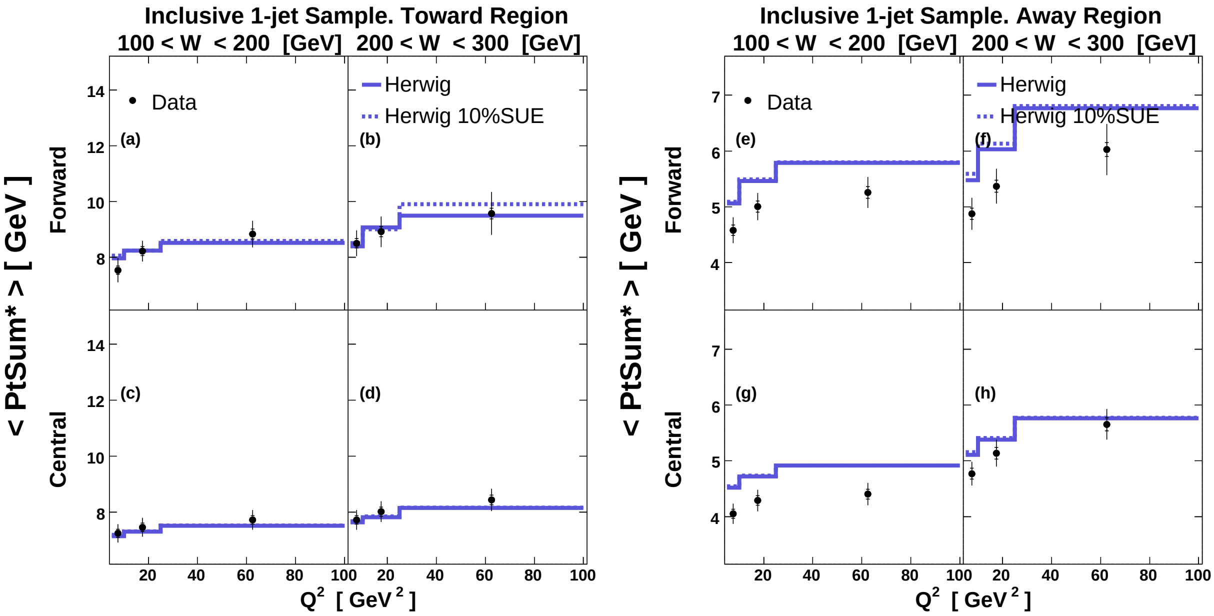


Figure 5.52: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. . Data are compared to the predictions of Herwig and Herwig 10%SUE.*

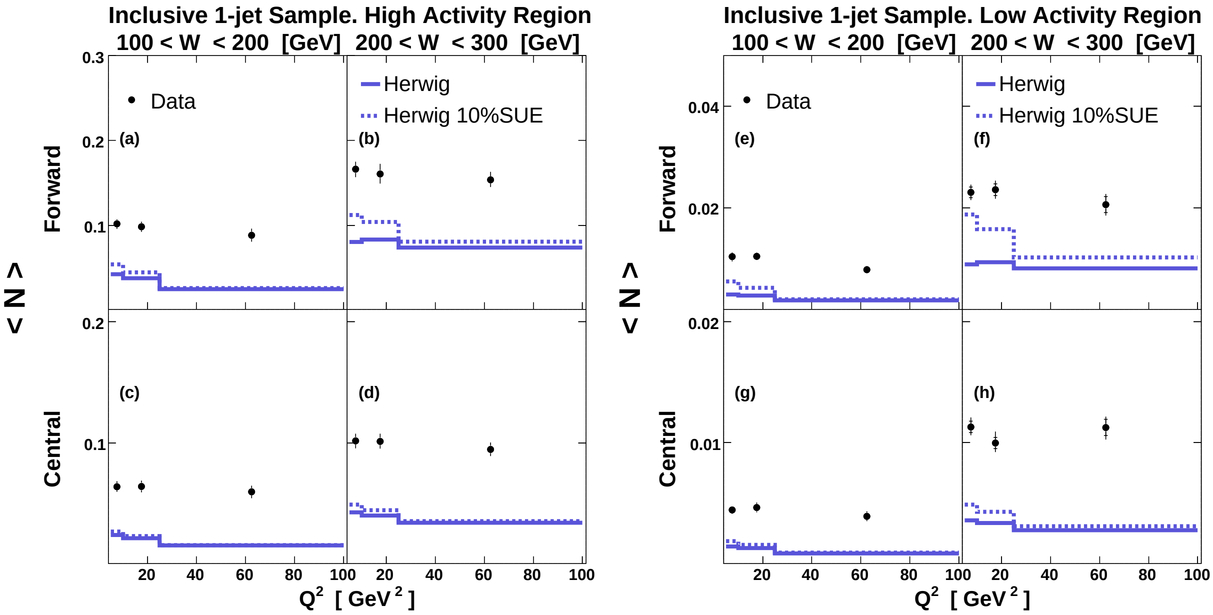


Figure 5.53: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Herwig and Herwig 10%SUE.*

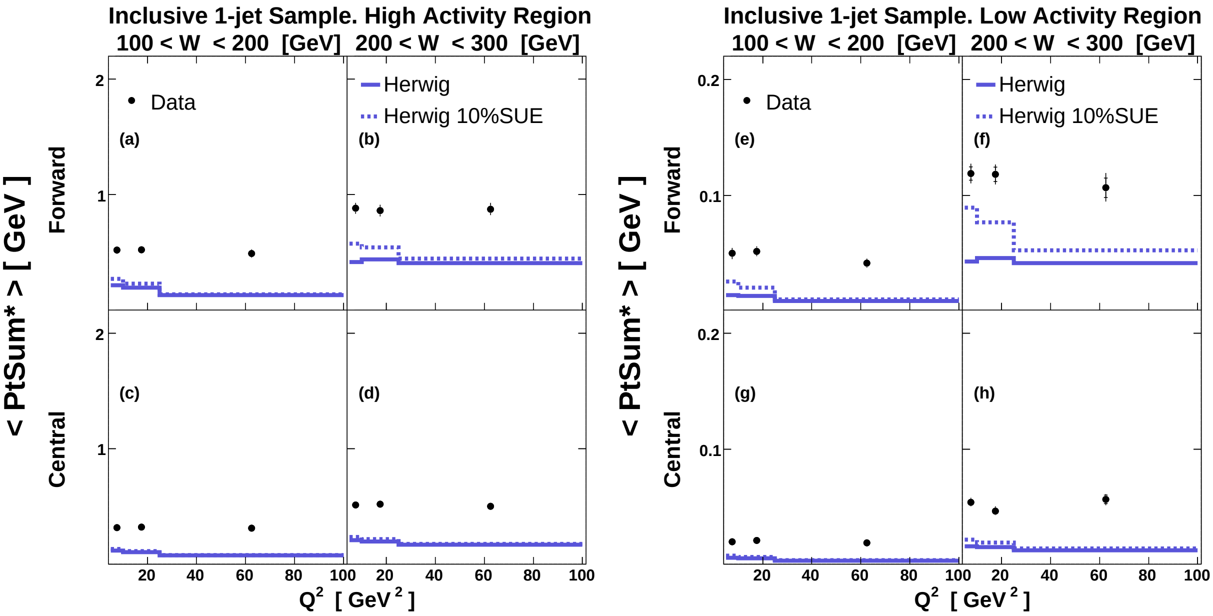


Figure 5.54: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Herwig and Herwig 10%SUE.*

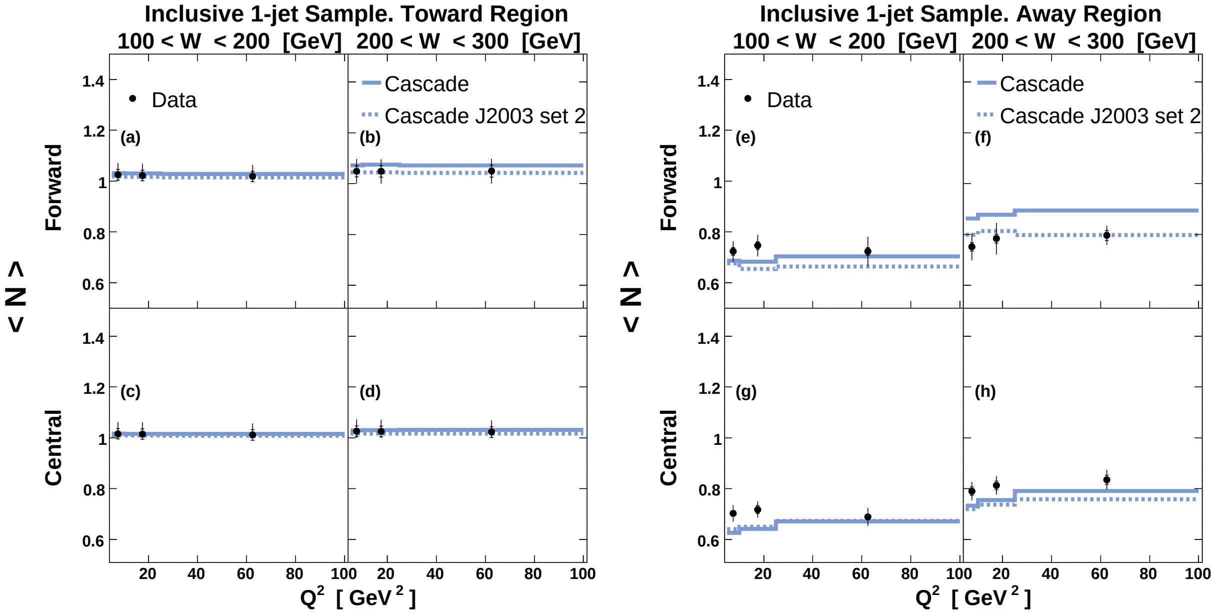


Figure 5.55: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Cascade and Cascade J2003 set2.*

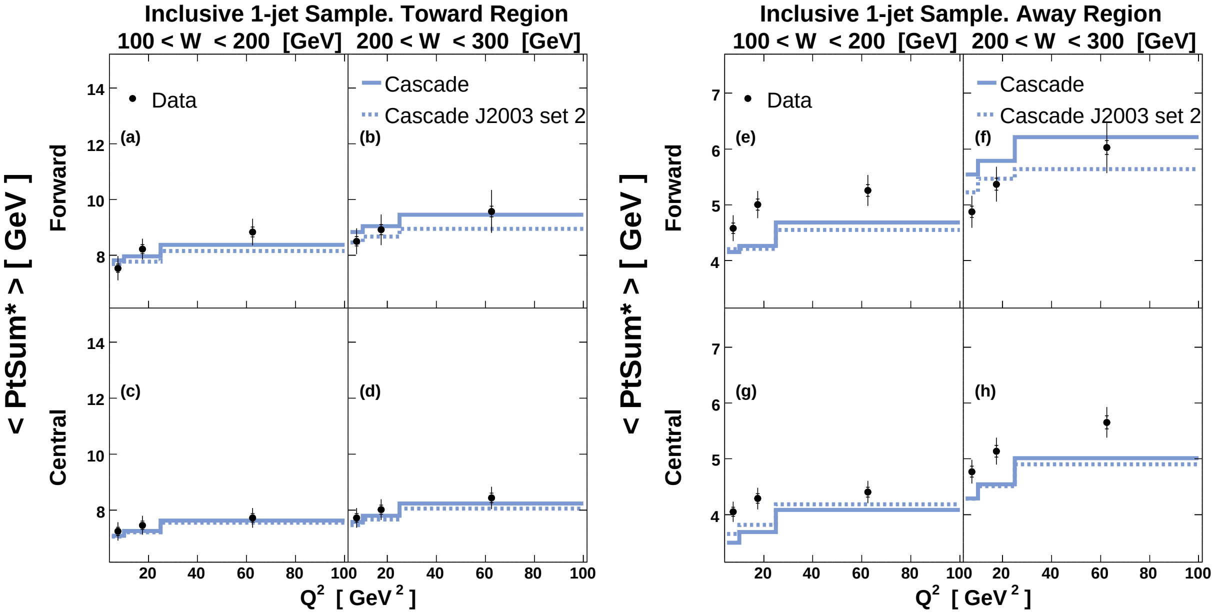


Figure 5.56: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$ as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. . Data are compared to the predictions of Cascade and Cascade J2003 set2.*

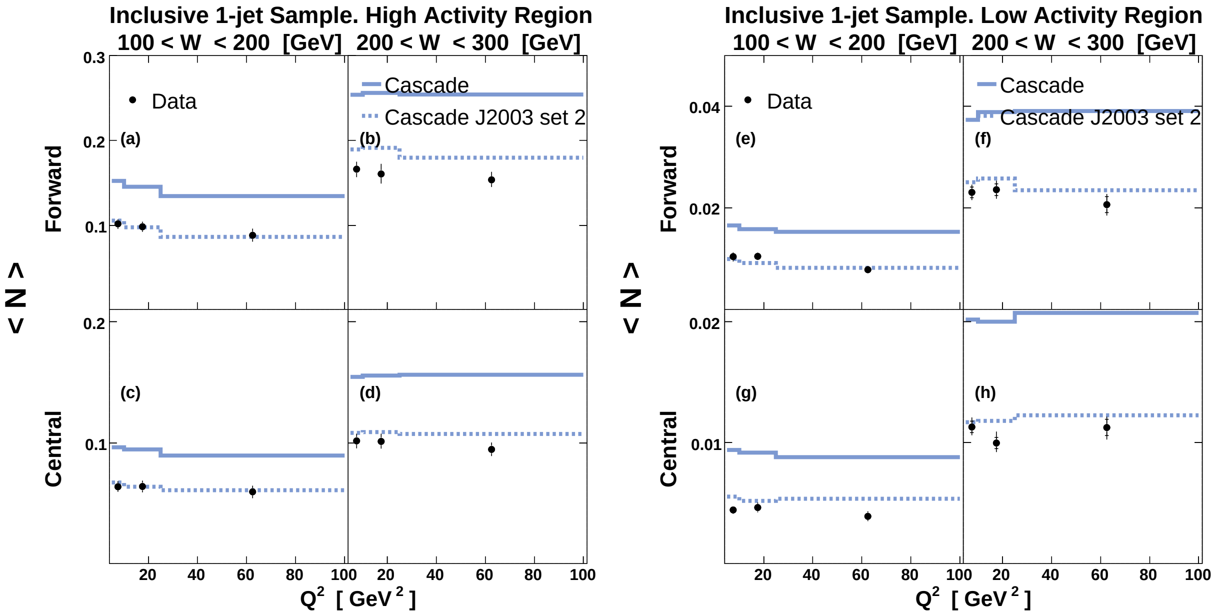


Figure 5.57: *The average multiplicity of mini-jets,* $\langle N \rangle$*, as a function of* Q^2 *for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of* W*. Data are compared to the predictions of Cascade and Cascade J2003 set2.*

