Faculty of Mathematics and Physics of Charles University
in Prague

# Production of hadrons in interactions of electrons with protons in the experiment H1

Diploma Thesis
1998

Author : Kamil Sedlák
Supervisor : Doc. Ing. Josef Žáček, DrSc.
Adviser : Ing. Jaroslav Cvach, CSc.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně s použitím uvedené literatury. Souhlasím se zapůjčením diplomové práce.

V Praze, 14.4.1998

*motto:*

Co se v mládí naučíš

V létě, když jsou dlouhé dny,
vlezu na noc do bedny.
Pomalu si zvykám,
že tak jednou ztuhle,
až natáhnu brka,
budu ležet v truhle.

*(Emil Jabčák Hofman*
*Sbírka básní "Hřbitovy svítí")*

# Introduction

The study of electron-proton scattering greatly contributes to our understanding of fundamental constituents of matter and the nature of interactions between them. Previous fixed-target experiments reached CMS energies about 30 GeV. At the accelerator HERA, the energy of 300 GeV was achieved, which significantly increased physics potential for lepton-nucleon scattering and resolution of proton constituents.

Cross sections of the electron-proton interactions are well calculable by QCD. These processes can be measured, and are successfully described, by predictions of Monte Carlo (MC) models based on the leading order QCD calculations. On the other hand, the soft underlying interactions, i.e. the interactions of the proton and photon remnants, are not yet fully understood. The description of these "underlying processes" by MC programs is still in progress. One of the main aims of this thesis is to contribute to the understanding of additional remnant interaction, to compare it with available MC simulation programs, and to tune some of the MC parameters which correspond to the remnant interactions.

Collisions with a very low virtuality of the photon (i.e. $Q^2 < 0.01$ at H1) are historically called photoproduction, while interactions with $Q^2$ larger than a few GeV$^2$ correspond to deep inelastic scattering (DIS). The hadron production at HERA, for both the photoproduction and DIS, has been intensively studied by the H1 collaboration. The single particle spectra and jet production has already been treated and presented in several publications (e.g. in [21, 22, 23, 24]). However, the energy flow outside jets, which is very closely related to the additional underlying interactions of proton and photon remnants, has not yet been analysed in detail. One publication partly devoted to this topic is: "Jets and Energy Flow in Photon-Proton Collisions at HERA" (see [19]), which presents the treatment of 1993 photoproduction data. In my thesis these previous analyses are extended to photoproduction, as well as deep inelastic scattering 1995 data. This enabled the study of the dependence of energy flow on $Q^2$. Monte Carlo generators HERWIG 5.9 and RAPGAP 2.9 were used for comparison with the measured data, since they were the only available simulation programs for the $e - p$ interactions in a very low $Q^2$ area. My results were presented at several meetings of the "Photoproduction Working Group" at DESY.

The time consuming computations were carried out on the Sun work station *Sparc 40* (sun103, Institute of Physics, AV ČR), Silicon Graphics stations *Indigo Entry* and *Indy R5000* (indigo-2, indigo-3, Nuclear Centre, MFF UK, ČR) and work group server *naxos* (DESY, Hamburg).

This text is available as a postscript file at the html address:
*http://indigo-3.troja.mff.cuni.cz/˜sedlak/*

# Contents

# Chapter 1

# The H1 Detector

## 1.1 HERA collider

The high energy accelerator HERA is situated at the German national laboratory DESY in Hamburg.

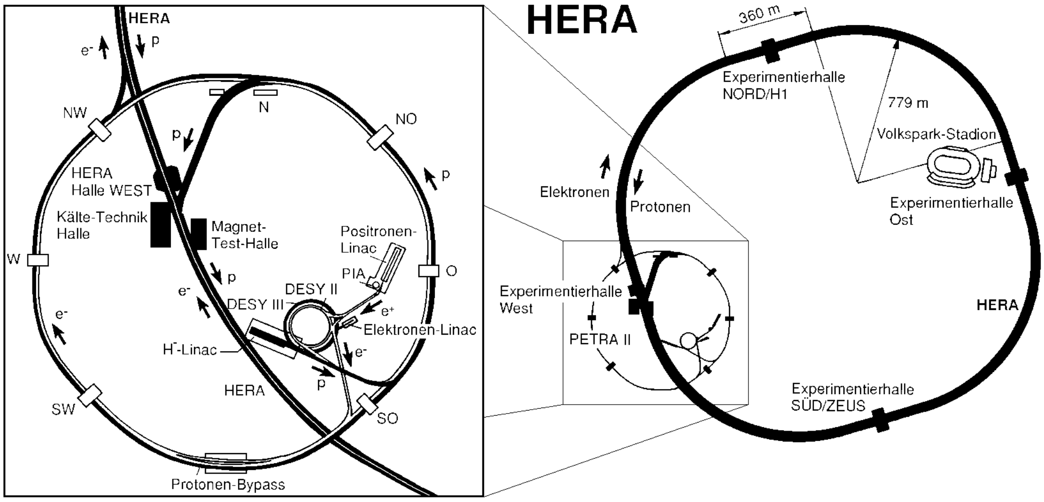


Figure 1.1: HERA accelerator

HERA consists of two storage rings, one for protons of the energy 820 GeV and the other one for electrons or positrons of 27.5 GeV. The two beams circulate in the 6.3 km long underground tunnel and collide at two interaction points in which the reaction products are measured by two large detectors – ZEUS and H1. HERA was the first ever constructed e–p storage ring and enabled to study a new kinematic region of e–p collisions.

Beam injection into HERA is done in several steps involving several preaccelerators (see Fig. 1.1). In total, 1170 main dipole and quadrupole magnets are guiding the electrons and 646 superconducting magnets the protons. Accelerated particles are kept in bunches, the time between two subsequent bunch crossings is 96 ns. The bunch length, width and height is about 10 mm, 0.3 mm and 0.04 mm respectively for electrons and 110 mm, 0.3 mm and 0.08 mm for protons.

## 1.2 The layout of the H1 detector

A hard $e-p$ interactions usually contain several leptons accompanied by quark and gluon jets. Therefore the H1 detector was designed for good identification of leptons (especially electrons, but also muons) and good resolution for jets.

Because of the accelerated protons have much higher kinetic energy than the electrons, the centre of mass is boosted along the proton direction with $\gamma_{cm} \doteq 2.9$ (that corresponds to a centre of mass energy of about 300 GeV). Thus the detector is asymmetric and more massive in the proton direction[1].

The layout of the H1 arrangement is displayed in Fig. on page 7. The tracking devices $\boxed{2}$ and $\boxed{3}$ are located next to the beam pipe $\boxed{1}$. Each of them (central, forward and backward) contains different layers of drift chambers and trigger proportional chambers. Behind the backward drift chamber the SPACAL $\boxed{12}$ (a lead scintillating fibre calorimeter) is placed. All these detectors are surrounded by the liquid argon calorimeter $\boxed{4}$ and $\boxed{5}$, which consists of two parts: the electromagnetic section made of lead absorber plates and the hadronic section made of steel plates.

A superconducting cylindrical coil $\boxed{6}$ with a diameter of 6 meters provides the analysing magnetic field of 1.15 T. This field is compensated by another superconducting coil $\boxed{7}$ in order not to influence the HERA beam.

The iron return yoke of the magnet is laminated and filled with limited streamer tubes. A small fraction of hadronic energy leaking out of the liquid argon calorimeter is registered here, and muon tracks are measured. Muons are also identified and measured by muon chambers $\boxed{9}$ inside and outside the iron. Measurement of muon tracks in the forward region is done in the toroid equipped with drift chambers $\boxed{9}$ and $\boxed{11}$.

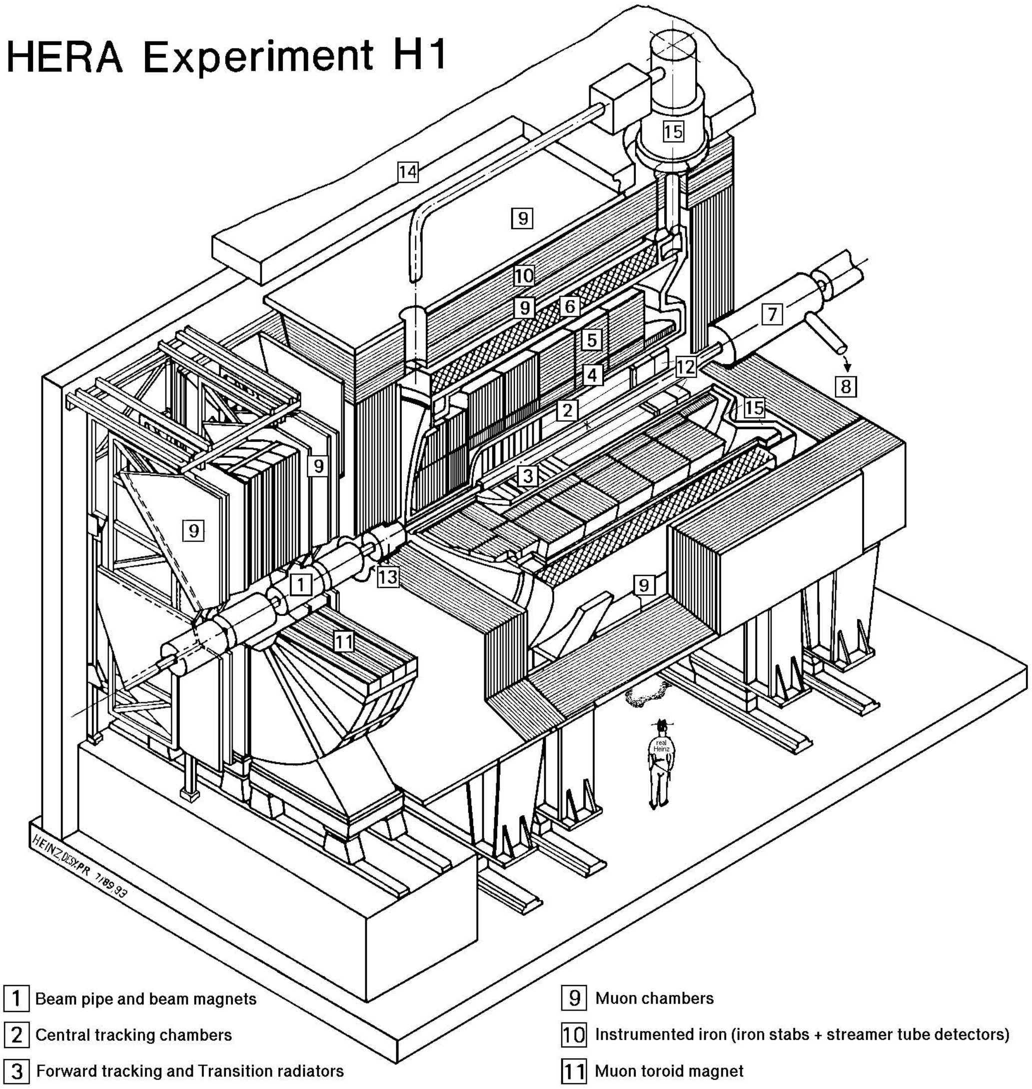
The H1 detector is completed with the small angle electron and photon detectors placed at 33.4 and 102.9 meters respectively from the nominal interaction point in the HERA tunnel.

The major components of the H1 relevant for the physical analysis presented in this thesis will be discussed in the next sections. More details can be found in [1, 2].

---

[1] The convention at H1 detector is that the + z-axis is oriented in the direction of proton beam and the origin is placed to the interaction point.

# HERA Experiment H1



| | | |
|---|---|---|
| 1 | Beam pipe and beam magnets | |
| 2 | Central tracking chambers | |
| 3 | Forward tracking and Transition radiators | |
| 4 | Electromagnetic calorimeter (lead) | Liquid Argon |
| 5 | Hadronic calorimeter (stainless steel) | |
| 6 | Superconducting coil (1.2T) | |
| 7 | Compensating magnet | |
| 8 | Helium cryogenics | |
| 9 | Muon chambers | |
| 10 | Instrumented iron (iron stabs + streamer tube detectors) | |
| 11 | Muon toroid magnet | |
| 12 | Warm electromagnetic calorimeter | |
| 13 | Plug calorimeter (Cu, Si) | |
| 14 | Concrete shielding | |
| 15 | Liquid Argon cryostat | |

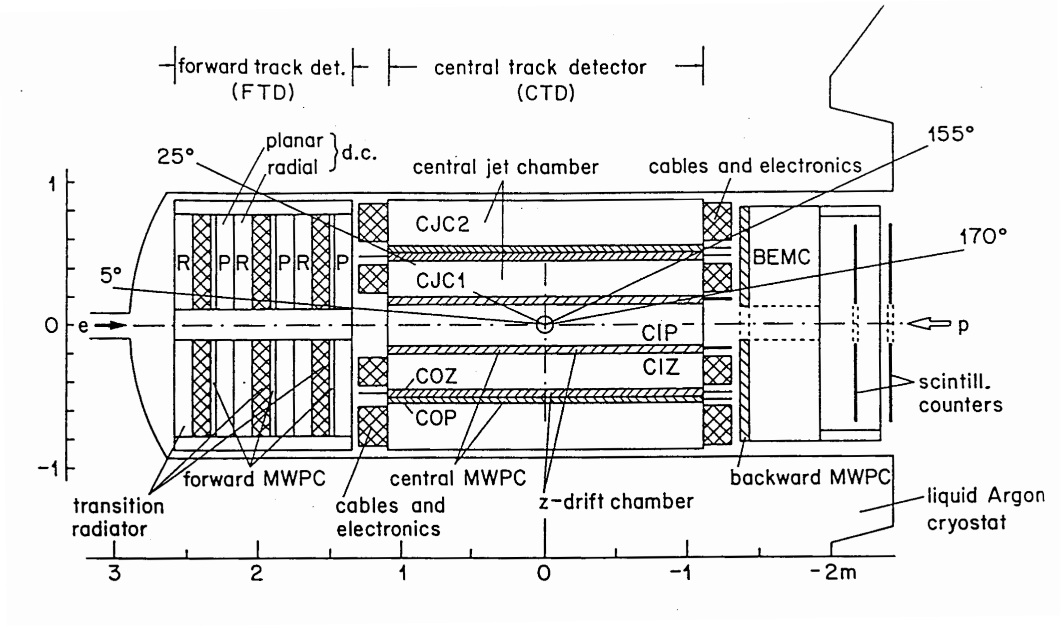


Figure 1.2: The H1 tracking system (r – z view). Figure was taken from [1, 2].

### 1.2.1 Tracking

The tracking system of H1 (see Fig. 1.2) provides simultaneously triggering, track reconstruction and particle identification. To obtain a good spatial resolution, tracking system was divided into three mechanically distinct parts – the central tracker (CTD), the forward tracker (FTD) and the backward drift chamber (BDC). Each of them is optimised for tracking and triggering in its angular region.

**Track reconstruction in the central region** is based on the two large concentric drift chambers, CJC1 and CJC2 (see Fig. 1.3), covering the polar angle range $15° < \theta < 165°$. The chambers have wires strung parallel to the beam axes (z-direction). The space resolution is about 170 $\mu$m for $x, y$ (or $r, \phi$) coordinates and 22 mm in the $z$ direction. Because of the high magnetic field, only tracks with transverse momentum $\geq$ 0.15 GeV will pass through both chambers. The momentum and angular resolutions are $\sigma_p/p^2 \leq 0.01$ GeV$^{-1}$ and $\sigma_\theta = 21\,\mathrm{mrad}$.

Two thin drift chambers, the central inner (CIZ) and central outer (COZ) z-chambers complement the measurement of charged track momenta. The sense wires are perpendicular to the beam axis. These two chambers provide track resolution about 300 $\mu$m in $z$, and 30 – 60 mm in $x, y$ direction. Linking these track elements to those obtained from the CJCs gives the final accuracy on both the longitudinal and the transverse momentum components.

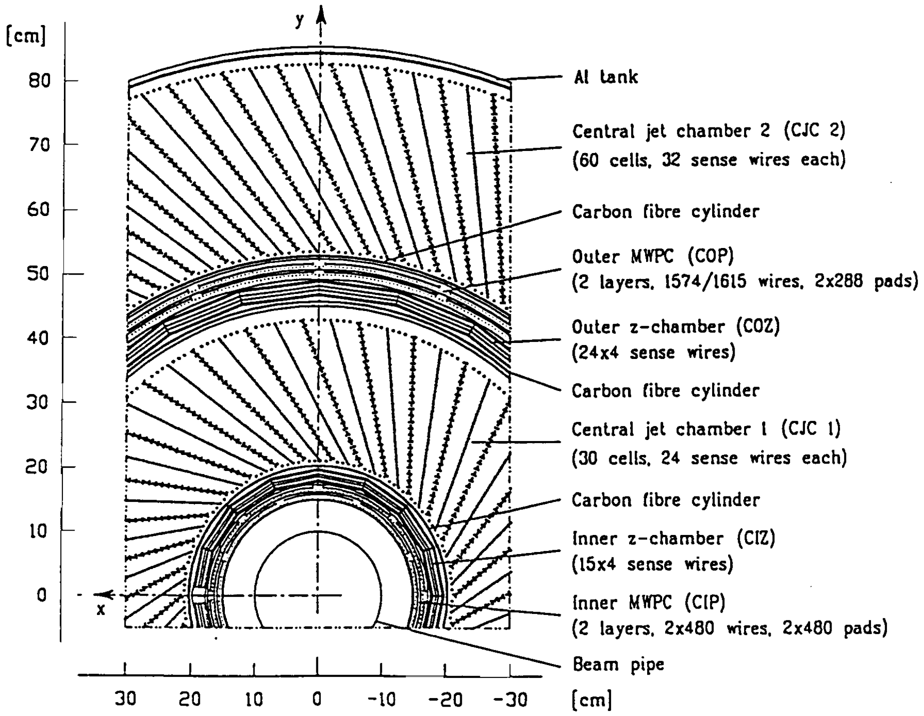


Figure 1.3: Central tracking system, section perpendicular to the beam. Figure was taken from [1, 2].

**The forward tracking detector (FTD)** provides an accurate measurement of charged particles in the forward direction $5^\circ < \theta < 30^\circ$. It consists of three supermodules. Each of them includes three differently oriented planar wire drift chambers designed to provide accurate $\theta$ measurements, a multiwire proportional chamber (FWPC) for fast triggering, a passive transition radiator and a radial wire drift chamber which provides accurate $r, \phi$ information. The planar chambers contain wires parallel to the beam direction and provide homogeneous spatial resolution in the plane perpendicular to the beam direction, whereas in the radial chambers the wires radiate outwards from the beam pipe – all wires are perpendicular to the beam direction and serve for precise momentum measurement.

**The backward drift chamber (BDC)** is designed to provide track segments for final state electrons that enter the backward region. It consist of four double layers of drift chambers in eight octants with sense wires strung parallel to the radial direction. The double layers are rotated with respect to one another by $11.5^\circ$ in $\phi$ for better spatial resolution (which is about $\sigma_r = 0.4\,\mathrm{mm}$ and $\sigma_\phi = 0.8\,\mathrm{mm}$).

### 1.2.2 Calorimetry

There are four subdetectors which provide energy measurement at the H1 – the liquid argon calorimeter, SPACAL, tail catcher and plug calorimeter.

**The liquid argon calorimeter (LAr)** provides the primary measurement of hadronic energy in H1. The scattered electron with the energy above $\simeq 100$ GeV is also detected in this device. The liquid argon technique has been chosen since it makes possible to reach a good stability, simplicity of calibration, fine transverse granularity and homogeneity of the signal response. The LAr covers the polar angle region $4^\circ < \theta < 154^\circ$ and is housed in a large cryostat inside the magnet solenoid. Placing the cryostat inside the magnet has the advantages of reducing both the size and weight of the calorimeter and the amount of dead material through which particles have to pass. The LAr is divided into inner electromagnetic and outer hadronic section, and is segmented into wheels which sit on rails inside the cryostat. Each wheel is azimuthally further segmented into eight identical octants. The total thickness of the electromagnetic section varies between 20 and 30 radiation lengths and the depth of hadronic part varies from 4.7 to 8 interaction lengths.

The LAr calorimeter is highly segmented in both sections with total of approximately 45 000 cells. The energy resolution of the electromagnetic part has been measured in beam tests to be $\sigma_{em}(E)/E \simeq (11\%/\sqrt{E}) + 1\%$, where $E$ is the energy of the incident particle in GeV. Similarly the resolution of the hadronic section was determined to be $\sigma_{had}(E)/E \simeq (50\%/\sqrt{E}) + 2\%$. The fine granularity of the LAr allows efficient $e/\pi$ separation, with an observed contamination of less than $10^{-3}$.

