RHEINISCH
WESTFÄLISCHE
TECHNISCHE
HOCHSCHULE
AACHEN

PITHA 98/44

Dezember 1998

Power Corrections to Event Shape Variables measured in ep Deep-Inelastic Scattering

Klaus Rabbertz

I. Physikalisches Institut der Technischen Hochschule Aachen

PHYSIKALISCHE INSTITUTE
RWTH AACHEN
52056 AACHEN, GERMANY

Power Corrections to Event Shape Variables measured in *ep* Deep-Inelastic Scattering

Von der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinisch-Westfälischen Technischen Hochschule Aachen
genehmigte Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften

vorgelegt von

Diplom-Physiker
Klaus Rabbertz

aus Mönchengladbach

Berichter	:	Universitätsprofessor Dr. Ch. Berger
	:	Universitätsprofessor Dr. S. Bethke
Tag der mündlichen Prüfung	:	23.12.1998

Meinen Eltern

Ceci n'est pas une pipe.—René Magritte

Abstract

Deep-inelastic ep scattering data, taken with the H1 detector at HERA, are used to study event shape variables over a large range of "relevant energy" Q between 7 GeV and 100 GeV. Previously published analyses on thrust, jet broadening, jet mass and C parameter are substantially refined and updated; differential two-jet rates treated as event shapes are presented for the first time.

The Q dependence of the mean values is fit to second order calculations of perturbative QCD applying power law corrections proportional to $1/Q^p$ to account for hadronization effects. The concept of these power corrections is tested by a systematic investigation in terms of a non-perturbative parameter $\overline{\alpha}_{p-1}$ and the strong coupling constant.

Kurzfassung

Ereignisformvariablen in der tiefinelastischen ep-Streuung, gewonnen aus Daten des H1-Detektors bei HERA, werden über einen großen Bereich der „relevanten Energieskala" Q von 7 GeV bis 100 GeV untersucht. Bereits publizierte Studien zu *thrust*, Jetbreite, Jetmasse und C Parameter werden erheblich verbessert und aktualisiert; Ergebnisse zu differentiellen Zweijetraten als Ereignisformen werden erstmals vorgestellt.

Bei der Anpassung der Q-Abhängigkeit der Mittelwerte an Berechnungen der perturbativen QCD in zweiter Ordnung sollen potenzartige Korrekturterme der Form $1/Q^p$ Hadronisierungseffekte berücksichtigen. Eine Überprüfung des Konzepts solcher Potenzkorrekturen erfolgt im Rahmen der Bestimmung eines nichtperturbativen Parameters $\overline{\alpha}_{p-1}$ und der Kopplungskonstanten der starken Wechselwirkung.

Contents

Chapter 1

Introduction

Curiosity is one of mankind's elementary driving forces — the more so when scientists are concerned. In European history the appearance of the first "professional scientists" is usually dated back to antique Greece, where, besides others, philosophers laid the foundations for a great part of European culture. Two of them living around 500 B.C., Leukippos and Demokritos, suggested first that matter consists of tiny indivisible particles called *atoms*. They may be considered the forefathers of today's elementary particle physicists.

Their reasoning was of a basically theoretical nature. Turning to experiment, a "quantum leap" leads us to Antoni van Leeuwenhoek (1632–1723). Employing microscopes he constructed himself, he was able to observe "creatures that swarm and multiply in a drop of water"[1] for the first time.

With the advent of the modern microscope **H**adron-**E**lektron-**R**ing-**A**nlage HERA at the **D**eutsches **E**lektronen-**SY**nchrotron DESY in Hamburg three centuries later, the resolution power could be increased by about twelve orders of magnitude. Now it is possible to study "gluons that propagate and split in a droplet of nuclear matter called proton."

The theory that is believed to describe the *strong* interaction, responsible for the structure of hadronic (nuclear) matter, is called **Q**uantum **C**hromo**D**ynamics (QCD) in analogy to **Q**uantum **E**lectro**D**ynamics (QED) dealing with the *electromagnetic* interaction. Both are part of a more comprehensive theory usually referred to as the *Standard Model* (for introductory textbooks s. e.g. [1, 2, 3, 4]). It unites three, the *weak*, electromagnetic and strong interaction, out of the four fundamental forces observed in nature into a common framework. Gravity so far does not fit in. Comprising merely twelve elementary particles of spin 1/2, six quarks and six leptons, and the spin 1 exchange quanta of the three forces, i.e. the Z^0 and $W^\pm$ bosons, the photon γ and eight gluons g, the Standard Model gives a very good account of a tremendous amount of data gathered by experiments during the last decades [5].

The HERA collider offers the unique possibility to investigate the structure of the proton and its interacting constituents in a completely new kinematic domain unreached

[1] H.G. Wells, *The War of the Worlds*, 1898.

by fixed-target experiments. With respect to this analysis, the large accessible range in "relevant energy" Q is of great advantage. Observables can that way be studied and compared to theory in dependence of the available energy Q in a single experiment. To that goal, the thesis is organized as follows:

Chapter two describes the machinery necessary to perform ep scattering experiments at the required high energies with special attention to the parts relevant for this analysis. The third chapter explains the basic kinematics of the scattering process. Subsequently, the observables to measure are defined in chapter four. The series of chapters five, six and seven deals in detail with the selection of data, their comparison with simulations and the correction for detector imperfections. Fully corrected data are presented. The two following chapters eight and nine are dedicated to the derivation of the necessary theoretical input employing **p**erturbative **QCD** (pQCD). Finally, experimental and theoretical results are compared and combined to test the power correction approach to hadronization effects that was initiated by Dokshitzer and Webber [6]. The last chapter presents a summary and an outlook.

Chapter 2

HERA and H1

2.1 The Storage Ring HERA

Scattering experiments are the basic tools to investigate the elementary building blocks of matter, i.e. quarks and leptons, and the fundamental forces acting between them. Two rather straightforward set-ups are e^+e^- and $p\bar{p}$ colliders. In the last decades, a lot of valuable information has been accumulated with such machines [5]. However, for the detailed study of a strongly bound, hadronic object like the proton, it is preferable to use a probe that does not itself interact strongly. Fixed-target experiments employ electrons and myons as well as neutrinos to determine the structure of nucleons. To enlarge the resolving power, it is necessary to increase the energy available in the collisions by accelerating the "targets," i.e. a collider experiment is asked for.

The Hadron-Elektron-Ring-Anlage HERA, shown in fig. 2.1, is situated at DESY in Hamburg, Germany, and is currently the world's only facility where electrons[1] and protons are accelerated in two separate storage rings to final energies of 27.5 GeV and 820 GeV respectively. The resulting center of mass energy of about 300 GeV corresponds to electron beams of 50 TeV for fixed-target experiments. For a recent overview of the knowledge gained on the structure of nucleons consult e.g. [7].

Along the circumference of 6.3 km two locations, the north hall and the south hall, are assigned to the study of ep scattering. The interaction regions, where during operation every 96 ns particle bunches may collide, are surrounded almost hermetically by complex detectors, H1 [8] and ZEUS [9]. They are dedicated to the task of measuring as many collision products as precisely as possible. Two other experiments, HERMES and HERA-B, are situated in the east and west halls and are committed to spin and B meson physics respectively.

[1] Due to a considerably higher lifetime at large currents, the electron beam has been replaced by a positron beam in July 1994. Since for the purpose of this study it does not make a difference, the term *electron* will henceforth be used synonymously for positrons, too.

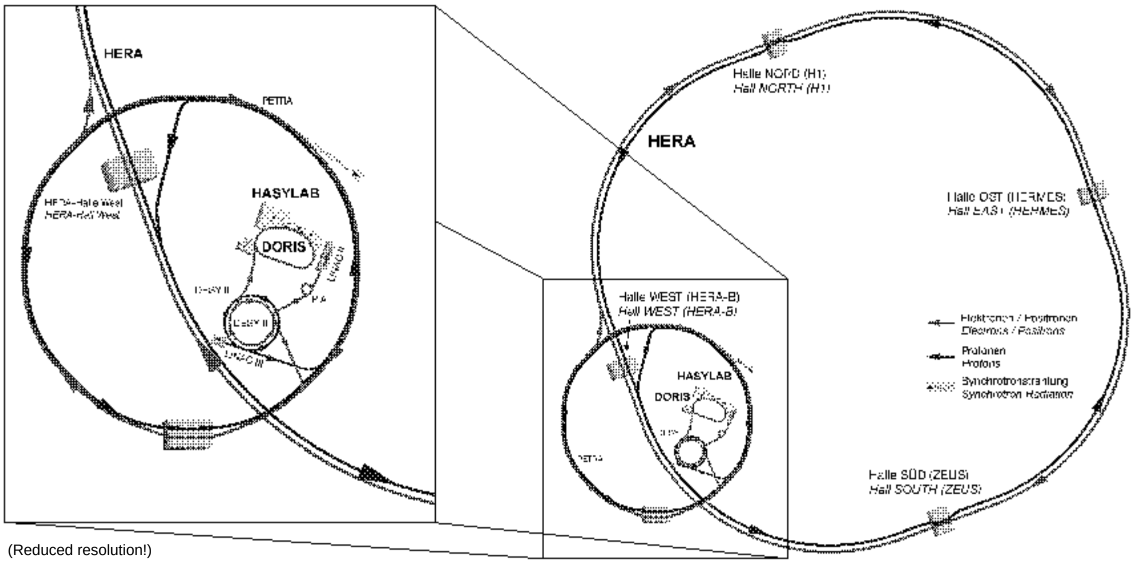


Figure 2.1: The storage ring HERA at the DESY laboratory in Hamburg, Germany.

2.2 The H1 Detector

The basis of this experimental analysis are data collected with the H1 detector, which is located at the northern interaction point of HERA. Only a brief overview of the total system, shown in fig. 2.2, will be given in the next section. The parts that provide the main information needed in this study will subsequently be discussed in more detail. A complete description of the detector and its performance in the first three years of operation (1992–1994) can be found in [8, 10].

During the winter shutdown 1994/95, a major upgrade of the H1 backward region was undertaken [11]. However, none of the components relevant to this analysis were significantly affected. Therefore, these improvements will not be covered here, although data from 1994 up to 1997 are used. Events where the scattered electron was found in the backward calorimeter are taken from 1994 data only.

2.2.1 Overview of the H1 Detector

The most obvious feature of the H1 detector in fig. 2.2 is its asymmetric design. Because the momentum of the protons is about 30 times higher than that of the electrons, the ep center of mass system is moving along the direction of the proton beam. As a result, the density of collision products hitting the equipment in the "forward" region is very high which is accordingly reflected in the more massive as well as finer granulated material of that part. The term "forward" refers to the conventional coordinate system used in the H1 collaboration where the proton beam direction is defined to be the $+z$-axis. The x-axis (y-axis) points from the nominal interaction point at the center of the HERA ring (upwards). Subsequently, the detector components are briefly described proceeding from the innermost parts outwards:

- The tracking system:

  Directly surrounding the beam pipe and beam magnets, a tracking system is installed to measure the momentum of charged particles. It consists of two main parts: the **C**entral **T**racking **D**evice (CTD) covering the region around the nominal interaction vertex and the **F**orward **T**racking **D**evice (FTD) supplementing it in the $+z$-direction.

- The calorimeters:

  To complement the momentum measurement and to detect neutral particles, the CTD and FTD are enclosed in the forward and central region by a sampling calorimeter (LAr) with **L**iquid **Ar**gon as active material. For the innermost absorber stacks, lead has been chosen to ensure a good containment and energy determination of electromagnetic showers produced by electrons and photons (**E**lectromagnetic **CAL**orimeter, ECAL). The outer part of the LAr (**H**adronic **CAL**orimeter, HCAL) predominantly measures hadronic showers and is equipped with steel absorber plates, which also serve as mechanical support structure.

  The remaining holes of the LAr around the beam pipe are closed with a silicon-copper calorimeter for polar angles below $4°$ (PLUG) and a lead-scintillator calorimeter in the backward direction (**B**ackward **E**lectro**M**agnetic **C**alorimeter, BEMC). During the upgrade in 1994/95 the BEMC was replaced by a lead-scintillating fibre **Spa**ghetti **Cal**orimeter (SpaCal) [12], which is subdivided into a first electromagnetic section and an additional second part to determine energy leakage and to improve the containment of hadronic showers. For the purpose of this analysis, it was only employed to supplement eventual energy deposits in the backward direction not due to the scattered electron.

- The superconducting coil:

  A cylindrical superconducting coil providing the magnetic field of 1.15 T for the trackers envelops the calorimeters. Thereby, the amount of dead material in front of the calorimeters is reduced and the time of flight of myons within the magnetic field is increased improving the resolution of their momentum measurement.

- The instrumented iron yoke and myon system:

  The **IRON** return yoke (IRON) of the magnet, enclosing almost completely all other parts of the detector, is sandwiched with streamer tubes for the measurement of myons and energy leakage from the inner calorimeters (**T**ail **C**atcher, TC). Supplementary chambers inside the IRON further improve the evaluation of myon tracks. Myons with high momenta in the forward direction are analyzed by a spectrometer consisting of four drift chambers in the magnetic field of 1.5 T of a toroïdal coil in front of the IRON.