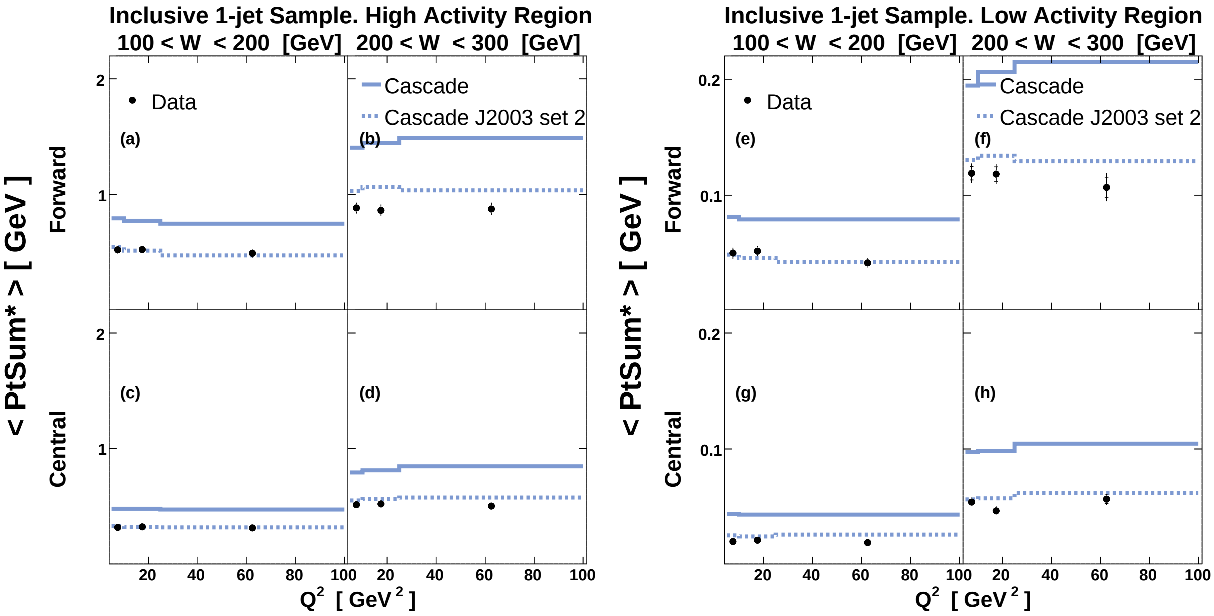


Figure 5.58: *The average of the total transverse momentum of mini-jets,* $\langle PtSum^* \rangle$ *as a function of* Q^2 *for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of* W*. Data are compared to the predictions of Cascade and Cascade J2003 set2.*

5.5.2 The Inclusive 2-jet Sample

For the inclusive 2-jet sample, results on $\langle N \rangle$ and $\langle PtSum^* \rangle$ are presented for two regions in x_γ, where $x_\gamma > 0.7$ is dominated by direct photon interactions and $x_\gamma < 0.7$ by resolved photon scattering. The same two regions in W as for the inclusive 1-jet sample are covered. Comparisons of data to the predictions of all models for the toward and away regions are shown in Figures 5.59-5.66. Except for some deviations by Rapgap dir+res and Lepto(CDM), mainly for resolve photon processes, the agreement for all other models is excellent in these regions. Consistent with the observation from the inclusive 1-jet sample, the influence of MPI and SUE in these regions is very small.

In the Figures 5.67-5.74 comparisons between data and all models are provided for the transverse regions. From Figures 5.67 and 5.68 the predictions on $\langle N \rangle$ and $\langle PtSum^* \rangle$ of Rapgap dir+res and Lepto(CDM) it is seen that direct photon processes ($x_\gamma > 0.7$) are described well by both MC models. Data in the resolved photon region are substantially higher than predicted by the models, although Lepto(CDM) is closer to data than Rapgap dir+res.

The results of Phythia in the transverse regions are compared to data in Figures 5.69 and 5.70. It is noticable that Phythia is not able to describe the direct photon interactions in all bins, even where the influence of MPI is small. This is consistent with the comparatively bad description of the inclusive 1-jet data by Pythia. The predictions of Pythia clearly fall below data for resolved processes. The inclusion of MPI gives a significant contribution as expected for resolved interactions but it is not enough to provide agreement.

Herwig also fails to describe data in the direct photon region in the same bins as Pythia, which can be seen from Figures 5.71 and 5.72. The Herwig predictions for resolved photons are well below data and the addition of 10% SUE does not drastically change this situation.

The predictions of Cascade and Cascade J2003 set2 in the transverse regions are shown in Figures 5.73 and 5.74 together with experimental data. There is substantially more jet activity in the direct photon region predicted by Cascade compared to data, whereas Cascade J2003 set2 agrees well with data. Also jets produced in resolved processes are predicted to be more frequent by Cascade than what is the case. Cascade J2003 set2 gives prediction in the resolved region for the high activity region, which are somewhat high but gives on the other hand a good description of the low activity region.

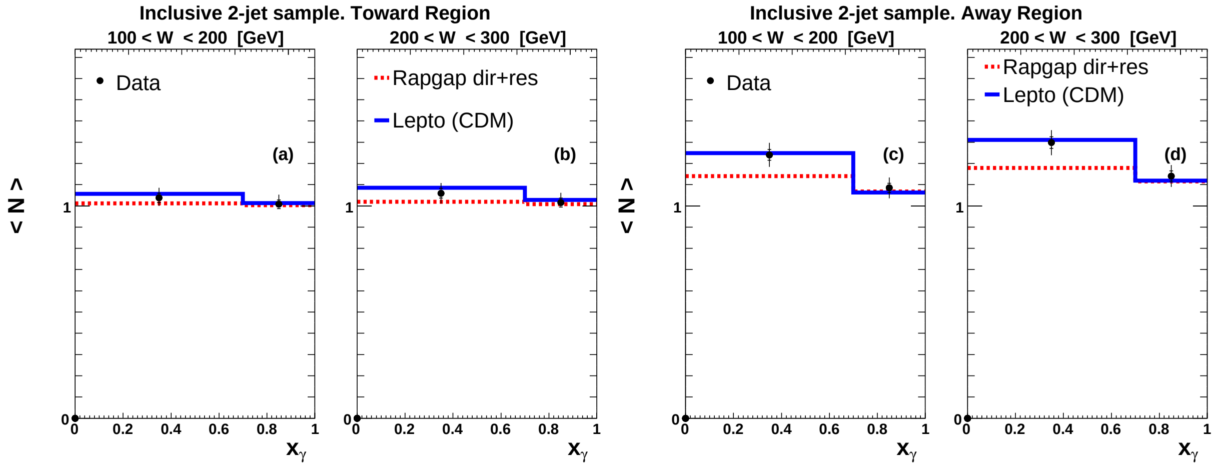


Figure 5.59: *The average multiplicity of mini-jets,* $\langle N \rangle$*, as a function of* x_γ *for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of* W*. Data are compare to the predictions of Lepto(CDM) and Rapgap dir+res.*

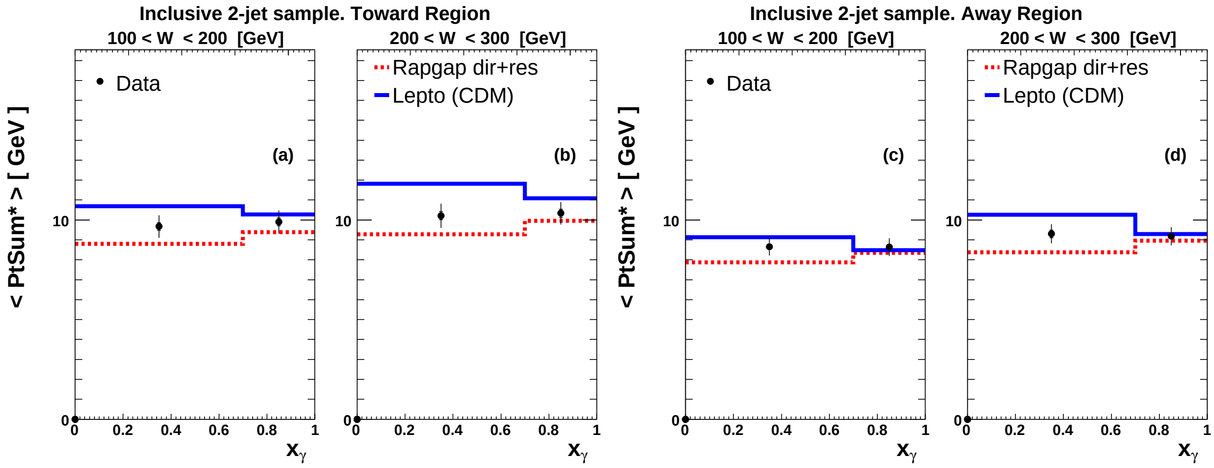


Figure 5.60: *The average of the total transverse momentum of mini-jets,* $\langle PtSum^* \rangle$*, as a function of* x_γ *for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of* W*. Data are compare to the predictions of Lepto(CDM) and Rapgap dir+res.*

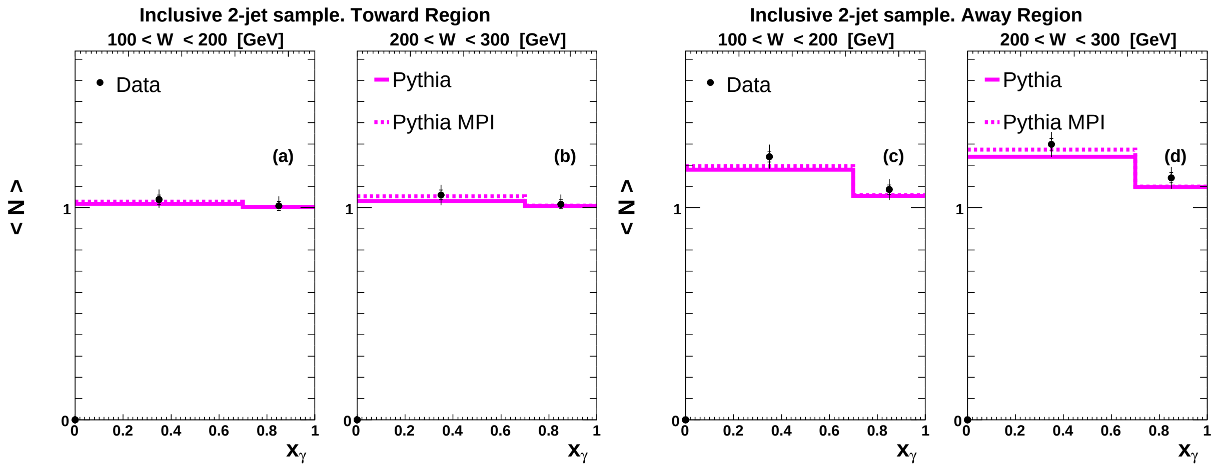


Figure 5.61: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Pythia and Pythia MPI.*

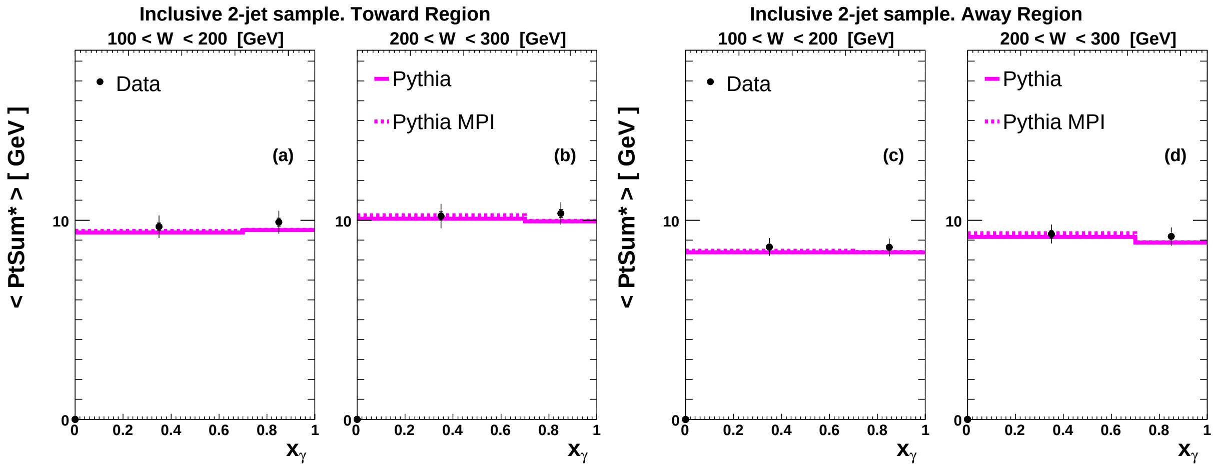


Figure 5.62: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Pythia and Pythia MPI.*

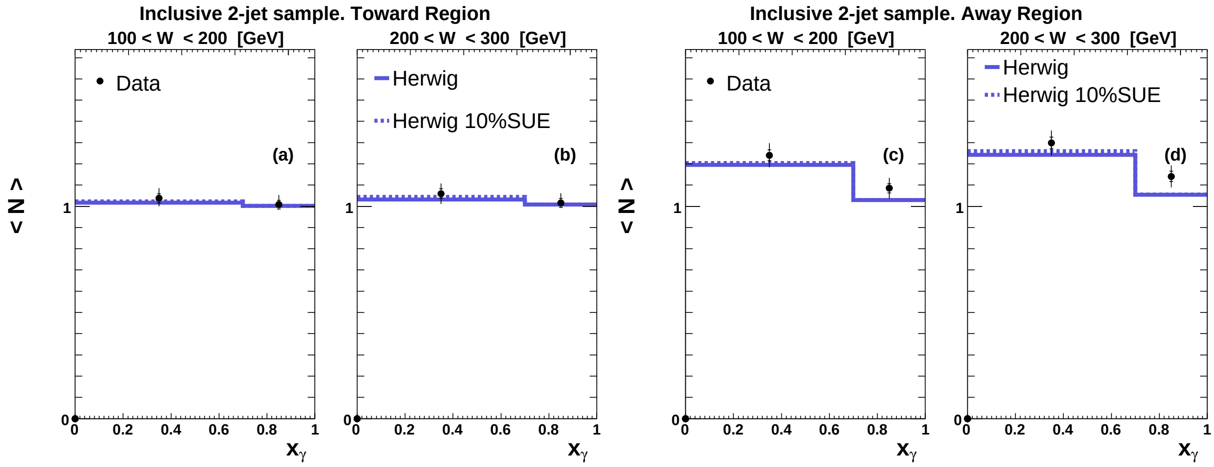


Figure 5.63: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Herwig and Herwig 10%SUE.*

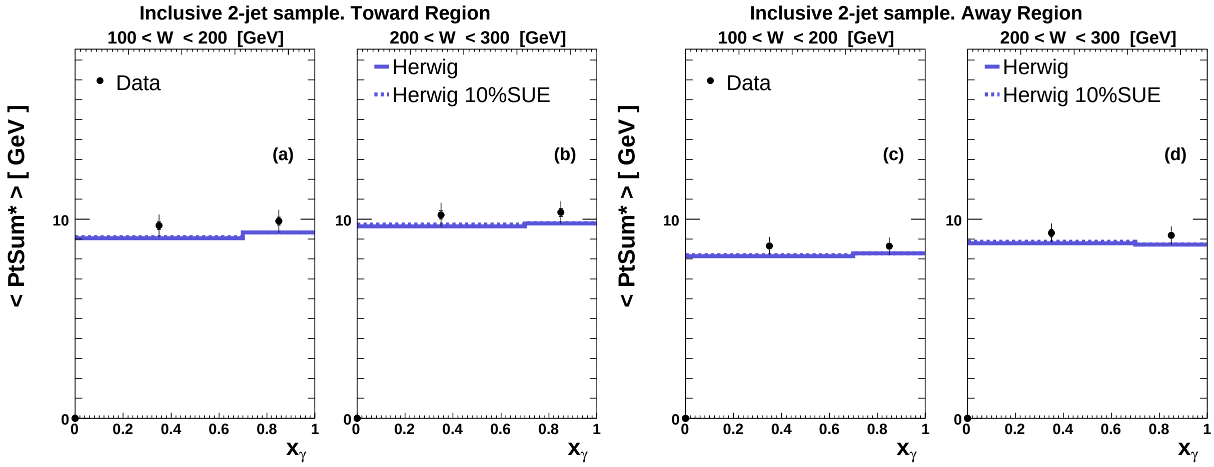


Figure 5.64: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Herwig and Herwig 10%SUE.*