**The SPACAL** (backward lead scintillating fiber calorimeter) has replaced the BEMC (backward electromagnetic calorimeter) during the winter shutdown 1994-95. It consists of two sections, electromagnetic and hadronic. The primary goal of the finegrained electromagnetic section is the measurement of the energy and impact point of scattered electrons or positrons with high precision in the angular range $153^\circ \leq \theta \leq 178^\circ$. The hadronic section aims to measure electronic energy leakage from the electromagnetic section and to determine hadronic energy flow in the backward region. The depth of approximately $\sim 1\lambda$ of the hadronic section[2] together with $\sim 1\lambda$ of the electromagnetic part improves the $e/\pi$ separation capabilities. The $\pi^{\pm}$ rejection factor $\epsilon_e/(1-\epsilon_\pi)$ is larger than 100 ($\epsilon_e$ and $\epsilon_\pi$ are the electron and the charged pion identification efficiency, respectively). The resolution of the electromagnetic part is $\sigma_{em}(E)/E \simeq (10\%/\sqrt{E}) + 1.7\%$ and of the hadronic one $\sigma_{had}(E)/E \simeq (30\%/\sqrt{E}) + 2\%$. More information about SPACAL can be found in [3, 4].

**The tail catcher (TC)** uses the instrumented iron of the return yoke to provide a coarse measurement of the energy of hadronic showers leaking out of the LAr. The iron itself forms the absorber and covers almost the entire $4\pi$ solid angle. The energy resolution of the TC has been measured at the beam tests with pion beams to be $\sigma(E)/E \simeq 100\%/\sqrt{E}$.

**The plug calorimeter (PLUG)** covers the extreme forward direction $0.6^\circ < \theta < 3^\circ$ and is designed to close the gap in acceptance for energy flow measurements between the pipe and most forward part of the LAr calorimeter. Its primary task is to minimise the loss in transverse momentum due to hadrons emitted close to the beam pipe. The design value of the resolution of plug is $\sigma(E)/E \simeq 150\%/\sqrt{E}$.

### 1.2.3 Calibration of the LAr calorimeter

Because the main topic of the next chapters is related to the study of the distribution of the energy flow in the H1 detector, the calibration of the LAr calorimeter appears to be important.

The LAr calorimeter is non compensating. The charge output for hadrons is about 30% smaller than for the electrons of the same energy. Therefore additional corrections have to be applied to the hadronic response. The electromagnetic and hadronic showers can be distinguished by comparing their shapes (e.g. the fraction of the cluster energy deposited in the electromagnetic part of the calorimeter, and in the four most energetic cells of a cluster).

---

[2] $\lambda$ is the nuclear interaction length.

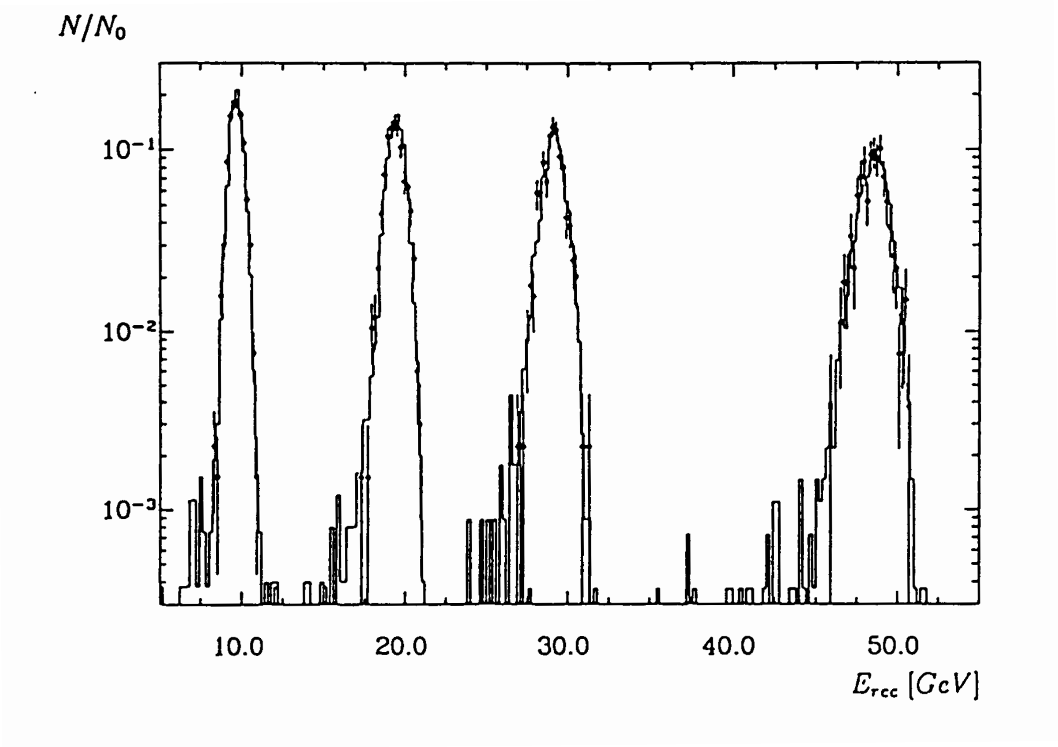


Figure 1.4: Reconstructed energy for data (histogram) and MC (points) for electron energies of 10, 20, 30 and 50 GeV. Figure was taken from [1, 2].

The energy calibration of the calorimeter modules was done in a test beam of $e^-$, $\pi^-$ and $\mu^-$ at CERN (see Fig. 1.4 or for more details [1, 2, 5, 6, 7]). Only one of each type of the calorimeter stacks was tested, since the achieved mechanical precision at the 1% level allowed to transfer calibration constants from one module to another. Data were taken with beam energies in the range 3.7 to 80 GeV for electrons and 3.7 to 205 GeV for pions.

As a result two calibration constants ($c^{exp}_{EMC}$, $c^{exp}_{HAC}$) were determined for each wheel[3], which transform the measured electric charge $Q_i$ per cell into energy deposited by electron showers:

$$E^{exp}_{rec} = c^{exp} \sum_i Q_i$$

The energy $E^{exp}_{rec}$ should be equal to the energy $E^{MC}_{rec}$ obtained from the MC simulation program which simulates the detector response of the initial energy of the beam particle:

$$E^{MC}_{rec} = c^{MC}\, E^{beam}$$

---

[3] The shortcut EMC (HAC) denotes the electromagnetic (hadronic) part of LAr calorimeter.

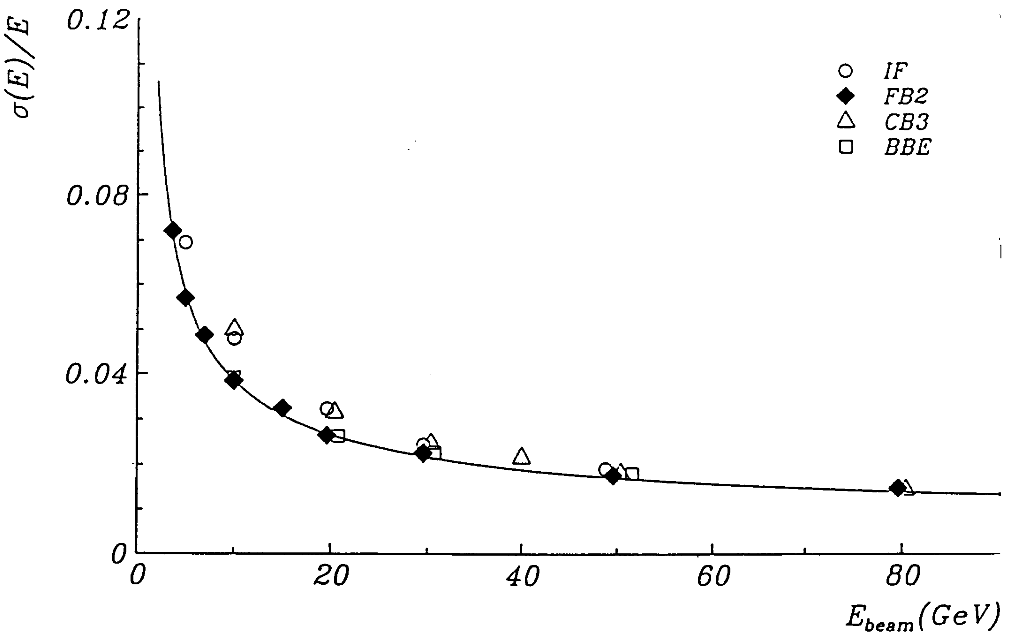


Figure 1.5: Energy resolution as a function of the electron energy for different wheels (BBE, CB, FB2 and IF). Solid line: parametrisation for FB2 wheel. Figure was taken from [1, 2].

Thus the corresponding experimental constant $c^{exp}$ can be expressed as:

$$c^{exp} = \frac{c^{MC}\ E^{beam}}{\sum_i Q_i} \tag{1.1}$$

From the test runs at CERN the following calibration constants were obtained: $c^{exp}_{EMC} = 3.55$ GeV/pC and $c^{exp}_{HAC} = 7.1$ GeV/pC. Energy resolution of the LAr calorimeter is plotted in Fig. 1.5.

There are also another ways how to check both the overall electromagnetic and hadronic energy scales by experimental data. The first method uses the cosmic muons, the second one the deep inelastic scattering events (see [1, 2]):

1. A fraction of the cosmic muons traversing the detector generates electrons. Their momentum $p$ can be measured in the central tracker CJC and can be compared with the energy $E$ measured in the LAr calorimeter. The overall response to muons varied from October 1991 to November 1994 by less than ±3%.

2. Deep inelastic scattering events for which both the hadronic jet and scattered electron are detected in the LAr calorimeter can be used for a direct comparison of the hadronic and electromagnetic energy measurement exploiting $p_t$ balance. The transverse components of the momentum of the

hadronic system $p_{t\,h}$ and electron $p_{t\,e}$ are calculated by summing vectorially the calorimeter cell energies. The measured energy balance is compatible with the simulation.

### 1.2.4 The luminosity system

The luminosity system serves to several purposes. Its main task is a fast relative luminosity measurement. It also provides electron beam monitoring for the HERA machine, absolute luminosity measurement in the interaction region, tagging of photoproduction events and energy measurement for electrons scattered under small angles and for photons from initial state radiation. The luminosity is determined from the rate of Bethe-Heitler events $ep \rightarrow ep\gamma$. The main source of the background is bremsstrahlung from the residual gas in the beam pipe, $eA \rightarrow eA\gamma$. At design luminosity these events are expected to be 10% of the $ep \rightarrow ep\gamma$ rate, but can be subtracted using data from electron pilot bunches. The luminosity is calculated as

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

where $R_{tot}$ is the total rate of the bremsstrahlung events, $R_0$ is the rate from the electron pilot bunches, $I_{tot}$ and $I_0$ are the corresponding electron beam currents and $\sigma_{vis}$ is the visible part of the $ep \rightarrow ep\gamma$ cross-section with acceptance and trigger efficiency included. The luminosity system is in more detail described in [1, 2].

### 1.2.5 Trigger

The purpose of the trigger system is to select interesting $e-p$ collision events and to reject background events, which are orders of magnitude higher than rate of $e-p$ collisions. Background collisions originate from the interactions of beam particles with pipe walls and with residual gas, from the synchrotron radiation and cosmic rays.

Because of the low $e-p$ cross section, large proton and electron beam currents are needed. This is realized by running in a multi-bunch mode. The bunch crossing time is very short ($\simeq 96\,ns$), and therefore a fast complex trigger using several levels with the increasing complexity has been achieved. Each subsequent level is only started if a given event has been accepted by the preceding one.

The trigger consists of "subtriggers[4]", which are controlled by electronic signals from different parts of the detector. The subtrigger may be activated for example if there are some tracks reconstructed in the track chambers, if the energy deposited in some calorimeter is higher than some threshold value, if the

---

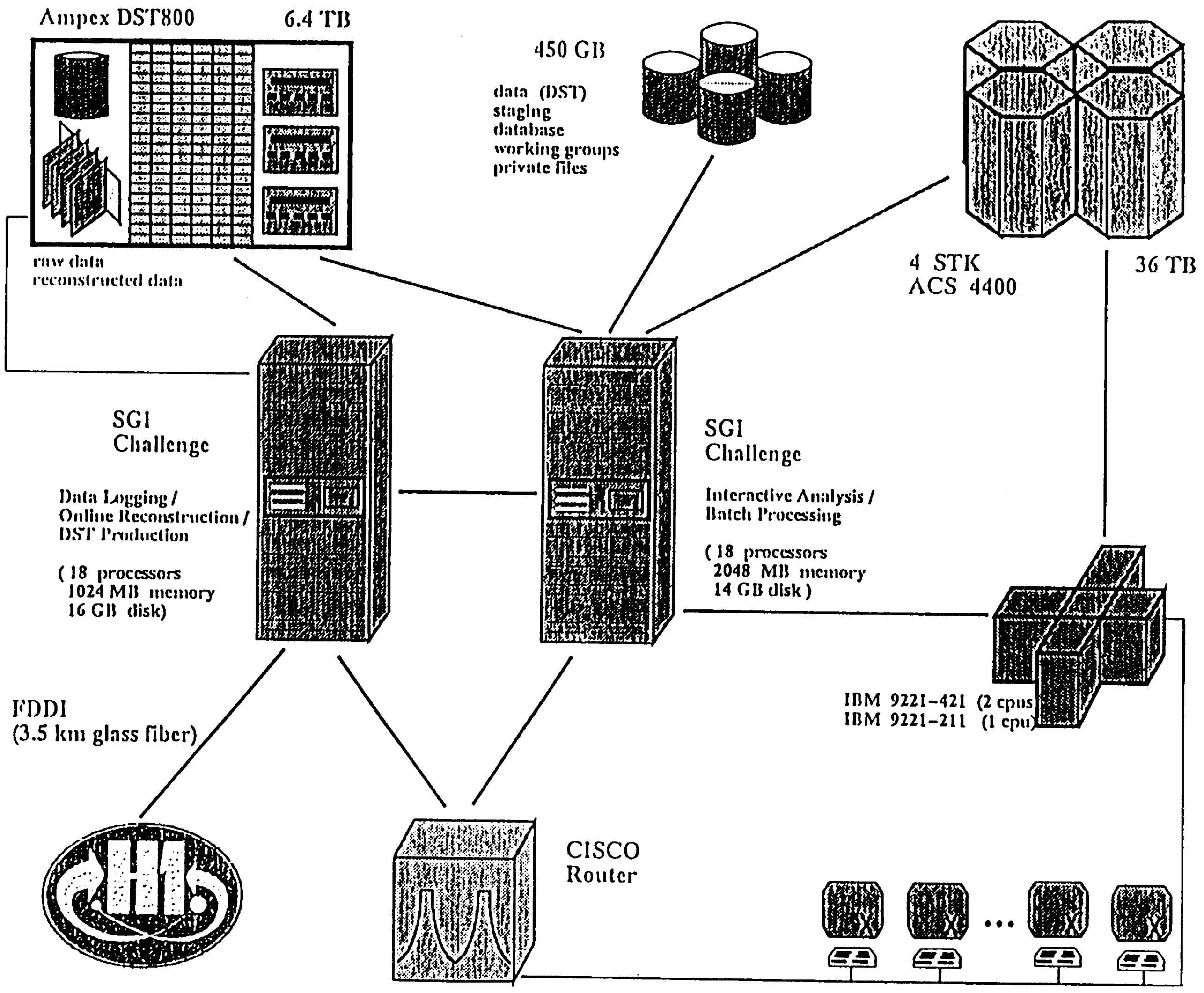
[4] There is 128 subtriggers at HERA denoted *s1, s2, ... s128.*

vertex of interaction is successfully reconstructed, etc. By an appropriate set of subtriggers different physical processes may be distinguished.

### 1.2.6 Data acquisition and computing environment

One triggered event results in some 3 Mbyte of raw digitised information. Data compression and formatting reduce this number to 50 – 100 Kbyte so the final data recording rate is about 1 Mbyte/s.

An overview of the H1 offline computing environment is given in the figure on page 17. The H1 experiment is using two multiprocessors SGI Challenge series computers for all main computing purposes. They are connected to the DESY computing infrastructure via fast UltraNet and Ethernet.


Ampex DST800
6.4 TB
raw data
reconstructed data
450 GB
data (DST)
staging
database
working groups
private files
4 STK
ACS 4400
36 TB
SGI
Challenge
Data Logging /
Online Reconstruction /
DST Production
( 18 processors
1024 MB memory
16 GB disk)
SGI
Challenge
Interactive Analysis /
Batch Processing
( 18 processors
2048 MB memory
14 GB disk )
IBM 9221-421 (2 cpus)
IBM 9221-211 (1 cpu)
FDDI
(3.5 km glass fiber)
H1
CISCO
Router
...

# Chapter 2

# Photoproduction and DIS

One of the basic characteristics of photon-proton processes is the virtuality $Q^2$ of the photon. If this quantity is close to zero, the photon is quasi-real (nearly on mass shell), and the process is historically called photoproduction. Interactions which involve a $Q^2$ larger then a few GeV$^2$ are named DIS (deep-inelastic scattering) processes. The main difference between the photoproduction and DIS consist in the different behaviour of the photon.

In the first generation of photon-nucleon fixed target scattering experiments photoproduction seemed to be very similar to hadron-hadron collisions. The photon looked like a hadron with the same quantum numbers as photon, i.e. $J^{PC} = 1^{--}$. This behaviour was quantitatively included in the VDM (Vector Meson Dominance Model). If the transverse energy of hadrons in the final state is small, the VDM describes well observed measurements.

The QCD theory predicts significant modifications to the VDM. According to QCD there exist processes, in which the photon couples directly to quarks (see Fig. 2.1). This leads to the hard scattering of a parton from photon with a parton from proton and observation of jets in the final state.

Results from the experiments with sufficiently high centre of mass energies have confirmed QCD predictions. Thus there are two possible processes (see Fig. 2.2):

**direct** photon interacts directly with a quark from the proton. In this case no internal structure of the photon is observed.

**resolved** photon splits into a $q\bar{q}$ pair before interacting with the parton in the proton. The photon structure functions of the photon can be defined and measured.

The characteristic time in which photon splits into a $q\bar{q}$ pair is $\sim 1/Q^2$. Therefore, as the $Q^2$ increases, the photon hadronic structure is less and less observable and direct processes dominates. Thus DIS can be used for investigating the internal structure of the proton by a point-like virtual photon.

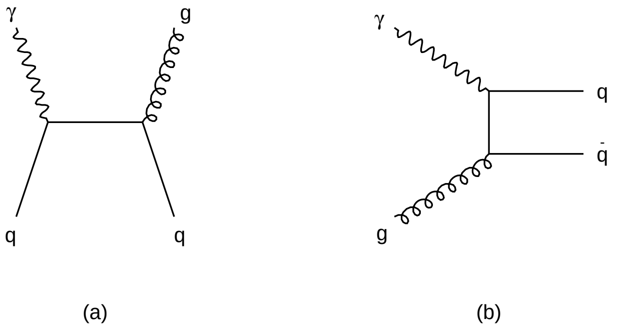


Figure 2.1: Diagrams of direct photon-nucleon interactions in leading order QCD: (a) the QCD Compton scattering, (b) the photon-gluon fusion.

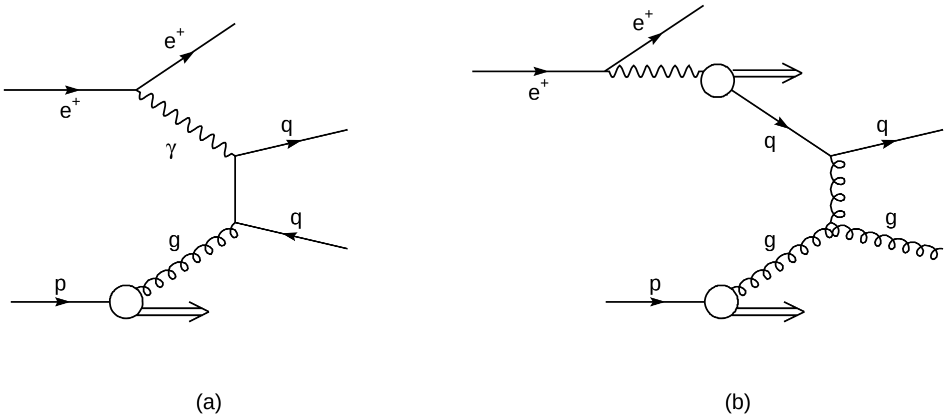


Figure 2.2: Diagrams of leading order QCD dijet production: (a) direct, (b) resolved.