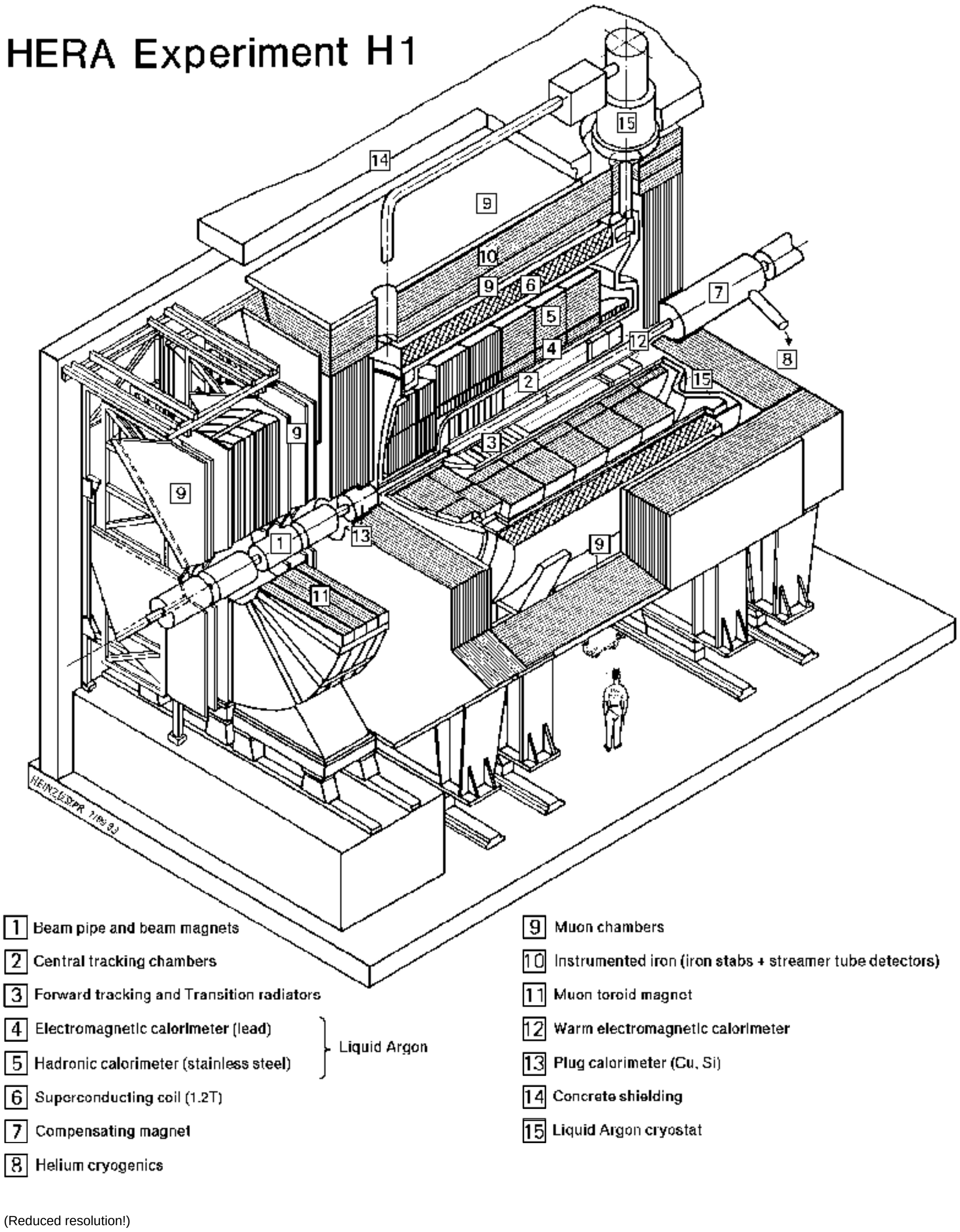


Figure 2.2: Schematic layout of the H1 detector.

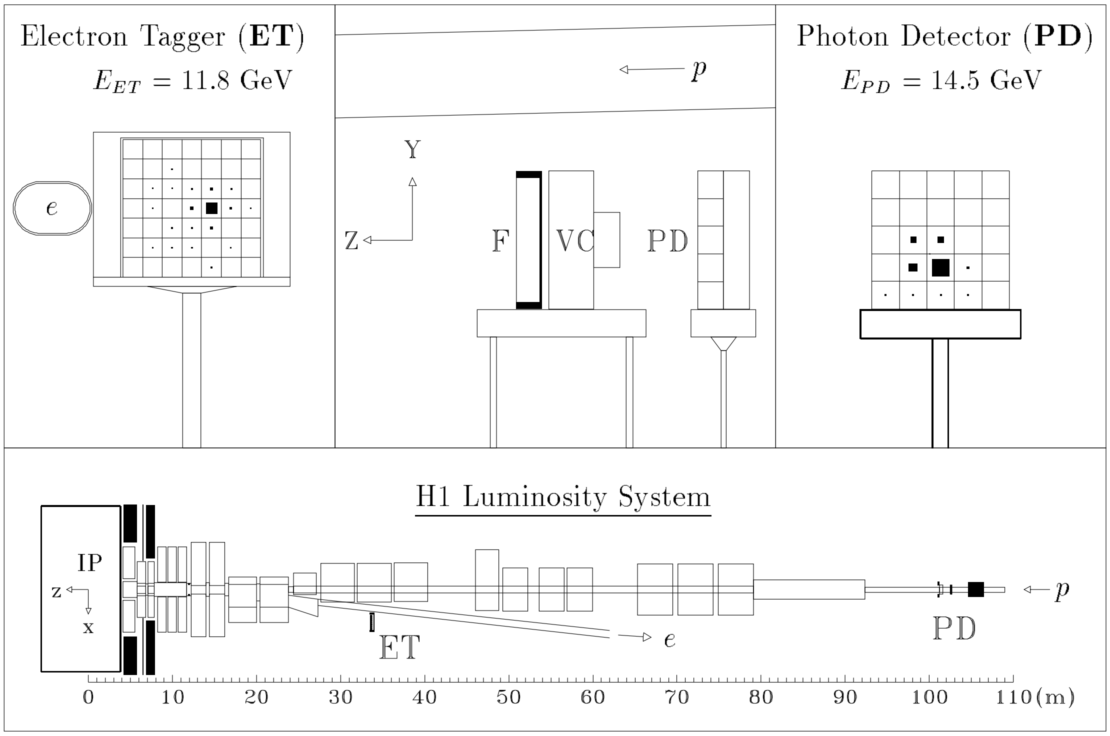


Figure 2.3: The layout of the H1 luminosity system [8].

- The luminosity system:

The luminosity system sketched in fig. 2.3 utilizes the Bethe-Heitler process $ep \to ep\gamma$. The electrons and photons are detected in coincidence by a hodoscope of crystal Cherenkov counters. Whereas the electrons from these grazing collisions, i.e. $\theta_{e'} \approx 180°$, are deflected by magnets from the beam line to hit the **E**lectron **T**agger (ET) at $z = -33.4\,\mathrm{m}$, this is not possible for the photons, of course. They leave the proton beam through a window at $z = -92.3\,\mathrm{m}$ instead, where the beam pipe bends upwards, and reach the **P**hoton **D**etector (PD) at $z = -102.9\,\mathrm{m}$.

The main background is caused by bremsstrahlung from residual gas atoms $eA \to eA\gamma$. The technique of electron pilot bunches that do not have colliding proton counterparts allows to correct for it.

2.2.2 Details on the Main Components

The basic ingredient for this analysis is the calorimetric information provided by the LAr and BEMC in the form of electromagnetic and hadronic clusters reconstructed from the primary energy depositions. The fragmentation region of the proton remnant, which is related to "soft" physics, can be excluded more easily after performing a Lorentz transformation into the Breit frame (s. section 4.1). This procedure requires a good electron identification. On the one hand, the electron cluster has to be separated from the hadronic final state, and on the other hand, it is necessary for the extraction of the event kinematics which is crucial in the calculation of the boost to the Breit frame. Therefore, additional information of the tracking system, especially the CTD, is exploited to determine the properties of the scattered electron and the actual event vertex.

The main attention in this analysis rests on mean values and normalized distributions. Hence, the luminosity system needed to measure absolute cross sections is not of prevalent importance.

Concluding, the CTD, LAr and BEMC are the most important subdetectors for the observables considered and will therefore be described in more detail [10]:

- The central tracking device CTD:

  The CTD is subdivided into two major units, the **C**entral **J**et **C**hambers CJC1 and CJC2 as shown in fig. 2.4 in a view perpendicular to the beam. The track reconstruction in the central region, $25° < \theta < 155°$, depends primarily on these concentric drift chambers with wires strung parallel to the beam pipe. They provide a good space point resolution in the (r, ϕ)-plane of $170\,\mu$m and, by comparing the signals at both ends, furthermore supply information about the z-coordinate of the hits. To avoid completely insensible paths for almost straight tracks, the sequence of the sense wires is inclined with respect to the radial direction.

  The CJC1 is sandwiched between the **C**entral **I**nner and **O**uter **Z**-chambers CIZ and COZ with wires oriented in the (r, ϕ)-plane. They complement the measurement in the CJC's with data on the z-coordinate of track elements with a precision of $\approx 300\,\mu$m.

  Mainly for the purpose of fast timing (< 96 ns) and triggering, multiwire proportional chambers are added for polar angles between 5° and 175°. Two of them, the **C**entral **I**nner and **O**uter **P**roportional chambers CIP and COP, are indicated in fig. 2.4. For electrons in the backward direction, only short track segments can be seen in the CTD. The **B**ackward **P**roportional **C**hamber BPC, consisting of four sensible layers of wires oriented along azimuthal angles of 0°, 45°, 90° and 135°, substantially improves the tracking for polar angles of $155.5° < \theta \leq 174.5°$. Installed directly in front of the BEMC, fig. 2.5, it also helps to distinguish between photons and electrons in this region.

  In general, the tracking system was designed to determine the momentum and angles of charged particles to a precision of $\sigma_p/p^2 \approx 0.003\,\mathrm{GeV}^{-1}$ and $\sigma_\theta \approx 1\,$mrad.

- The liquid argon calorimeter LAr:

  As mentioned, the LAr, covering polar angles of $4° < \theta < 154°$, is a sampling calorimeter where layers of absorber material and sensitive gaps, filled with liquid argon, alternate. Since only a sample of all energy deposits can be measured that way, the energy resolution of this technique is worse than that of calorimeters built of only one medium. To achieve nevertheless a maximal precision, care has been taken that the orientation of the absorber plates, shown in the upper part of fig. 2.6, ensures angles near the normal direction of the stacks for particles originating from the vertex. Another drawback is the difference in response on electromagnetic and hadronic showers produced by impacts of particles of the same original energy. The LAr is *non-compensating*.

  This is counterbalanced by stable calibration and homogeneous response properties as well as the possibility of a compact but still finely segmented construction. The structure of the read-out cells, again in an (r, z)-view, is presented in the lower half of fig. 2.6.

  Apart from the general subdivision in an **E**lectromagnetic and a **H**adronic section, the LAr is composed of eight wheels each further partitioned into eight octants in ϕ. Due to the use of weak lead for the ECAL, a pointing geometry, blind for hits straight on the border between two octants, could not be avoided in contrast to the CTD. Particles reaching these ϕ-cracks are detected only in the non-pointing HCAL. All wheels are, with respect to the nominal interaction point, combined to four aggregates: a **B**ackward, **C**entral and **F**orward **B**arrel and the **F**orward end cap with an **I**nner and **O**uter ring. The abbreviations in fig. 2.6 translate e.g. as "outer hadronic ring of forward end cap wheel 1" for OF1H.

  Test beam results yield energy resolutions of $\approx 11\%/\sqrt{E} \oplus 1\%$ for electromagnetic and $\approx 50\%/\sqrt{E} \oplus 2\%$ for hadronic showers.[2] The absolute energy scales are known to 1–3% [13] and 4% for electrons and hadrons respectively.

- The backward electromagnetic calorimeter BEMC:

  A transverse view of the warm lead-scintillator calorimeter and its segmentation into 88 stacks aligned parallel to the beam pipe can be seen in fig. 2.5. The stacks are multi-layer sandwich structures with active sampling units made of plastic scintillators. The front face directly behind the BPC is located at $z = -144\,\mathrm{cm}$. The angular region covered extends from polar angles of 151° up to 176°.

  The electromagnetic energy resolution was derived from test beam measurements to be $\approx 10\%/\sqrt{E} \oplus 1.7\%$, the absolute energy scale is known to a precision of 1%.

[2] The "$\oplus$" indicates quadratic addition.

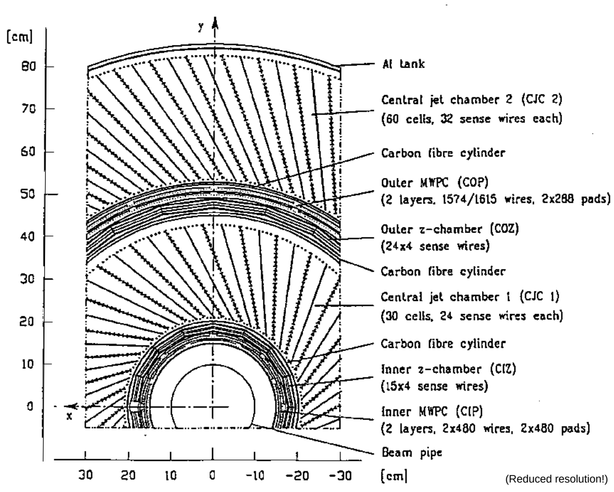


Figure 2.4: Central tracking system, section perpendicular to the beam [10].