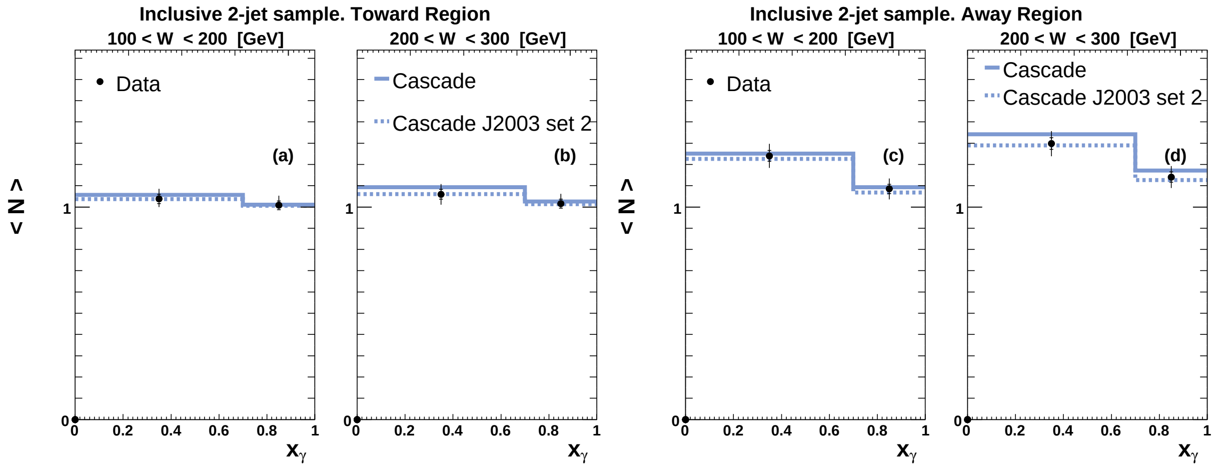


Figure 5.65: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Cascade and Cascade J2003 set2.*

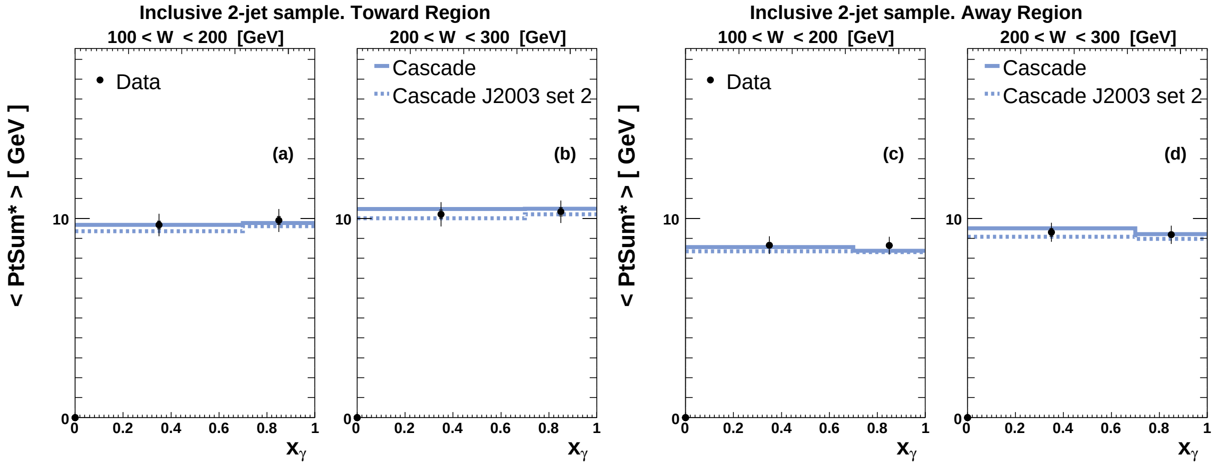


Figure 5.66: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the toward and away regions in two bins of W. Data are compare to the predictions of Cascade and Cascade J2003 set2.*

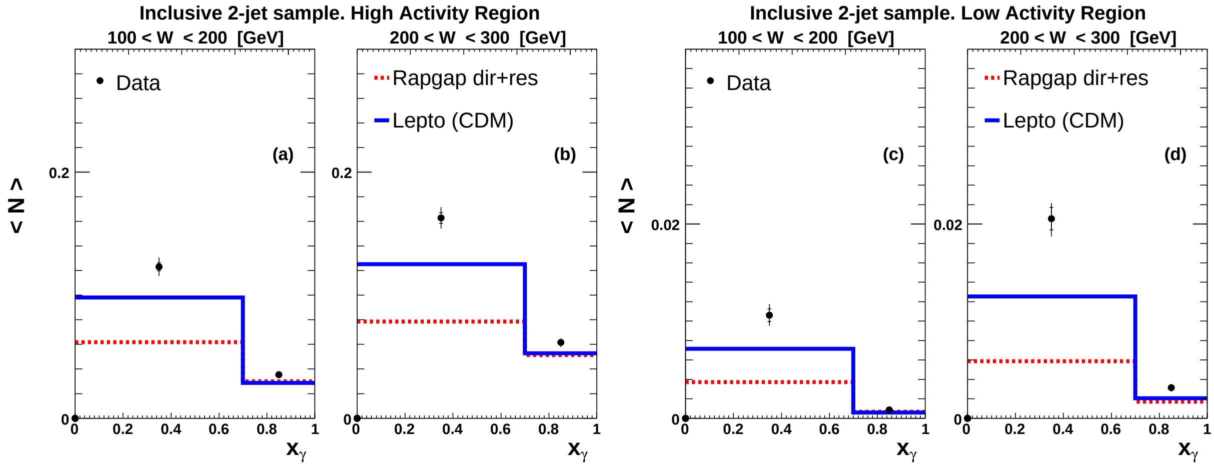


Figure 5.67: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Lepto(CDM) and Rapgap dir+res.*

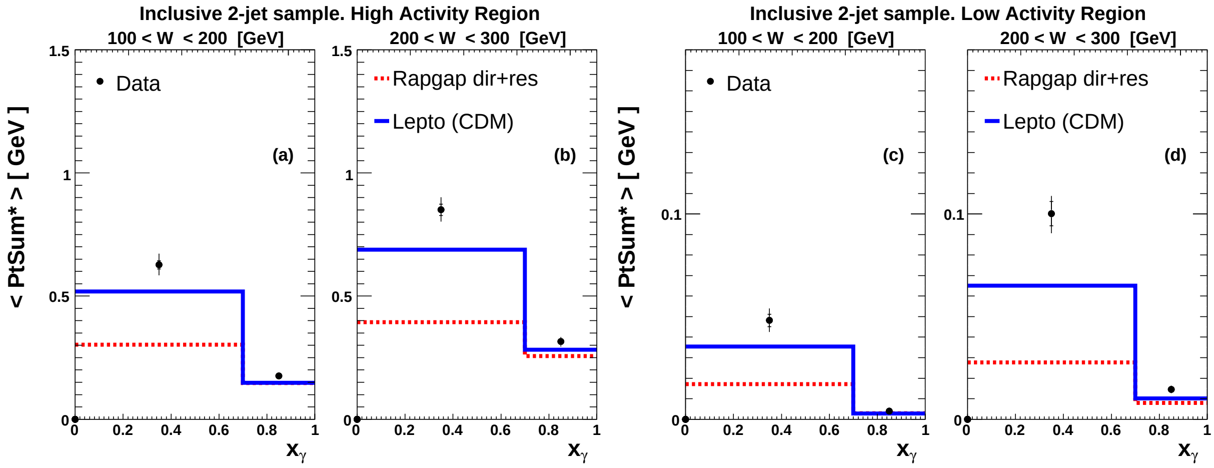


Figure 5.68: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Lepto(CDM) and Rapgap dir+res.*

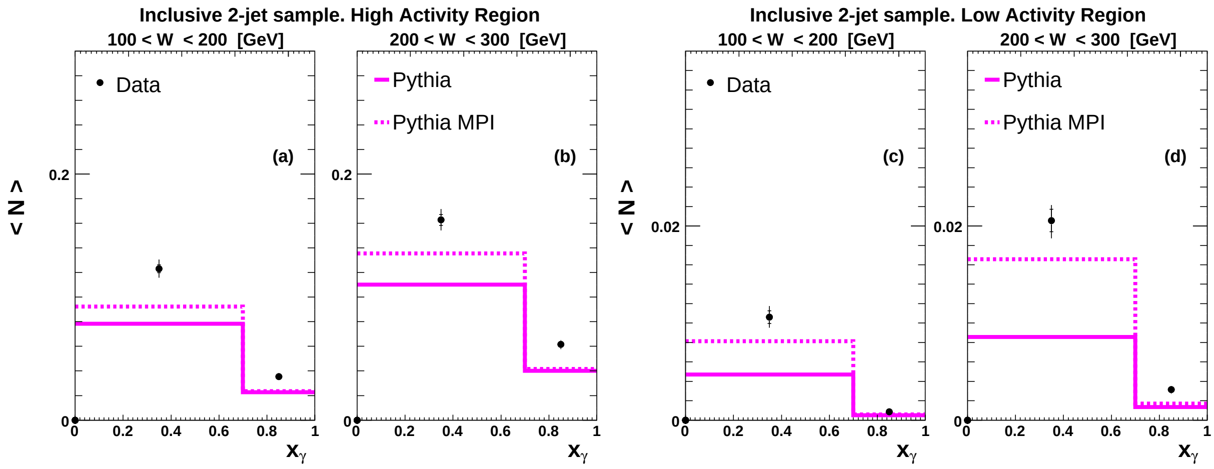


Figure 5.69: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Pythia and Pythia MPI.*

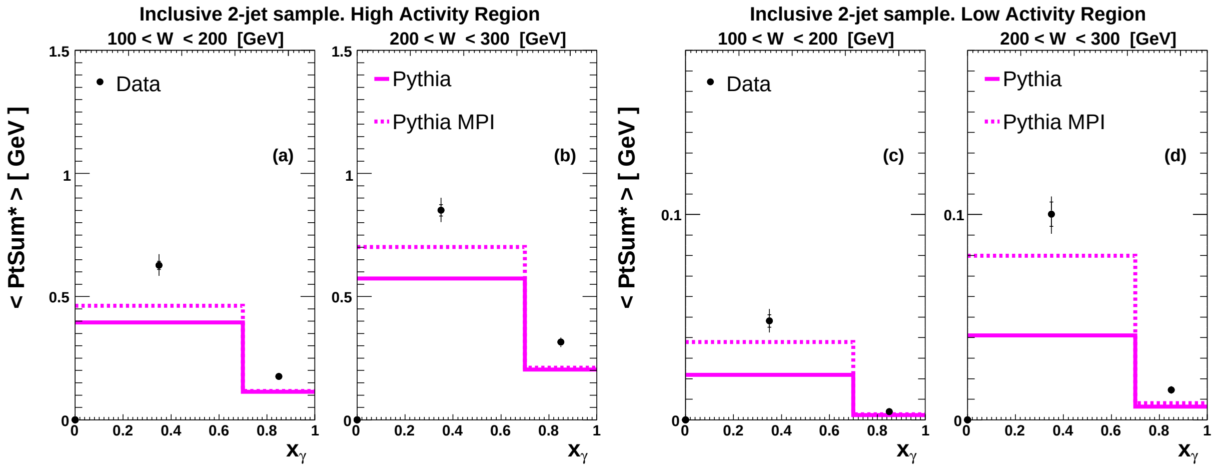


Figure 5.70: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Pythia and Pythia MPI.*

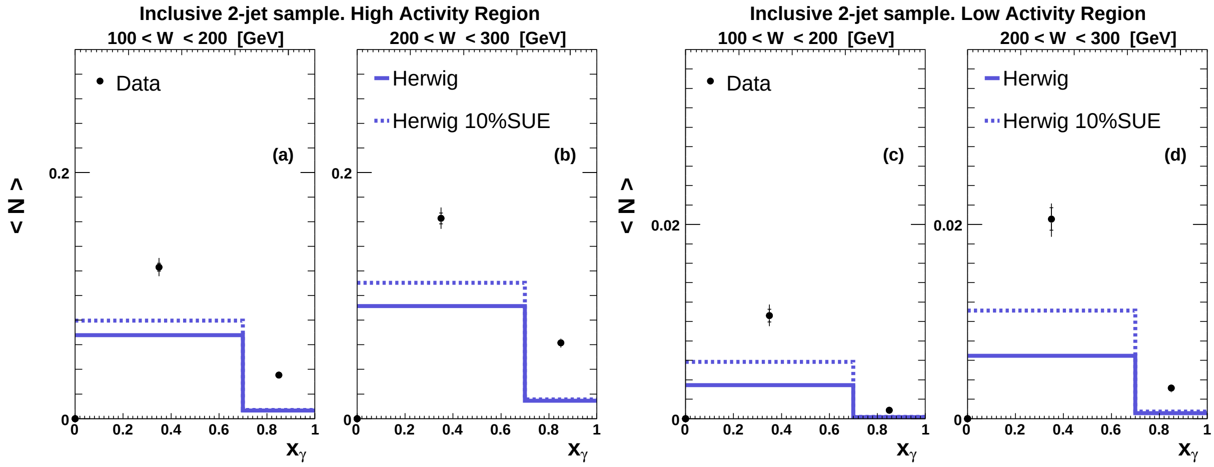


Figure 5.71: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Herwig and Herwig 10%SUE.*

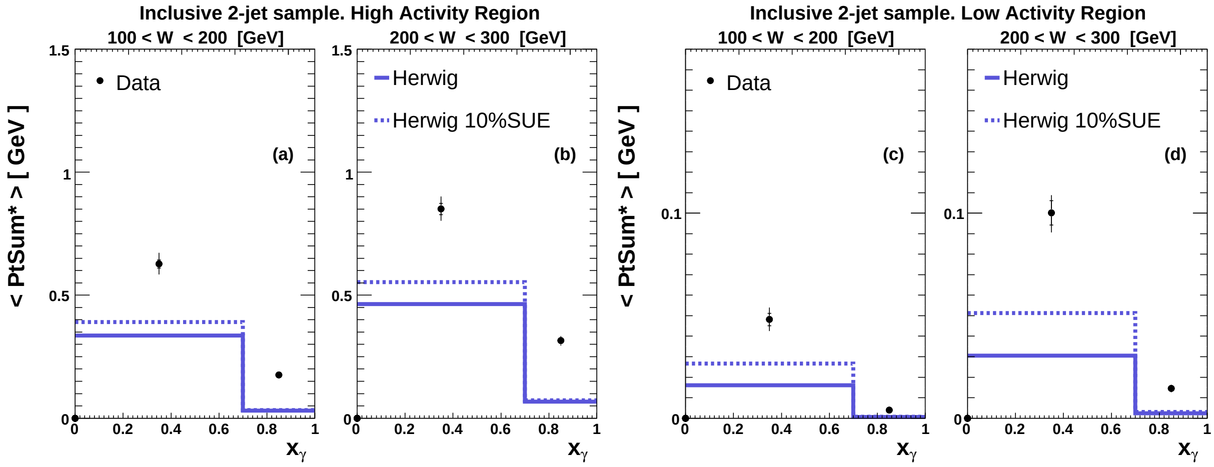


Figure 5.72: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Herwig and Herwig 10%SUE.*