Of course, there is no fixed boundary between the photoproduction and DIS. Recent DIS measurements have shown that the data in the low Bjørken-$x$ region have similar behaviour as the photoproduction. Therefore it would be very interesting to treat the transition region between this two classes of events.

## 2.1 Structure of the photon

The splitting of a photon into a quark-antiquark pair can be calculated in the QPM (Quark Parton Model). The probability $f_{q/\gamma}$ of finding a quark in the photon is given by:

$$f_{q/\gamma} = e_q^2 \, \frac{\alpha}{\pi} \, (x_\gamma^2 + (1 - x_\gamma)^2) \ln \frac{Q^2}{m_q^2} \tag{2.1}$$

where $e_q$ and $m_q$ are the charge and mass of a free quark, $x_\gamma$ is the energy fraction of the photon carried by the quark (see [8]). The probabilities $f_{q/\gamma}$ must be summed over all colours and flavours. As a result, the photon structure function $F_2^\gamma$ is obtained:

$$F_2^\gamma(x_\gamma, Q^2) = 2x_\gamma \sum_{n_f, n_c} e_q^2 \, f_{q/\gamma}(x_\gamma, Q^2) \tag{2.2}$$

QCD corrections to this simple QPM photon structure function can be calculated from the DGLAP evolution equations. The leading order of the QCD prediction is then given by [8]:

$$F_2^\gamma(x_\gamma, Q^2) = 3 \sum_{n_f} e_q^4 \, \frac{\alpha}{\pi} \, x_\gamma \, (x_\gamma^2 + (1 - x_\gamma)^2) \ln \frac{Q^2}{\Lambda_{QCD}^2} \tag{2.3}$$

The photon structure function $F_2^\gamma$ can be measured by deep inelastic electron-photon scattering experiments, from which the QCD parameter $\Lambda_{QCD}$ or $\alpha_s$ can be determined.

## 2.2 Electron-proton cross section

Electron-proton cross section mediated by photon exchange can be expressed in terms of virtual photon-proton cross sections:

$$\frac{d^2\sigma_{ep}(s)}{dy \, dQ^2} = A \, \sigma_{\gamma^\star p}^T(W^2, Q^2) + B \, \sigma_{\gamma^\star p}^L(W^2, Q^2) \tag{2.4}$$

where $\sigma_{\gamma^\star p}^T$ and $\sigma_{\gamma^\star p}^L$are the cross sections for transversely and longitudinally polarised photons, and $A$ and $B$ are known functions of $y$ and $Q^2$.

There exists a scale $\Lambda_\gamma^2$ such that for processes with $Q^2 > \Lambda_\gamma^2$ the virtual photon-proton cross sections decrease quickly with increasing $Q^2$. When $Q^2 < \Lambda_\gamma^2$, the virtual photon-proton cross sections can be approximated by the real photon-proton ones. The origin of $\Lambda_\gamma^2$ can be either kinematic (e.g. $W^2$) or dynamical (e.g. $m^2_{virt.meson}$). For $Q^2 < \Lambda_\gamma^2$ the Weizsäcker-Williams approximation (WWA) can be used for calculating the electron-proton cross section:

$$\left.\frac{d^2\sigma_{ep}(s)}{dy\ dQ^2}\right|_{WWA} = \frac{\alpha}{2\pi Q^2}\ \frac{1+(1-y)^2}{y}\ \sigma_{\gamma p}^{tot}(W^2). \tag{2.5}$$

# Chapter 3

# Monte Carlo Generators

The comparison of the experimental data with the theoretical predictions is one of the most important and interesting aims of the detector physics. Usually, the most convenient way to do so, is to take advantage of so called "event generators", which model physical processes based on the theoretical knowledge in the lowest order (LO) of perturbative QCD.

Because of most contemporary detectors such as H1 are rather complicated systems, it is impossible to compare measured data directly with the output of such event generators. Therefore, a program that simulates the response of the detector has to be used in addition to the event generators. Then, if the modelling of the physical processes in an event generator and the detector simulation program are correct, we obtain the same information as from the detector itself. Due to the technique used for event generators and detector simulation programs, they are usually called Monte Carlo (MC) simulations.

The information obtained from the Monte Carlo simulations is then treated in the same way as data from the detector. At the end of this process we are left with two final sets of events – MC and data. A comparison of the data with the theoretical expectations is now straightforward.

All Monte Carlo generators can be divided into two basic steps:

- **The hard scattering process** is the main part of MC program. It calculates the interaction of partons according the LO QCD rules. This part is basically the same for all MC programs, the difference may consist in the number of processes included in the model. Some generators also contain simulation of the initial and final radiation and additional interactions of remnants of incoming particles.

- **The hadronisation process**, in which the coloured partons are transformed into colourless hadrons, can't be calculated in the context of perturbative QCD and has to be treated phenomenologically in a MC fragmentation procedure. Mainly three different models of hadronisation are in use:

- The "Lund string model", which is used in ARIADNE, AROMA, LEPTO, PYTHIA and in most other generators for special processes. According to this model partons are connected by strings, which than fragment into hadrons.
- The formation and decay of colourless clusters (HERWIG). After the perturbative stage of parton generation, all final gluons are split into light quark and antiquark pairs, which are then combined into colourless clusters which decay into hadrons.
- A modified independent fragmentation model with subsequent global energy and momentum conservation is part of the programs COJETS, EUROJET and ISAJET, which do not simulate $e$-$p$ physics however.

For the purposes of the further analyses the HERWIG 5.9 and RAPGAP 2.9 generator were used.

## 3.1 HERWIG

HERWIG[1] is a general purpose generator for high energy physics processes. It includes the simulation of hard lepton-lepton, lepton-hadron and hadron-hadron scattering, soft hadron-hadron collisions and parton-shower approach for initial- and final-state QCD radiation.

A process of the simulation of an event can be factorised into the following subprocesses:

1. **Final state emission.** An outgoing virtual parton generates a shower of partons with lower virtuality. The amount of the emission depends on the upper limit of the virtual mass of the initiating parton. A branching algorithm correctly describes also the next-to-leading corrections to the distributions of soft partons, i.e. partons with momentum small compared with the hard scale $Q$.

2. **Initial state emission.** The theoretical analysis of this process is more complex than for the previous case of the final-state emission. Even to leading order, the structure function and associated radiation have been analysed quite recently for small $x$, which is the energy fraction of the incoming parton after the emission of initial state radiation. For lepton-hadron processes $x$ corresponds to the so called "Bjørken variable" (defined in eq. 4.5).

   Fortunately the initial state emission process factorises for any value of $x$ and can be described as a branching process suitable for Monte Carlo simulations.

---

[1] Hadron Emission Reaction With Interfering Gluons

3. **Elementary hard subprocess.** This can be computed exactly to a finite order in the perturbation theory. The hard subprocess momentum transfer scale $Q$ sets the boundary conditions for the initial and final state parton showers. There is a fairly large library of QCD and electroweak elementary subprocesses in HERWIG. Important function of the elementary subprocess is to set up the polarisations of any electroweak bosons or gluon jets that may be involved. These polarisations give rise to angular asymmetries and correlations in boson decays and jet fragmentation.

4. **Hadronisation process.** In order to construct a realistic simulation it is necessary to convert the partons into hadrons. This process takes place at a low momentum transfer scale, for which the strong coupling $\alpha_S$ constant is large and perturbation theory is not applicable. Thus a phenomenological hadronisation model has to be added to the above perturbative QCD calculations.

   There are three types of nonperturbative contributions to be considered: (a) the representation of the incoming partons as constituents of the incident hadrons; (b) the conversion of the emitted partons into outgoing hadrons; (c) the 'soft underlying event' associated with the presence of spectator partons. All of these can be referred as aspects of "hadronisation".

An important consequence of the factorisation theorem is that the distributions for any hard QCD process are obtained from the same four subprocesses described above. In lepton-lepton collisions we have to consider only the elementary hard subprocess and final state emission, while in interactions of hadrons (with leptons or hadrons) we have to consider also the emission from incoming parton and soft emission due to the presence of 'spectator' partons.

A big theoretical advantage of such an universal program as HERWIG is that the phenomenological parameters in the hadronisation model can be tuned simultaneously by fitting data at different machines and energies. This enhances the predictive power of the program.

### Cluster hadronisation model

A cluster hadronisation model in HERWIG is independent of the hard process and the energy. After the perturbative parton branching process, all outgoing gluons are split nonperturbatively into light ($u$ or $d$) quark-antiquark or diquark-antidiquark pairs. Thus each jet consists of a set of outgoing quarks and antiquarks, which can be formed into colour-singlet clusters.

Then the clusters fragment into hadrons. If a cluster has a low mass, it forms the lightest single hadron of its flavour. Its mass is shifted to the appropriate value by an exchange of momentum with a neighbouring cluster in the jet. If the cluster is massive enough, it can either decay into pairs of hadrons, or they are fragmented using an iterative fission model (which is very similar to the string fragmentation) until the masses of the fission products fall below the fission threshold.

### Soft underlying event – sue

In a typical resolved event[2] without the soft underlying interaction the photon and proton remnants escape most often undetected into the beam pipe. HERWIG enables to simulate a soft collision of these remnants. This leads to a redistribution of their originally high longitudinal energy among a larger number of secondary clusters, which decay according to a normal HERWIG prescription. In this way, *sue* may provide sizable pedestals under the hard jets also in the central region.

The necessity of adding an soft underlying event to the hard emission was analysed from the "pedestal height" in hadronic jet production. The determination of the amount of the soft underlying events in lepton-proton collisions is one of the main topics of this diploma work.

There exist several approaches for the underlying interaction. In HERWIG it is modelled as a collision of remnant clusters, which can be tuned by two parameters:

- ENSOFT – determines the multiplicity of a soft collision of beam remnants to be the same as in $p\bar{p}$ interaction at the energy ENSOFT * $E_{soft}$, where $E_{soft}$ is the total CMS energy of these remnants. In this thesis a default value ENSOFT$= 1$ has been used.
- PRSOF – determines the fraction of events, in which the soft underlying collision is generated. Since at H1 the parameter PRSOF is often called *"sue"*, this notation is used also in this thesis.

Eventual soft underlying interaction of remnants doesn't influence the basic hard scattering process cross section, but it may significantly change the number and properties of produced soft particles and consequently also such quantities as $E_t$ flow outside jets, etc. It may even distort the properties of hard jets. This is nicely illustrated in the internal H1 note [11].

The produced quarks that define the flavour of the soft clusters are taken to be $u$ or $d$ only. The cluster masses are chosen from the distribution

$$P(M) \propto (M - M_0)e^{-a(M-M_0)}$$

where $M_0 = 1\,\mathrm{GeV}$ and $a = 2\,\mathrm{GeV}^{-2}$.

The main features of the HERWIG generator are in detail described in [10] and [11].

---

[2] In all the following text the term "resolved event" denotes an event with the resolved gamma interaction (and similarly the "direct event").

## 3.2 RAPGAP

The program RAPGAP generates events in $e-p$ collisions where the interaction between proton and electron is mediated by the pomeron. It is applicable to photoproduction as well as DIS. Higher order gluon emission is simulated with the Colour Dipole Model (ARIADNE) and the hadronisation is performed using the "Lund string fragmentation model" (see below).

In the standard $e-p$ inelastic scattering a parton removed from a proton carries a colour charge. Therefore, the fragmentation string is spread between the proton remnant and a parton outgoing from the hard interactions. In diffractive scattering the proton stays intact, it radiates only colour neutral pomeron. The string then connects a parton from the interaction with the pomeron remnant.

The RAPGAP implementation and some phenomenological consequences are discussed in [13].

### 3.2.1 Lund string model of fragmentation

The lattice QCD studies for $q\bar{q}$ two-jet events produced in $e^+e^-$ annihilation support a linear confinement picture, i.e. the energy stored in the colour dipole field between a charge and an anticharge increases linearly with the separation between the charges. This assumption leads to a string model (see [12]). As the $q$ and $\bar{q}$ partons move in opposite direction to each other, a "colour flux tube" is stretched between them.

In order to obtain a Lorentz covariant description of the energy flow, the most straightforward way is to use the dynamics of massless relativistic string with only the longitudinal degree of freedom. Therefore an one-dimensional string is used for the description of the axes of a cylindrically symmetric flux tube. The string constant, i.e. the amount of energy per unit length, was deduced to be $\kappa \simeq 1\,\mathrm{GeV/fm}$. The constant $\kappa$ effectively corresponds to a "mass density" along the string.

As the $q$ and $\bar{q}$ move apart, the potential energy stored in the string increases, and the string may break by the production of a new $q'\,\bar{q'}$ pair. The system splits into two colour-singlet parts – the string $q\,\bar{q'}$ and $q'\,\bar{q}$. If the invariant mass or either of this string pieces is large enough, further breaks may occur. This procedure is repeated until only on-mass-shell hadrons remain. Here the hadron corresponds to a small piece of string with a quark in one end and an antiquark in the other.

The Lund model also involve the suppression of heavy quark production, u:d:s:c $\simeq$ 1:1:0.3:$10^{-11}$. Charm and heavier quarks are not expected to be produced in the soft fragmentation, but only in perturbative parton shower branching $g \to q\bar{q}$.

When the quark and antiquark form a meson, it is necessary to invoke an algorithm to choose between the pseudoscalar and vector mesons. Here the string model is not predictive. Qualitatively one expects a 1:3 ratio, from counting the number of spin states, multiplied by some normalisation factor, which should disfavour heavier states.

There exist two possible explanations of the baryon production in a string models. The first one is based on the idea, that the string may break not only into quark-antiquark, but also into diquark-antidiquark pair. According the second approach (often called “popcorn scenario”) the diquarks doesn't exist, but rather quark-antiquark pairs are produced one after the other. This second picture gives a less strong correlation in flavour and momentum space between the baryon and antibaryon.

The Lund string model is in more details described in [12].

# Chapter 4

# Jets and Reconstruction of Kinematic Variables

## 4.1 Jets and jet algorithms

One of the most important goal of the subnuclear physics is the description of parton-parton scattering processes. Unfortunately, the confinement property of the QCD makes the direct observation of coloured partons impossible. However, the properties of partons can be studied. The standard way how to do so is to work with "jets" – bunches of collimated, colourless particles which are the result of the hadronisation of scattered partons.

To obtain a good energy and space resolution, both the information from calorimeters and trackers were combined for the jet reconstruction. Low energy particle is hardly detected in the LAr calorimeter. If its energy is lower than about 200 MeV, it is usually not detected at all. Even for higher energetic particle the LAr calorimeter doesn't measure its whole energy. Thus the concept of "clim"[1] has been settled to improve the measurement of momentum at H1. The clim vectors are defined as momentum vectors determined from:

1. all calorimeter clusters ($\vec{p}_{clus}$)
2. all good tracks ($\vec{p'}_{trk}$) but with limited momentum, e.g. if

$$
\begin{aligned}
|\vec{p}_{trk}| \leq 400 \text{ MeV} \quad &\Longrightarrow \quad \vec{p'}_{trk} = \vec{p}_{trk} \\
|\vec{p}_{trk}| > 400 \text{ MeV} \quad &\Longrightarrow \quad \vec{p'}_{trk} = 400 \text{ MeV} \cdot \frac{\vec{p}_{trk}}{|\vec{p}_{trk}|}
\end{aligned}
$$

The comparison of the jet reconstructions using clims and clusters in a low $Q^2$ region has been studied by H. Rick in [14].

---

[1] CLIM is the abbreviation for Cluster + LImited Momentum tracks.

Many quantitatively different definitions of jet exist, but all of them aim to determine the basic properties of jet (i.e., energy and direction) as close as possible to the same quantities of the parton that has initiated the given jet.

One of the most often used jet-finding algorithm at low $Q^2$ region at HERA is the cone one.

### 4.1.1 Cone algorithms

Cone-type algorithms are mainly used in hadron-hadron physics, but also in photoproduction at HERA. According to these algorithms[2] jets are constructed from hadrons (or, in experimental language, calorimeter cells or clims) placed inside a cone which can be defined in several ways. There also exist various approaches of starting jet finding and merging overlapping jets. It makes the difference between various cone-type algorithms.

#### CDF algorithm

If not mentioned otherwise, the CDF algorithm was used all through this thesis. The distance $R_i$ of the hadron $i$ (or calorimeter cell $i$) from the jet centre is defined by the formula:

$$R_i = \sqrt{(\eta^i - \eta^{jet})^2 + (\phi^i - \phi^{jet})^2} \tag{4.1}$$

and has to be smaller than some given diameter $R$. Here $\phi^i$ denotes the azimuthal angle in radians and $\eta^i = -ln(tg\frac{\theta}{2})$ the pseudorapidity of the hadron $i$. The transverse energy $E_t^{jet}$, pseudorapidity $\eta^{jet}$ and azimuthal angle $\phi^{jet}$ are calculated as:

$$\begin{aligned} E_t^{jet} &= \sum_i E_t^i \\ \eta^{jet} &= \frac{1}{E_t^{jet}} \sum_i E_t^i \eta^i \\ \phi^{jet} &= \frac{1}{E_t^{jet}} \sum_i E_t^i \phi^i \end{aligned} \tag{4.2}$$

in which the sums run over all hadrons belonging to the jet.

The "seed" which initialise the process of jet finding is the hadron (or single calorimeter cell) with the highest transverse energy. Than all the hadrons from the selected cone of radius $R$ are excluded from the search for further seeds. This procedure is repeated until no hadrons remain. The jet quantities are initially calculated using the equations (4.2), where the seed stands instead of the jet at the first step. Equations (4.2) are then applied iteratively until a stable jet is found.

---

[2] A standard for cone algorithms was established at Snowmass meeting in 1990 ([15]).

It may happen that two stable jets overlap. If the overlapping transverse energy is higher than 75% of the smallest jet, jets are merged, otherwise the overlapping hadrons are split to the closest jet.

### 4.1.2 DECO algorithm

DECO is an algorithm developed by G. Knies at DESY. According to DECO jets are made by splitting the event into parts in a suitable way, by so called "DECOmposition".

The basic concept of this method is based on "event topological function $F$". In this thesis $F$ has been defined as

$$F = 1 - \frac{1}{M} \sum_{J=1}^{L} \sqrt{(P_J P)^2 / M^2 - P_J^2} \tag{4.3}$$

Here $P$ and $P_J$ are the 4-momenta (in the event CMS) of the full event and of the jet J, respectively. $M$ is the mass of all jets, $L$ is a number of jets into which the event will be decomposed. The quantity $F$ may be defined also in another ways, but it always has to be equal to zero by definition. Since the detector doesn't measure the momentum of particles precisely, the function $F$ is not exactly zero. Jet algorithm is based on finding a decomposition, for which the topological function achieves its minimal value. More details about this algorithm can be found in [16].

The main advantage of the DECO algorithm comparing to cone-type algorithms is that also jets flying close to the beam pipe are well recognised. This ability is not important in this analysis since jets are restricted to the central pseudorapidity area $-3 < \eta^{\star} < 0$. However a usage of more than one jet algorithm enables to determine the dependence of our results on the jet definition.

### 4.1.3 Correlation between jets and partons

The parton properties given by MC generator can be compared with the same properties of jets obtained after hadronisation procedure and simulation of detector response. Firstly, the parton have to be matched to a corresponding jet. This is done by calculating the distance between the given parton and jets. It is defined as

$$R^i = \sqrt{(\eta^{jet\,i} - \eta^{parton})^2 + (\phi^{jet\,i} - \phi^{parton})^2} \tag{4.4}$$

The minimal distance $R^i$ is then find out and the parton is matched to the jet $i$.

Fig. 4.1 shows the jet parton correlation for the transverse energy $E_t$, pseudorapidity $\eta$ and azimuthal angle $\phi$ for the DIS sample. The photoproduction behaves similarly (not shown). One can see that the correlation are not perfect, but allow a reasonable treatment of the experimental data.