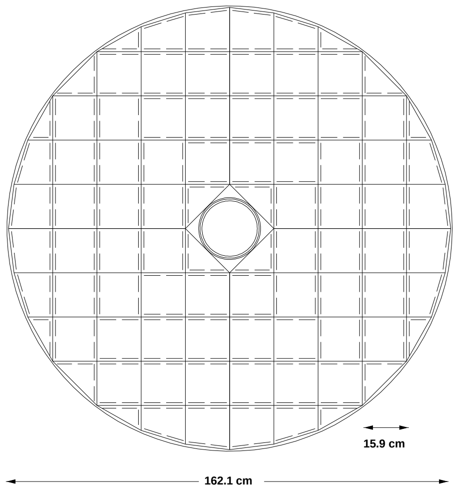


Figure 2.5: Transverse view of the stack segmentation of the BEMC [10].

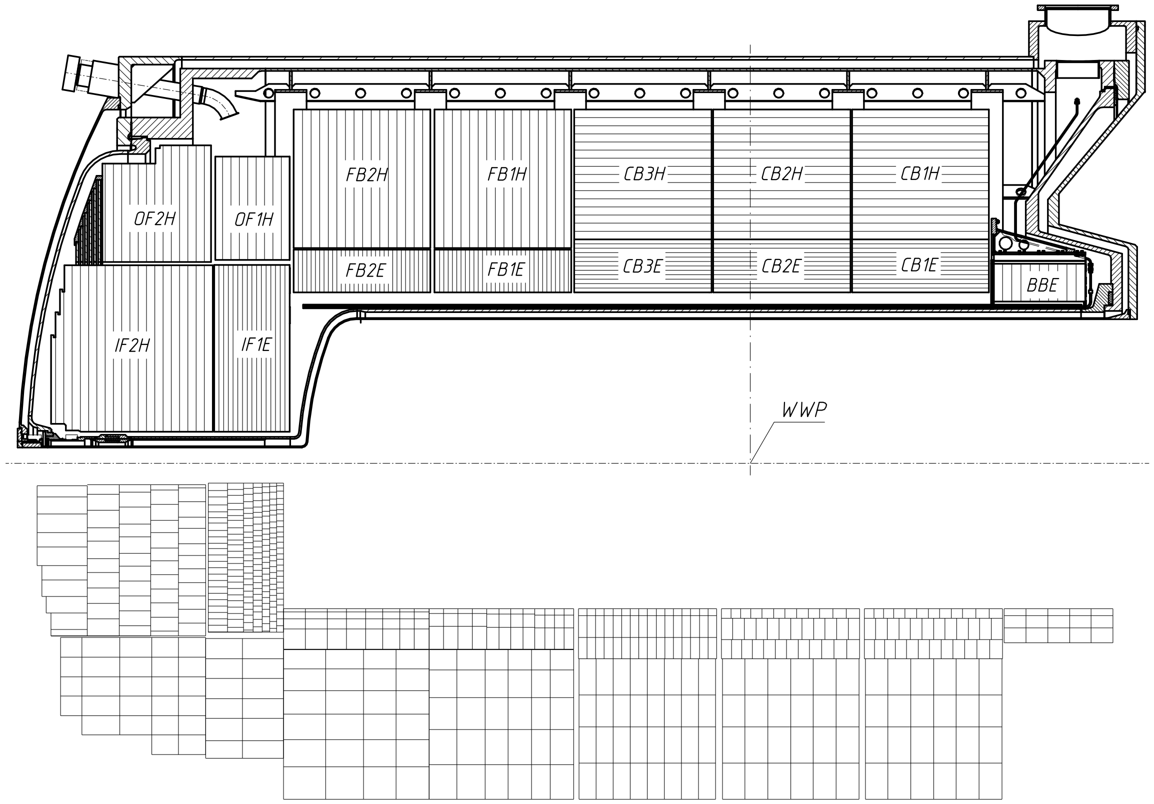


Figure 2.6: The orientation of the absorber plates (top) and the cell structure of the LAr in (r, z)-view (bottom).

Chapter 3

Deep-Inelastic Scattering

3.1 The Kinematics

Due to the high beam energies of $E_e = 27.5\,\mathrm{GeV}$ and $E_p = 820\,\mathrm{GeV}$ available at HERA, the masses of the electron and proton can safely be neglected for most purposes. This will be done throughout this study. Their four-momenta,[1] indicated in fig. 3.1, can therefore be written as $k = (E_e, 0, 0, p_{z_e}) = (27.5, 0, 0, -27.5)\,\mathrm{GeV}$ and $P = (E_p, 0, 0, p_{z_p}) = (820, 0, 0, 820)\,\mathrm{GeV}$ with $k^2 = P^2 = 0$. Thus, the center of mass energy $\sqrt{s}$ follows from

$$s := (k + P)^2 = 2k \cdot P = 4E_e E_p = 90200\,\mathrm{GeV}^2 \tag{3.1}$$

to be $300.33\,\mathrm{GeV}$.

The **N**eutral **C**urrent (NC) interaction displayed in fig. 3.1 is mediated via the exchange of a virtual γ or Z^0 boson with four-momentum $q = k - k'$ and mass $q^2 < 0$ where k', the four-momentum of the outgoing electron, may be chosen to be $k' = (E_{e'}, E_{e'} \sin\theta_{e'}, 0, E_{e'} \cos\theta_{e'})$, i.e. $\phi_{e'} := 0$. **C**harged **C**urrent (CC) reactions involving $W^\pm$ bosons would yield an outgoing neutrino ν_e and will not be considered here. In the case of elastic scattering $ep \to ep$, the Lorentz invariant momentum transfer squared

$$Q^2 := -q^2 = -(k - k')^2 = 2k \cdot k' = 2E_e E_{e'}(1 + \cos\theta_{e'}) \tag{3.2}$$

would suffice to characterize the process. Protons, however, are not point-like particles, but have a complex internal structure revealing itself at distances of $\lesssim 1\,\mathrm{fm}$, equivalent to $Q^2 \gtrsim 0.04\,\mathrm{GeV}^2$. As consequences the proton, represented by the lower blob in fig. 3.1, on the one hand has to be described in terms of *structure functions*, and on the other hand it usually breaks up since elastic reactions are strongly suppressed compared to inelastic ones with increasing Q^2 [1]. This inevitably leads to the need of another quantity to define the global outcome of an event, i.e. without differentiating the hadronic final state.

[1] To disentangle four-momenta like p from three-momenta, the latter are set in Roman font with arrows on top: $\vec{\mathrm{p}}$.

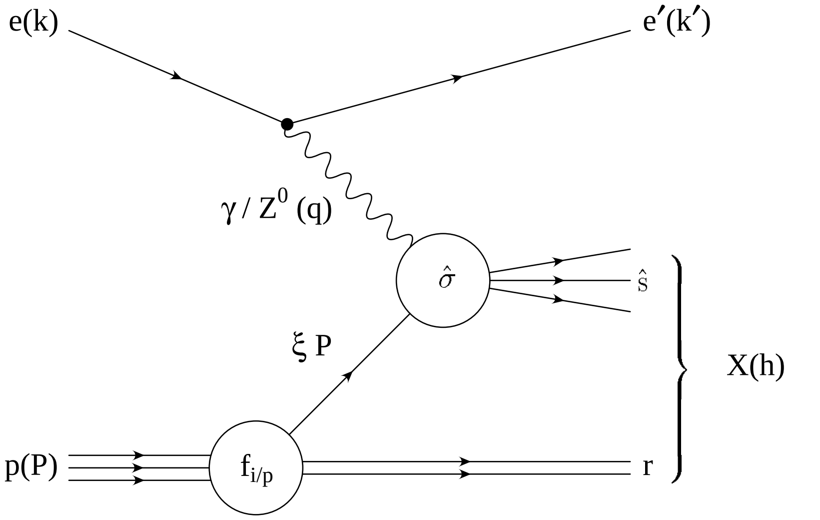


Figure 3.1: Diagram of the basic *ep* DIS process via neutral currents. The letters in parentheses label the corresponding four-momenta.

Looking at the photon-proton ($\gamma^* p$) interaction by itself, we have only one additional invariant at our disposal:

$$W^2 := (q+P)^2 = 2P \cdot q - Q^2, \tag{3.3}$$

which is limited by $0 \leq W^2 \leq s$.

Other popular choices are the two dimensionless variables[2]

$$x := \frac{Q^2}{2P \cdot q}, \tag{3.4}$$

$$y := \frac{P \cdot q}{P \cdot k} \tag{3.5}$$

with the neat property that $0 \leq (x, y) \leq 1$. In the context of the *infinite momentum frame*, where the proton is conceived of as a collinear stream of fast moving *partons* and masses are negligible, x can be interpreted as the fraction of the total momentum carried by the struck constituent as seen from the virtual boson. This reference frame has implicitly been adopted in fig. 3.1 as the hard scattering with a parton, labelled $\hat{\sigma}$ with mass $\sqrt{\hat{s}}$, is assumed to be incoherent and well separated from ensuing soft processes. In the simplest situation of a boson-parton collision where

$$\hat{s} := (q + \xi P)^2 = 2\xi q \cdot P - Q^2 = \left(\frac{\xi}{x} - 1\right) Q^2 \tag{3.6}$$

vanishes, x is identical to ξ of fig. 3.1 according to

$$\xi = \left(1 + \frac{\hat{s}}{Q^2}\right) x\,. \tag{3.7}$$

[2]The *scaling* variable x is also called Bjørken x and y may be denoted as *inelasticity*.

Under more complex circumstances with $\hat{s} > 0$, it can only be concluded that $x \leq \xi \leq 1$.

In fixed-target experiments, y is easily interpreted as relative energy loss of the scattered lepton because in the rest frame of the target denoted by $^\bullet$'s:

$$y = \frac{m_p(E_e^\bullet - E_{e'}^\bullet)}{m_p E_e^\bullet} = \frac{E_e^\bullet - E_{e'}^\bullet}{E_e^\bullet}. \tag{3.8}$$

Of course, only two of the four introduced kinematic quantities are independent of each other. The conversion formulae for the pairs $(Q^2, W^2) \leftrightarrow (x, y)$ are:

$$\begin{aligned} x &= \frac{Q^2}{Q^2 + W^2}, & y &= \frac{Q^2 + W^2}{s}, \\ Q^2 &= sxy, & W^2 &= s\,(1 - x)\,y. \end{aligned} \tag{3.9}$$

3.2 Reconstruction of the Kinematical Quantities

When confronted with real data, one has to reconstruct the kinematical quantities of an event from the measured energy depositions. Basically, we have four measurements at our disposal for the determination of two unknowns: the energies and polar angles of the scattered electron and the *current jet*, i.e. a collimated shower of hadrons produced by the struck parton: $E_{e'}$, $\theta_{e'}$, E_j, θ_j. Depending on the choice of input variables, several reconstruction methods exist, each with specific advantages and drawbacks.

Fig. 3.2 presents the lines of constant energy and polar angle for the scattered electron and the current jet in the (x, Q^2)-plane. Lines of constant $y = Q^2/(sx)$ are displayed for $y = 1, 10^{-1}, 10^{-2}$ and 10^{-3}. For reconstruction purposes it is best if the isolines of a quantity are closely staggered and intersect the isolines of the corresponding second variable predominantly under large angles. Henceforth, the electron data alone are sufficient for a good determination of Q^2 in the complete range shown. The large gaps, however, between energies $E_{e'} = 25\,\mathrm{GeV}$ and $E_{e'} = 40\,\mathrm{GeV}$ and angles $\theta_{e'}$ larger than 90° demonstrate that the electron method is not very reliable for large x or low y respectively. Despite the fact that hadronic energies and angles have much larger uncertainties than electromagnetic ones, the current jet data can provide better estimates for x or y in that region.

Yet, because we are interested in the shape of the hadronic final state, we have to restrict ourselves mainly to the electron quantities in order not to bias our results. The application of a *jet algorithm* like the ones described in section 4.2.3 to obtain E_j and θ_j is not advisable. Fortunately, there is also an inclusive method to derive the kinematics from hadronic measurements first proposed by Jacquet and Blondel [14]. Alternatively, one could rely on the angular measurements of the electron and the hadronic system alone [15]. These three methods will be presented in the next sections. For a comparison ref. [16] may be consulted.