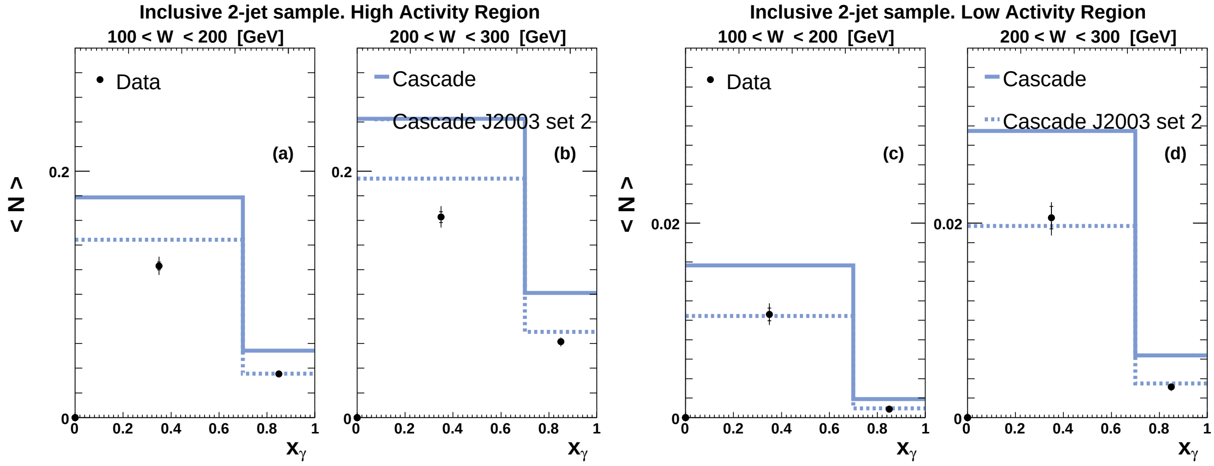


Figure 5.73: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Cascade and Cascade J2003 set2.*

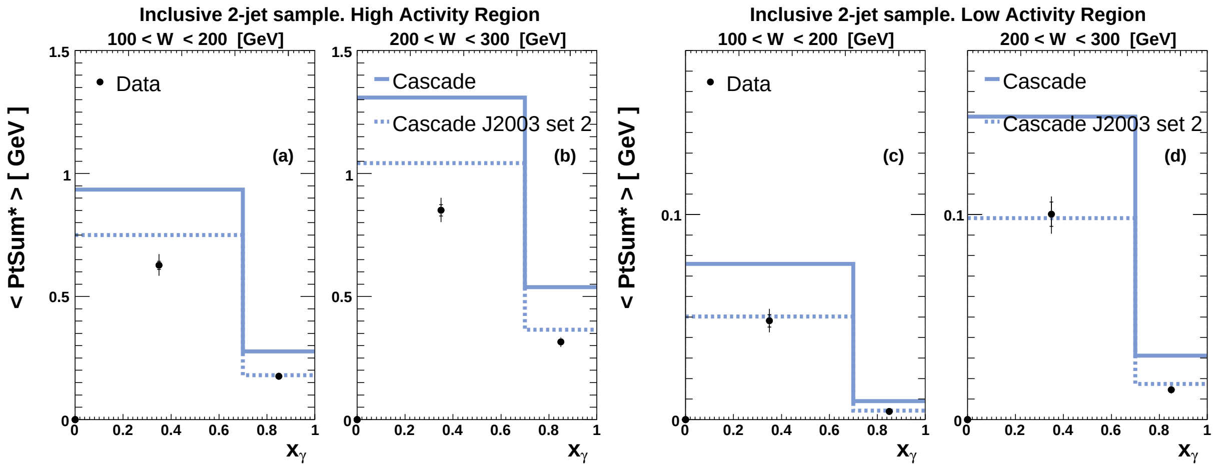


Figure 5.74: *The average of the total transverse momentum of mini-jets, $\langle PtSum^* \rangle$, as a function of x_γ for the inclsuive 2-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compare to the predictions of Cascade and Cascade J2003 set2.*

5.6 Interpretation of the Results

The philosophy of the analysis is to define regions in azimuth such that the hard scattering products should essentially be contained in the toward and away regions. Thereby a possible signal of MPI should be visible in the transverse regions. However, as already discussed higher order processes of the hard scattering may produce mini-jets, which contaminate the transverse regions. Such contributions are expected to be most prominent in the high activity region. Model predictions of MPI are thus influenced by the way mini-jets are generated by higher order processes in the primary interaction. Since calculations can only be performed to finite orders, the uncertainties related to the higher order corrections, not included in the calculations, are estimated by variations in the renormalization and factorization scales. The dependence on the scales of the mini-jet production from the hard scattering are studied in section 5.6.1.

Another source of uncertainty is the choice of parton density functions describing the structure of the proton and the photon. The PDF of the proton is determined to a high degree of accuracy in the kinematic region covered by this analysis, and its uncertainty is not expected to produce a large effect on the results of MPI. The uncertainty of the photon PDF, on the other hand, is much larger and therefore two different parametrisations have been used in Pythia to study their influence on MPI. Some typical results from this study are presented in section 5.6.2.

A discussion and study of mini-jet production from MPI in Pythia and SUE in Herwig is presented in section 5.6.3. Especially the different properties of mini-jets produced by Pythia and Herwig are investigated.

In order to interpret the results as correctly as possible in the light of the predictions by the various models, basic event variables have been investigated in more detail. Thus, it was studied how the different models describe multiplicity distributions of charged particles and their P_T spectra in section 5.6.4.

Finally, some possible interpretations of the predictions delivered by the Monte Carlo generators are discussed in section 5.6.4. The sensitivity of the data to parton showering models is discussed.

5.6.1 Scale Dependence

The choice of renormalization scale, μ_r, and factorization scale, μ_f, is not defined by first principles in QCD. In DIS the choice is especially difficult since the scattering process involves several scales. There are theoretical motivations for using different scales like for example Q^2 or P_T of the hardest jet, or some combination of the two. A change of the scale is equivalent to a change in the coefficient of the higher order terms in the perturbative expansion. The scale dependence is reduced by including more and more of the higher order terms. As it is not obvious what parameter should be used as the scale, it is also not obvious over which range it should be varied in order to account for the uncertainties from higher order corrections. A general consensus has, however, been developed that the scale should be varied by a factor of 2 up and a factor 1/2 down. Pyhtia uses the mass squared of the propagating partons in the ladder as scale for both the renormalization and factorization.

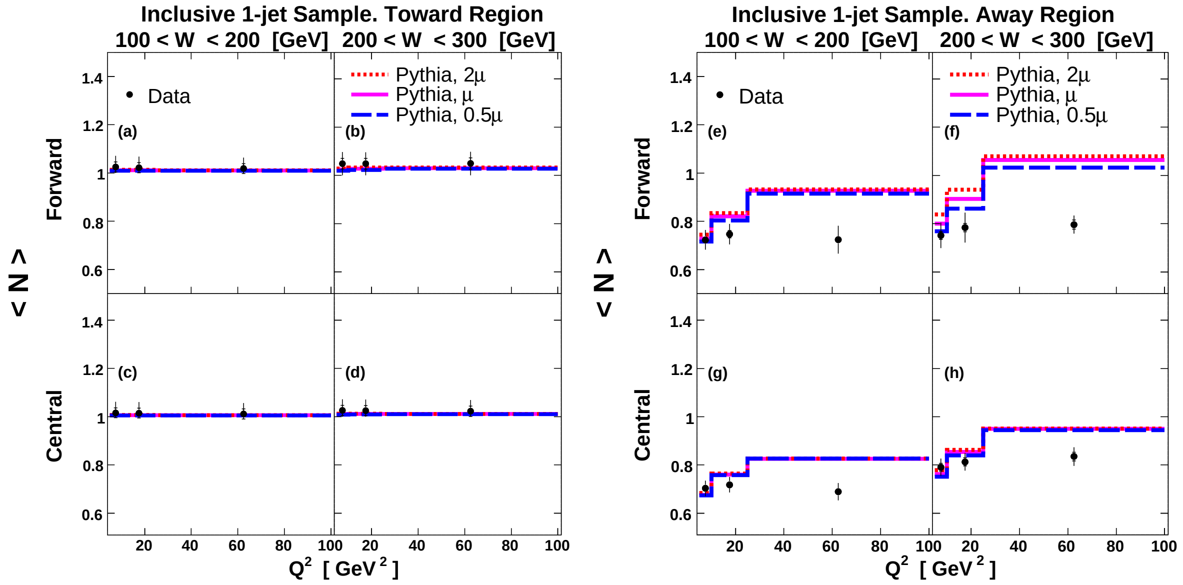


Figure 5.75: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the toward and away regions in two bins of W. Data are compared to the predictions of Pythia using three different values of the evolution scale, 2μ (dotted line), μ (solid line) and 0.5μ (broken line).*

In Figure 5.75 the mini-jet multiplicity is shown for the 1-jet inclusive sample in the toward and away regions with the leading jet forward and central, respectively, and for low and high values of W. As observed from the figure the predictions of the mini-jet production in the toward and away regions are not sensitive to the choice of the scale. The largest dependence of about 5% is observed when the leading jet is going in the forward region and at high values of W.

The predictions of Pythia on mini-jets in the low and high activity regions, for the same sample as above, are shown in Figure 5.76. The dependence on the scale is significant, especially in the regions of high W.

The scale dependence of the mini-jet activity of the inclusive 2-jet sample is shown in Figure 5.77 for the low and high activity regions in two bins of x_γ. A scale dependence of about 10% is observed.

It is clear from the results obtained that the scale dependence is not large enough to account for the deviation of the Pythia predictions from data.

5.6.2 Photon PDF Dependence

Data have been compared to the predictions of Pythia using two different parametrizations of the photon PDF. In Figure 5.78 the mini-jet multiplicity of the inclusive 1-jet sample is shown for the low and high activity regions, where the leading jet is in the forward and central regions and for low and high W values. The solid line represents the photon PDF parametrization of SaS1D, whereas the dotted line corresponds to the Drees-Grassie [115]

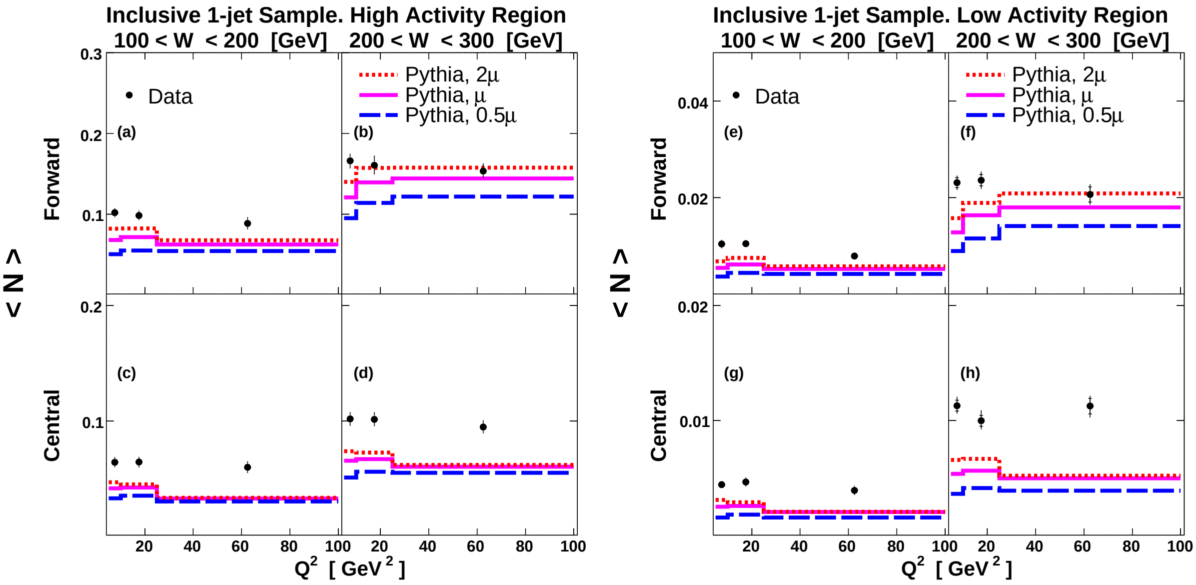


Figure 5.76: *The average multiplicity of mini-jets,* $\langle N \rangle$*, as a function of* Q^2 *for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the high and low activity regions in two bins of* W*. Data are compared to the predictions of Pythia using three different values of the evolution scale,* 2μ *(dotted line),* μ *(solid line) and* 0.5μ *(broken line).*

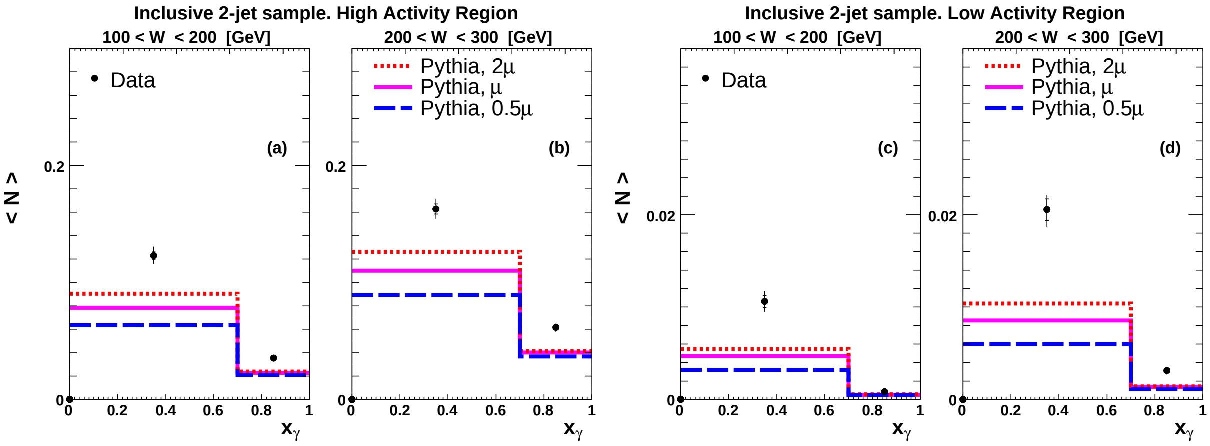
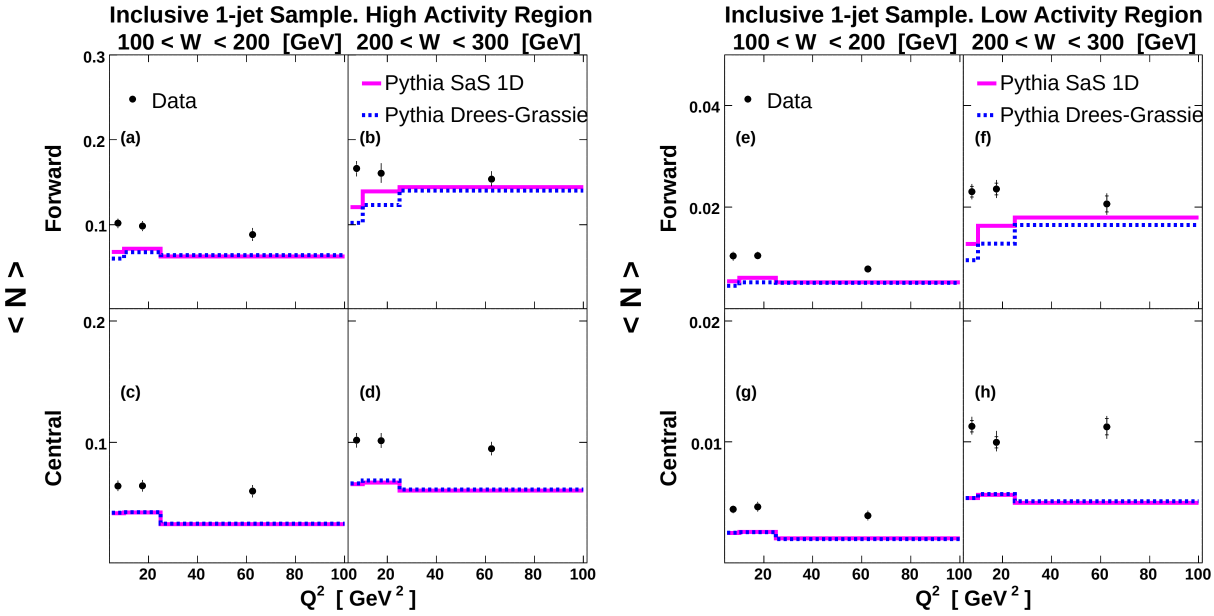


Figure 5.77: *The average multiplicity of mini-jets,* $\langle N \rangle$*, as a function of* x_γ *for the inclusive 2-jet sample, shown separately for the high and low activity regions in two bins of* W*. Data are compared to the predictions of Pythia using three different values of the evolution scale,* 2μ *(dotted line),* μ *(solid line) and* 0.5μ *(broken line).*


Inclusive 1-jet Sample. High Activity Region

Inclusive 1-jet Sample. Low Activity Region

Figure 5.78: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of Q^2 for the inclusive 1-jet sample, with the leading jet being forward and central, respectively, shown separately for the high and low activity regions in two bins of W. Data are compared to the predictions of Pythia using two different photon PDF's, the SAS 1D (solid line) and the Drees-Grassie (dotted line) parametrizations.*

parametrization. As can be seen there is a sensitivity of 10-20 % to the choice of PDF in events with a forward going leading jet and having high W values.