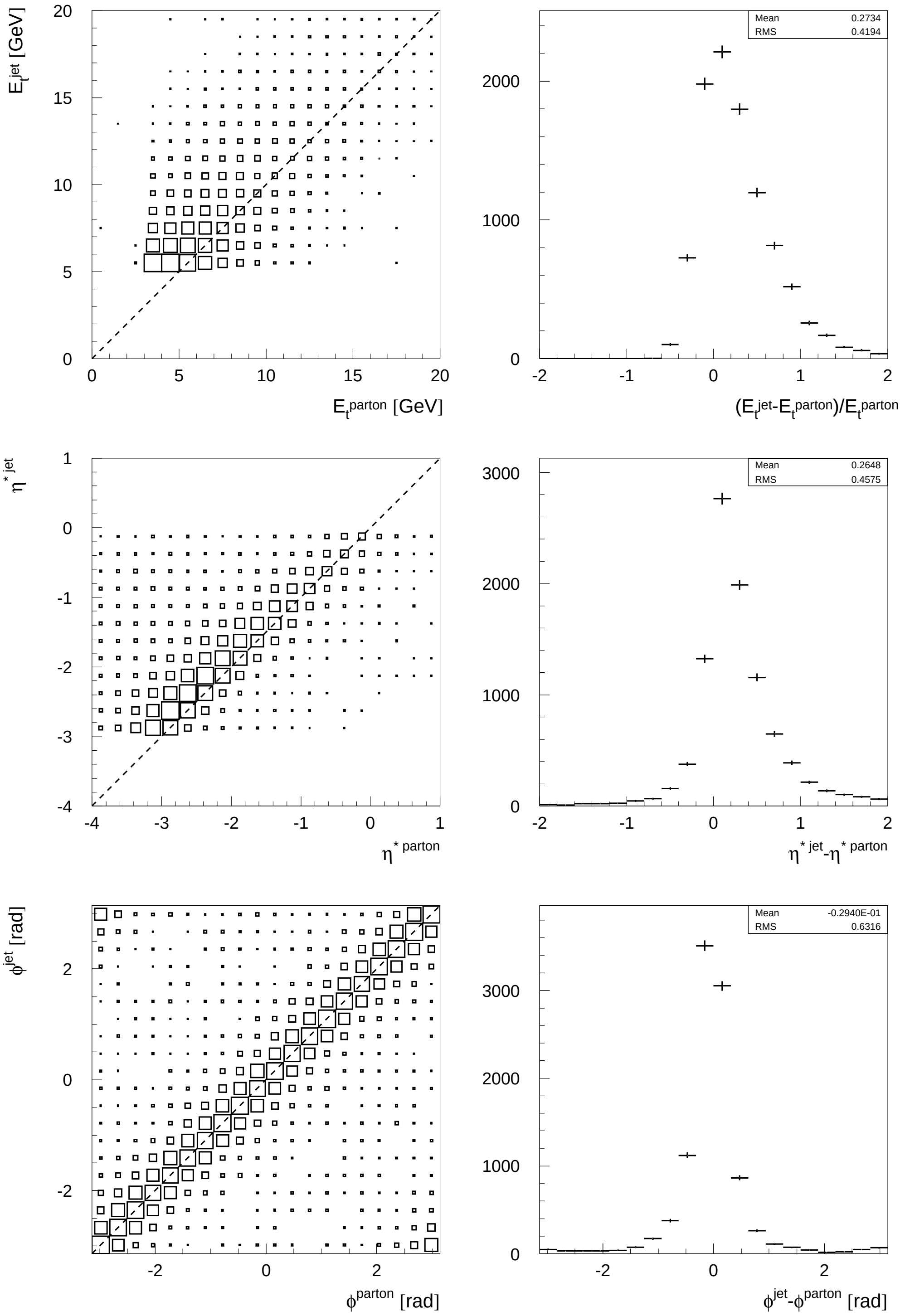
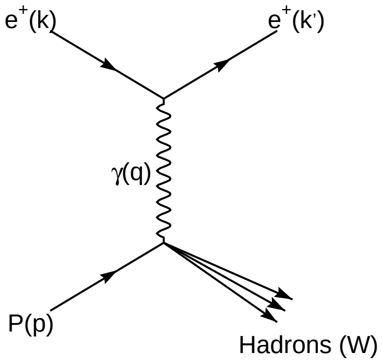


Figure 4.1: Correlation between partons and jets for transverse energy $E_t$, pseudorapidity $\eta$ and azimuthal angle $\phi$ for the DIS sample. CDF algorithm has been used for jet identification.

## 4.2 Kinematics of electron proton scattering

The kinematics of electron-proton deep inelastic scattering can be fully described by three independent variables. One of them, the centre-of-mass energy $\sqrt{s_{ep}}$, is fixed at HERA to the value of 296 GeV. Thus we are left with only two independent variables. The most often used are Bjørken $x$, which can be interpreted as the momentum fraction of the parton in the proton[3], $y$, the relative energy loss of the electron in the proton rest frame, $Q^2$, the negative four-momentum squared of the gauge boson exchanged between the electron and proton, and $W$, which corresponds to the sum of four-momentums of the hadronic final state (see Fig. 4.2):



$$\begin{aligned} s_{ep} &= (p+k)^2 \\ Q^2 &= -q^2 = -(k-k')^2 \\ x &= \frac{Q^2}{2(p\cdot q)} \\ y &= \frac{p\cdot q}{p\cdot k} \\ W &= (p+q)^2 \end{aligned} \tag{4.5}$$

Figure 4.2: Kinematics of $e-p$ scattering.

## 4.3 Reconstruction of the kinematic variables

$Q^2$, $x$ and $y$ can be expressed in directly measurable quantities such as energy of incoming electron $E_e$, energy of scattered electron $E'_e$ and the angle $\theta$ between the scattered electron and the initial proton direction:

$$\begin{aligned} Q^2 &= 4E_eE'_e\cos^2\frac{\theta}{2} \\ y &= 1-\frac{E'_e}{E_e}\sin^2\frac{\theta}{2} \end{aligned}$$

---

[3] This is satisfied as long as $Q^2 \gg m_q^2$.

$$x = \frac{Q^2}{y s_{ep}}$$

There are some other ways of extracting these basic variables from measured quantities of the hadronic final state. Moreover, it is possible to combine more methods of extraction $Q^2$, $x$ or $y$ and increase the precision of its determination or even to calculate the amount of the initial state radiation.

## 4.4 Determination and resolution of $\mathbf{x}_\gamma$

The photon splits into two partons in resolved interactions. Thus the momentum fraction $x_\gamma$ of the parton which interacts with the parton from the proton appears to be particularly important quantity. It can be determined by several methods. Two of them were used for purposes of this thesis:

**"Pseudorap. method"**

$$x_\gamma^{jets} = \frac{E_t^{jet1} e^{-\eta^{jet1}} + E_t^{jet2} e^{-\eta^{jet2}}}{2E_\gamma} \tag{4.6}$$

or

**"$\mathbf{E - p_z}$ method"**

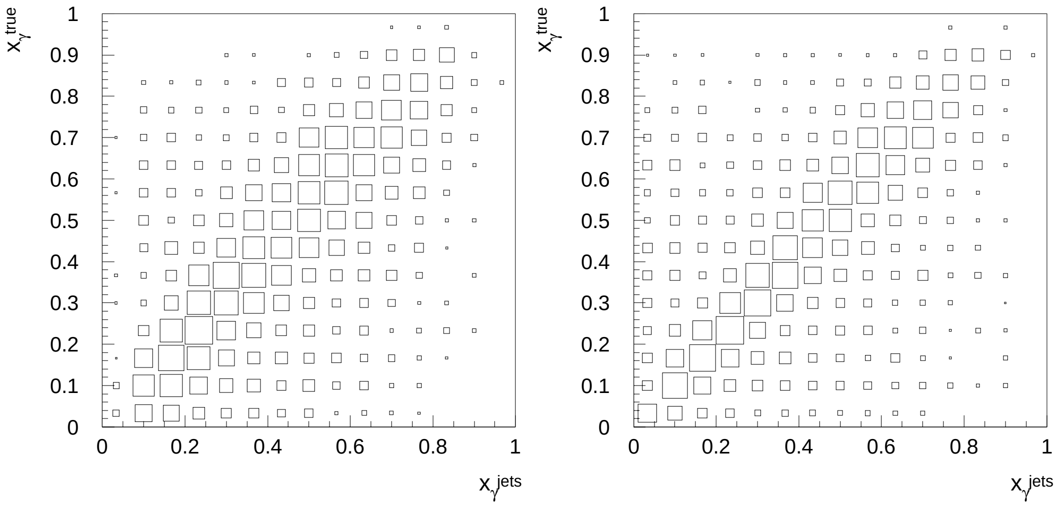
$$x_\gamma^{jets} = \frac{E^{jet1} - p_z^{jet1} \quad + \quad E^{jet2} - p_z^{jet2}}{\sum_{hadrons}(E_i - p_{z_i})} \tag{4.7}$$

Here superscript *"jet1"* (respectively *"jet2"*) denotes the jet with the highest (second highest) transverse energy.

To prove which of the two previous definitions is more suitable for the data treatment, a correlation between true $x_\gamma^{true}$, which was generated by HERWIG, and the value of $x_\gamma^{jets}$ determined by the "pseudorapidity" and "$E - p_z$" method after the simulation of hadronisation and the detector response was made. Fig. 4.3 shows the qualitative picture of this correlation.

The resolution of $x_\gamma$ can be expressed also quantitatively. In Fig. 4.4, $x_\gamma^{true}$ has been divided into five intervals and for each of them the $x_\gamma^{true} - x_\gamma^{jets}$ is shown. The mean values and the mean root squares (RMS) are summarised in Tab. 4.1.

In the ideal case, the mean value of $x_\gamma^{true} - x_\gamma^{jets}$ together with the RMS should be equal to zero. Since the detection efficiency and precision are not perfect, this requirement is not well satisfied and the value of $x_\gamma^{jets}$ depends on the jet algorithm and definition formula used for its computation. One can see in table 4.1 that the best correlation between $x_\gamma^{true}$ and $x_\gamma^{jets}$ is achieved for CDF algorithm and $x_\gamma$



(a)

(b)

Figure 4.3: Correlation between the $x_\gamma^{true}$ generated in HERWIG for DIS and $x_\gamma^{jets}$ computed after simulation of the detector response using two different methods: "pseudorapidity" one in the picture (a), and "$E - p_z$" one in the picture (b).

defined by the "$E-p_z$ method". Therefore the CDF algorithm and formula (4.7) are used in the following text.

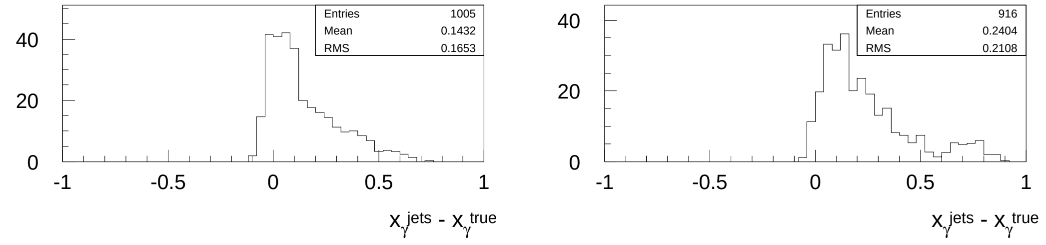
Another conclusion obtained from table 4.1 (or Fig. 4.4) is the fact, that the mean root squares reach values typically from 0.15 to 0.20. It means, that the width of $x_\gamma$-bins in any further mentioned distribution shouldn't be greater than 0.2.

## 4.5 Correction for the detector response

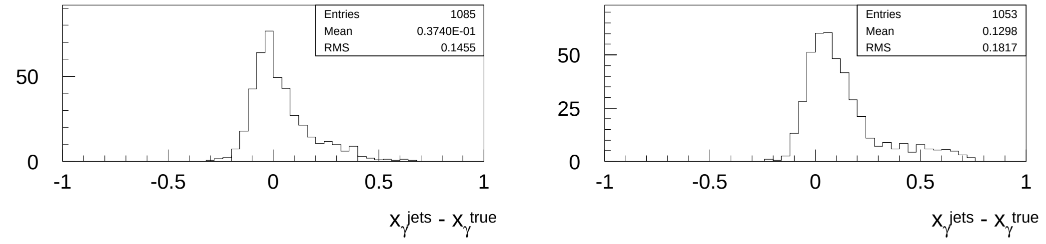
The relationship of the measured quantities (e.g. $p_t^{jet}$, $x_\gamma^{jet}$) with the directly unobservable parton quantities (e.g. $p_t^{parton}$, $x_\gamma^{parton}$) can be achieved only with the help of a Monte Carlo simulation program which models the predictions of a theory together with the response of the detector.

Experimental physicists usually correct the detector response with the help of MC simulations. Since there are many different fragmentation and hadronisation models available, theoreticists can choose any of them for comparison with the experimental data.

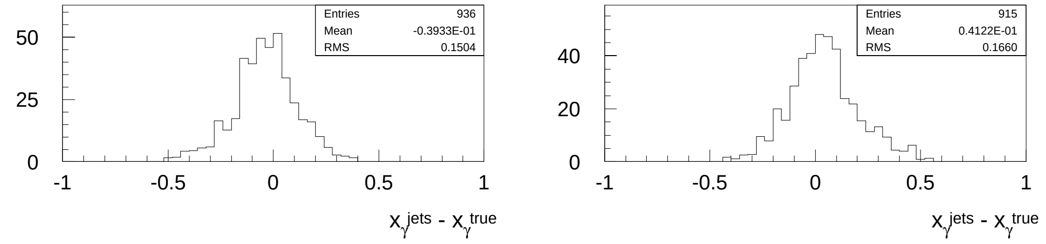
The simple and widely used method which corrects measured results for detector effects is a "bin-by-bin correction". It's main idea is based on the fact, that the influence of a measurement by a detector can be estimated by a comparison of the distributions (or quantities) obtained before and after MC simulation of the detector response. Thus, the bin-by-bin correction factors $c^i$ for a given



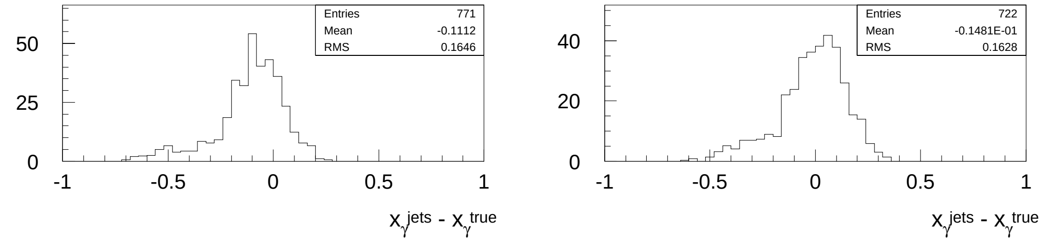
CDF, $0.0 < x_\gamma^{true} < 0.2$ DECO, $0.0 < x_\gamma^{true} < 0.2$



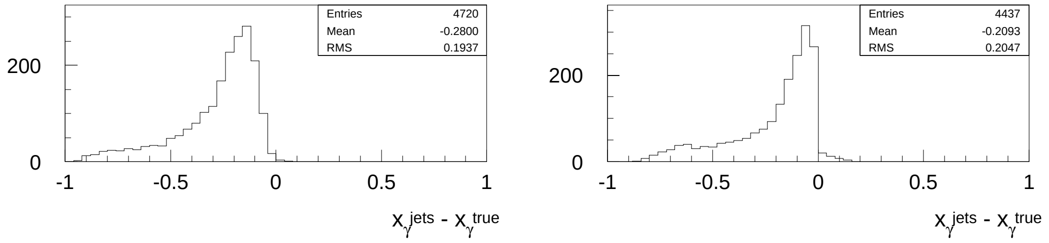
CDF, $0.2 < x_\gamma^{true} < 0.4$ DECO, $0.2 < x_\gamma^{true} < 0.4$



CDF, $0.4 < x_\gamma^{true} < 0.6$ DECO, $0.4 < x_\gamma^{true} < 0.6$



CDF, $0.6 < x_\gamma^{true} < 0.8$ DECO, $0.6 < x_\gamma^{true} < 0.8$



CDF, $0.8 < x_\gamma^{true} < 1.0$ DECO, $0.8 < x_\gamma^{true} < 1.0$

Figure 4.4: The distribution of $x_\gamma^{true} - x_\gamma^{jets}$ in five intervals of $x_\gamma^{true}$ generated in HERWIG for DIS. The variable $x_\gamma^{jets}$ is calculated using "$E - p_z$ method" and the CDF jet algorithm (left histograms) or DECO algorithm (right histograms).

|  |  | $x_\gamma^{true}$ | 0.0-0.2 | 0.2-0.4 | 0.4-0.6 | 0.6-0.8 | 0.8-1.0 |
|---|---|---|---|---|---|---|---|
| C | “Pseudorapidity method” | mean | 0.16 | 0.06 | -0.01 | -0.08 | -0.27 |
| D |  | RMS | 0.18 | 0.16 | 0.17 | 0.19 | 0.22 |
| F | “$E-p_z$ method” | mean | 0.14 | 0.04 | -0.04 | -0.11 | -0.28 |
|  |  | RMS | 0.17 | 0.15 | 0.15 | 0.16 | 0.19 |
| D | “Pseudorapidity method” | mean | 0.15 | 0.04 | -0.07 | -0.14 | -0.36 |
| E |  | RMS | 0.16 | 0.14 | 0.15 | 0.17 | 0.22 |
| C | “$E-p_z$ method” | mean | 0.24 | 0.12 | 0.04 | -0.01 | -0.21 |
| O |  | RMS | 0.21 | 0.18 | 0.17 | 0.16 | 0.20 |

Table 4.1: The mean values and mean root squares of the distribution $x_\gamma^{true} - x_\gamma^{jets}$ (i.e. the resolution of $x_\gamma$).

distribution are defined from two corresponding histograms and are defined as:

$$c^i = \frac{N^i_{MC\,Had}}{N^i_{MC\,Det}} \tag{4.8}$$

where $N^i_{MC\,Had}$ ($N^i_{MC\,Det}$) is a number of events before (after) the simulation of the detector response in the bin $i$. The correction factors are then applied on the measured data using the formula:

$$N^i_{DT\,Had} = c^i \cdot N^i_{DT\,Det} \tag{4.9}$$

Here $N^i_{DT\,Had}$ and $N^i_{DT\,Had}$ denotes the same quantities as $N^i_{MC\,Had}$ and $N^i_{MC\,Det}$ but for experimental data.

### 4.5.1 Purity

An important quantity characterising the efficiency of the bin-by-bin correction is “purity”. It is defined (e.g. in [18]) as

$$P^i = \frac{N^i_{stay}}{N^i_{stay} + N^i_{leave}} \tag{4.10}$$

Here $P^i$ denotes the purity of the bin $i$, $N^i_{stay}$ is a number of events which have stayed after the simulation of detector response in the same bin as they were generated in, while $N^i_{leave}$ is a number of events which have migrated from the bin $i$ to another ones. Thus, the higher is the value of $P^i$, the better is the correlation between the measured and corrected quantity.

Purity of $x_\gamma$ for different jet algorithms and definitions of $x_\gamma$ is compared in Fig. 4.5.

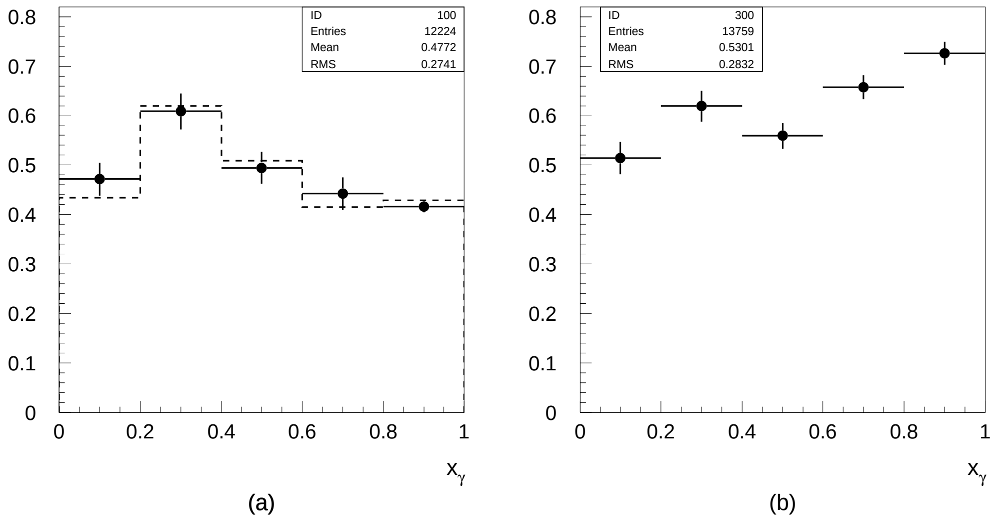


Figure 4.5: Purity of $x_\gamma$: (a) between the measured and generated $x_\gamma$ , (b) between the measured $x_\gamma$ and $x_\gamma$ corrected for the detector response. Full lines correspond to the the "$E - p_z$ method", while dashed lines denote the "pseudorapidity method". The CDF jet algorithm was used.

# Chapter 5

# Data Selection

Events used for further physics analysis were taken during the 1995 running period, after the SPACAL calorimeter had been installed. The integrated luminosity achieved the value of 1.045 $\text{pb}^{-1}$.