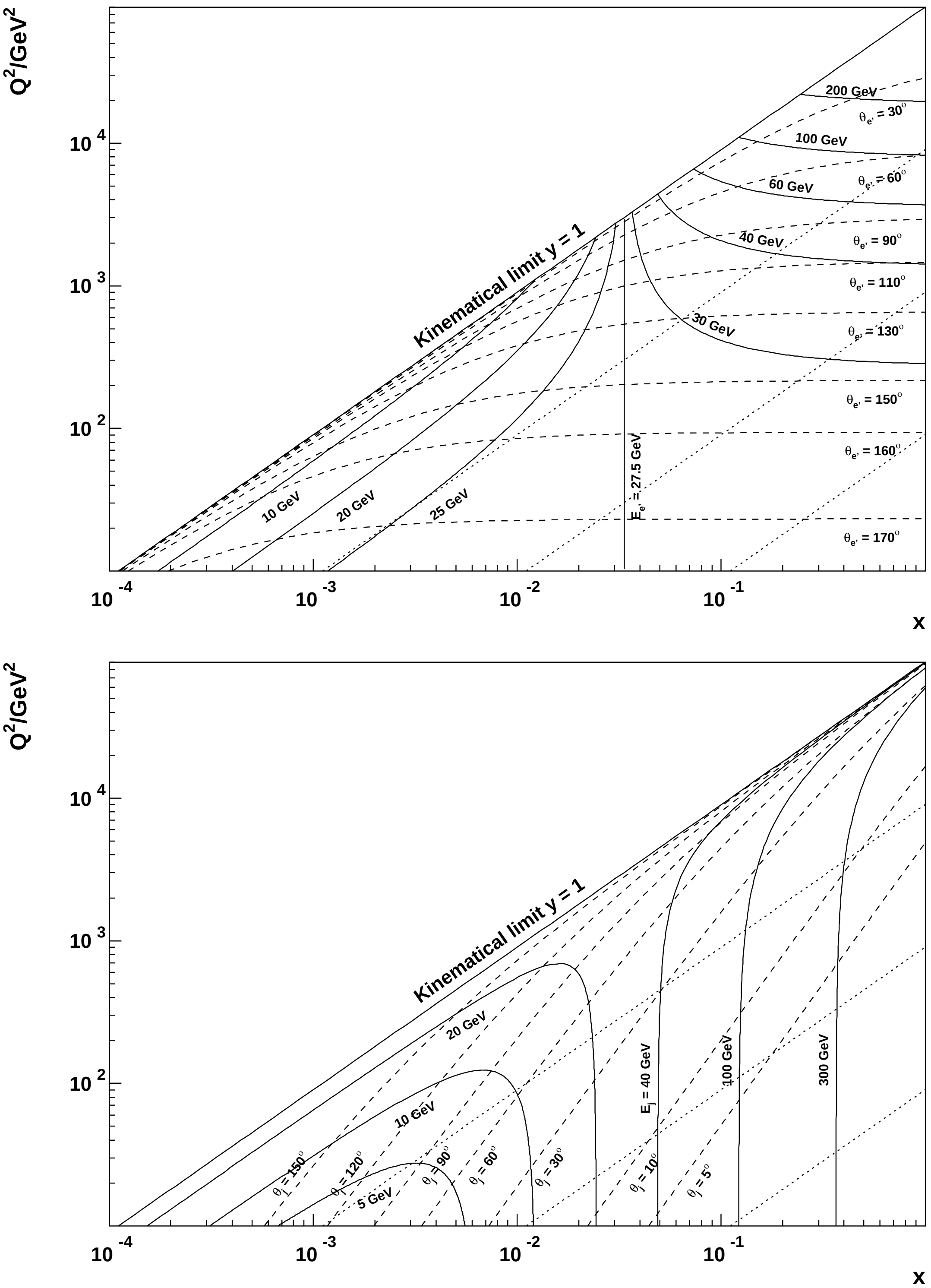


Figure 3.2: The two plots show the lines of constant energy (full lines) and polar angle (dashed) for the scattered electron (top) and the current jet (bottom) in an (x, Q^2)-plane. Lines of constant y (dotted except for $y = 1$) are drawn for $y = 1, 10^{-1}, 10^{-2}$ and 10^{-3}.

3.2.1 The Electron Method

The electron method is the simplest possibility to extract y and Q^2 and employs the scattered electron only. x and W^2 can then be calculated according to eqs. (3.9). Using the definitions (3.2) and (3.5), it follows directly:

$$y_e = 1 - \frac{P \cdot k'}{P \cdot k} = 1 - \frac{E_{e'}}{E_e}\frac{1-\cos\theta_{e'}}{2} = 1 - \frac{E_{e'} - p_{z_{e'}}}{2E_e}, \tag{3.10}$$

$$Q_e^2 = 2k \cdot k' = 2E_e E_{e'}(1+\cos\theta_{e'}) = 2E_e(E_{e'} + p_{z_{e'}}) = \frac{p_{t_{e'}}^2}{1-y_e}. \tag{3.11}$$

3.2.2 The Jacquet-Blondel Method

With h representing the four-momentum of the complete hadronic final state, it is clear from fig. 3.1 that $q = k - k' = h - P$. Furthermore, p_t-balance enforces $p_{t_h} = p_{t_e}$ so that

$$y_h = \frac{P \cdot (h-P)}{P \cdot k} = \frac{\sum\limits_i (E_i - p_{z_i})}{2E_e}, \tag{3.12}$$

$$Q_h^2 = \frac{p_{t_h}^2}{1-y_h} = \frac{\left(\sum\limits_i p_{x_i}\right)^2 + \left(\sum\limits_i p_{y_i}\right)^2}{1-y_h} \tag{3.13}$$

where i loops over all hadronic objects.

3.2.3 The Double Angle Method

Requiring $y_h = y_e$ and $p_{t_h} = p_{t_{e'}}$, the dependence on $E_{e'}$ and $\sum_i E_i$ for all hadrons i can be eliminated such that:

$$y_{da} = \frac{\sin\theta_{e'}(1-\cos\theta_h)}{\sin\theta_{e'}(1-\cos\theta_h) + \sin\theta_h(1-\cos\theta_{e'})}, \tag{3.14}$$

$$Q_{da}^2 = 4E_e^2 \frac{\sin\theta_h(1+\cos\theta_{e'})}{\sin\theta_{e'}(1-\cos\theta_h) + \sin\theta_h(1-\cos\theta_{e'})}. \tag{3.15}$$

3.3 The Born Cross Section

To derive a cross section formula for deep-inelastic $ep \to eX$ scattering, it will be required that the total process can be separated into a two step procedure. In the first part, involving very small space-time scales of $\mathcal{O}(1/Q)$ where the strong interaction is weak, the basic kinematic outcome of a reaction is fixed by the incoherent elastic scattering of the boson probe and a proton constituent. The ensuing soft hadronization of the struck parton and the proton remnant takes place at a time scale typically of the order of inverse

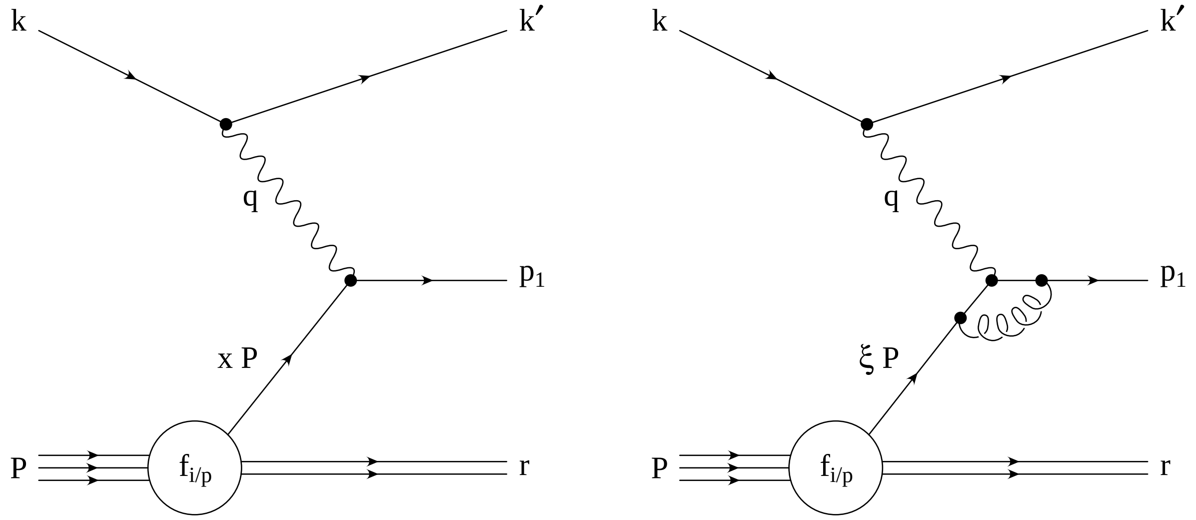


Figure 3.3: Quark-Parton-Model Feynman graph (left) and an example of a virtual correction to it (right).

hadron masses $1/M$ with $M \approx 200\,\mathrm{MeV} \ll Q$ instead and merely affects the detailed structure of the hadronic final state.

Without resolving such details, it is therefore possible to calculate a cross section by applying perturbation theory in lowest or **L**eading **O**rder (LO, here: $\mathcal{O}(\alpha^2\alpha_s^0)$) to the hard subprocess $\hat{\sigma}$ of fig. 3.1. The blob is then replaced by a single boson-parton vertex as shown in the **Q**uark-**P**arton-**M**odel (QPM) Feynman diagram of fig. 3.3. If the proton itself is left out, essentially an elastic two-body scattering reaction remains for which the cross sections have been calculated, s. e.g. [1]. One possibility to derive a general cross section formula is to assume partons with spin 1/2 and 0 [1]. Depending on the spin, one takes over the result from $e\mu$ respectively $e\pi$ scattering:

$$\frac{d\sigma^{e\mu}}{dQ^2} = \frac{2\pi\alpha^2}{Q^4}(1+(1-y)^2)\,, \tag{3.16}$$

$$\frac{d\sigma^{e\pi}}{dQ^2} = \frac{4\pi\alpha^2}{Q^4}(1-y)\,. \tag{3.17}$$

Here, α denotes the electromagnetic coupling strength and NC contributions from Z^0 exchange, which are suppressed at least $\propto Q^2/(Q^2+M_Z^2)$, are neglected.

At this stage, we must again consider the internal arrangement of partons in the proton that can be parameterized in the form of **p**arton **d**ensity **f**unctions (pdfs) $f_{i/p}(x)$ which represent the probability to find a constituent i with a momentum fraction in the interval $[x, x+dx]$. Denoting the spin-1/2 pdfs with $f_{i/p}(x)$ and the spin-0 ones with $f_{j/p}(x)$, the differential cross section can be written as

$$\frac{d\sigma^{ep}}{dQ^2} = \frac{4\pi\alpha^2}{Q^4}\left(\frac{1+(1-y)^2}{2}\sum_i q_i^2 f_{i/p}(x)dx + (1-y)\sum_j q_j^2 f_{j/p}(x)dx\right)\,, \tag{3.18}$$

where q_i and q_j are the corresponding electromagnetic charges of the constituent flavours i and j in units of the positron charge. Historically, the derivation of the cross section implied the evaluation of a leptonic tensor $L^{\mu\nu}$ prescribed by QED and its contraction with the most general form of a hadronic tensor $W_{\mu\nu}$ describing the hadronic vertex. This approach, employed e.g. in [2, 4], led to the definition of structure functions $F_1(x)$ and $F_2(x)$ which here translate to

$$F_1(x) = \sum_i \frac{1}{2} q_i^2 f_{i/p}(x)\,, \tag{3.19}$$

$$F_2(x) = \sum_i q_i^2 x f_{i/p}(x) + \sum_j q_j^2 x f_{j/p}(x)\,. \tag{3.20}$$

The double differential cross section (3.18) now reads

$$\frac{d^2\sigma^{ep}}{dQ^2dx} = \frac{4\pi\alpha^2}{xQ^4}\left((1-y)F_2(x) + xy^2F_1(x)\right)\,. \tag{3.21}$$

So far the partons served as a term for point-like constituents within the proton that lead to the experimentally observed $1/Q^4$-dependence of the inelastic cross section in contrast to the dramatic $1/Q^{12}$-decrease for elastic reactions. Combining this with the quark model by identifying partons and spin-1/2 quarks, one immediately concludes that $f_{j/p}(x) \equiv 0\,\forall j$ and henceforth

$$F_2(x) = 2xF_1(x)\,. \tag{3.22}$$

This is the experimentally well established Callan-Gross relation which demonstrates the fermionic nature of the charged proton constituents. In addition, one can conclude from eq. (3.21) that the cross section normalized to the corresponding one for point-like particles depends on the scaling variable x only. The observation of deviations can be attributed to the neglect of masses, intrinsic transverse momenta and — most importantly — the strong interaction.

3.4 The NLO Cross Section

Up to now, the proton was treated like a stream of collinear non-interacting quarks. However, in **N**ext-to-**L**eading **O**rder (NLO), i.e. $\mathcal{O}(\alpha^2\alpha_s^1)$, quarks can emit and absorb gluons which again may split up into two gluons or $q\bar{q}$ pairs and so forth. The rather simple picture involving structure functions $F_1(x)$, $F_2(x)$ to describe the composition of the proton in terms of quark densities alone has to be modified accordingly to include a gluon density. Two of the most important consequences are *scaling violations*, i.e. a dependence of $F_1(x)$, $F_2(x)$ on Q^2, and non-vanishing contributions to the longitudinal structure function

$$F_L := F_2 - 2xF_1\,. \tag{3.23}$$

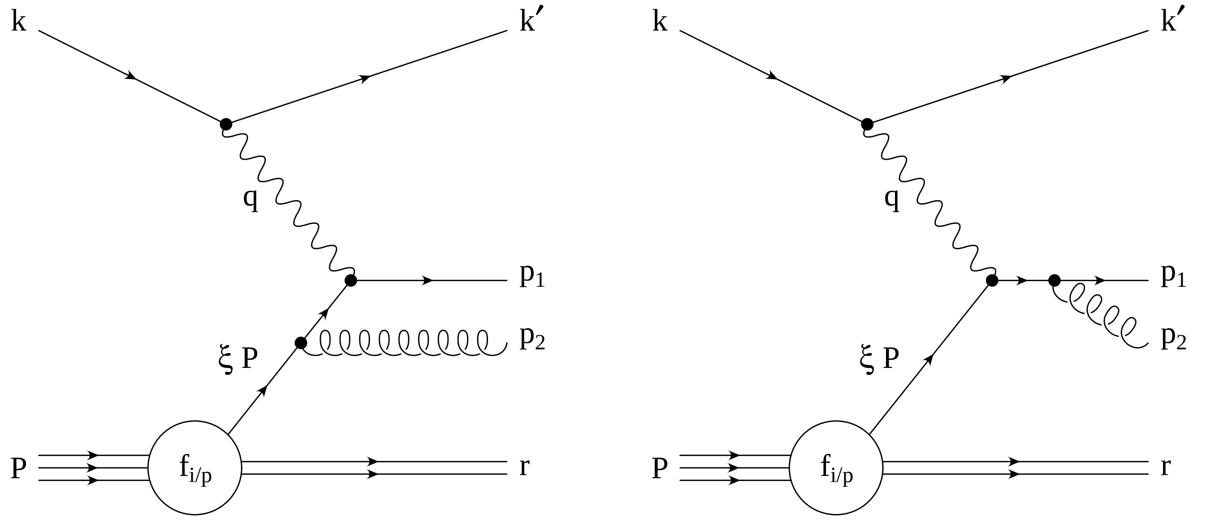


Figure 3.4: QCD-Compton Feynman graphs.