The corresponding results from the inclusive 2-jet sample are shown in Figure 5.79. In this case the uncertainty due to the choice of the photon PDF on the mini-jet activity is about 20 % in the low x_γ bin.

It is obvious that the failure of Pythia to reproduce the data is not due to the uncertainty in the description of the photon PDF.

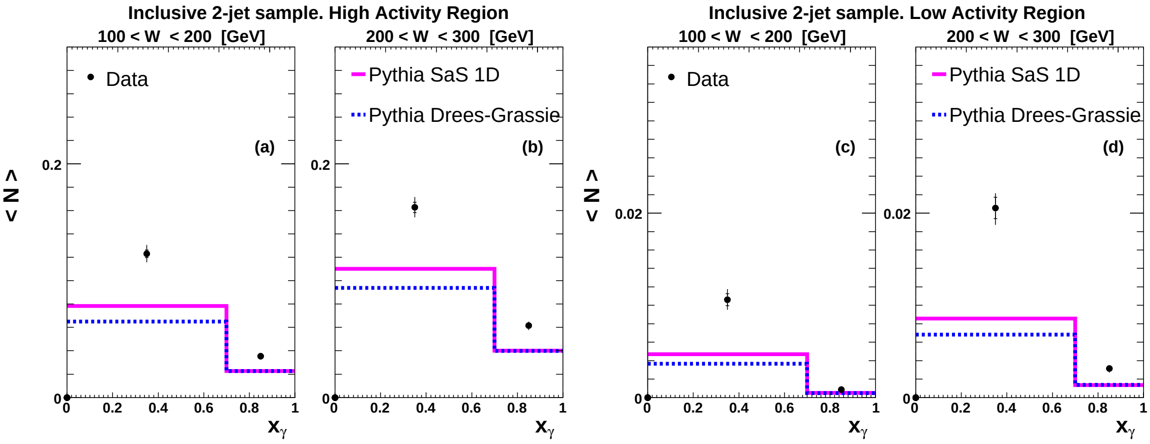


Figure 5.79: *The average multiplicity of mini-jets, $\langle N \rangle$, as a function of x_γ for the inclusive 1-jet sample, shown separately for the high and low activity regions in two bins of W. Data are compared to the predictions of Pythia using two different photon PDF's, the SAS 1D (solid line) and the Drees-Grassie (dotted line) parametrizations.*

5.6.3 Mini-jet Production in Pythia and Herwig

The fact that the inclusion of SUE to Herwig gives rise to additional jets in the transverse regions might be surprising, since an isotropic emission of soft particles should in principle not produce jets. In order to better understand this an investigation and comparison of the properties of mini-jets in the transverse regions for the inclusive 1-jet sample reconstructed from events generated by Herwig, Herwig 10%SUE, Pythia and Pythia MPI, respectively. An interesting and obvious question in this context is whether mini-jets produced in Herwig by adding SUE to the events can be rejected by making appropriate cuts in the data sample.

Figure 5.80 shows the rapidity distribution and transverse momentum spectrum of mini-jets reconstructed from the MC data in the azimuthal regions transverse to the leading jet. The rapidity distribution clearly shows that SUE mainly contribute to mini-jets in the forward (proton) region, which is not so pronounced for MPI in the case of Pythia. From the $P_{T,MiniJets}$ spectrum we can conclude that both the inclusion of SUE and MPI give an increased number of mini-jets over the full $P_{T,MiniJets}$ range of the mini-jets. Thus, an increase of the $P_{T,MiniJets}$ cut in the jet reconstruction is not expected to provide any suppression of SUE in favor of MPI

The next step is to look at the particle composition of the mini-jets, where the particle distributions shown have been normalised to the number of mini-jets in the transverse regions. In Figure 5.81a the particle multiplicity of the mini-jets is shown. It can be observed that mini-jets reconstructed from events generated by Pythia have the same particle multiplicity of about 8 on the average, independent of whether MPI is included or not. Mini-jets reconstructed from events generated by Herwig without SUE give an average particle multiplicity similar to what Pythia gives, but if SUE is included the average particle multiplicity increases by nearly one unit. This means that the jets coming from additional interactions (MPI) have the same particle content as those from the primary interaction, which is not surprising since they are produced in the same way but from different gluon ladders. In contrast to this the additional jets from the inclusion of SUE include more particles. There might be two explanations for this. One explanation for the increased number of mini-jets caused by the SUE might be that mini-jets from initial or final state radiation, which were originally rejected for not exceeding the $P_{T,MiniJets}$ cut of 3.5 GeV, might gain one or more particles from the SUE, which make them fulfill the $P_{T,MiniJets}$ requirement. That is, no additional jets are produced but instead jets from higher order processes might become visible when the jet reconstruction algorithm merges additional particles from SUE into the jet. Another possible explanation could be that the jet algorithm simply forces uncorrelated soft particles into jets. It might also be possible that there are contributions from both sources.

The average multiplicity of the particles in the mini-jets, $\langle N_{PartInJet} \rangle$, as a function of particle transverse momenta are shown in Figure 5.81b. $\langle N_{PartInJet} \rangle$ with and without MPI from Pythia are identical, as expected since the jets produced from one gluon ladder should look the same as those produced by another. For Herwig the $\langle N_{PartInJet} \rangle$ spectrum shows a significant excess of low momentum particles as SUE is included. Thus, SUE mainly contributes particles with transverse momenta lower than 700 MeV. From the spectrum of the total particle momenta, shown in Figure 5.81c, it is evident that the shapes are the same over the full momentum range. This means that the higher the total momentum of a particle from SUE is, the more it has to go into the forward region in order to have a transverse momentum lower than 700 MeV. This is confirmed by Figure 5.81d, which shows the polar angle distribution of $\langle N_{PartInJet} \rangle$ from which it is clearly seen that particles from SUE are peaked in the forward region ($\theta < 40°$).

It is obvious that a cut of $P_{T,PartInJet} > 700$ MeV is most efficient to remove particles produced by the SUE, as seen in Figure 5.81b. On the other hand this is a fairly hard cut, which severely reduces the Pythia mini-jet sample to about 20% of the original. The effects of a cut on the particle momentum at 700 MeV can be further appreciated from Figure 5.82(a-d). The particle multiplicity in the mini-jets is naturally decreased by this cut, as shown in Figure 5.82a, since now only few particles are needed to add up to the $P_{T,MiniJets}$ cut of the jet. As now essentially the full contribution from SUE has been removed, the average particle multiplicity of mini-jets from the samples generated with and without SUE is the same. Moreover, it is the same as the particle multiplicity of mini-jets reconstructed from events generated by Pythia with and without MPI. Also the distributions of $\langle N_{PartInJet} \rangle$ as a function of the total jet particle momenta, shown in Figure 5.82c, agree independently on whether the samples with or without SUE or MPI have been used. Although the polar angle distributions of $\langle N_{PartInJet} \rangle$ are different for Herwig and Pythia, the Herwig distribution no longer contains any contribution from SUE, as seen in Figure 5.82d.

The behavior of the rapidity and the transverse momentum distributions of the mini-jets, after having applied the $P_{T,PartInJet} > 700$ MeV cut on the mini-jet particles, is shown in Figure 5.83a and b. Now, the rapidity distribution of Herwig does not exhibit any extra contribution from SUE, whereas for Pythia there is still a remaining excess of mini-jets from MPI. This becomes more evident from Figure 5.83b where it is clearly seen that there is essentially no contribution to the transverse momentum spectrum of mini-jets from SUE (except for a very small excess in the lowest $P_{T,MiniJets}$ bin). On the other hand, a significant contribution from MPI is noticed up to the highest mini-jet momenta. Due to the fact that the particles from SUE are mainly directed into the forward region, the number of mini-jets in this region gets strongly suppressed from the $P_{T,PartInJet} > 700$ MeV cut. This is not the case for mini-jets from MPI, where the rapidity distribution has the same shape with or without the $P_{T,PartInJet} > 700$ MeV cut.

Although, contributions from MPI can be separated from those of SUE, the required cuts could not be applied due to statistical reasons.

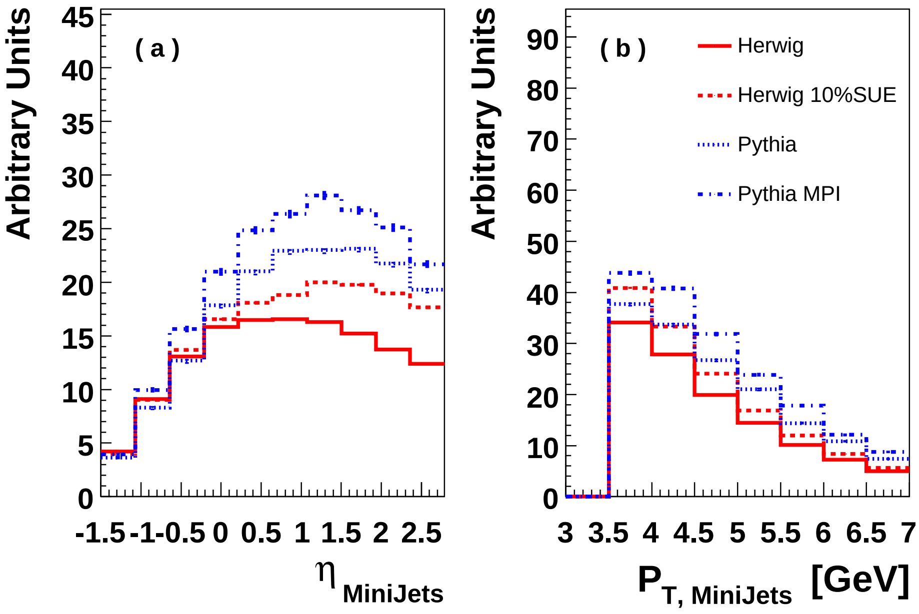


Figure 5.80: *The a)* $\eta_{MiniJets}$ *and b)* $P_{T,MiniJets}$ *distributions of mini-jets in the transverse regions. The predictions of the Monte Carlo generators Herwig, Herwig 10%SUE, Pythia and Pythia MPI.*

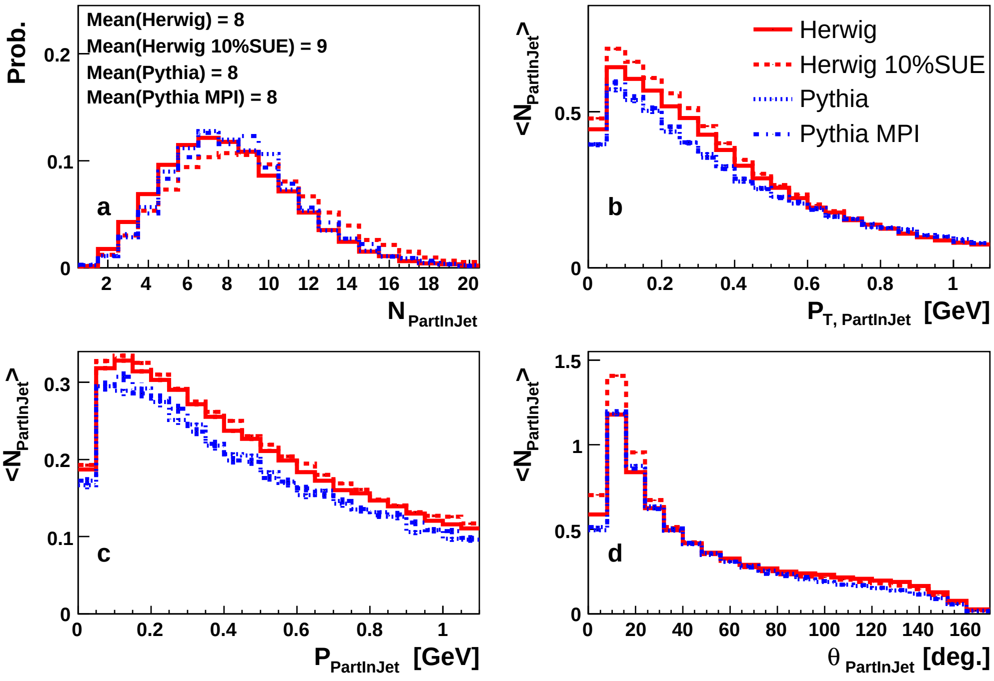


Figure 5.81: *The a) distribution of particle multiplicty in the mini-jets,* $N_{PartInJet}$*, and the average value ,* $\langle N_{PartInJet} \rangle$*, as a function of b) the transverse momentum* ($P_{T,PartInJet}$)*, c) the total momentum* ($P_{PartInJet}$)*, and d) the polar angle* ($\theta_{PartInJet}$) *of the particle jets in the transverse regions. The predictions of the Monte Carlo generators Herwig, Herwig 10%SUE, Pythia, and Pythia MPI are shown.*