One typical two-jet photoproduction event is shown for illustration in the figure on the next page.

Data selection can be divided into two chronological steps:

- *online selection* is performed by a sophisticated trigger system (see [1, 2] or section 1.2.5). Its basic task is to save as much physically interesting data as reasonably possible and to remove background events. Online selection is carefully checked by physicists-experts, because of the bad triggering could negatively influence measured data and consequently physics results obtained from its treatment.
- To remove remaining background events, an *offline selection* has to be applied.

## 5.1 Event selection criteria

In following, two classes of events are treated:

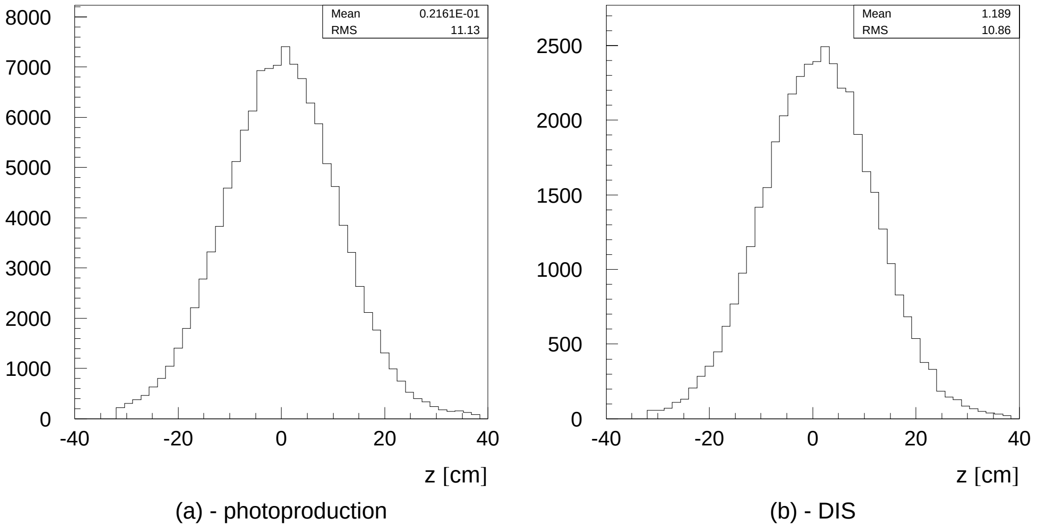
**Photoproduction sample** – electron has to be measured in the electron tagger detector. This requirement restricts the virtuality of photon $Q^2$ to below $0.01\,\text{GeV}^2$. An photoproduction event is selected only if there was activated at least one of the subtriggers s82 and s83.

**DIS sample** – electron has to be detected in SPACAL. Its acceptance restricts the $Q^2$ to the region $0.9 < Q^2 < 150\,\text{GeV}^2$. To avoid any uncertainty that could appear on the edges of the SPACAL calorimeter, a cut on $Q^2$ was applied: $1.4 < Q^2 < 50\,\text{GeV}^2$. At least one of the subtriggers s2, s3, s4, s12, s53, s56, s60 has to be active.

Events in both previous samples were selected only if they fulfilled the following criteria:

1. The distance of the reconstructed event vertex from the nominal interaction position along the beam axis (which is denoted as $z$-axis at H1) is less than 35 cm:
$$|z_{vtx} - z_{nom}| < 35\,\mathrm{cm} \tag{5.1}$$
The distribution of the z-coordinate of the event vertex is approximately Gaussian with $\sigma \simeq 10\,\mathrm{cm}$ (see Fig. 5.1).

   This cut removes events coming from the satellite bunches and proton-gas background occurring outside of the beam crossing area.



(a) - photoproduction (b) - DIS

Figure 5.1: Distributions of z-vertexes for the photoproduction (a) and DIS (b).

2. The quantity $y$ defined in the equation (4.5) is in the interval
$$0.25 \leq y \leq 0.7 \tag{5.2}$$
This cut restricts the energy in the electron tagger to $8\,\mathrm{GeV} < E_{tag} < 20\,\mathrm{GeV}$ for photoproduction and thus selects the $y$ (or $E_{tag}$) region, where the acceptance of the electron tagger is reasonably large. It reduces the possible background, which is mostly caused by random coincidences of two different events in one bunch crossing – the Bremsstrahlung process and the interaction of the proton beam with residual gas in the vertex region[1]. The photon from Bremsstrahlung is misidentified in the electron

---

[1] Background events induced by electrons are almost completely suppressed by requiring at least one reconstructed track in the central tracker already at the L4-level of online selection.

tagger as scattered electron and the beam gas interaction is then considered to be an photoproduction event. The background distribution peaks at higher energy than the data distribution and so the cut $E_{tag} < 20\,\mathrm{GeV}$ suppresses most of the background events. The scattered electron energy $E_{tag}$ for DIS case is displayed in Fig. 5.2.

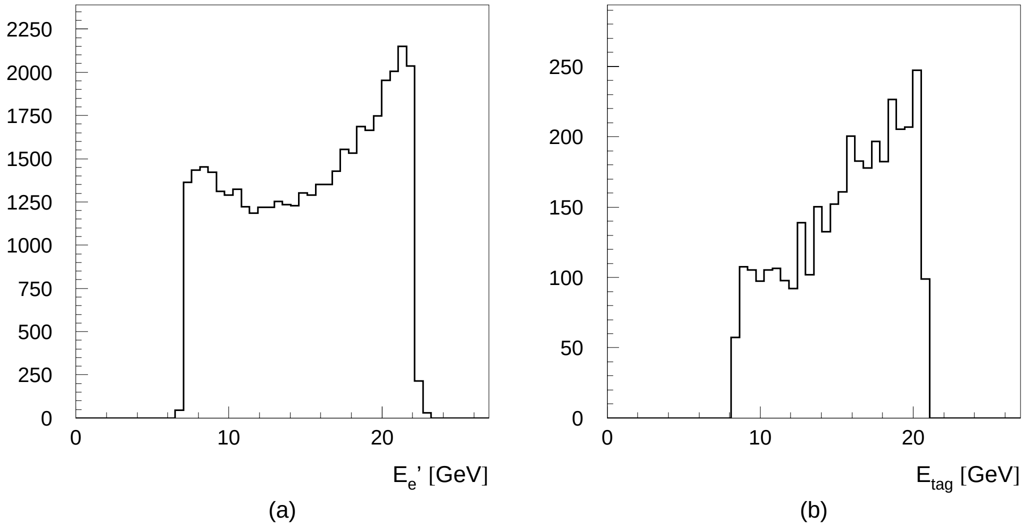


Figure 5.2: Distribution of the scattered electron energy $E_{tag}$ deposited in the electron tagger detector for the case of DIS. Picture (a) shows the data before applying all cuts, (b) after applying them.

3. The difference of the total energy $E$ of the event and the z-component of the total momentum $p_z$ (see Fig. 5.3) is between

$$35\,\mathrm{GeV} < E - p_z < 75\,\mathrm{GeV} \tag{5.3}$$

This cut is based on the conservation of the quantity $(E - p_z)$ of the whole event before and after collision:

$$\begin{aligned}
(E - p_z)^{initial} &= (E - p_z)^{final} \\
E_{e^-} - p_{z_{e^-}} + E_p - p_{z_p} &= \sum_{i=all} (E^i - p_z^i) \\
\{27.5 - (-27.5) + 920 - 920\}\ \mathrm{GeV} &= \sum_{i=all} (E^i - p_z^i) \\
55\ \mathrm{GeV} &= \sum_{i=all} (E^i - p_z^i)
\end{aligned}$$

Here the sums $\sum_{i=all}$ run over all particles in the final state (calorimeter cells).

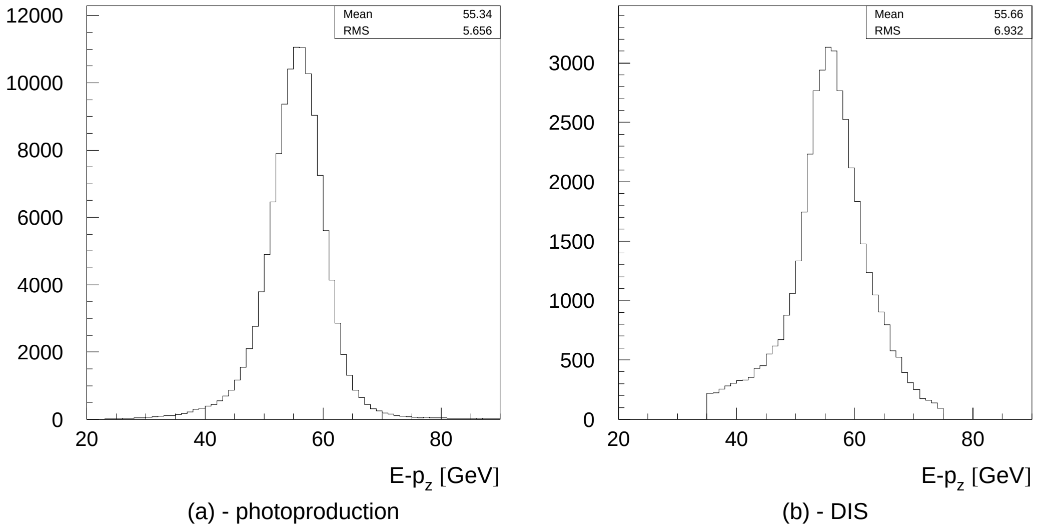


Figure 5.3: Distribution of $E - p_z$ of the whole event for photoproduction (a) and DIS (b).

The main aim of this cut is to remove interactions of proton with the beam gas, since in this case the $(E - p_z)$ of the whole event equals zero (there is no electron contribution to the $(E - p_z)$ term).

Particles (and also whole jets) which moves close to the proton direction have positive values of $p_z$ and thus their $(E - p_z)$ is close to zero, while particles that go close the beam axis in the electron direction, approximately fulfil relation $(E - p_z) = 2E$ and so their contribution to the global $\sum_{i=all}(E^i - p_z^i)$ is large. It means, that the cut (5.3) also removes events for which a big amount of energy has escaped in the beam pipe in the electron direction or through a space gap between the electron tagger and SPACAL.

4. There are at least two jets with the transverse energy higher than 5 GeV

$$E_t^{jet} > 5\,\mathrm{GeV} \tag{5.4}$$

and the pseudorapidity in between

$$-3 < \eta^{\star\,jet} < 0 \tag{5.5}$$

Here $\eta^{\star\,jet}$ denotes the pseudorapidity of a jet in the $\gamma$-proton centre of mass system[2]. Jets were found using the CDF cone algorithm. This cut on $\eta^{\star\,jet}$ was chosen since the comparison of the scattered partons and the measured jets in MC simulation has shown, that there are many “fake” jets

---

[2] In all this text $\eta$ corresponds to pseudorapidity in the laboratory system, while $\eta^\star$ in the $\gamma$-proton centre of mass system.

that doesn't correspond to any parton and are created artificially by the H1 detector, if the pseudorapidity is out of the region $-3 < \eta^{\star jet} < 0$. Only two jets with highest $E_t$ were used for computation of physical quantities.

The same selection criteria were also applied to the Monte Carlo samples. The HERWIG simulation included 100 000 events of the gamma-resolved generation and 100 000 events of the direct generation for both photoproduction and DIS cases. In RAPGAP, the same numbers of direct and resolved events (again 100 000) were generated, but only for the DIS sample. All on-line selected data are in H1 experiment saved on DST tapes. After the data-taking period all events have to be reprocessed by H1PHAN and H1REC programs, which reconstruct the physical quantities (e.g. tracks, momenta, particle identification, ...) from the electronic signal obtained from the detector, correct measurement to some detector effect and also identify the kind of physical process observed in a given event (e.g. photoproduction, high $Q^2$, exotic data, ...). 123 748 of photoproduction and 41 504 of DIS events were chosen from these DSTs and saved on the N-tuples. After applying previous selection criteria the following numbers of events were analysed:

| | Number of events |
|---|---:|
| Photopr. – data | 8 450 |
| Photopr. – HERWIG | 35 606 |
| DIS – data | 2 947 |
| DIS – HERWIG | 8 456 |
| DIS – RAPGAP | 14 989 |

Table 5.1: Numbers of events which left after applying all the selection criteria.

## 5.2 Monte Carlo samples

The direct and resolved events were generated separately, which allows to mix them in different weights.

The transition of the quark distribution functions of the real photon ($f_q^{real}$) to the virtual photon one ($f_q^{virt}$) is so far incalculable. However, there exist some models which interpolate between $f_q^{real}$ and $f_q^{virt}$. The method used in our Monte Carlo samples can be expressed in the following formula (for more details see [17]):

$$f_q^{virt}(x, P^2, Q^2) = f_q^{real}(x, P^2) L(P^2, Q^2, \omega) \tag{5.6}$$

where $L$ is a suppression factor defined as:

$$L \equiv \frac{\ln[(P^2+\omega^2)/(Q^2+\omega^2)]}{\ln[(P^2+\omega^2)/(\omega^2)]} \tag{5.7}$$

Here $P^2$ denotes the hard scattering scale, $Q^2$ the virtuality of the photon and $x$ is the Bjørken variable ($Q^2$ and $x$ were already defined in equation (4.5)). An analogous relation can be written for gluon distribution functions – in that case $L$ is replaced by $L^2$ in the equation (5.6).

The parameter $\omega$ in the equation (5.7), which is usually called *"phomas"*, specifies the behaviour of the suppression factor $L$. The smaller is the $\omega$, the stronger is the suppression factor $L$ already for weekly off-mass shell photons and the smaller is the value of $Q^2$ for which the direct processes start to dominate. In general, $\omega$ is a function of $x$. In our region of photon virtuality, $Q^2 < 50\text{GeV}^2$, parameter $\omega$ can be considered in the first approach as $x$ independent.

Besides the $\omega$, which is used in both HERWIG and RAPGAP, there is another tunable parameter in HERWIG called *"sue"* (see section 3.1 for more details). The *sue* defines the fraction of resolved events that contain additional soft underlined interaction. The HERWIG MC was originally generated with $sue = 0.25$ (it means that 25% of resolved events contain the additional interaction of $\gamma$ and proton remnants). Using a weighting procedure, any other fraction of *sue* can be obtained from already generated Monte Carlo sample.

As one can see from the formula (5.7), the factor $L$ becomes insensitive to $\omega$ as the virtuality $Q^2$ falls down. Thus the photoproduction sample, for which $Q^2 < 0.01$, can't be influenced by this parameter and only *sue* can be tuned.

## 5.3 Comparison of MC with Data

To obtain reasonable results from a treatment of measured data, the data have to be well described by the MC simulation. Thus, the comparison of both this samples appears to be particularly important.

If not mentioned otherwise, the following values of parameters will be used for the comparison of the MC samples with the measured data:

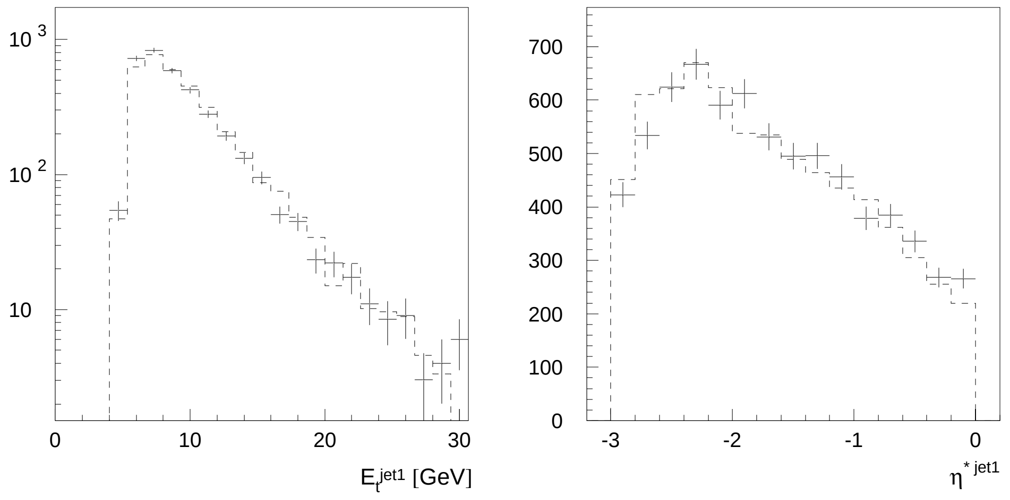
**HERWIG:**

- Photopr.: $sue = 0.15$
- DIS: $\omega = 0.2$, $sue = 0.15$

**RAPGAP:**

- DIS: $\omega = 0.2$

In the following sections only selected events left after applying all cuts described in section 5.1 are used.



(b) - DIS (d) - DIS

Figure 5.4: Distribution of $E_t^{jet}$ and $\eta^{\star\,jet}$ for DIS.

### 5.3.1 Distribution of $\mathbf{E_t^{jet}}$ and $\eta^{\star\mathbf{jet}}$

The transverse energy and the pseudorapidities of the two highest jets are shown in Fig. 5.4. The shape of $p_t^{jet}$ and $\eta^{jet}$ distributions are very well described by the MC simulation – the differences may be explained as statistical fluctuations.

### 5.3.2 Jet profiles

A very illustrative picture of two-dimensional jet profile in $(\eta^\star, \phi)$ space for photoproduction case is shown in Fig. 5.5(a). For each clim in a given event (see section (4.1)) for the definition of clim) the quantity $\Delta\phi = \phi^{clim} - \phi^{jet}$ and $\Delta\eta^\star = \eta^{\star\,clim} - \eta^{\star\,jet}$ were computed and then the transverse energy of this clim was plotted in the $(\Delta\phi, \Delta\eta^\star)$ space. Therefore the clims which are placed very close to the jet axis lie in the centre of the histogram. Here the superscript *“jet”* denotes the jet with the highest transverse energy $E_t$.

Fig. 5.5 (b) and (c) show again the profile of the jet with highest $E_t$, but only as one-dimensional function of $\Delta\eta^\star$ or $\Delta\phi$. This histograms were obtained as projection of the previous histogram (Fig. 5.5(a)) onto one of its axes.

For the last picture, Fig. 5.5(d), an additional cut $|\Delta\eta^\star| < 1$ has been applied.

Note that the distributions in Fig. 5.5 (c) and (d) increase as the $|\Delta\phi|$ approaches to the value of $\pi$. This is caused by the clims from the second jet which is usually oriented in the opposite direction in respect of the first jet because of the balance of momentum.

Jet profiles of the measured data are well described by the MC simulation. A small discrepancy appears only in DIS case around the centre of the jet ($\Delta\eta^\star$ and $\Delta\phi$ is close to zero), where the HERWIG simulation slightly overshoots the data.

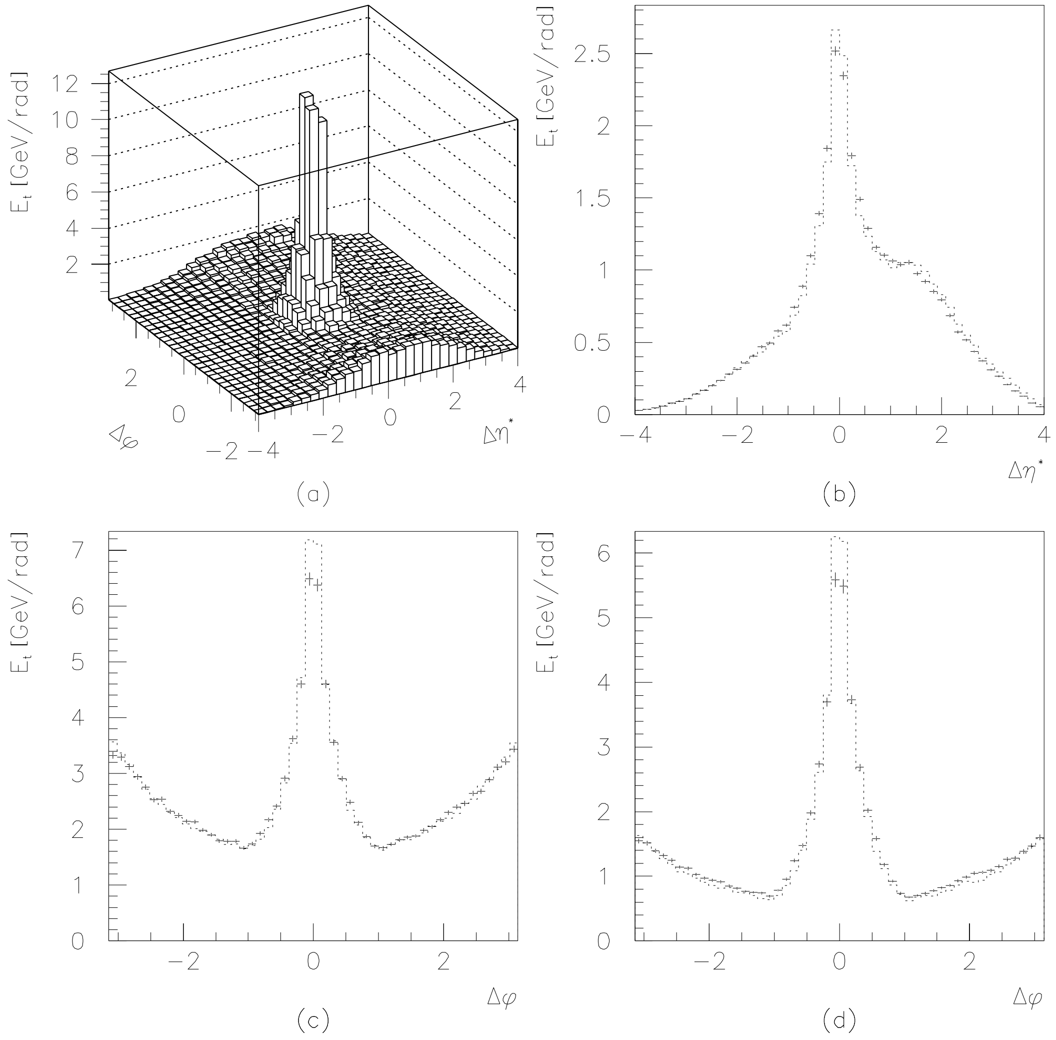


Figure 5.5: Jet profiles for photoproduction sample. Full line denotes the data, dotted line the MC HERWIG simulation.