For low x, i.e. $x \lesssim 10^{-3}$, approximations reveal the effects to be proportional to $\alpha_s(Q^2)$ times the gluon density $g(x, Q^2)$:

$$\frac{dF(x,Q^2)}{d\log Q^2} \propto \alpha_s(Q^2) \cdot g(x,Q^2)\,, \tag{3.24}$$

$$F_L(x,Q^2) \propto \alpha_s(Q^2) \cdot g(x,Q^2)\,. \tag{3.25}$$

Both are intensively studied in DIS experiments and can be exploited in inclusive measurements to gain information on the strong coupling constant as well as the gluon density. For H1 publications on this topic consult refs. [17, 18].

In this analysis all hadronic final states are considered but will be differentiated with respect to suitably chosen characteristic variables (s. section 4.2), i.e. the measurement is semi-inclusive.

To $\mathcal{O}(\alpha_s^1)$ there are two kinds of processes resulting in two instead of one final state parton, where "parton" from now on is used synonymously for quarks, anti-quarks and gluons. Since they may lead to "real" effects like an additional *jet* they are called *real corrections*. In the first case of the **QCD**-**C**ompton graphs (QCDC), an additional gluon is emitted by the struck quark. The two possible Feynman diagrams are presented in fig. 3.4. The second kind with diagrams shown in fig. 3.5 involves the production of a $q\bar{q}$ pair from a reaction between the boson and a gluon emitted from the proton. They are expressively labelled **B**oson-**G**luon-**F**usion graphs (BGF).

To complete the set of contributions to *ep* DIS to $\mathcal{O}(\alpha_s^1)$, one also has to consider loop diagrams. One example is given on the right-hand side of fig. 3.3. The diagram itself results in an $\mathcal{O}(\alpha_s^2)$ add-on to the matrix element squared. However, since the final states of both graphs in fig. 3.3 are indistinguishable, an interference term of $\mathcal{O}(\alpha_s^1)$ has to be taken into account. For the purpose of differentiating the hadronic final state, this *virtual correction* is, of course, irrelevant and merely changes the total cross section.

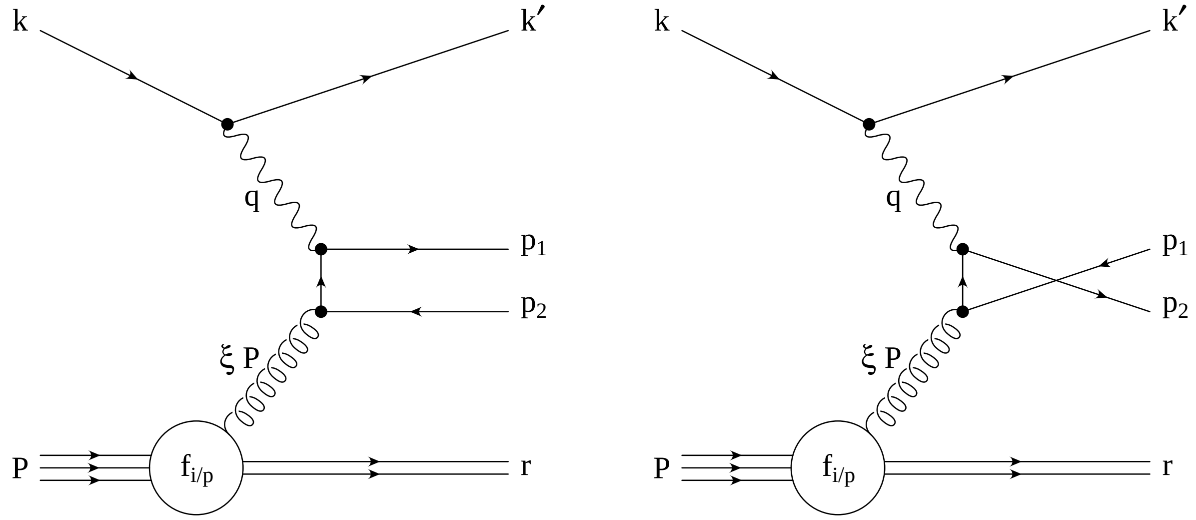


Figure 3.5: Boson-Gluon-Fusion Feynman graphs.

Focusing on the two-parton QCDC and BGF processes, three new degrees of freedom arise in the matrix elements corresponding to energy, azimuthal and polar angle of one parton. The other parton is then determined by energy-momentum conservation. Usually, the now five-fold differential cross section is written employing the variables

$$x_p \;:=\; \frac{Q^2}{2\,(\xi P)\cdot q} = \frac{x}{\xi}\,, \tag{3.26}$$

$$z_p \;:=\; \frac{P\cdot p_1}{P\cdot q} = 1 - \frac{P\cdot p_2}{P\cdot q}\,, \tag{3.27}$$

and ϕ, where ϕ denotes the angle between the planes fixed by the two outgoing partons on the one hand, and the scattered electron on the other hand in a suitable reference frame. This may be e.g. the *Breit frame* defined in the next chapter or any other reference system connected to it by a Lorentz boost along its z-direction. Concerning the characterization of the hadronic final state, no reference will be made to the electron. Therefore, the integration over ϕ can be performed beforehand such that for our purposes the NLO cross section may be labelled as

$$\frac{d^4\sigma^{ep}}{dQ^2 dx dx_p dz_p}\,. \tag{3.28}$$

The allowed ranges for x_p and z_p are $x \le x_p \le 1$ and $0 \le z_p \le 1$.

Yet, it does contain singularities! For the QCDC and BGF processes they can be extracted from ref. [19], which presents the complete $\mathcal{O}(\alpha_s^1)$ corrections to electroweak NC and CC ep DIS cross sections, to be:

$$
\begin{aligned}
d\sigma^{\mathrm{QCDC}} &\propto \frac{1+x_p^2 z_p^2}{(1-x_p)(1-z_p)}\,, && (3.29)\\
d\sigma^{\mathrm{BGF}} &\propto \frac{[x_p^2+(1-x_p)^2][z_p^2+(1-z_p)^2]}{z_p(1-z_p)}\,. && (3.30)
\end{aligned}
$$

Partially, they cancel against corresponding divergent terms of the virtual corrections. The remaining initial state mass singularities can be absorbed in a proper redefinition of the parton densities. Nevertheless, one has to be careful about the precise definition of quantities that shall be calculated by pQCD. In order to allow the divergences to compensate each other, the investigated variables must agree in the limits where the real and virtual corrections go to infinity. The *number of final state particles* for example would always be two in the first case and one in the latter.

To be more definite, we make use of eqs. (4.22)–(4.26) and can refer the limits $x_p \to 1$ and $z_p \to 0, 1$ to configurations where either p_1 or p_2 are soft (*infrared divergence*) or any pair of $\{p_1, p_2, r\}$ is collinear (*collinear divergence*). Hence, any quantity $F_n(\vec{\mathrm{p}}_1, \ldots, \vec{\mathrm{p}}_n)$ to be meaningful in pQCD has to fulfil the conditions

$$
\begin{aligned}
F_n(\vec{\mathrm{p}}_1, \ldots, \lambda\vec{\mathrm{p}}_i, \ldots, \vec{\mathrm{p}}_j = (1-\lambda)\vec{\mathrm{p}}_i, \ldots, \vec{\mathrm{p}}_n) &= F_{n-1}(\vec{\mathrm{p}}_1, \ldots, \vec{\mathrm{p}}_i, \ldots, \vec{\mathrm{p}}_{j-1}, \vec{\mathrm{p}}_{j+1}, \ldots, \vec{\mathrm{p}}_n)\\
&\text{and} \qquad\qquad (3.31)\\
\lim_{\lambda\to 0} F_n(\vec{\mathrm{p}}_1, \ldots, \lambda\vec{\mathrm{p}}_i, \ldots, \vec{\mathrm{p}}_n) &= F_{n-1}(\vec{\mathrm{p}}_1, \ldots, \vec{\mathrm{p}}_{i-1}, \vec{\mathrm{p}}_{i+1}, \ldots, \vec{\mathrm{p}}_n)
\end{aligned}
$$

for $0 < \lambda < 1$. That is, collinear splittings and soft particles do not affect F; it is *collinear and infrared safe*. One example for such a variable is *thrust*, or rather 1−thrust, defined by eq. (4.9). It is invariant with respect to collinear splittings and varies smoothly for one momentum approaching zero.

Chapter 4

Event Shapes

4.1 The Breit Frame

4.1.1 Introduction

Initially, all energy deposits measured with the H1 detector are given in the laboratory system. Since we are, however, merely interested in specific properties of the hadronic final state caused by the hard interaction, two problems arise:

1. The transverse momentum of the scattered electron is balanced by the hadronic system. Hence, from longitudinal and transverse momenta of the hadrons one can reconstruct the global event kinematics, but they are not characteristic of the underlying hadronic process.

2. Somehow we have to differentiate between the products of the hard reaction and the proton remnant. A maximal separation, which is not given in the laboratory system, is desirable.

For illustration fig. 4.1 shows a comparatively simple NC event with a clearly identifiable electron and a lot of hadronic activity on the opposite side with respect to ϕ. In addition, the proton rest manifests itself in the form of some clusters in the forward direction.

Here already, it is not too obvious how to define characteristic properties of the hadronic energy deposits. In the case of much more complicated average events, it becomes forbidding. The solution to the two problems is to apply a Lorentz transformation into the Breit frame of reference [20]. For an event of QPM type as shown in fig. 4.2, it is defined as the reference system where the incoming parton with momentum $xP^\star$ in $+z^\star$-direction[1] is back-scattered by a purely space-like boson of momentum $q^\star$.[2]

[1] Note that this is in contrast to e.g. [20] and [21] where the $+z^\star$-axis has been chosen for $q^\star$.

[2] To distinguish non-invariant quantities in the Breit system from those in the laboratory, they will be marked by a *.

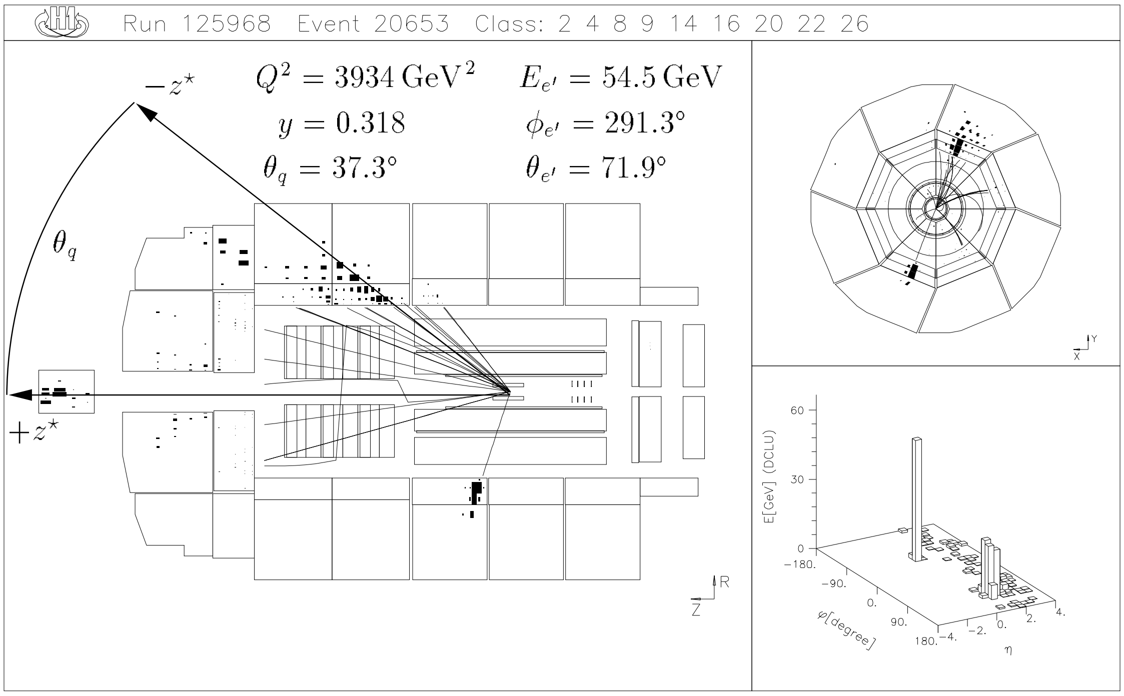


Figure 4.1: Example of an ep collision measured with the H1 detector which resembles a QPM-like configuration. The isolated electromagnetic cluster in the LAr with one track linked to it represents the scattered electron, whereas the broader energy deposits in the upper half of the LAr balance the electron p_t and correspond to a jet of hadrons produced by the struck quark. Some energy measured in the very forward direction indicates the proton remnant. The basic kinematic quantities deduced from the electron are given.