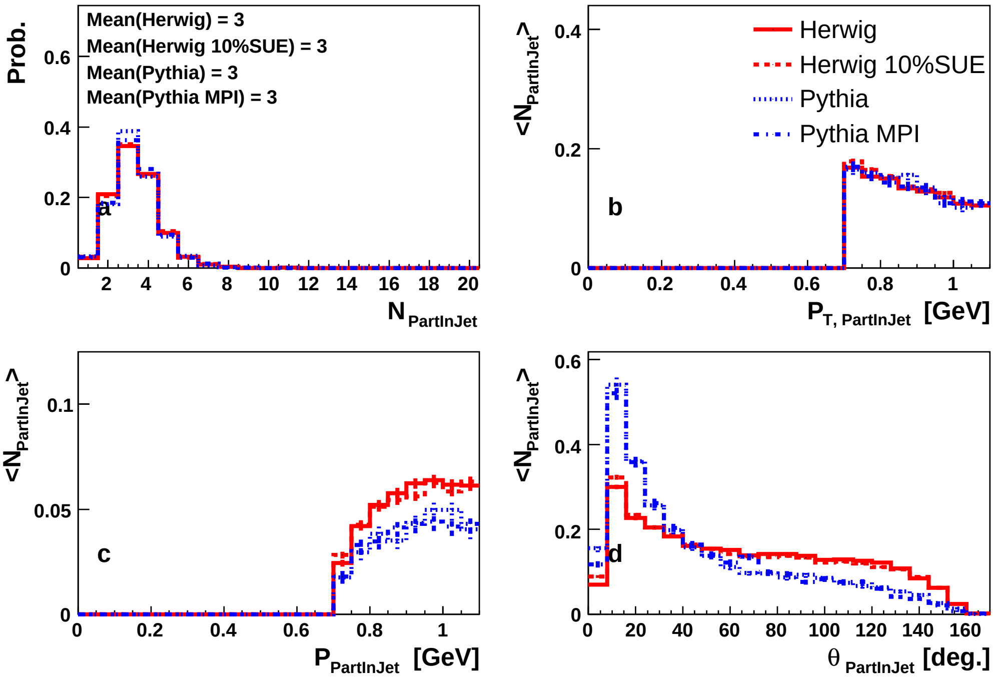


Figure 5.82: *The a) distribution of particle multiplicty in the mini-jets,* $N_{PartInJet}$ *and the average value ,* $\langle N_{PartInJet} \rangle$*, as a function of b) the transverse momentum* ($P_{T,PartInJet}$)*, c) the total momentum* ($P_{PartInJet}$)*, and d) the polar angle* ($\theta_{PartInJet}$) *of the particle jets in the transverse regions. The predictions of the Monte Carlo generators Herwig, Herwig 10%SUE, Pythia and Pythia MPI are shown for mini-jets reconstructed from particles with* $P_T > 0.7$ *GeV.*

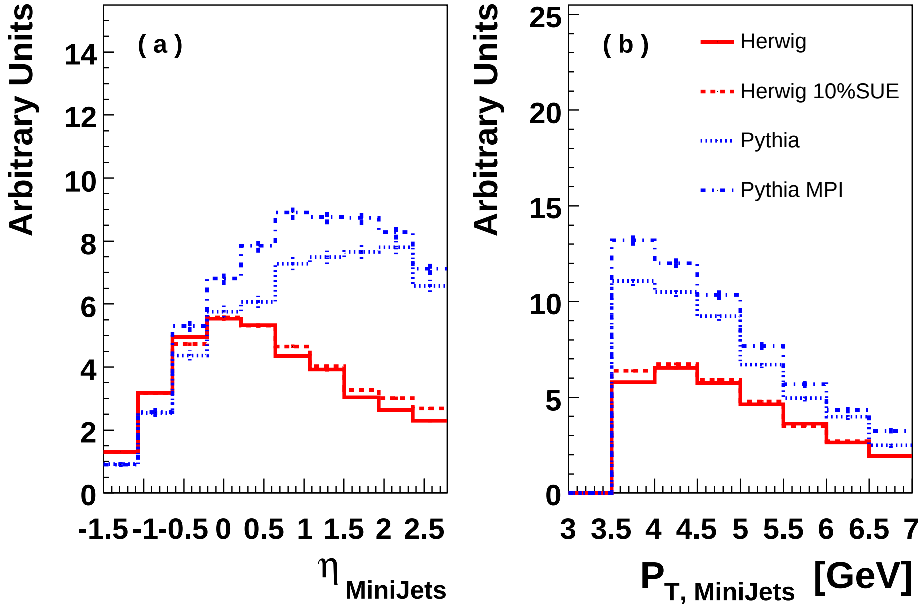


Figure 5.83: *The a)* $\eta_{MiniJets}$ *and b)* $P_{T,MiniJets}$ *distributions of mini-jets in the transverse regions from particles with* $P_T > 0.7$ *GeV. The predictions of the Monte Carlo generators Herwig, Herwig 10%SUE, Pythia and Pythia MPI are shown.*

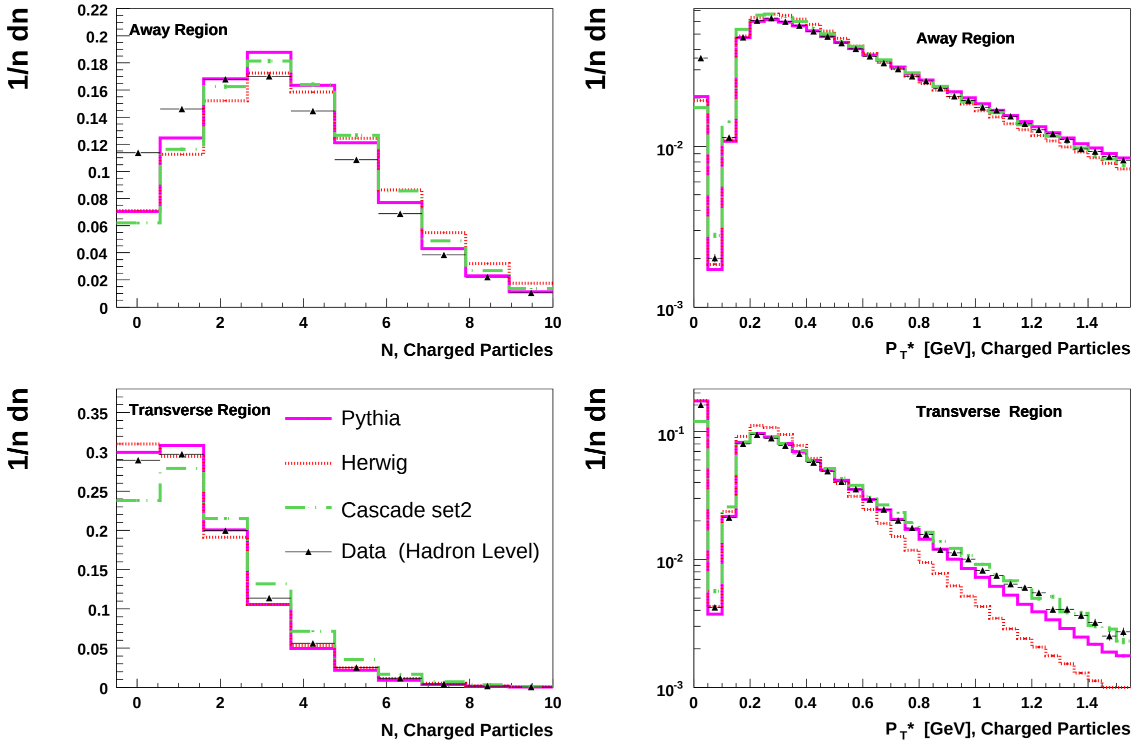


Figure 5.84: *The P_T spectra and the multiplicity distributions of charged particles from the inclusive 1-jet sample in the away and transverse regions. Data on hadron level are compared to the predictions of Pythia, Herwig and Cascade set2.*

5.6.4 Discussions and Comments

The claim of a signal from MPI in this analysis has to rely upon an observation of jet activities in excess of what is predicted by standard Monte Carlo generators. This, however, requires that the event generator is able to give a reasonable description of data in regions where the influence of MPI is expected to be small. As is noticed from the results of the inclusive 1-jet sample some of the MC predictions exhibit significant deviations from the data in these regions and in order to better understand the reason for this, a first comparison in terms of some basic event parameters has been performed. Thus, the multiplicity distributions of charged particles with $P_T > 0.15$ GeV and their P_T spectra have been studied. Such distributions have been extracted separately for the toward and away regions, whereas the transverse regions have not been split into high and low activity regions. In order to discuss the discrepancies in detail such plots should be produced for each of the kinematic bins that have been investigated, which, however, could not been realized due to time limitations. Therefore, only the most spectacular deviations between data and MC results can be expected to be understood at this stage. Figure 5.84 shows the P_T spectra and the multiplicity distributions of charged particles for experimental data and MC data from three of the event generators used in this analysis, when the leading jet is central. Furthermore, the charged particles are restricted to fall within the rapidity range of $|\eta^{lab}| < 1.5$. This is not quite consistent with the allowed range of mini-jets in the analysis which is $-1.5 < \eta_{jets} < 2.7$. However, in order to study the description of the hard scattering this should not have a big impact.

It was observed that although Pythia gives excellent agreement with data in the toward region, there are large deviation in the away and transverse regions for the higher Q^2 bins. In the away region the high mini-jet multiplicity is consistent with the high multiplicity of charged particles predicted by Pythia and the excess of $\langle PtSum^* \rangle$ is a consequence of the harder P_T spectrum of charged particles. The charged particle multiplicity in the transverse regions is in good agreement with data and the P_T is also reasonably well described, the deviations leaving room for contributions from MPI, consistent with the predicted rate in the low Q^2 bins.

Herwig gives approximately the same description of the multiplicity and transverse momenta of charged particles as Pythia in the away region, which makes it overshoot data on mini-jets somewhat for both $\langle N \rangle$ and $\langle PtSum^* \rangle$. In the transverse regions the charged particle multiplicity is well reproduced but the P_T spectrum is too soft. This probably explains some of the deviation compared to data but it is unlikely that it will account for the whole gap.

The predictions of Cascade J2003 set2 give the best overall agreement with mini-jets data and especially it is very successful in the transverse regions. This suggests that the data might be explained by using unintegrated parton densities and applying non k_t ordered emissions, which consequently would mean that no MPI is needed. An inspection of the charged particle multiplicity, however, shows that this is in general significantly higher than what is the case for the data so this could be the reason for the good agreement with mini-jet data in the transverse regions.

A more complete study would clearly be necessary in order to understand the details of the deviations between model predictions and data, but the general conclusion is that there is no model that gives a good description of data in all phase space regions and consequently a claim of MPI from the 1-jet sample is not straight forward.

5.7 Summary and Outlook

A study of mini-jets in deep inelastic electron proton scattering has been performed with the aim of finding evidence for hadronic activities in excess to those expected from the primary interaction. The multiplicity of mini-jets has been measured in four different regions of the azimuthal angle with respect to the hardest (leading) jet of the event. The region around the leading jet, called 'the toward region' and the region opposite to this, 'the away region', are expected to contain most of the transverse momentum which is needed to restore the momentum balance of the primary interaction. In the regions perpendicular to the leading jet the signal from additional hadron activities are expected to be most visible. For each event these regions have been separated into 'the high activity region', which contains the highest transverse momentum of the two, and 'the low activity region'.

Two samples have been used, one of which contains inclusive 1-jet events and the other inclusive 2-jet events, where the second hardest (sub-leading) jet is required to proceed in 'the away region'. The inclusive 1-jet sample has been subdivided into three Q^2, two W and two rapidity regions, corresponding to the leading jet being produced in the central or forward region. This subdivision has not been possible for the inclusive 2-jet sample due to statistical reasons but instead a cut on x_γ has been applied to separate direct and resolved photon processes.

The observation for the inclusive 1-jet sample can be summarized in the following way:

The toward region

The average jet multiplicity, $\langle N \rangle$ is well described by all models, whereas the average transverse momentum carried by the mini-jets, $\langle PtSum^* \rangle$, is not so well reproduced by Lepto(CDM) and Pythia exhibits large deviations in the higher Q^2 bins.

The away region

$\langle N \rangle$ is reasonably well described by most models except Pythia, which fails in the high Q^2 region. The model predictions of the $\langle PtSum^* \rangle$-distributions in general show somewhat larger deviation from the data and again Pythia fails to describe the high Q^2 bins.

The high activity region

All DGLAP based models produce $\langle N \rangle$-distributions which undershoot the data. Cascade using CCFM evolution predicts much higher values of $\langle N \rangle$ and Cascade J2003 set2 is in agreement with data. However, Lepto(CDM), which also produces parton emission non-ordered in k_t, gives significantly smaller $\langle N \rangle$-values. Pythia with MPI essentially agrees with data in the lowest Q^2 bin, whereas 10% SUE is not enough to bring Herwig in agreement with data. The same conclusions can be drawn from the $\langle PtSum^* \rangle$-distributions.

The low activity region

The DGLAP based models again give $\langle N \rangle$-distributions, significantly lower than data, which is also true for Lepto(CDM). Both Cascade models are more or less in agreement with data as is also Pythia MPI in the lowest Q^2 bin, while Herwig with 10% SUE still undershoots the data. Again essentially the same conclusions are true for the $\langle PtSum^* \rangle$-distributions.

Thus, the predictions on $\langle N \rangle$ and $\langle PtSum^* \rangle$ of the DGLAP based models and of Lepto(CDM) exhibit deviations in the transverse regions which leave room for MPI. Although the predictions by Cascade J2003 set2, with CCFM evolution and unintegrated gluon densities, indicate that no MPI is needed, the agreement with data may be explained by the too high particle multiplicity produced by this model. Herwig including SUE on a reasonable level is clearly not able to reproduce data. The inclusion of MPI in Pythia significantly improves the agreement with data in the lowest Q^2 bin.

The observations from the inclusive 2-jet sample are summerized in the following:

The toward region

The agreement of all models with data on $\langle N \rangle$ is very good. The $\langle PtSum^* \rangle$-distributions are also well described by all models except for Rapgap dir+res and Lepto(CDM), especially for resolved photon processes ($x_\gamma < 0.7$).

The away region

The $\langle N \rangle$ and $< PtSum >$-distributions are well desribed, with the exception of Rapgap dir+res for low x_γ.

The high activity region

The DGLAP based models and Lepto(CDM) predict distributions for $\langle N \rangle$ and $\langle PtSum^* \rangle$, which are significantly too low and would thus call for additional parton activity like MPI. The inclusion of MPI in Pythia makes a significant constribution to the resolved photon bin, bringing it much closer to data. Again 10% SUE is too little to account for the deviation to data. The Cascade model predictions even overshoot the data in some regions, consistent with a too high particle multiplicity.

The low activity region

The same conclusion are valid as for the high activity region.

The results from the inclusive 2-jet sample are consistent with additional jet activities in the transverse regions for resolved photon processes, which are not predicted with standard MC models but might be explained by MPI.

Altogether, model comparisons to data from the inclusive 1-jet and 2-jet samples consistently point at jet activities in addition to what is expected for the hard scattering process, in the regions transverse to the leading jet(s) in the event. The much improved agreement for resolved photon events due to the inclusion of MPI in Pythia supports the assumption that this might come from MPI.

Although it is clear that the discussion of MPI is by no means concluded as a result of this analysis, it nevertheless provides valuable experimental information, which can be used by theorists to test and improve their models.

The significantly higher statistics collected by H1 in the years 2001-2007 would allow a much more detailed analysis of MPI based on an inclusive dijet sample, for which the dependence on Q^2 and x_{Bj} could be studied and the sample could be subdiveded two η regions, as in the case of the inclusive 1-jet sample. The forward region is especially interesting since this is where the MPI signal is expected to be the strongest.

Acknowledgement

First and foremost I would like to thank my supervisor, Prof. Leif Jönsson for all help and guidance during my time as a Ph.D student. My warmest thanks for your patience and encouragement whenever I was having a hard time with the analysis. I am also very grateful to your help, comments and patience during my tough time on writing this thesis.