According to the CDF algorithm jets were computed for radius $R = 1.0$ (see section 4.1.1). Clims deposited further than $R$ from the jet axis define "jet pedestal". As will be shown in chapter 6, the $(\eta^\star, \phi)$ area far from the jet axis is very sensitive to the amount of *sue*. Fig. 5.6 shows, that in DIS case the HERWIG with $sue = 8\%$ very well agrees with the measured data.

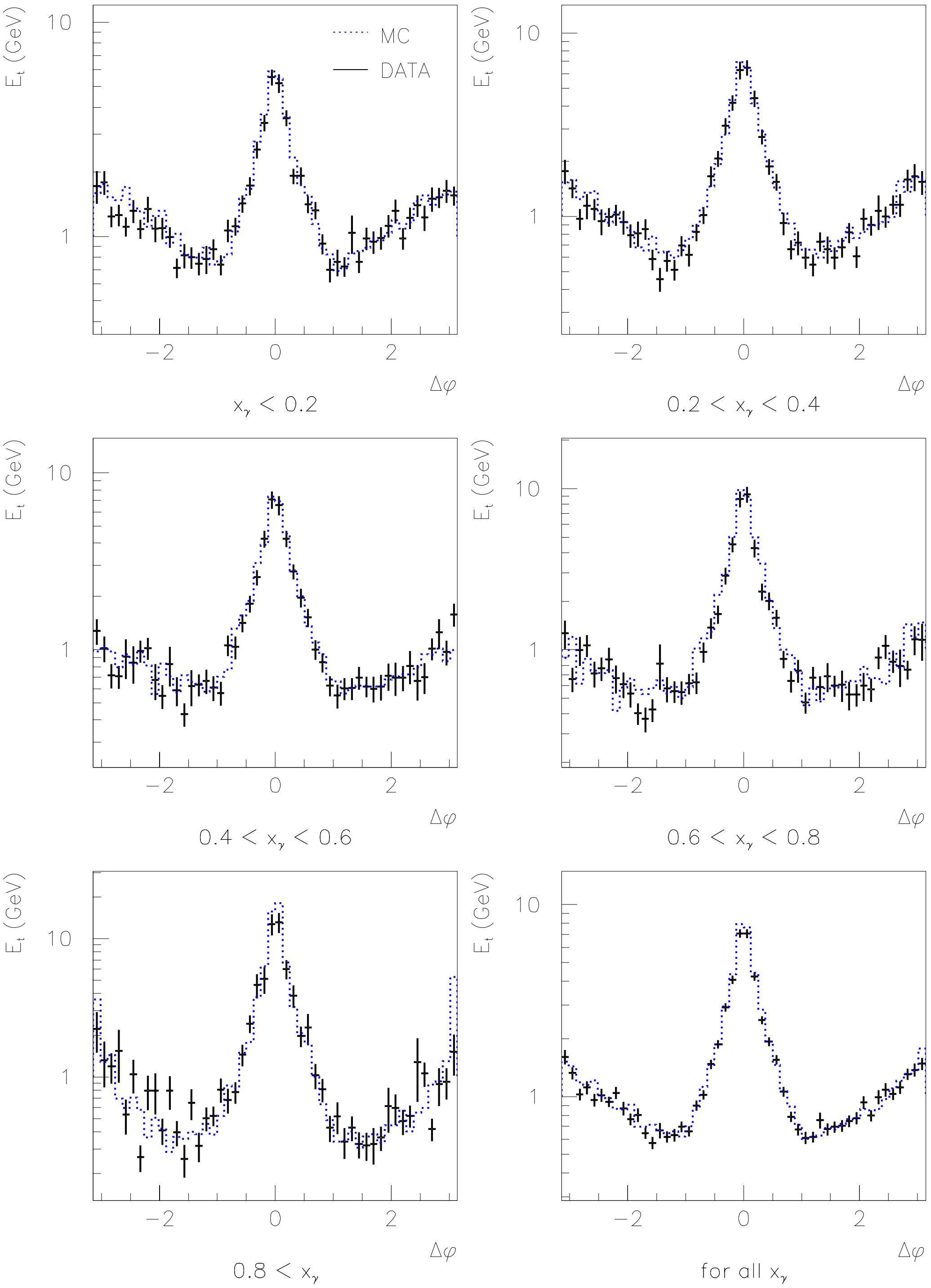


Figure 5.6: Jet profiles for **DIS** sample. Distributions are corrected for detector effects, parameter $sue = 8\%$ in HERWIG.

# Chapter 6

# Transverse Energy Flow Outside Jets

In this chapter the transverse energy flow outside jets normalised per a unit area in $(\eta^{\star},\phi)$ space, $\langle E_t \rangle$ , is analysed. This quantity is particularly important for understanding of the additional interactions of photon and proton remnants. Although the Monte Carlo programs usually very well describe hard processes (e.g. hard jets), the energy flow outside jets wasn't jet fully described. Especially in $e-p$ interactions at low $Q^2$ region there is a lack of studies of this quantity. At the time, MC programs simulating the interactions at HERA for low $Q^2$ region are still being developed. Results presented in this chapter should help to further understand to the energy deposited outside jets for measured data and MC samples.

The sources of this energy flow consist of:

- The hard scattering process. Most of it's energy should be included into the jets that correspond to the partons outgoing from the interaction. Nevertheless, some fraction of it's energy could be also deposited rather far from the jet axes.
- Initial state radiation.
- Final state radiation.
- An interaction of the proton and photon remnants (so called spectator partons).
- Non-interacting spectator partons.

The last two contributions are essentially uncorrelated with the hard parton scattering process.

The resolved events contain all the previously mentioned sources of $E_t$, while in the direct case there is no initial state radiation from the photon side, no interaction between the spectator partons, and no photon remnant. Thus the direct and resolved events should differ in their $\langle E_t \rangle$ flow outside jets. In general, the $\langle E_t \rangle$ flow should be higher for resolved events than for the direct ones, which will be shown and discussed in sections 6.3 and 6.4.

## 6.1 Underlying event energy

Two possible definitions of the transverse energy flow outside jets, $\langle E_t \rangle$ , were studied. The first one, which will be referred as "underlying event energy", has been adopted from the publication [19].

According to this method the transverse energy is summed in some $(\eta^\star, \phi)$ region and then normalised to the unit area of $(\eta^\star, \phi)$ space and averaged over all events in the sample. Energy deposited around the axes of two most energetic jets[1] within the distance $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 1.3$ is excluded from the energy summation. Fig. 6.1 transparently illustrates the definition of the underlying event energy.

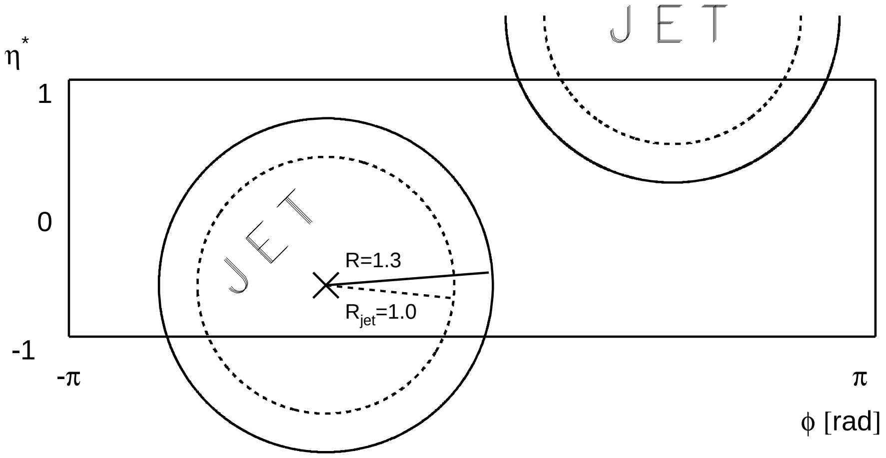


Figure 6.1: The definition of "underlying event energy".

Fig. 6.2 shows the dependence of the underlying event energy $\langle E_t \rangle$ on $x_\gamma$ for the photoproduction sample. The distributions are corrected for detector effects using the bin-by-bin correction method (see section (4.5). My results are compared with the analysis of 1993 data published in [19]. For this purpose the cuts from the section (5.1) were slightly redefined to get an event sample

---

[1] The terms "most energetic jet" or "highest jet" always mean a jet with the highest transverse energy through all this text.

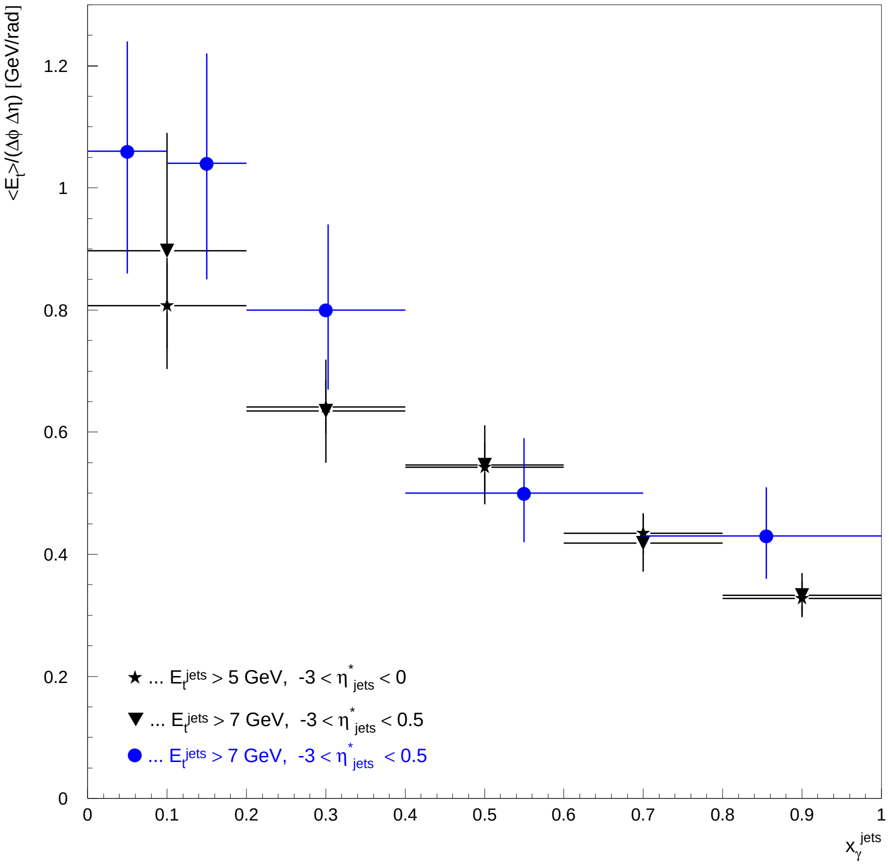


Figure 6.2: Transverse "underlying event energy" in the central pseudorapidity region $-1 < \eta^{\star clim} < 1$ for the photoproduction data, corrected for the detector effects (stars). Triangles represent our results obtained by using cuts from reference [19], circles are published points taken from [19], Fig. 4.

.

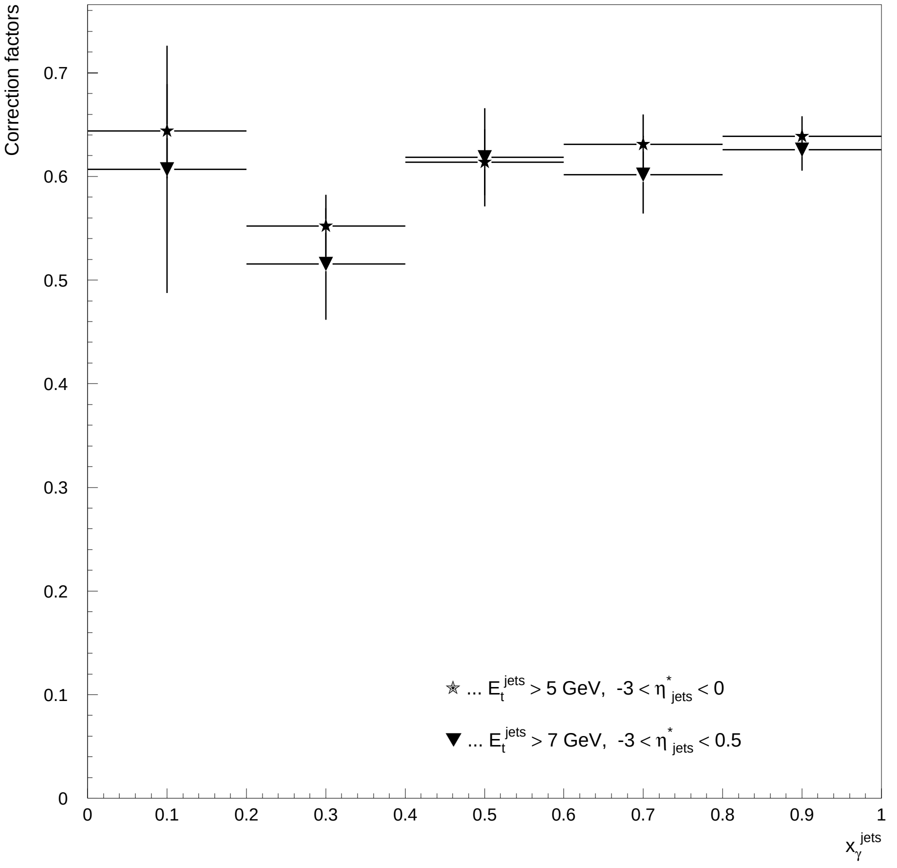


Figure 6.3: The correction factors for reducing the detector effects for distribution plotted in Fig. 6.2. The value in a given $x_\gamma$ bin of the uncorrected distribution has to be multiplied by the corresponding correction factor of the same $x_\gamma$ bin. This factors were calculated with the help of HERWIG simulation with $sue = 15\%$.

comparable with that of 1993 used in (5.1). Accordingly, two highest jets had to have transverse energy above 7 GeV and had to lie in the pseudorapidity range $-3 \leq \eta^{\star jet} \leq 0.5$. The published points are slightly higher than my results. This may be caused by different jet reconstruction, correction procedure (which depends on the used MC simulation program) or it can also partly be a statistical fluctuation.

The correction factors are displayed in Fig. 6.3. In general, the correction factors for all the distributions computed in the central pseudorapidity region $-1 < \eta^{\star clim} < 1$ are close to the value of 0.6, which means, that the transverse energy outside jets measured in the LAr calorimeter is $1/0.6 \simeq 1.7$ times higher than the true value. The detector response is much less biased in the backward pseudorapidity region $-3 < \eta^{\star clim} < -1$, where the correction factors are around

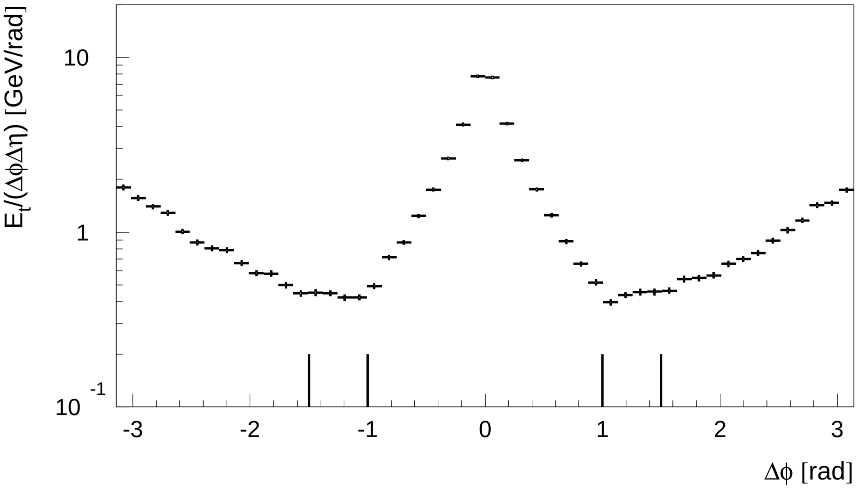


Figure 6.4: The definition of "pedestal energy".

0.9. On the other hand, the detector produces artificial energy flow in the forward region $\eta^{\star\,clim} > 1$. There the correction factors are very low, which leads to a big uncertainty of final results[2].

## 6.2 Jet pedestal energy

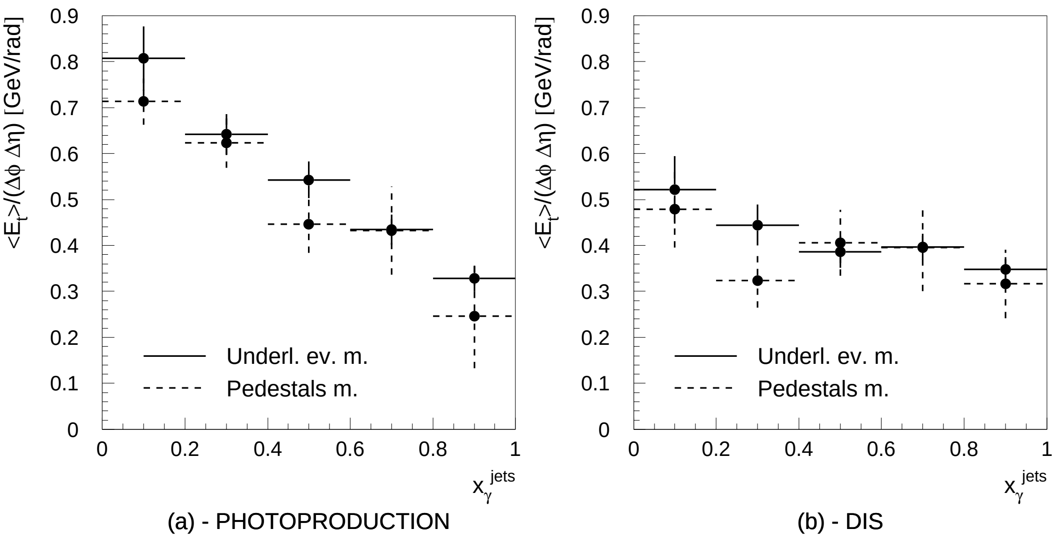
The second approach to study energy flow deposited outside jets is based on the jet pedestal energy, which can be determined from the jet shape introduced in the section (5.3.2). The "jet pedestal energy" is calculated as the transverse energy deposited in the area $|\Delta\phi| = |\phi^{jet} - \phi^{clim}| \in (1, 1.5)$ and $|\Delta\eta^{\star}| = |\eta^{\star\,jet} - \eta^{\star\,clim}| < 1$ (see Fig. 6.4). This sum is then normalised to the unit $(\eta^{\star}, \phi)$ area and averaged over all events in the sample. Since the energy flow outside jets depends on the pseudorapidity area in which it is calculated, the pseudorapidity of jets has been restricted to the interval $\eta^{\star\,jet} \in (-1, 0)$. This requirement makes this method easily comparable with the "underlying event" one.

The main difference compared to the previously defined underlying method concerns to the position of the jets in the $(\eta^{\star}, \phi)$ phase space. The area, in which the transverse energy is summed, is fixed to some $(\eta^{\star}, \phi)$ region in the first case, while according to the second method, the $(\eta^{\star}, \phi)$ area depends on the pseudorapidity of the jet.