From $q^{\star 2} = q^2 = -Q^2$ and $xP^\star + q^\star = -xP^\star$, it follows that

$$q^\star = (0, 0, 0, -Q) \tag{4.1}$$

and

$$P^\star = \left(\frac{Q}{2x}, 0, 0, \frac{Q}{2x}\right) \tag{4.2}$$

where Q is defined to be

$$Q := \sqrt{-q^2} = \sqrt{Q^2}\,. \tag{4.3}$$

Thereby, the event and also the "soft" and "hard" physics is separated by the $(x^\star, y^\star)$-plane into a $+z^\star$ **R**emnant **H**emisphere (RH) and a $-z^\star$ **C**urrent **H**emisphere (CH). Simultaneously, the transverse momentum of the scattered quark has been eliminated such that the CH with its available energy of $Q/2$ is very similar to one half of an $e^+e^- \to q\bar{q}$ event with purely "time-like" energy $\sqrt{s}/2$, relating the "relevant energies" Q and $\sqrt{s}$.

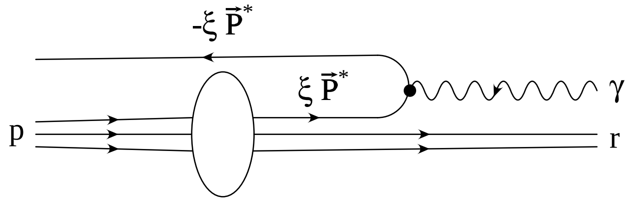


Figure 4.2: Diagram of a QPM-type ep collision as seen in the Breit frame.

Technically, the necessary Lorentz transformation is decomposed into a pure boost demanding that

$$2xP^\star + q^\star = \begin{pmatrix} Q \\ \vec{0} \end{pmatrix} \tag{4.4}$$

and a rotation afterwards to readjust $P^\star$ into the $+z^\star$-direction. In addition, the incoming and outgoing electron is usually required to lie in the $(x^\star, z^\star)$-plane. Their boosted and rotated four-momenta read

$$k^\star = \left(\frac{Q}{2y}(2-y), \frac{Q}{y}\sqrt{1-y}, 0, -\frac{Q}{2}\right) \tag{4.5}$$

and

$$k^{\star\prime} = \left(\frac{Q}{2y}(2-y), \frac{Q}{y}\sqrt{1-y}, 0, \frac{Q}{2}\right) . \tag{4.6}$$

Thus, another feature of the Breit frame is that the energy loss of the lepton, $E_e^\star - E_e{}^{\star\prime}$, vanishes.

4.1.2 Properties

For a better clarification of what happens to the clusters in fig. 4.1, we introduce another variable θ_q which corresponds to the polar angle of the scattered quark as seen by the electron:

$$\cos\theta_q = \frac{Q_e^2\,(1-y_e) - 4y_e^2 E_e^2}{Q_e^2\,(1-y_e) + 4y_e^2 E_e^2} . \tag{4.7}$$

A comparison of θ_q, drawn also in fig. 4.1, with the polar angle of the most energetic hadronic energy deposition in the LAr obviously demonstrates that they are approximately equal. When looked at it from the Breit frame, the $+z$-axis remains at its position, but the new $-z^\star$-direction is given by θ_q! As a result, the angular region of $0° \leq \theta \leq \theta_q$ is stretched until $0° \leq \theta^\star \leq 180°$ and, together with the remaining $180° - \theta_q$, the complete polar angular range for opposite ϕ, including the electron and most of the proton remnant, is squeezed. Intermediate ϕ angles other from ϕ_q and ϕ_e with $\phi_q - \phi_e = 180°$ constitute the transitional domain between the two cases.

For all variables defined in the next section, the energies and polar angles of the transformed four-vectors are the important input quantities. Azimuthal angles with respect to that of the scattered electron, i.e. $\phi_e^\star = 0°$, are rather insignificant. In fig. 4.3 we

therefore restrict ourselves to four plots showing the $\theta^{\star}(\theta,\phi)$-, $\theta^{\star}(\theta,E)$-, $E^{\star}(E,\phi)$- and $E^{\star}(E,\theta)$-functions for the sample boost of fig. 4.1. For simplicity, an exact balance in p_t of the electron and the hadronic final state is assumed so that $E = 85.5\,\text{GeV}$ is taken as default.
From these plots it can be concluded that:

1. $\theta^{\star}(\theta,\phi)$ peaks sharply at (θ_q,ϕ_q).

2. Only clusters opposite of the electron hemisphere in the laboratory system appear in the CH, i.e. $\theta^{\star} > 90°$.

3. $\theta^{\star}$ is independent of the energy of the boosted vector.

4. $E^{\star} \propto E$ with the slope strongest at ϕ_e and $\theta = 180°$.

The first property may lead to a serious deterioration of the resolution in polar angle in the Breit frame and is the motivation for the cut-off no. 5 introduced in section 5.5.

4.2 Definition of the Event Shapes

According to fig. 4.2, the hadronic final state in the simplest case consists of merely one parton with longitudinal momentum only and no mass. The hard interaction is of a purely electromagnetic nature. QCD induces deviations from this constellation. By investigating variables that are sensitive to these deviations, it is possible to learn more about perturbative and non-perturbative aspects of QCD.

All quantities, introduced in the next sections and generically labelled as F, are *event shapes* and have the following properties in common:

1. They are dimensionless.

2. For ep DIS they are defined in the Breit frame of reference.

3. In the limit of a QPM-like event $F = 0$, otherwise $F \geq 0$.

4. They are infrared and collinear safe, which is most important for a valid comparison with pQCD.

5. Soft fragmentation and hadronization processes cause discrepancies between fixed-order pQCD calculations and measured data that generally can be parameterized to be $\propto 1/Q^p$ with Q being the relevant energy scale and $p = 1$ or $p = 2$ in our circumstances.

The event shapes discussed in the next two subsections assume that the hard subprocess takes place in the CH alone. However, for $\xi > 2x$ it is possible for the CH to be completely empty, although experimentally this is improbable due to hadronization, backscattering, noise, etc. In order to be insensitive to such effects and to remain infrared

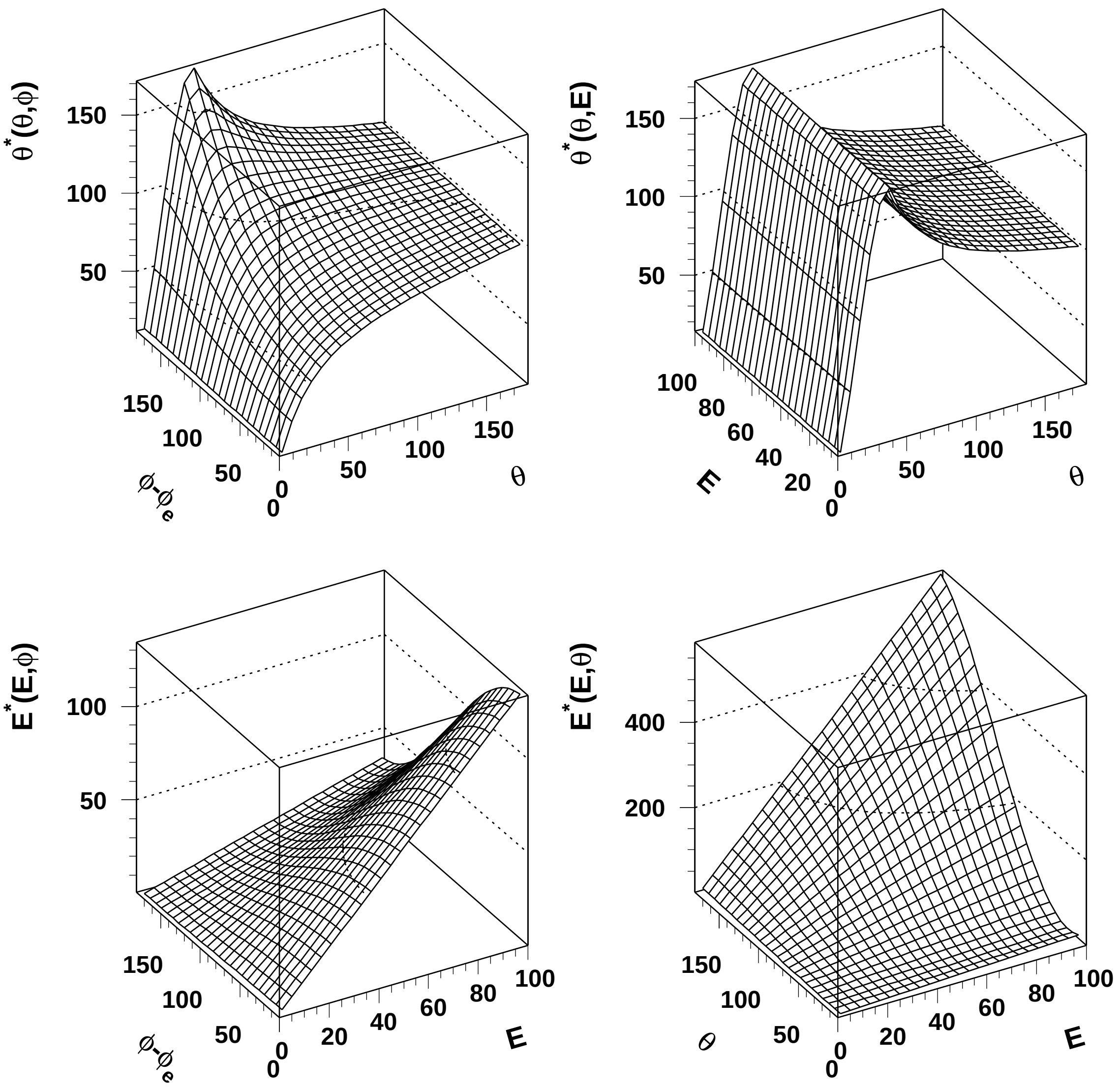


Figure 4.3: The $\theta^{\star}(\theta,\phi)$- (top left), $\theta^{\star}(\theta,E)$- (top right), $E^{\star}(E,\phi)$- (bottom left) and $E^{\star}(E,\theta)$-functions (bottom right) for the sample boost of fig. 4.1. The third input variable is set to $E = 85.5\,\mathrm{GeV}$, $\phi = \phi_q = 111.3°$, $\theta = \theta_q = 37.3°$ and again $\phi = \phi_q$ respectively.

safe [22], one has to specify what is meant by an "empty" or, vice versa, a "full" CH. We adopt the prescription that the total energy available in the CH $E^\star := \sum_{i\in\mathrm{CH}} E_i^\star$ has to exceed 20% of the value it should have according to QPM:

$$E^\star > 0.2 \cdot Q/2 = Q/10\,. \tag{4.8}$$

Otherwise the event is ignored. The precise value of the cut-off is motivated by a study of the measured energy flow in both hemispheres performed for ref. [23].

To keep event shapes dimensionless, it was originally suggested in [24] to normalize energies, momenta and masses to $Q/2$. "Empty" events would then lead to $F = 0$. Following the proposal in [25] to use for experimental reasons the actually present total energy $E^\star$ or momentum $P^\star := \sum_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star|$ instead, one would get ill-defined expressions. Therefore, F is set to zero in [21] analogously for this normalization in contrast to our definition, thereby affecting the total cross section σ and the left-most bin of the differential cross section $d\sigma/dF$. Except for the difference in σ, mean values $\langle F\rangle$, however, are not altered. In this work, we will mainly restrict ourselves to the study of the event shapes normalized to $E^\star$ or $P^\star$.

Note that in contrast to the first published experimental results on event shapes in *ep* DIS [25, 23] we adopt the modified naming scheme from ref. [21]. Except for τ_C defined below, the subscript indicates the quantity (E, P or Q) used for the normalization. The event shapes $1-T_C$, $1-T_Z$, B_C and ρ_C investigated in the above-mentioned publications will be labelled τ_C, τ_P, B_P and ρ_Q.

4.2.1 Event Shapes employing $z^\star$ as Event Axis

Choosing the direction of the exchanged boson $\vec{\mathrm{q}}^{\,\star} = (0, 0, -Q)$ as event axis $\vec{\mathrm{n}}$, one can, with fig. 4.2 in mind, easily deduce quantities F being zero for QPM-like configurations and $F > 0$ otherwise. The simplest event shapes thrust (or rather 1−thrust $\tau := 1 - T$) and jet broadening B can be written as [26, 24, 25]

$$\tau_P := 1 - \frac{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star \cdot \vec{\mathrm{n}}|}{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star|} = 1 - \frac{\sum\limits_{i\in\mathrm{CH}} |p_{li}^\star|}{P^\star}\,, \qquad \tau_Q := 1 - \frac{\sum\limits_{i\in\mathrm{CH}} |p_{li}^\star|}{Q/2}\,, \tag{4.9}$$

$$B_P := \frac{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star \times \vec{\mathrm{n}}|}{2\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star|} = \frac{\sum\limits_{i\in\mathrm{CH}} |p_{ti}^\star|}{2P^\star}\,, \qquad B_Q := \frac{\sum\limits_{i\in\mathrm{CH}} |p_{ti}^\star|}{Q} \tag{4.10}$$

where p_l and p_t denote the longitudinal respectively the transversal momentum components of $\vec{\mathrm{p}}$. The factor of 1/2 for B is conventional.