Many thanks to Dr. Hannes Jung for his supervision, support and insistence in pushing me forward with my analysis. To my former Ph.D. colleagues and office roommates Albert Knutsson and Magnus Hansson: thanks for the helpful exchange on the consistency of the analysis and the enjoyable discussions about physics and life.

I wish to thank the whole H1 collaboration, especially the HaQ working group. Very special thanks to my referees Guenter Grindhammer and Dmitri Ozerov for their great job. Many thanks to the members of the Calorimeter Analysis Task Force working group Emmanuel Sauvan, Katja Krueger and Armen Bunyatyan for their supervision.

I'm also very grateful to Gösta Gustafson, Leif Lönnblad and Torbjörn Sjöstrand from the theory department for their help in explaining the physics behind multiple parton interaction.

I would like to thank the thesis inspector, Hans-Åke Gustafsson for the comments and feedback on the thesis. Also, thanks to Stephen J. Maxfield, Dave Milstead, Magnus Hansson and Hannes Jung for taken the time to read this thesis and making useful comments and suggestions.

Many thanks to the members of the Lund Department of Elementary Particle Physics for their help during these years. Especially Ulf Mjörnmark, Anders Oskarsson, Evert Stenlund, Lennart Österman, Paula Eerola and Bozena Wlosinska.

Bibliography

[1] J.J. , ThompsonPhilosophical Magazine, 44, 293 (1897)

[2] H. Geiger and E. Marsden, 1909, Proc. R. Soc. London **82**, 495

[3] E. Rutherford, 1911, Philos. Mag. **21**, 669

[4] J. Chadwick, Nature **129** (1932) 312.

[5] M. Gell-Mann, Phys. Lett. **8** (1964) 214-215.

[6] G. Zweig, CERN preprints TH401 (January 17, 1964) and TH412 (January 21, 1964) (unpublished)

[7] O. W. Greenberg, Phys. Rev. Lett. **13** (1964) 598.

[8] Y. Nambu, *In *Lichtenberg, D. B. (Ed.), Rosen, S. P. (Ed.): Developments In The Quark Theory Of Hadrons, Vol. 1*, 112-121. (In *De-shalit: Preludes In Theoretical Physics*, 133-142)*

[9] M. Y. Han and Y. Nambu, Phys. Rev. **139** (1965) B1006.

[10] B.H Wiik, Proc. Intern. Neutrino Conf. (Bergen, 1979) p. 113 P
P. söding, Proc. EPS Intern. Conf. on high energy physics (Geneva, 1979) p.271
R. Brandelik *et al.* [Tasso Collaboration], Phys. Lett. **86B** (1979) 243
D.P. Barber *et al.* [MARK-J Collaboration], Phys. Rev. Lett. **43** (1979) 830
Ch. Berger *et al.* [PLUTO Collaboration], Phys. Lett. **86B** (1979) 418
W. Bartel *et al.* [JADE Collaboration], Phys. Lett. **91B** (1980) 142

[11] R. E. Taylor, Rev. Mod. Phys. **63** (1991) 573

[12] H. W. Kendall, Rev. Mod. Phys. **63** (1991) 597

[13] J. I. Friedman, Rev. Mod. Phys. **63** (1991) 615

[14] F. Gross, “Relativistic Quantum Mechanics and Field Theory”

[15] F. Halzen, Alan D. Martin, “Quarks and Leptons”

[16] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, Phys. Rept. **97** (1983) 31

[17] B. Andersson, G. Gustafson and B. Soderberg, Nucl. Phys. B **264** (1986) 29

[18] B. R. Webber, Nucl. Phys. B **238** (1984) 492.

[19] J. D. Bjorken and E. A. Paschos, Phys. Rev. **185** (1969) 1975

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

[21] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **30** (2003) 1 [arXiv:hep-ex/0304003].

[22] S. Aid *et al.* [H1 Collaboration], Nucl. Phys. B **470** (1996) 3 [arXiv:hep-ex/9603004].

[23] V. Gribov and L. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438 and 675

[24] L. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 94

[25] G. Altarelli and G. Parisi, Nucl. Phys. **B 126** (1977) 298

[26] Y. Dokshitzer, Sov. Phys. JETP **46** (1977) 641

[27] S. Moch, J. A. M. Vermaseren and A. Vogt, Nucl. Phys. B **688** (2004) 101 [arXiv:hep-ph/0403192].

[28] E. Kuraev, L. Lipatov and V. Fadin, Sov. Phys. JETP **44** (1976) 443

[29] E. Kuraev, L. Lipatov and V. Fadin, Sov. Phys. JETP **45** (1977) 199

[30] Y. Balinsky and L. Lipatov, Sov. J. Nucl. Phys. **28** (1978) 822

[31] M. Ciafaloni, Nucl. Phys. B **296** (1988) 49

[32] S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. B **234** (1990) 339

[33] S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. B **336** (1990) 18

[34] G. Marchesini, Nucl. Phys. B **445** (1995) 49

[35] A.E. Chudakov: Izv. Akad. Nauk SSSR, Ser. Fiz. 19, 650 (19551) C. vonWeizs icker: Z. Physik 88, 612 (1934);
E. Williams:Phys. Rev. 45, 729 (1934); Kgl. Dansk. Vid. Selsk. 13, 4 (1935)
B.I. Ermolaev, V.S. Fadin: ZhETF Pis'ma, 33, 285 (1981)
A.H. Mueller: Phys. Lett. 104B, 161 (1981)
A.H. Mueller: Nucl. Phys. B213, 85 (1983)

[36] G. Gustafson, Phys. Lett. B **175** (1986) 453

[37] G. Gustafson and U. Pettersson, Nucl. Phys. B **306** (1988) 746

[38] B. Andersson *et al.*, Z. Phys. C **43** (1989) 625

[39] B. Andersson G. Gustafson and L. Lönnblad, Nucl. Phys. B **339** (1990) 393

[40] The Color-dipole model and the ARIADNE program at high Q**2 arXiv:hep-ph/9908368 L. Lönnblad,

[41] M. Klasen and G. Kramer, Z. Phys. C **76** (1997) 67 [arXiv:hep-ph/9611450].

[42] B. W. Harris and J. F. Owens, Phys. Rev. D **56** (1997) 4007 [arXiv:hep-ph/9704324].

[43] S. Frixione, Nucl. Phys. B **507** (1997) 295 [arXiv:hep-ph/9706545].

[44] M. Klasen and G. Kramer, Phys. Lett. B **366** (1996) 385 [arXiv:hep-ph/9508337]

[45] S. Frixione and G. Ridolfi, Nucl. Phys. B **507** (1997) 315 [arXiv:hep-ph/9707345]

[46] A. Banfi and M. Dasgupta, JHEP **0401** (2004) 027 [arXiv:hep-ph/0312108]

[47] V. A. Abramovsky, V. N. Gribov and O. V. Kancheli, Yad. Fiz. **18** (1973) 595 [Sov. J. Nucl. Phys. **18** (1974) 308].

[48] J. Bartels and M. G. Ryskin, Z. Phys. C **76** (1997) 241 [arXiv:hep-ph/9612226].

[49] J. Bartels, M. Salvadore and G. P. Vacca, Eur. Phys. J. C **42** (2005) 53 [arXiv:hep-ph/0503049].

[50] Borras, K., QCD and collider physics II Lecture, SoSe 24/5-2006, University Hamburg, Diffaction,

[51] S. Donnachie, G. Dosch, P. Landshoff and O. Nachtmann, Pomeron Physics and QCD, Cambridge University Press, 200212, ISBN-10: 0-521-78039-X

[52] V. Barone and E. Predazzi, High-Energy Particle Diffraction, ISBN 3-540-42107-6

[53] A. Donnachie and P. V. Landshoff, Phys. Lett. B **296** (1992) 227 [arXiv:hep-ph/9209205]. Alternativ

[54] S. Aid *et al.* [H1 Collaboration], "Jets and Energy Flow in Photon-Proton Collisions at HERA," Z. Phys. C **70** (1996) 17-30 [arXiv:hep-ex/9511012].

[55] J. Breitweg *et al.* [ZEUS Collaboration], "ZEUS results on the measurement and phenomenology of F2 at low x and low Q**2," Eur. Phys. J. C **7** (1999) 609 [arXiv:hep-ex/9809005].

[56] A. Donnachie and P. V. Landshoff, Z. Phys. C **61** (1994) 139 [arXiv:hep-ph/9305319].

[57] K. J. Golec-Biernat and M. Wusthoff, Phys. Rev. D **60**, 114023 (1999) [arXiv:hep-ph/9903358].

[58] K. J. Golec-Biernat and M. Wusthoff, Phys. Rev. D **59** (1999) 014017 [arXiv:hep-ph/9807513].

[59] J. Bartels, K. J. Golec-Biernat and H. Kowalski, Phys. Rev. D **66** (2002) 014001 [arXiv:hep-ph/0203258].

[60] J. R. Andersen *et al.* [Small x Collaboration], "Small x phenomenology: Summary of the 3rd Lund small x workshop in 2004," Eur. Phys. J. C **48** (2006) 53 [arXiv:hep-ph/0604189].

[61] A. Petrukhin, "New measurement of the structure function $F_2(X, Q^2)$ at low Q^2 with initial state radiation data," Proceeding of DIS 2004, Slovakia.

[62] Rick Field analysis 1 T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65** (2002) 092002

[63] D. E. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **70** (2004) 072002 [arXiv:hep-ex/0404004].

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

[65] T. Sjostrand, Comput. Phys. Commun. **39** (1986) 347

[66] T. Sjostrand, Comput. Phys. Commun. **82** (1994) 74

[67] A. Kwiatkowski *et al.* Proc. of the Workshop on Physics at HERA,

[68] L. Lönnblad, Comp. Phys. Comm. **71** (1992) 15

[69] G. Ingelman, A. Edin and J. Rathsman, Comput. Phys. Commun. **101** (1997) 108 [arXiv:hep-ph/9605286]

[70] K. Charchula, G. A. Schuler and H. Spiesberger, Comput. Phys. Commun. **81** (1994) 381.

[71] T. Sjostrand, Phys. Lett. B **157** (1985) 321.

[72] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour and L. Stanco, Comput. Phys. Commun. **67** (1992) 465.

[73] H. Jung and G. P. Salam, Eur. Phys. J. C **19** (2001) 351 [arXiv:hep-ph/0012143]

[74] H. Jung, Comp. Phys. Comm. **143** (2002) 100 [arXiv:hep-ph/0109102]

[75] H. Jung, arXiv:hep-ph/0411287

[76] M. Hansson and H. Jung, arXiv:hep-ph/0309009

[77] H. Jung, Acta Phys. Polon. B **33** (2002) 2995 [arXiv:hep-ph/0207239]

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

[79] R. Brun *et al.*, "GEANT User's guide", CERN-DD/EE/84-1, 1987

[80] B. Andrieu *et al.* [H1 Calorimeter Group], Nucl. Instrum. Meth. A **336** (1993) (460-498).

[81] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386** (1997) (310-347)

[82] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386** (1997) (348-396)

[83] B. Andrieu *et al.* [H1 Calorimeter Group], Nucl. Instrum. Meth. A **344** (1994) 492.

[84] J.P.Kubenka, H.Oberlack, P.Schacht, H.P.Wellisch, MPI-PhE/94-03, January 1994, *HADRONIC CALIBRATION OF THE H1 LAr CALORIMETER USING SOFTWARE WEIGHTING TECHNIQUES*

[85] A. Walther and R. Barschke [H1 Collaboration], *Prepared for Workshop on Scintillating Fiber Detectors (SCIFI 93), Notre Dame, IN, 24-28 Oct 1993*

[86] G. Muller [H1 Collaboration], *Prepared for 5th International Conference on Calorimetry in High-energy Physics, Upton, NY, 25 Sep - 1 Oct 1994*

[87] M. Weber [H1 Collaboration], *In *Beijing 1994, Proceedings, Calorimetry* 107-115, and Hamburg DESY - DESY-95-067 (95/04,rec.Apr.) 11-15*

[88] J. Hladky [H1 Collaboration], *Prepared for International Workshop on Elementary Particle Physics: Present and Future, Valencia, Spain, 5-9 Jun 1995*

[89] R. D. Appuhn *et al.* [H1 SPACAL Group],

[90] B. List, Nucl. Instrum. Meth. A **549** (2005) 33.

[91] D. Pitzl *et al.*, Nucl. Instrum. Meth. A **454** (2000) 334 [arXiv:hep-ex/0002044].

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

[93] S.Egli et al DESY, Zuerich, Dortmund. *Calculating Event Weights in Case of Donwscaling on Trigger Levels 1-4*, H1 internal note H1-04/97-517.

[94] R. Pöschl, Ph.D. Thesis, Dortmund (2000)

[95] U. Bassler and G. Bernardi, Nucl. Instrum. Meth. A **361** (1995) 197 [arXiv:hep-ex/9412004].

[96] U. Bassler and G. Bernardi, Nucl. Instrum. Meth. A **426** (1999) 583 [arXiv:hep-ex/9801017].

[97] M. Peez, B. Portheault and E. Sauvan, *An Energy Flow Algorithm for Hadronic Reconstruction In OO: Hadroo2*, H1 internal note h1-0105-616

[98] S. Catani, Y. L. Dokshitzer and B. R. Webber, Phys. Lett. B **285** (1992) 291

[99] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406** (1993) 187

[100] S. D. Ellis and D. E. Soper, Phys. Rev. D **48** (1993) 3160 [arXiv:hep-ph/9305266]

[101] J. Huth *et al.*, FERMILAB-CONF-90 249-E (1990)

[102] Humboldt-Univ, Berlin. V.V.ARKADOV *Measurement of the Deep-Inelastic ep Scattering Cross Section using the Backward Silicon tracker at the H1 Detector at HERA*

[103] M. Jacquet1, Z. Zhang1,V. Brisson1, S. Kermiche2, C. Vallee2 *Absolute hadronic jet calibration of the H1 liquid Argon calorimeter*, H1 internal note ,h1-0499-571

[104] A. Aktas *et al.* [H1 Collaboration], "Forward jet production in deep inelastic scattering at HERA," Eur. Phys. J. C **46**, 27 (2006) [arXiv:hep-ex/0508055].

[105] F. D. Aaron *et al.* [H1 Collaboration], "Three- and Four-jet Production at Low x at HERA," Eur. Phys. J. C **54**, 389 (2008) [arXiv:0711.2606 [hep-ex]].

[106] A. Knutsson, *Forward jet production in ep collisions at HERA* , LUNFD6-NFFL-7225-2007

[107] V. Blobel, Lectures given at 1984 CERN School of Computing, Aiguablava, Spain, Sep 9-22, 1984. Published in CERN Comp.School 1984:0088, "Unfolding Methods In High-Energy Physics Experiments,"

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

[109] G. D'Agostini, A Multidimensional Unfolding Method Based on Bayes' Theorem, DESY 94-099

[110] G. D'Agostini, Comments on the Bayesian unfolding, ZEUS-Note 95-166

[111] M. TAŠEVSKÝ, Ph.D. Thesis, Virtual photon structure from low Q^2 dijet production at HERA, 1999

[112] K. Sedlak, Ph.D. Thesis, Measurement of Dijet Production at Low Q^2 at HERA, 2004

[113] T. Kluge, Ph.D. Thesis, Measurement and QCD Analysis of Event Shape Variables in Deep-Inleastic Electron-Proton Collisions at HERA, DESY-THESIS-2004-024

[114] M. Hansson, Ph.D. Thesis, Azimuthal Correlations in Dijet Events from Deep Inelastic Positron-Proton Scattering at HERA, 2007

[115] M. Drees and K. Grassie, Z. Phys. C **28** (1985) 451.