The comparison of the "pedestal method" and previously described "under-

---

[2] Furthermore, the reconstruction of jets with low $E_t$ doesn't work properly in this pseudorapidity area. For this reason the $\langle E_t \rangle$ flow wasn't analysed in the forward region.



(a) - PHOTOPRODUCTION

(b) - DIS

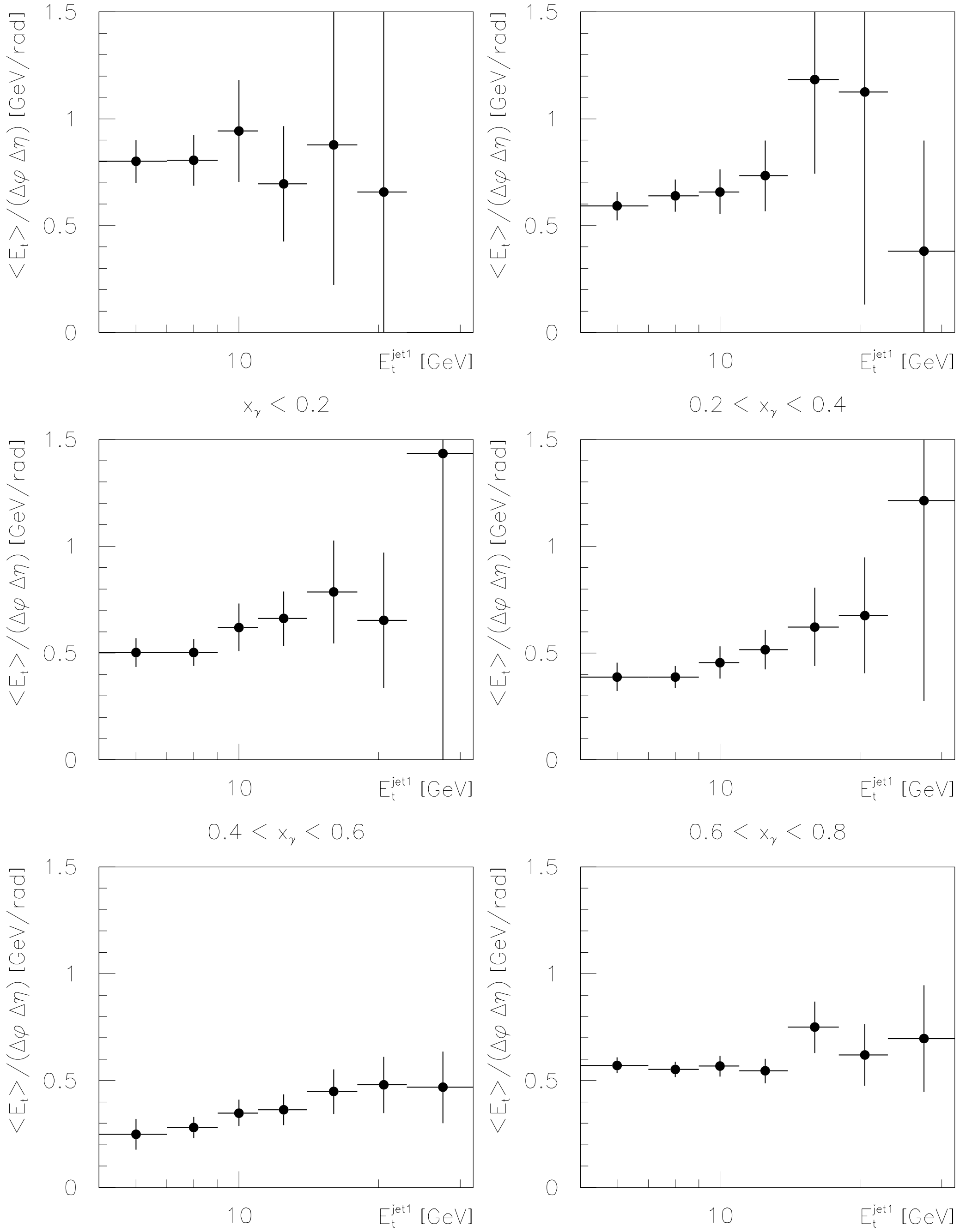
Figure 6.5: Comparison of the "underlying event" (full line) and "pedestal" (dashed line) methods of computing $\langle E_t \rangle$ flow outside jets in the central pseudorapidity region. The photoproduction (a) and DIS (b) corrected data are shown.

lying event method" is shown in the Fig. 6.5. One can see that there is a visible difference between both of these approaches. However, the shape of the $x_\gamma$ dependence tends to be the same, which reflects the similarity and rationality of both definitions. Since the difference between these distributions is not big (it is almost always within the range of statistical errors), only one method will be used for further study of $E_t$ flow outside jets. The "underlying event method" has been chosen, since it is computed from larger $(\eta^\star, \phi)$ area and its statistical errors are lower than that of the pedestal approach.

## 6.3 The dependence of the underlying event energy on $\mathbf{E_t^{jet}}$

An important quantity that characterises an event is $E_t^{jet}$ – the transverse energy of the jets that correspond to the partons outgoing from the hard scattering interaction. $E_t^{jet}$ is often used as a QCD scale which plays an important role in theoretical studies and models.

The dependence of $\langle E_t \rangle$ flow outside jets on the transverse energy of the jet with the highest $E_t$ is plotted in Fig. 6.6 for the photoproduction sample and in Fig. 6.7 for DIS sample. Data are divided into five $x_\gamma$ bins. It is well visible that the amount of underlying event energy decreases as the $x_\gamma$ rises. This reflects the fact that the energy flow outside jets is higher for the resolved events (which are characterised by low $x_\gamma$), while the direct events contain lower amount of the



$x_\gamma < 0.2$ $0.2 < x_\gamma < 0.4$

$0.4 < x_\gamma < 0.6$ $0.6 < x_\gamma < 0.8$

$0.8 < x_\gamma$ for all $x_\gamma$

Figure 6.6: $\langle E_t \rangle$ flow outside jets for the **photoproduction** corrected data in the central pseudorapidity region $-1 < \eta^{\star\, clim} < 1$. $E_t^{jet1}$ denotes the transverse energy of the jet with the highest $E_t$.

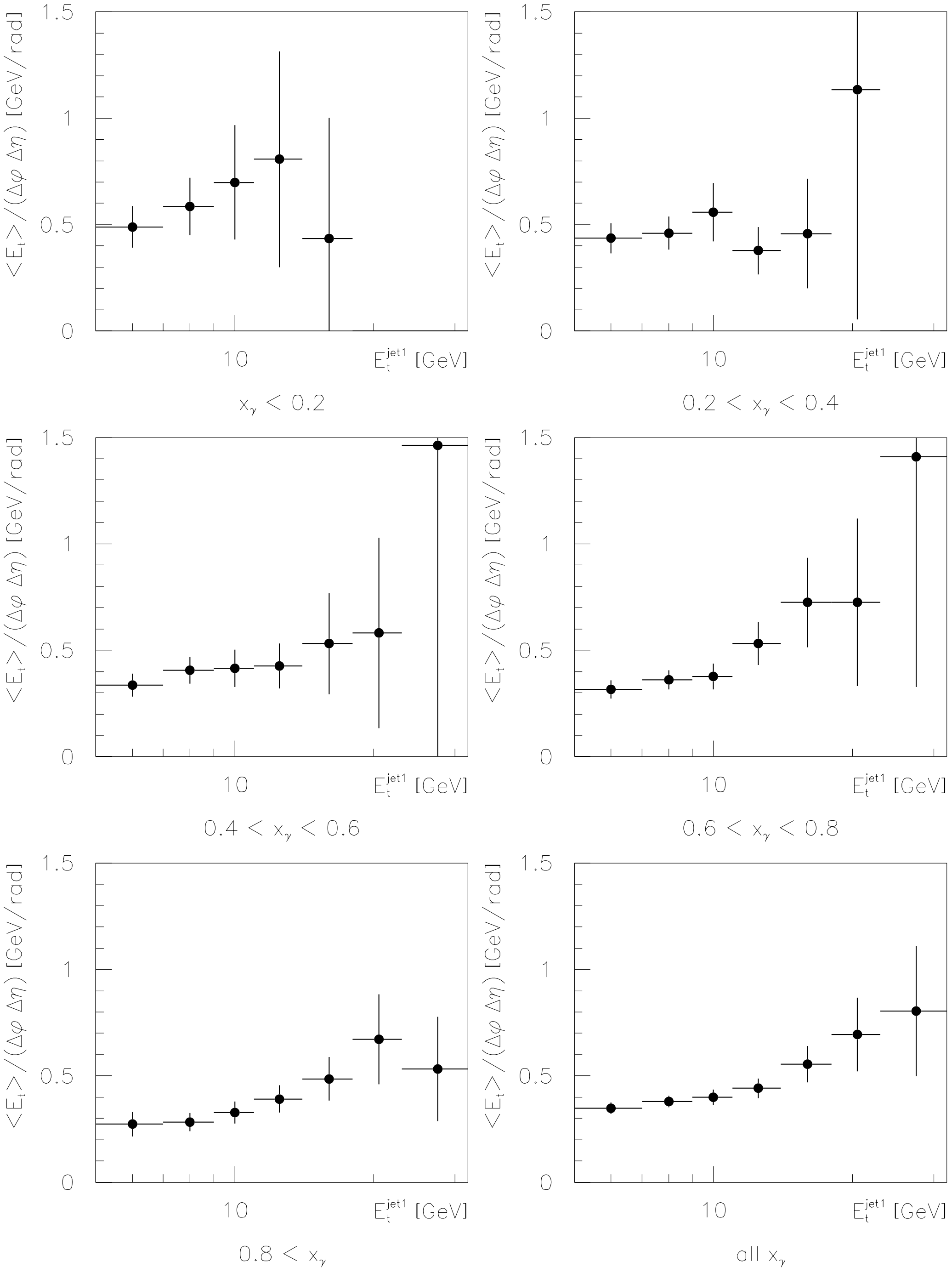


Figure 6.7: $\langle E_t \rangle$ flow outside jets for the **DIS** corrected data in the central pseudorapidity region $-1 < \eta^{\star\,clim} < 1$. $E_t^{jet1}$ denotes the transverse energy of the jet with the highest $E_t$.

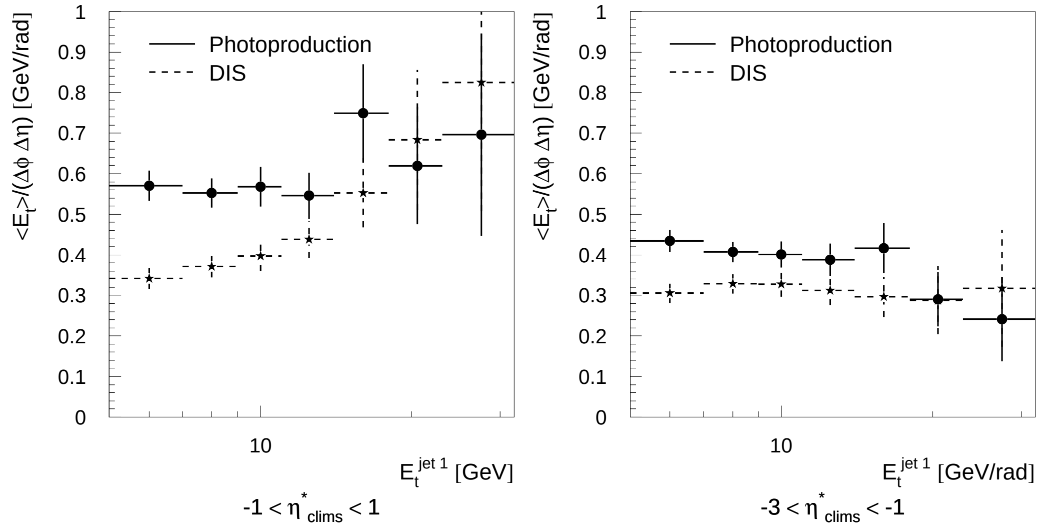


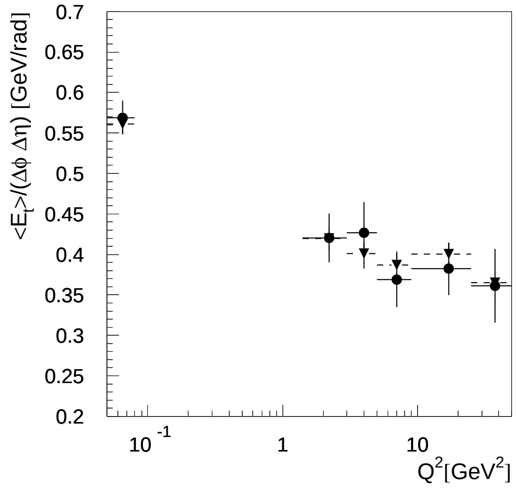
Figure 6.8: Comparison of the $\langle E_t \rangle$ flow outside jets for the DIS and photoproduction data in the central pseudorapidity region $-1 < \eta^{\star clim} < 1$ (left picture) and in the backward region $-3 < \eta^{\star clim} < -1$ (right picture) corrected on the detector effects.

transverse energy distributed outside jets. The resolved events, in contrast to the direct ones, contain the photon remnant and sometimes also an additional interaction of photon and proton remnants, which leads to the observation of higher underlying energy than in the direct case.

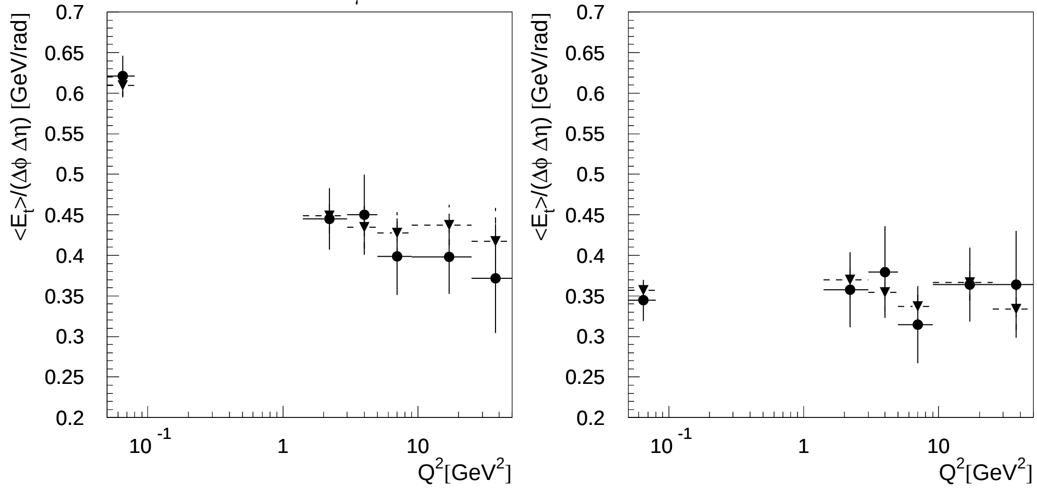
Another figures, which compare the $\langle E_t \rangle$ flow for photoproduction and DIS samples, are shown in the Fig. 6.8. The distribution of $\langle E_t \rangle$ is lower for DIS events than for the photoproduction ones, which again reflects the difference between the direct and resolved interactions. The number of resolved events decreases with the $Q^2$ rising. So there is a higher fraction of resolved events in photoproduction sample than in the DIS one.

## 6.4 The dependence of the underlying event energy on Q$^2$

Together with $E_t^{jet}$ also $Q^2$ characterises an interaction. The dependence of $\langle E_t \rangle$ flow outside jets on $Q^2$ is plotted in Fig. 6.9(a). As the $Q^2$ decreases, the underlying energy decreases as well. This is quantitatively included in the formula (5.6) and (5.7): the suppression of the resolved interactions becomes more significant if $Q^2$ rises and thus the direct interactions dominate. For low $Q^2$, on the other



(a) - all $x_\gamma$



(b) - $x_\gamma < 0.75$

(c) - $x_\gamma > 0.75$

Figure 6.9: Dependence of underlying energy on $Q^2$ : (a) – all $x_\gamma$, (b) – $x_\gamma < 0.75$, which corresponds to resolved case, (c) – $x_\gamma > 0.75$, which signifies the direct interactions. Experimental data are denoted by full lines and circles, the HERWIG simulation by dashed lines and triangles. Distributions are shown in the central pseudorapidity region ($-1 < \eta^{\star clim} < 1$) and are corrected on the detector effects.

hand, there is a high possibility for interaction to be the resolved one. Since the resolved events have in general higher amount of underlying energy than the direct ones, $\langle E_t \rangle$ flow outside jets decreases with $Q^2$ rising.

In Fig. 6.9(b) and (c) the same distributions are shown for two different $x_\gamma$ intervals. The conventional criterion in photoproduction analyses at H1 which distinguish the resolved event from the direct one is usually based on the reconstructed $x_\gamma$. If it is higher than some value (typically 0.75), the interaction is considered to be direct, otherwise it is qualified as a resolved one. Thus the Fig. 6.9(b) corresponds to resolved events, while the Fig. 6.9(c) shows direct interactions.

The underlying event energy for resolved interactions (Fig. 6.9(b)) depends on the fraction of energy carried by the spectator parton from the photon, i.e. $1 - x_\gamma$. The lower is $x_\gamma$, the higher fraction of the photon initial energy carries the spectator parton and consequently also the higher underlying activity (energy flow outside jets) is expected. This is really very clearly visible in Fig. 6.9(b).

The boundary value of $x_\gamma = 0.75$ is of course not sharp. There is a wide $x_\gamma$ region where both the direct and resolved components overlay and thus can't be separated perfectly. Especially in case of higher $Q^2$ , for which the resolved processes are suppressed, there is a relatively high fraction of direct events even for low $x_\gamma$ . On the other hand, resolved processes dominates in the photoproduction case, therefore even a small fraction from a high number of resolved events above the value 0.75 may significantly influence the behaviour of distributions in this region.

Fig. 6.10 shows the dependence of underlying energy on $x_\gamma$ for several $Q^2$ intervals. The shape of these distributions corresponds to the previous ones in Fig. 6.9. No dependence of $\langle E_t \rangle$ flow outside jets is observed for $x_\gamma$ close to one, while as $x_\gamma$ decreases $\langle E_t \rangle$ flow rises for low $Q^2$ (the lower is the $Q^2$ , the steeper is the rise).

The HERWIG MC simulation for phomas parameter $\omega = 0.2$ shows that the number of direct events in all $x_\gamma$-bins is higher than the number of the resolved ones if $Q^2$ is in between $15\,\mathrm{GeV}^2 < Q^2 < 50\,\mathrm{GeV}^2$. This partly explains, why the underlying energy for low $x_\gamma$ is significantly lower for high $Q^2$ intervals than for the low $Q^2$ ones.