4.2.2 Event Shapes without Reference to $z^\star$ as Event Axis

A possibility for differentiation without reference to $z^\star$ as event axis is given by the jet mass ρ [27]:

$$\rho_E := \frac{\left(\sum\limits_{i\in\mathrm{CH}} p_i^\star\right)^2}{4\left(\sum\limits_{i\in\mathrm{CH}} E_i^\star\right)^2} = \frac{M^2}{4E^{\star 2}}\,, \qquad \rho_Q := \frac{M^2}{Q^2} \tag{4.11}$$

with M being the total mass in the CH.

Another way is the evaluation with respect to a new axis to be defined. Originally, event shapes were introduced in the context of e^+e^- annihilations, where, in distinction to ep DIS, a preferred direction like $z^\star$ is not given beforehand. The first definition deals with the *thrust axis* $\vec{\mathrm{n}}_T$ which is characterized as the normalized vector that maximizes the sum of all projections of momenta (absolute values) onto it. With regard to that, $\tau = 1-$thrust can now be described as

$$\tau_C := 1 - \max_{\vec{\mathrm{n}},\vec{\mathrm{n}}^2=1} \frac{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star \cdot \vec{\mathrm{n}}|}{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star|} = 1 - \frac{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i{}^\star \cdot \vec{\mathrm{n}}_T|}{P^\star}\,. \tag{4.12}$$

In contrast to e^+e^- physics, the maximizing procedure is applied in the CH alone. For merely one momentum vector, τ_C always equals zero.

A second kind of event axis brings the momentum tensor

$$\Theta^\star_{jk} := \frac{\sum\limits_{i\in\mathrm{CH}} \frac{p^\star_{j_i} p^\star_{k_i}}{|\vec{\mathrm{p}}_i^\star|}}{\sum\limits_{i\in\mathrm{CH}} |\vec{\mathrm{p}}_i^\star|} \tag{4.13}$$

into play. This real symmetric matrix is positive semi-definite with trace $Tr(\Theta) = \lambda_1 + \lambda_2 + \lambda_3 = 1$. For $0 < \lambda_3 \leq \lambda_2 \leq \lambda_1 < 1$ it describes an ellipsoïd with pairwise orthogonal axes named minor, semi-major and major with increasing eigenvalue. The major axis is similar but not identical to $\vec{\mathrm{n}}_T$. If $\lambda_3 = 0$, then the ellipsoïd degenerates into an elliptical cylinder with all momentum vectors in one plane. Is $\lambda_2 = 0$ as well, then all momenta are collinear and the corresponding normal area consists of one or two parallel planes. Utilizing the eigenvalues, we can define the C parameter [28]:

$$C_P := 3(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1)\,, \qquad C_Q := \frac{4P^{\star 2}}{Q^2} C_P\,. \tag{4.14}$$

The conventional factor of three ensures a maximal value of one for C_P.

4.2.3 Event Shapes employing Jet Algorithms for the Separation of the Remnant

Another approach of characterizing an event with regard to deviations from the QPM type does not make use of the CH. As depicted in section 3.1, the elementary reaction yields one hard parton moving along the $-z^{\star}$-direction. The ensuing production of soft partons during fragmentation and the final hadronization distort this picture to a limited extent only since no large transverse momenta with respect to the original one are involved. Basically, the one parton gets transformed into a tight stream of hadrons, flying along the original direction, which one refers to as a *jet*.

The inclusion of more complex processes into the pQCD calculation facilitates the production of two or more hard partons (s. section 3.4), and henceforth the events can acquire more than one jet in addition to the remnant jet. QPM-like events may also be called to be of a $(1+1)$-jet type in contrast to $(n+1)$-jet events with $n \geq 2$. In order to decide to which category a given constellation belongs, precise instructions on how to combine jets from an assortment of four-momenta are needed. The most basic of these *jet algorithms* [29] makes use of angular *cones* around seeds given by the input four-vectors. The two schemes we employ are of another type called *cluster* algorithms. Both, the Durham- or k_t- [30, 31] and the JADE-algorithm [32, 33] are applied in a modified form adapted to ep DIS in the Breit frame.

The central procedure is almost the same for both. Two distance measures are defined, one for distances between two four-vectors, y_{ij}, and another one for the separation of each from the remnant, y_{ir}. When all combinations are evaluated, the minimal value determines which are the closest two in y. If an y_{ij} was smallest, then these two are recombined to one new four-vector. In case of y_{ir} to be minimal, i is ascribed to the remnant. The whole routine is repeated until either all y's are larger than a lower bound y_{cut}, or until a certain number of jets is reached. The first prescription is used to divide a sample of reactions into sets of $(1+1)$-, $(2+1)$- and so forth events. The second approach is taken to employ y as an event shape variable. Here, y_{k_t} and y_{fJ} always denote the y-value where the transition $(2+1) \to (1+1)$ occurs. The respective distance measures are

$$y_{ij} := \frac{2E_i^{\star}E_j^{\star}(1-\cos\theta_{ij}^{\star})}{Q^2}, \tag{4.15}$$

$$y_{ir} := \frac{2E_i^{\star}xE_p^{\star}(1-\cos\theta_i^{\star})}{Q^2} \tag{4.16}$$

for the factorizable JADE-algorithm and

$$y_{ij} := \frac{2\min(E_i^{\star 2}, E_j^{\star 2})(1-\cos\theta_{ij}^{\star})}{Q^2}, \tag{4.17}$$

$$y_{ir} := \frac{2E_i^{\star 2}(1-\cos\theta_i^{\star})}{Q^2} \tag{4.18}$$

for the Durham-algorithm.

4.3 Event Shapes to $\mathcal{O}(\alpha_s)$

As explained above, the event shapes are designed to distinguish between a QPM-like topology of the hadronic final state with the relevant cross section given by eq. (3.21) and deviations from it that are described in lowest order pQCD by eq. (3.28). In the first case, F always equals zero, whereas in the latter, F depends on the two additional degrees of freedom x_p and z_p.

To derive explicit formulae, we first determine $p_1^\star$ and $p_2^\star$ from

$$p_1^\star + p_2^\star = \xi P^\star + q^\star = \begin{pmatrix} \frac{Q}{2x_p} \\ 0 \\ 0 \\ \frac{Q}{2x_p} - Q \end{pmatrix} \tag{4.19}$$

according to eqs. (4.1) and (4.2). Remembering that azimuthal angles are irrelevant, we define the y-components to be zero such that

$$p_1^\star := \frac{Q}{2} \begin{pmatrix} z_0 \\ z_t \\ 0 \\ z_3 \end{pmatrix}, \qquad p_2^\star := \frac{Q}{2} \begin{pmatrix} \bar{z}_0 \\ -z_t \\ 0 \\ \bar{z}_3 \end{pmatrix}. \tag{4.20}$$

Using $p_1^{\star 2} = p_2^{\star 2} = 0$ and

$$z_p = \frac{p^\star_{1_0} - p^\star_{1_3}}{Q} = \frac{z_0 - z_3}{2} = 1 - \frac{\bar{z}_0 - \bar{z}_3}{2}, \tag{4.21}$$

one can calculate z_0, $\bar{z}_0$, z_t, z_3 and $\bar{z}_3$ to be

$$z_0 = \frac{2x_p z_p - x_p - z_p + 1}{x_p}, \tag{4.22}$$

$$\bar{z}_0 = \frac{-2x_p z_p + x_p + z_p}{x_p}, \tag{4.23}$$

$$z_t = 2\sqrt{\frac{z_p}{x_p}(1 - x_p)(1 - z_p)}, \tag{4.24}$$

$$z_3 = \frac{1 - x_p - z_p}{x_p}, \tag{4.25}$$

$$\bar{z}_3 = \frac{z_p - x_p}{x_p}. \tag{4.26}$$

In the limit of $x_p \to 1$ and $z_p \to 1$ they evaluate to $z_0 = 1$, $z_3 = -1$ and $\bar{z}_0 = \bar{z}_3 = z_t = 0$ corresponding to a QPM-like event with only one final parton. Taking into account the mismatch in the sign of the $z^\star$-direction and noting that x_p is named ξ in [21], the formulae coincide with those of [20] and [21].

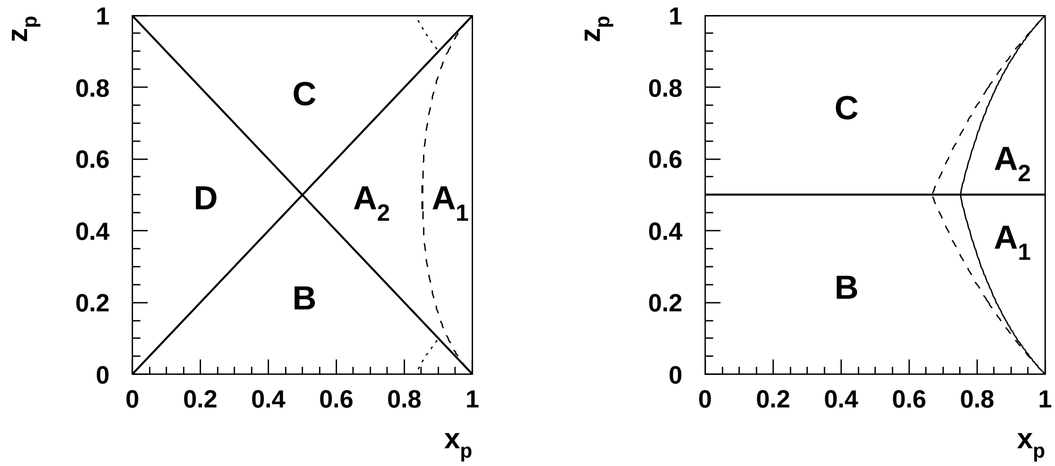


Figure 4.4: (x_p, z_p) phase space region for $\mathcal{O}(\alpha_s)$ corrections to DIS. See text for an explanation of the labels.

Finally, to achieve results for the event shapes, the available phase space in (x_p, z_p) has to be subdivided. Fig. 4.4 shows on the left-hand side the appropriate subregions for the definitions of sections 4.2.1 and 4.2.2 involving the CH only. The four triangles A to D correspond to:

A: Both partons are in the CH: $z_3, \bar{z}_3 \leq 0$.

B: Only parton $p_1^\star$ is in the CH: $z_3 \leq 0, \bar{z}_3 > 0$.

C: Only parton $p_2^\star$ is in the CH: $z_3 > 0, \bar{z}_3 \leq 0$.

D: The CH is empty: $z_3, \bar{z}_3 > 0$.

Except for subregion D which is excluded from our event shape definitions by the explicit requirement of a minimal energy of $Q/10$ in the CH, all results are listed in table 4.1. The dotted lines in fig. 4.4 point out two parts of regions B and C rejected in addition to D. Note that ρ_E, ρ_Q, C_P and C_Q vanish throughout B and C.

The additional separation into A_1 and A_2 indicated by the dashed line is necessary for τ_C. Depending on the angle enclosed by $\vec{\mathrm{p}}_1{}^\star$ and $\vec{\mathrm{p}}_2{}^\star$, the thrust axis $\vec{\mathrm{n}}_T$ is represented by

$$\vec{\mathrm{n}}_T = \begin{cases} \frac{\vec{\mathrm{p}}_1{}^\star + \vec{\mathrm{p}}_2{}^\star}{|\vec{\mathrm{p}}_1{}^\star + \vec{\mathrm{p}}_2{}^\star|}, & \text{for } \angle_{12} \leq 90^\circ\,(\mathrm{A}_1)\,\text{and} \\ \frac{\vec{\mathrm{p}}_1{}^\star - \vec{\mathrm{p}}_2{}^\star}{|\vec{\mathrm{p}}_1{}^\star - \vec{\mathrm{p}}_2{}^\star|}, & \text{for } \angle_{12} > 90^\circ\,(\mathrm{A}_2)\,. \end{cases} \tag{4.27}$$

In the first case, τ_C is equal to τ_P, in the latter, τ_C also measures momentum components perpendicular to the boson axis $\vec{\mathrm{n}}$ as can be seen from the appearance of z_t in the formula for A_2.

For the event shapes employing jet algorithms as explained in section 4.2.3, the complete phase space is accessed. As displayed on the right-hand side of fig. 4.4, it is split up into three main regions:

A: The two partons are merged together.

B: Parton $p_1^\star$ is clustered to the remnant.

C: Parton $p_2^\star$ is clustered to the remnant.

The full line designates the border between A, B and C for y_{k_t}, whereas the dashed line is valid for y_{fJ}. Here, the subdivision into A_1 and A_2 reflects the $\min(E_i^{\star 2}, E_j^{\star 2})$ condition of eq. (4.17).

It is remarkable that to $\mathcal{O}(\alpha_s)$ some of the definitions from section 4.2 lead to the same formulae, e.g. for τ_Q, ρ_Q and y_{fJ}. However, considering the complete phase space in (x_p, z_p), discrepancies appear. When higher orders are included, all event shapes will differ from each other. Table 4.2 gives an overview of the allowed ranges for the defined event shapes.

F	A_1	A_2	B	C
τ_P	$2(1-x_p)$		$1+z_3/z_0$	$1+\bar{z}_3/\bar{z}_0$
τ_Q	$(1-x_p)/x_p$		$1+z_3$	$1+\bar{z}_3$
B_P	$x_p z_t$		$z_t/2/z_0$	$z_t/2/\bar{z}_0$
B_Q	z_t		$z_t/2$	$z_t/2$
ρ_E	$x_p(1-x_p)$		0	0
ρ_Q	$(1-x_p)/x_p$		0	0
τ_c	$2(1-x_p)$	$1-\sqrt{4x_p^2 z_t^2+(1-2z_p)^2}$	0	0
C_P	$3(2x_p-1)^2 z_t^2/(z_0\bar{z}_0)$		0	0
C_Q	$3(2x_p-1)^2 z_t^2/(x_p^2 z_0\bar{z}_0)$		0	0
y_{fJ}	$(1-x_p)/x_p$		z_p	$(1-z_p)$
y_{k_t}	$z_0/\bar{z}_0\cdot(1-x_p)/x_p$	$\bar{z}_0/z_0\cdot(1-x_p)/x_p$	$z_p z_0$	$(1-z_p)\bar{z}_0$

Table 4.1: Formulae for the event shapes F in $\mathcal{O}(\alpha_s)$. For the definition of the phase space subdivisions A_1, A_2, B and C see fig. 4.4 and the explanations in the text.