[116] *QHQTRK maunal (Heavy flavour working at H1, Group track selection code be Lee West)*

Appendix A

Definition of Trigger Elements

The definitions of the physics sub-triggers (S0, S3, S61) and the monitor sub-triggers (S39, S64, S66, S67, S77) are presented here. The sub-trigger definitions changed during the 1999-2000 run period, therefore the definition of the main trigger elements is given here.

SPCLe_IET $>$ 2: An energy deposition above 5.7 GeV in one of the trigger Towers of the outer SPACAL.
BG: A combination of trigger elements used to reject background events in DIS.
SPCLe_ToF_E_2: Rejection of background events in DIS.
SPCLe_IET_Cen_3: An energy deposition of $E > 6$ GeV in the central SPACAL region.
DCRPh_THig: At least one high momentum track candidate in the CJC.
zVtx_sig : A signature for the z-vertex existence.
LAr_IF: Energy summed over all Big Towers lying in IF.
LAr_BR: Big Tower energy above a given threshold and a validated track from MWPC.
LAr_Etmiss $>$ 1(2): Missing transverse energy above a given threshold from the Big Towers calculation.
LAr_electron_1: An energy deposition in the electromagnetic part of the LAr calorimeter above 11 GeV.
LAr_electron_2: An energy deposition in the electromagnetic part of the LAr calorimeter above 6 GeV.
LAr_2or3_electrons: Two or three electrons found in the LAr.
FwdRay_T0 : At least one track found in the MWPCs.

Subtrigger		Definitions of the Trigger Elements
S0	=	SPCLe_IET $> 2 \wedge$ BG
S3	=	SPCLe_IET $> 2 \wedge$ SPCLe_ToF_E_2
S61	=	(SPCLe_IET $> 2 \vee$ SPCLe_IET_Cen_3) $\wedge$ DCRPh_THig $\wedge$ zVtx_sig
S39	=	LAr_BR $\wedge$ LAr_electron_2 $\wedge$ FwdRay_T0 $\wedge$ LAr_IF $\wedge$ LAr_2or3_electrons
S64	=	LAr_IF $> 1 \wedge$ LAr_Etrans > 2
S66	=	LAr_IF $> 1 \wedge$ LAr_Etmiss > 2
S67	=	LAr_electron_1
S77	=	LAr_Etmiss > 1

Appendix B

Final State Particle Classification

Selected Tracks: Three classes of tracks are defined: Central (tracks reconstructed only from CJC measurements), Forward (tracks reconstructed only from Forward tracker measurements) and Combined (tracks reconstructed from CJC and Forward tracker measurements) tracks. The main cuts are summarised in the tables below.

Central Track

	P_T	> 120 MeV
$20^o <$	θ	$< 160^o$
	$\lvert dca' \rvert$	≤ 2 cm
	R_S	≤ 50 cm
	R_{length}	≥ 10 cm for $\theta < 150^o$
	R_{length}	≥ 5 cm for $\theta > 150^o$

Forward Track

	P_T	> 1 MeV
	P	≥ 500 MeV
$6^o \leq$	θ	$\leq 25^o$

Combined Track

P_T	> 120 MeV
θ	$< 40^o$
$\lvert dca' \rvert$	≤ 5 cm
R_S	≤ 50 cm

The dca' is the distance of the closest approach of the track extrapolated to the measured vertex in the (x, y, z_{vertex}) plane. Projected onto the xy-plane, R_S is the distance from the vertex point at which the track starts, and R_{length} is the length of the track. More details how the tracks are obtained can be found in [116].

Isolated muons and electrons:

A muon is isolated if:

$$E^{em}_{r=35} + E^{h}_{r=75} < 5 \text{ GeV},$$
$$N^{sel.track.}_{R=0.5} = 1$$

$E^{em}_{r=35}$ and $E^{h}_{r=75}$ are the energy deposited in a cylinder with radius of 35 cm and 75 cm in the electromagnetic and hadronic LAr section, respectively, around the extrapolated muon track. $N^{sel.track.}_{R=0.5}$ is the number of selected tracks around the muon in a cone of

$$R = \sqrt{\Delta\phi^2 + \Delta\theta^2} = 0.5.$$

An electron is isolated if:

$$E_{R=0.5} < 0.03E_e$$

where $E_{R=0.5}$ is the total energy of clusters not belonging to any other electron in a cone of

$$R = \sqrt{\Delta\phi^2 + \Delta\theta^2} = 0.5$$

around the electron with energy E_e.

Appendix C

Paper I

To appear in the proceedings of the 15th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2007), April 16-20, 2007, Munich, Germany.

Mini-Jets in Deep Inelastic Scattering at HERA

S. Osman *

Lund University,
Box 118, SE 221 00, Sweden, E-mail: sakar.osman@hep.lu.se

The production of jets with low P_T, mini-jets, in deep inelastic electron-proton scattering is studied. Mini-jet multiplicities are presented as a function of the P_T of the leading jet in bins of η and Q^2. The analysis is performed for an inclusive jet sample, and for a dijet sample, where the second jet is required to have an azimuthal angle larger than 140 degrees with respect to the leading jet. The dijet sample is split into two samples which are enhanced in direct photon and resolved photon processes, respectively. The results are compared to various QCD based models. Here only the result for the inclusive jet sample will be presented.

1 Introduction

In electron-proton scattering the partonic content of the exchanged virtual photon may be resolved if the P_T of the interacting partons is larger than Q^2 and thereby the photon will behave like a hadronic object. Thus, similar to hadron-hadron scattering there will be a certain probability that collisions between the resolved photon and the proton involve more than one parton interaction, multiple interactions (MI). Previous measurement in photoproduction at HERA [2] have shown that only models with MI give a satisfactory description of the data. This analysis constitutes the first study of possible MI in DIS from measurements of low P_T jets produced in addition to the leading jet(s) of the event. The basic principle of the analysis is to define regions in phase space where contributions from the final state products originating from the primary hard interaction are expected to be small. The concept follows closely the one used by the CDF collaboration at the TEVATRON [3]

2 Analysis Method

The starting point is to define and isolate the leading jet(s) originated from the hard primary interaction and investigate the remaining regions for additional activities, which in this analysis comprise the presence of jets with low transverse momenta, mini-jets. Two different event samples are studied; *inclusive jet events* and *dijet events*, of which the latter constitutes a subsample of the inclusive sample. The analysis procedure is the following:

Inclusive jet sample: The leading jet is identified and reconstructed using the k_t-algorithm [4] in the h.c.m. rest frame. The jet with the highest transverse momentum in the h.c.m. rest frame is taken as the leading jet. The leading jet axis defines the azimuthal angle $\Delta\phi^*$=0 [a]. The region $|\Delta\phi^*| < 60^o$ is defined as the 'toward region', and is expected to contain all particles belonging to the leading jet. The angular region $|\Delta\phi^*| > 140^o$ is called the 'away region'. The transverse regions, $60^o < |\phi^*| < 120^o$ are those where contributions from the primary collision should be small and the effects from additional activities should be most visible. Event by event a 'high activity'- and 'low activity' region are defined, depending on

---

[*] On behalf of the H1 Collaboration.

[a] Observables in the h.c.m. frame are labeled with *.

which region contains the most and least transverse momentum, respectively. These four regions are shown. in Figure 2

Dijet sample: The dijet sample includes events having at least two jets, where the two reconstructed jets are required to be almost back-to-back. The leading jet is again defining $\Delta\phi^{*}$=0, whereas the jet axis of the jet with the second highest transverse momentum, the sub leading jet, is restricted to be inside the 'away region'. This leaves some angular space to accommodate the transverse spread of the jet within the 'away region'.

3 Event Selection

The analysis is based on data taken with the H1 detector in 1999/2000 using colliding positrons and protons at energies of 27.5 GeV and 920 GeV, respectively. The DIS events are selected by requiring a positron in the SPACAL calorimeter with $E_e^{'}$ $>$ 9 GeV , $156^\circ < \theta_e < 175^\circ$ where $E_e^{'}$ and θ_e are the energy and polar angle of the scattered positron, respectively. The photon inelasticity, y, and the virtuality, Q^2, are determined using the electron method and must fulfill $0.1 < y < 0.7$ and $5 < Q^2 < 100$ GeV^2. The invariant mass of the hadronic final state, W, is required to be higher than 200 GeV in order to enhance small x_{bj} contributions and to increase the probability of mini-jet productions.

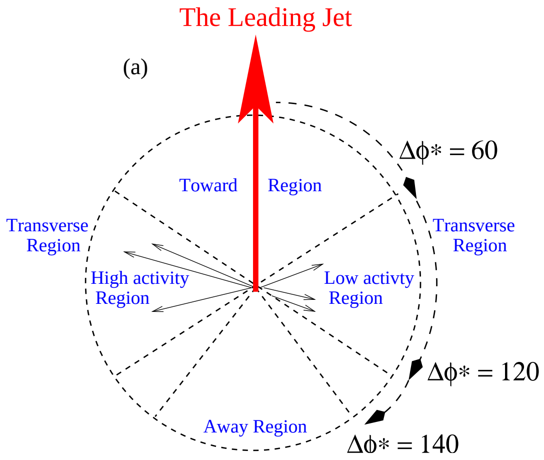


Figure 1: The transverse , toward and away regions.

The inclusive jet sample consists of events that contain at least one jet, whereas the dijet sample includes events with at least two jets. For both samples the jet with the highest transverse momentum is chosen as the leading jet and in the dijet sample the sub leading jet has to fulfill the requirement $|\Delta\phi^{*}_{ls}| = |\phi^{*}_{lj} - \phi^{*}_{sj}| > 140^\circ$, where ϕ^{*} is the jet azimuthal angle and the labels lj and sj denotes the leading and sub leading jets, respectively. The jets are reconstructed by the inclusive k_t-algorithm [4] in its p_t weighting scheme mode, applied to combined object of tracks and calorimetric clusters in the h.c.m. rest frame. To ensure a good jet reconstruction it is required that the leading and sub leading jets must fulfill -1.7 $< \eta^{lab} <$ 2.79 and $P_T^{(*)} > 5$ GeV. Here, the pseudo-rapidity is given by $\eta_j = -\ln(\tan(\theta_j/2))$, where θ_j is the polar angle of the jet in the lab frame, and P_T is the transverse energy of the jet. The pt cut are applied both in lab and h.c.m. rest frame.

Mini-jets are reconstructed with the same jet algorithm as the leading jets and within the same η^{lab} region. However, the minimum transverse momentum of a mini-jets is required to be larger than as 3 GeV both in lab and h.c.m rest frame .

Data is corrected for limited detector resolution and acceptance using a bin-by-bin procedure. Correction factors are determined using detector simulated events, generated by the Monte Carlo programs RAPGAP [5] and DJANGO [6] with ARIADNE [7], where QED radiation has been taken into account.

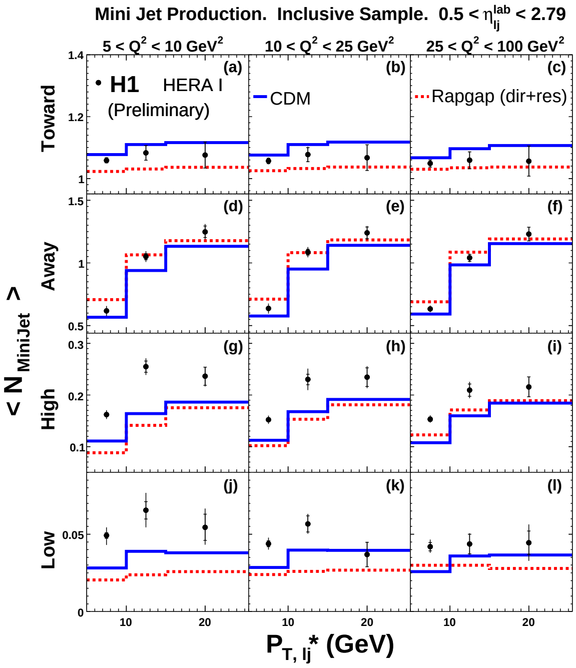


Figure 2: *The average mini-jet multiplicity at the different* $\Delta\phi^*$ *regions in bins of* Q^2 *as a function of* P_T^* *of the leading jet for the inclusive jet sample. The data is compared with the CDM model (solid line) and Rapgap (dashed line).*

4 Results

The average multiplicity of mini-jets, $< N_{minijet} >$, for the inclusive jet sample where the leading jet proceeds in the forward η region are shown in Figures 2 - 3 for three bins of Q^2. The data are presented as a function of P_T^* of the leading jet. Results have also been obtained for leading jets in the central region and for the dijet sample but they are not shown here, but can be found at [1]. The following can be observed:

The toward and away regions: All the different MC models, with or without MI, describe the 'toward region' well in all Q^2 bins , as expected. In the 'away region' there is an overall reasonable agreement for all models, Figures 2 - 3 (a-f).

The high activity region: The predictions of $< N_{minijet} >$ by the MC models including no MI are generally too low in all Q^2-bins, Figure 2 (g-i). PYTHIA+MI [8, 9] describes the data points fairly well in the lowest Q^2-bin and somewhat less well in the higher Q^2-bins, Figures 3 (g-i).

The low activity region: The MC models with no MI included, significantly undershoots the data, Figure 2 (j-l), in all Q^2 bins. These deviations clearly increase with decreasing Q^2-values. PYTHIA+MI gives a much, Figure 3 (j-l), better description of data, although the deviations are still large in the highest Q^2-bin.

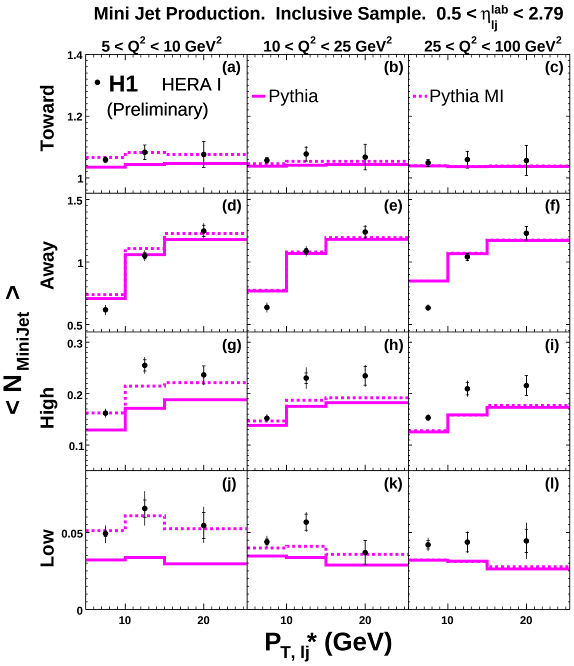


Figure 3: *The average mini-jet multiplicity at the different $\Delta\phi^*$ regions in bins of Q^2 as a function of P_T^* of the leading jet for the inclusive jet sample. The data is compared with Pythia (solid line) and Pythia with MI (dashed line).*

References

[1] Slides:
`http://indico.cern.ch/contributionDisplay.py?contribId=226&sessionId=6&confId=9499`

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

[3] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **65** (2002) 072005.

[4] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**(1993) 3160.

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

[6] K. Charchula, G. A. Schuler and H. Spiesberger, Comput. Phys. Commun. **81** (1994) 381.

[7] L. Lonnblad, Comput. Phys. Commun. **71** (1992) 15.

[8] T. Sjostrand, Phys. Lett. B **157** (1985) 321.

[9] T. Sjostrand, Comput. Phys. Commun. **82** (1994) 74.