## 6.5 Underlying event energy in HERWIG

In principle any distribution can be chosen for the comparison of the measured data and MC simulation. Some of such possible plots were already compared in section 5.3. In contrast to them, the distributions of energy flow outside jets, which are shown in Fig. 6.11, are highly sensitive to an additional photon and proton remnant collisions. For this reason they are very suitable for determination of Monte Carlo parameters that influence the amount of spectator partons

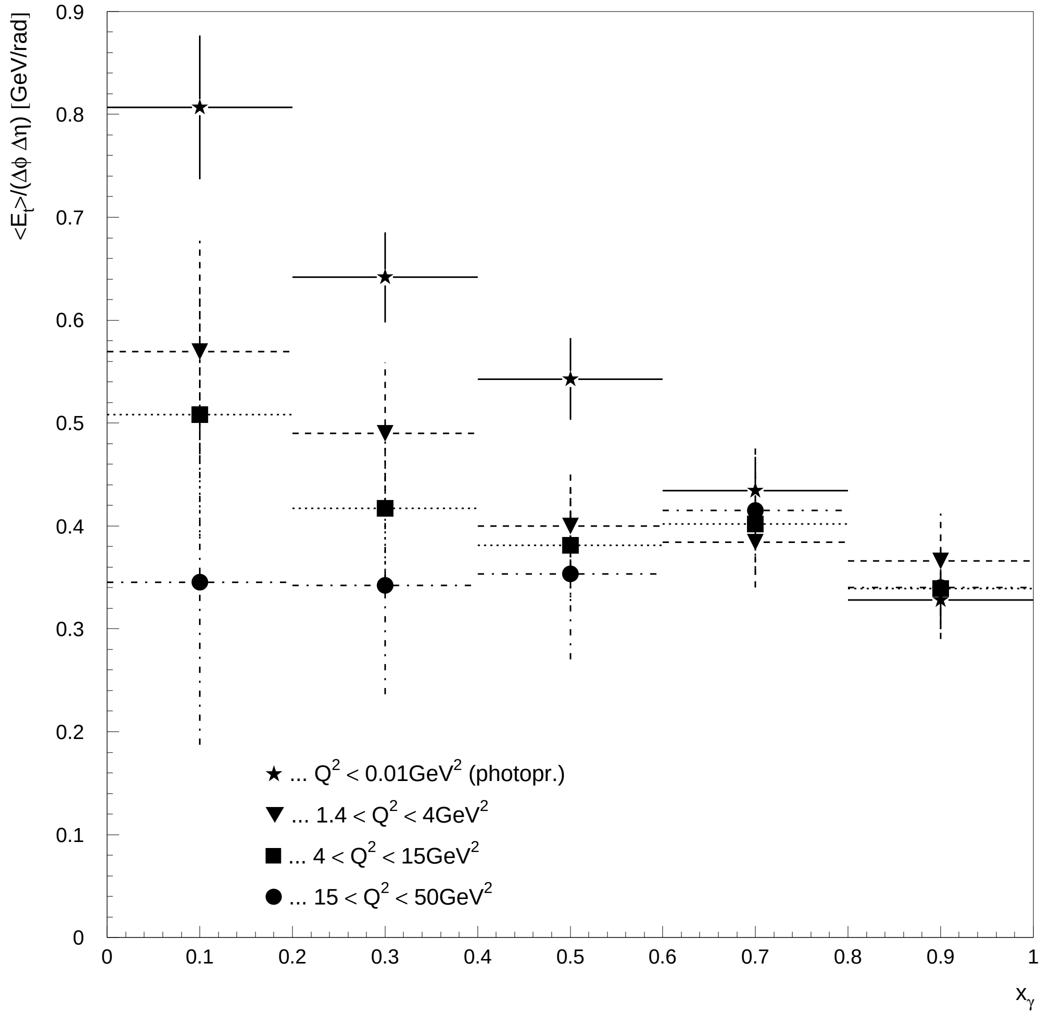


Figure 6.10: Dependence of underlying event energy on $x_\gamma$ for four $Q^2$ intervals. Distributions are shown in the central pseudorapidity region ($-1 < \eta^{\star clim} < 1$) and are corrected for detector effects.

interactions. The corresponding parameter in HERWIG is *sue* (see paragraph "Soft underlying event" on page 26 or section 5.2).

Fig. 6.11 demonstrates the sensitivity of HERWIG Monte Carlo on amount of *sue*. In the photoproduction case there is a good agreement between the data and HERWIG with *sue* about 15%. In this case the simulation differs from the data only in one $x_\gamma$ bin (in Fig. 6.11 (b)) for more than statistical error. In the DIS case, the best agreement is given for *sue* about 8%, and the differences between the simulation and data are almost negligible.

## 6.6 Comparison of the measured data with the RAPGAP simulation

In contrast to HERWIG, RAPGAP 2.9 simulation program doesn't contain any model describing interactions between photon and proton remnants. As a result, distributions concerning to $\langle E_t \rangle$ flow outside jets in RAPGAP sample doesn't agree with the measured data. The difference is relatively large, as one can see from Fig. 6.12. This is in a good agreement with the fact, that even HERWIG doesn't describe the data, if no soft underlying interactions are included (i.e. $sue = 0$). In this case there is much lower amount of energy deposited outside jets, which also well corresponds to results presented in [19]. The authors of that publication have showed that also another simulation program, PYTHIA 5.7, without additional interaction of photon and proton remnants doesn't describe observed $\langle E_t \rangle$ flow outside jets for the photoproduction case.

## 6.7 Low $\mathbf{p_t}$ particle production

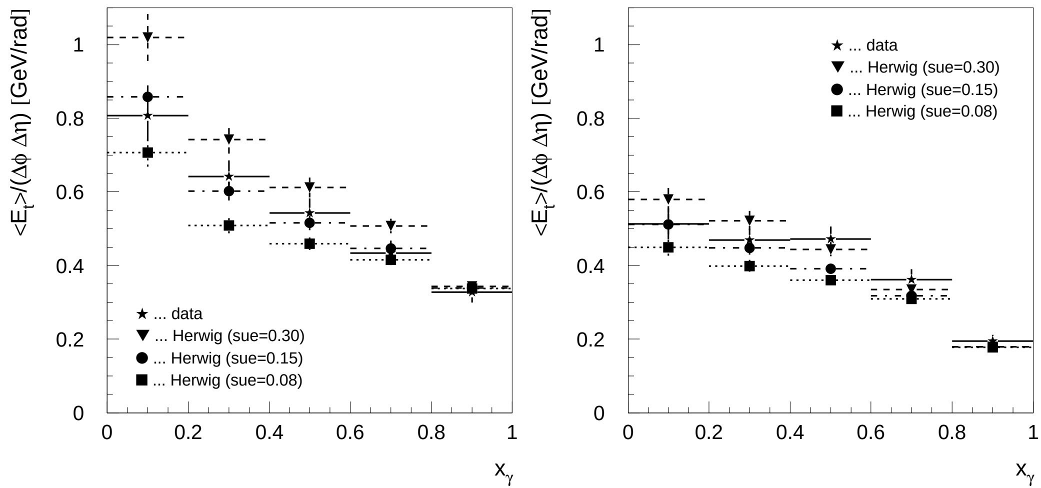
Recently, there was published an interesting prediction concerning to a low $p_t$ particle production in [20]. The authors predicted that there should be twice as much low $p_t$ particles produced in the resolved interactions than in the direct ones, if the particles are counted only in a small $(\eta^\star, \phi)$ region perpendicular to the jet direction in jet-jet CMS.

The dependence of numbers of clusters (which should more or less correspond to single particles) on the $p_t$ is displayed in Fig. 6.13 (a). The ratio of resolved and direct component is shown in Fig. 6.13 (b). According to a previously mentioned publication, the ratio should be equal approximately to 2 for low $p_t$. The values in Fig. 6.13 (b) are a little bit lower.

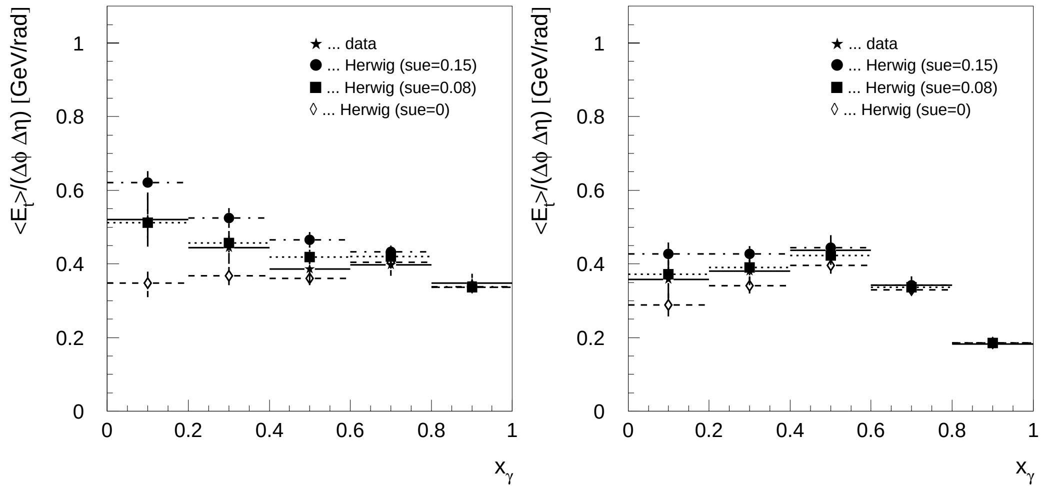
Unfortunately, this result is not very precise. The disputable question is how well a calorimeter cluster corresponds to a single particle[3]. The efficiency of detection and precision of measurement in LAr calorimeter is poor especially for

---

[3] This problem was also discussed with W. Ochs, one of the authors of publication [20].

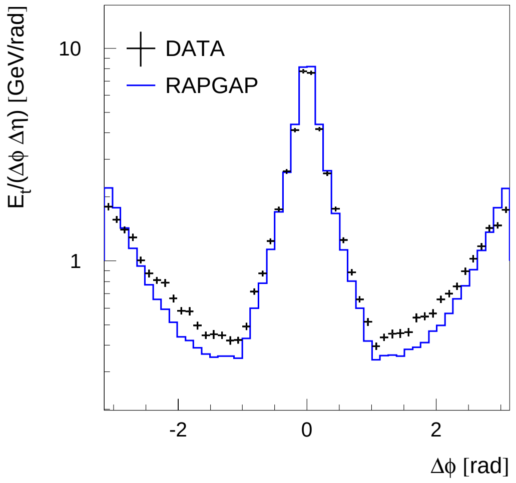


(a) - Photopr., $-1 < \eta^*_{clim} < 1$ (b) - Photopr., $-3 < \eta^*_{clim} < -1$

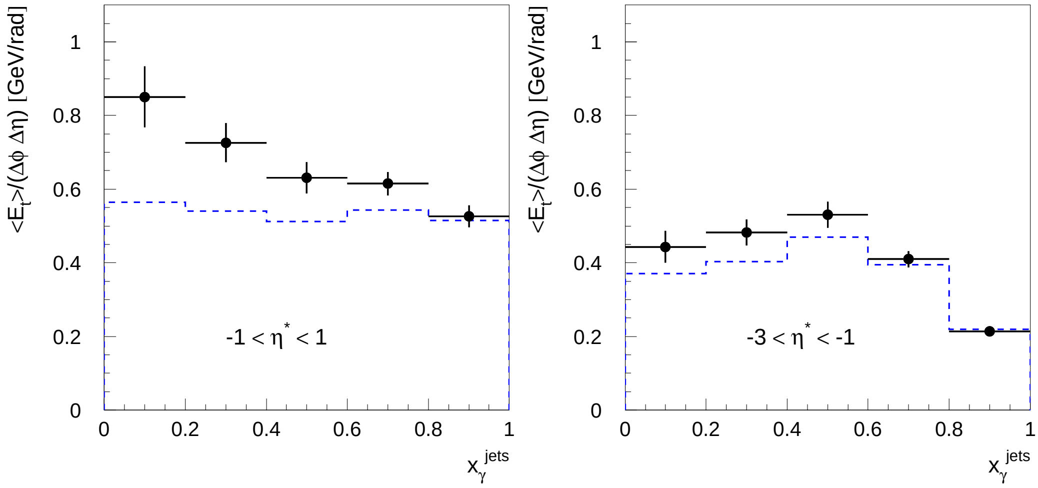


(c) - DIS, $-1 < \eta^*_{clim} < 1$ (d) - DIS, $-3 < \eta^*_{clim} < -1$

Figure 6.11: Comparison of the dependence of underlying event energy on $x_\gamma$ for the data and HERWIG simulations: (a) – photoproduction, underlying energy is calculated in the pseudorapidity interval ($-1 < \eta^{\star\, clim} < 1$), (b) – photoproduction, ($-3 < \eta^{\star\, clim} < -1$), (c) – DIS, ($-1 < \eta^{\star\, clim} < 1$), (d) – DIS, ($-3 < \eta^{\star\, clim} < -1$). Distributions are corrected for detector effects.



(a) - jet shapes



(b) - $<E_t>$ flow outside jets

(c) - $<E_t>$ flow outside jets

Figure 6.12: Comparison of the RAPGAP simulation and experimental data

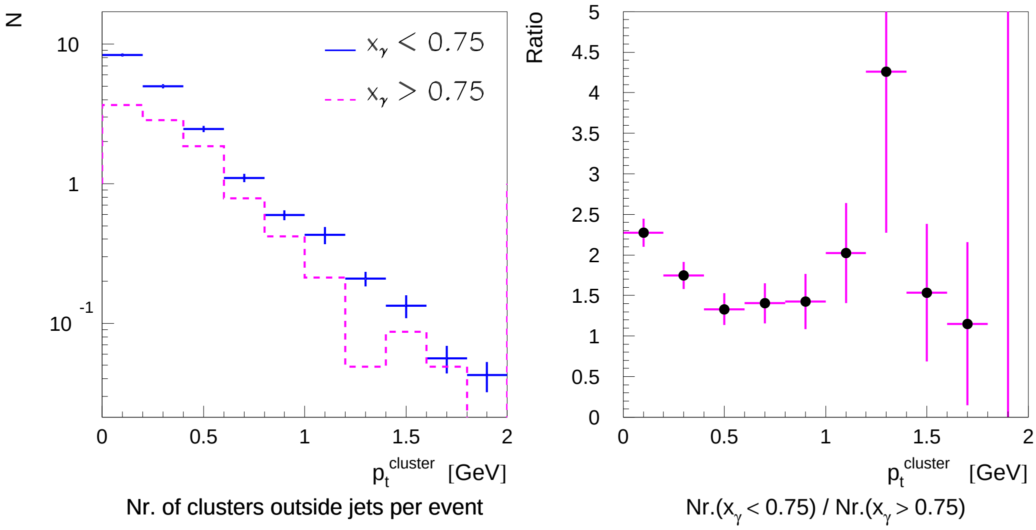


Figure 6.13: The dependence of numbers of clusters on their $p_t$ (left), and the ratio of resolved and direct contributions (right).

low energies, which we are interested in. To improve precision of obtained results, information from the track detectors should be used. This is planed to be done in future.

# Summary

The measurement of the hadron flow outside jets in the kinematic region restricted by requirements $-3 < \eta^{\star\, jets} < 0$, $\quad 0.25 < y < 0.7$, and $Q^2 < 50\,\mathrm{GeV}^2$ has been presented in this thesis. The 1995 measured data were compared with HERWIG 5.9 and RAPGAP 2.9 Monte Carlo programs. Different definitions of jet algorithms, $\langle E_t \rangle$ (the transverse energy flow outside jets) and $x_\gamma$ (the fraction of the photon energy coming into the interaction) were tested.

It has been shown that the underlying transverse energy strongly depends on $x_\gamma$. The underlying activity is lowest in direct interactions ($x_\gamma = 1$) and rises as $x_\gamma$ decreases. The dependence of the $\langle E_t \rangle$ flow outside jets on $Q^2$ also strongly depends on $x_\gamma$ – this distribution is almost flat for direct events, while for the resolved ones strongly rises with $Q^2$ decreasing.

A different behaviour of the distribution of underlying energy on $E_t^{jet}$ has been observed for DIS and photoproduction samples. $\langle E_t \rangle$ flow outside jets doesn't depend on $E_t^{jet}$ in photoproduction case, while for DIS it tends to rise with $E_t^{jet}$ increasing.

The first attempt to understanding of low $p_t$ particle spectra has been done. The number of low $p_t$ particles in resolved events is about 1.5 to 1.7 higher than for the direct ones. The predicted value of this ratio is approximately 2, which roughly corresponds to the observed result.

It has been shown that the simulations without an additional remnant interactions (HERWIG with $sue = 0$ or RAPGAP) doesn't describe observed energy flow outside jets. By comparison of the measured data with the HERWIG MC the amount of soft underlying events (as it is defined in HERWIG) has been determined to be about 8% for DIS and 15% for the photoproduction case. The determination of this fractions in $e - p$ collisions had never been published before.

# Bibliography

[1] H1 Collaboration: The H1 detector at HERA, (Nuclear Instruments and Methods in Physics Research A 368 (1997) 310-347)

[2] H1 Collaboration: The H1 detector at HERA, (Nuclear Instruments and Methods in Physics Research A 368 (1997) 348-396)

[3] H1 SPACAL group: H1 Backward Upgrade with a SPACAL Calorimeter: the Hadronic Section, (January 1996, DESY 96-013)

[4] H1 SPACAL group: The H1 Lead/Scintillating-Fibre Calorimeter, (August 1996, DESY 96-171)

[5] W. Braunschweig *et al.*: Results from a test of Pb-Fe liquid argon calorimeter, DESY 89-022

[6] H1 calorimeter group, B. Andrieu *et al.*: Results from pion calibration runs for the H1 liquid argon calorimeter and comparison with simulations, Nucl. Instr. and Meth. A336 (1993) 499

[7] H1 calorimeter group, B. Andrieu *et al.*: Beam tests and calibration of the H1 liquid argon calorimeter with electrons, Nucl. Instr. and Meth. A350 (1994) 57

[8] M. Erdmann: The Partonic Structure of the Photon: Photoproduction at the Lepton-Proton Collider HERA, Springer Tracts in Mod. Phys. 138, Heidelberg: Springer 1997

[9] UA5 Collaboration, G.J. Alner et al., Nucl. Phys. B291 (1987) 445

[10] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, L. Stanco: HERWIG - Hadron Emission Reactions With Interfering Gluons, Computer Phys. Commun. 67 (1992) 465

[11] J. Chýla: HERWIG 5.9: current status and H1 modifications; (November, 1997)

[12] Torbjörn Sjöstrand: PYTHIA 5.7 and JETSET 7.4 Physics and Manual, (August 1995, hep-ph/9508391)

[13] H. Jung: Hard Diffractive Scattering in High Energy *ep* Collisions and the Monte Carlo Generator RAPGAP, (December 1993, DESY 93-182)

[14] H. Rick: Messung der effektiven Partonverteilung im Photon, Dissertations Druck Darmstadt GmbH (Juni 1997, Dortmund)

[15] J.E. Huth *et al.*: Proc. of the workshop "Research Directions for the Decade; Summer Study on High Energy Physics", Snowmass (1990), 134

[16] A. Valkárová, G. Knies: Comparison of Cone and Event-Decomposition Jet Algorithms in Resolved Photon Reactions (March 1994, DESY 94-047)

[17] J. Chýla, J. Cvach: Virtual Photon Structure from Jet Production at HERA, Future Physics at HERA – Proceedings of the Workshop 1995/96 (September 1996, DESY) 545

[18] B. Laforge, L. Schoeffel: Elements of Statistical Methods in High Energy Physics Analyses, (H1 08/97-528)

[19] H1 Collaboration: Jets and Energy Flow in Photon-Proton Collisions at HERA, Z. Phys. C70 (1996) 17

[20] V.A. Khose, S. Lupia, W. Ochs: Perturbative Universality in Soft Particle Production, (November 1997, CERN-TH/97-199) hep-ph/9711392v1

[21] H1 Collaboration: Energy Flow and Charged Particle Spectra in Deep Inelastic Scattering at HERA, Z.Phys. C63 (1994) 377

[22] H1 Collaboration: Hadron Production in Diffractive Deep Inelastic Scattering, DESY-98-029

[23] H1 Collaboration: Photoproduction of $K^0$ and $\Lambda$ at HERA and a Comparison with Deep Inelastic Scattering, Z.Phys. C76 (1997) 213

[24] H1 Collaboration: Proton Dissociative Rho and Elastic Phi Electroproduction at HERA, Z.Phys. C75 (1997) 607

# Acknowledgements

I would like to express my gratitude to Josef Žáček, the supervisor of this diploma thesis, who gave me the opportunity to participate at the H1 experiment, enabled me to work for about six weeks at the DESY laboratory, introduced me into the practical high energy physics of nowadays and also spent a lot of time by explaining me physics more of less related to the topic of this thesis. He has helped me to correct many shortcomings and embarrassment of this text.

I greatly appreciate the support from Jaroslav Cvach who was a constant source of valuable advice and motivation. He spent a plenty of time by discussions with me and by helping me with whatever was needed – he checked my results, clarified me many physical and technical problems and helped me with the presentation of my results at meetings of the Photoproduction Working Group.

I am very grateful to Marek Taševský, Alice Valkárová Jiří Chýla, Tomáš Laštovička and Jaroslav Zálešák for their valuable comments, advice and discussions. They all illuminated me many problems concerning to physics at HERA. Marek Taševský helped me to overcome especially the difficulties of the computing equipment, Alice Valkárová clarified me some aspects of jet algorithms and RAPGAP MC program and Jiří Chýla helped me mainly with the interpretation of observed results.

I wish to thank to all the other people from the “Nuclear Centre, MFF[4]”, “Department of Nuclear and Subnuclear Physics, MFF[4]” and “High Energy Physics Department, AV ČR[5]” for their technical support and for all the knowledge they have given to me during previous 5 years.

I appreciate the financial support of one month visit at DESY from the grant Foundation of Czech Universities[6] No. 1278 and two one-week stays in DESY, which were supported by the grants GAUK 177/96 and GAČR 202/96/0214. These stays enabled me to present the results of this thesis at the meetings of the Photoproduction Working Group.

---

[4] Charles University, Praha, Czech Republic

[5] Institute of Physics, Academy of Sciences of the Czech Republic

[6] Fond vysokých škol