F	Upper bounds		F	Upper bounds	
	to $\mathcal{O}(\alpha_s)$	absolute		to $\mathcal{O}(\alpha_s)$	absolute
τ_P	1	1	τ_Q	1	1
B_P	1/2	1/2	B_Q	1	1^*
ρ_E	1/4	1/4	ρ_Q	1	1^*
τ_C	$1-\sqrt{2}/2$	1/2			
C_P	3/4	1	C_Q	$3(\sqrt{5}-1)^3(3-\sqrt{5})/4$	2.89^*
y_{fJ}	1/2	0.84^*	y_{k_t}	$\leq 1/(4x)$	276^*

Table 4.2: Upper bounds for the defined event shapes. The starred numbers are not absolute values, but the largest ones encountered.

Chapter 5

Data Selection

The very basis of every collider experiment is formed by the equation

$$\frac{dN}{dt} = \mathcal{L}\sigma\,. \tag{5.1}$$

It relates the observed event rate dN/dt with the corresponding cross section σ. The machine dependent proportional factor $\mathcal{L}$ is called *luminosity* and has to be measured e.g. via comparison to a theoretically well-known reaction like the Bethe-Heitler process $ep \to ep\gamma$ employed in H1 (s. section 2.2.1).

In order to gather as many events of a certain kind as possible, one would like to have a large luminosity. For Gaussian beam profiles with horizontal and vertical widths $\sigma^*_{x_1}$, $\sigma^*_{x_2}$, $\sigma^*_{y_1}$ and $\sigma^*_{y_2}$, it is given by

$$\mathcal{L} = \frac{1}{e^2 f_b} \frac{I_1 I_2}{2\pi\left(\sigma^*_{x_1}\sigma^*_{y_1} + \sigma^*_{x_2}\sigma^*_{y_2}\right)} \tag{5.2}$$

where I_1, I_2 are the beam currents and f_b is the bunch frequency. By increasing the currents, the HERA crew was able to improve the performance considerably over the years. Fig. 5.1 gives an overview of the integrated luminosity $\mathcal{L}_{\text{int}}$ that was produced and accumulated during the running periods from 1992 up to 1997.[1]

5.1 Background Sources

Nevertheless, it should be kept in mind that background processes are enhanced right along. In fact, interactions with atoms of the rest gas in the beam pipe are even dominating! Table 5.1 gives an impression of the rates to expect. Therefore, a careful consideration of possible background sources is necessary. Basically, they can be subdivided into "true" background not related to ep collisions and misinterpreted competing ep reactions. The first are common to all analyses, whereas the latter depend on the topic under study.

[1]Cross sections σ are usually measured in barns b, where $1\,\text{b} = 10^{-24}\,\text{cm}^{-2}$. The luminosity integrated over the time $\mathcal{L}_{\text{int}}$ may therefore be given in e.g. pb^{-1}.

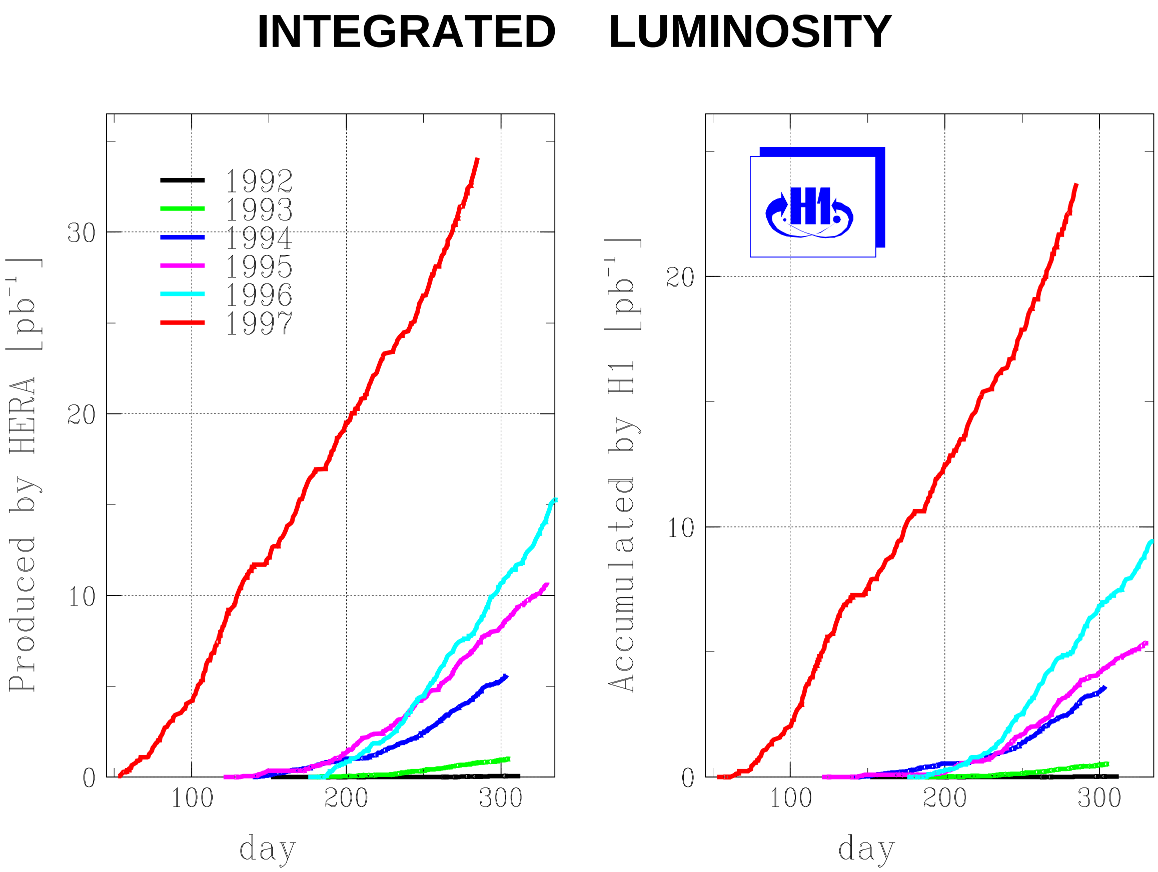


Figure 5.1: Integrated luminosity produced and accumulated during the running periods from 1992 up to 1997.

Cross sections and rates (at design luminosity)		
Beam gas interactions		50000 Hz
Cosmic myons		700 Hz
Photoproduction	1600 nb	25 Hz
NC DIS, $Q^2 < 100\,\mathrm{GeV}^2$	150 nb	2.2 Hz
NC DIS, $Q^2 > 100\,\mathrm{GeV}^2$	1.5 nb	0.022 Hz
CC DIS, $P_t > 25\,\mathrm{GeV}$	0.05 nb	0.001 Hz

Table 5.1: Cross sections and rates at a design luminosity of $1.5 \times 10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [8].

5.1.1 Non-*ep* Background

Beam Gas Interactions

Instead of colliding with particles from the other beam, it is also possible (and even probable) to hit residual gas atoms. Especially the high energetic protons are able to produce large numbers of particles that may be scattered into the detector. Most of these events can be rejected because there is no vertex or the tracks point to vertices outside the interaction region.

Cosmic Myons

The surface of the earth and henceforth the detector are constantly hit by myons of cosmic origin. Most of them cross the experiment out of time with respect to *ep* collisions and out of place, i.e. nowhere near the interaction region. Owing to their high rate, however, sometimes they pass in time right through the CTD and fake DIS events, despite the fact that they usually deposit little energy in the LAr. Still, they can be identified topologically by looking for back to back tracks in the IRON and CTD. In addition, the produced clusters are low energetic and imbalanced in transverse momentum.

Halo Myons

A second source of myons are stray protons interacting with material around the beam line. They always surround the proton beam (halo) and occasionally give rise to electromagnetic showers in the calorimeters, but since they are moving along the proton direction, they usually do not cause high p_t tracks in the CJC.

5.1.2 *ep* Background

Concerning our aim to investigate NC DIS, the following processes have to be considered as background.

Photoproduction

The term "photoproduction" derives from the picture of a quasi-real photon interacting with the proton, i.e. $Q^2 \approx 0$. Hence, the electron is only slightly deflected from its original direction and escapes the central detectors. Instead, it can be found with a certain efficiency in the electron tagger of the luminosity system. Nevertheless, the γp reaction may result in a hadronic final state involving large transverse momenta. The misidentification of a particle as the scattered electron is then, due to the very high rate (s. table 5.1), a main source of background.

Diffractive Events

Another name for these reactions is "**L**arge **R**apidity **G**ap events" (LRG's), where the rapidity is defined as

$$y_R := \frac{1}{2} \ln \frac{E + p_z}{E - p_z} . \tag{5.3}$$

In the case of massless four-momenta, it is identical to the pseudo-rapidity

$$\eta := -\ln\left(\tan\frac{\theta}{2}\right) . \tag{5.4}$$

In the region between the proton remnant and the current jet, which are normally connected via colour forces, an unusual gap without hadronic activity is exhibited. These events constitute around 5–10% of the NC DIS sample, yet they are not described by the pQCD calculations invoked and have to be excluded.

CC Reactions

At last, it may happen that clusters of the hadronic final state in one of the seldom CC events with a neutrino as scattered lepton are misidentified as electron. Since the neutrino, however, escapes detection completely, the total transverse momentum is imbalanced.

5.2 Trigger Scheme

To keep notwithstanding the high rates the dead time of the experiment low and specifically select true *ep* collisions, a trigger system has been set up in four stages [8]. The total read-out time of the detector of $\approx 1\,\mathrm{ms}$ is four orders of magnitude larger than that between two bunch-crossings of $\approx 0.1\mu\mathrm{s}$. Therefore, only a small part of the measured data, especially those from detector parts with short response times, is available for a fast decision.

The first level trigger L1 collects information from nine trigger systems attached to one subdetector each. These *trigger elements* are combined to form various subtriggers which provide a *KEEP* or *REJECT* signal within $2.5\,\mu\mathrm{s}$. A pipelining system stores the full data at the front end during the delay caused by L1 and ensures a dead time free running at this stage.

For future requirements intermediate trigger levels L2[2] and L3 are foreseen to operate during primary dead time of the read-out and are based on the same input as L1. However, they are able to evaluate a larger number of signals and their complex correlations. The decision times of these systems are designed to be around $20\,\mu\mathrm{s}$ and $800\,\mu\mathrm{s}$ respectively.

The last stage consists of the L4 software trigger, which has the raw data of the full event at its disposal. Depending on the time consumption, either fast algorithms specifically adapted to the requirements of L4 or parts of the standard offline reconstruction

[2] L2 was commissioned in 1996.

program H1REC [34] are applied to reach a quick decision. All events accepted by L4 are finally stored on tape. In addition, a small fraction of **L4 ReJecTed** events (L4RJT's) of $\approx 1\%$ are kept for monitoring purposes.

The complete scheme including a fifth level L5 discussed in the next section is illustrated in fig. 5.2.

5.3 Event Classification

Even after all four trigger stages, the reconstructed events kept for physics analysis contain to a large part unwanted or background reactions for most investigations. To save tape, network and computing resources, the program FPACK [35] used for platform independent data access has the possibility to skip events according to a classification word stored directly at the beginning of an event. Thereby, only the properly marked data tracks are completely read. Depending on the aim of a study, different classes, each matching to a set bit in the classification word, can be selected. Events falling in no physics class at all are **L5 ReJecTed** (L5RJT). They are kept as raw data, but do not appear fully reconstructed on **P**roduction **O**utput **T**apes (POT's), except for a small fraction for monitoring purposes as in the case of L4RJT's.

The chief requirement distinguishing NC DIS from background processes is the presence of a candidate for the scattered electron in the BEMC or LAr calorimeter. Therefore, loose criteria such as the existence of at least one compact mainly electromagnetic cluster with a minimal energy of several GeV are applied to define two DIS classes, a low Q^2 class (no. 11) in case of the BEMC and a high Q^2 one (no. 9) for the LAr.

Using additionally data from the tracking and myon systems, beam-induced and cosmic background is rejected more effectively.

5.4 Preselection

Due to the soft cuts applied in the event classification, the resulting classes are usually still too contaminated with background. In addition, some analyses might not need all parts of the **intended** content. Yet, physics studies are refined and iterated several times before being finished so it is recommendable to reduce the data sample further. This is done by a preselection which copies the selected events to a local disk for fast access.

Since the operation of such a complicated machinery like HERA and H1 is no easy task, running conditions are subject to variations. Any time a significant change in status like the (temporary) failure of an important detector component occurs, a new *run* is initiated. Only runs qualified as "medium" or "good" are allowed by the preselection. Additionally, the most important detector components for this analysis, i.e. the CJC1/2 and LAr for class 9 and also the BEMC and BPC for class 11, are required to be operational.

Relying on the fact that at least the scattered electron should cause a well measured track, the existence of a reconstructed vertex is mandatory.

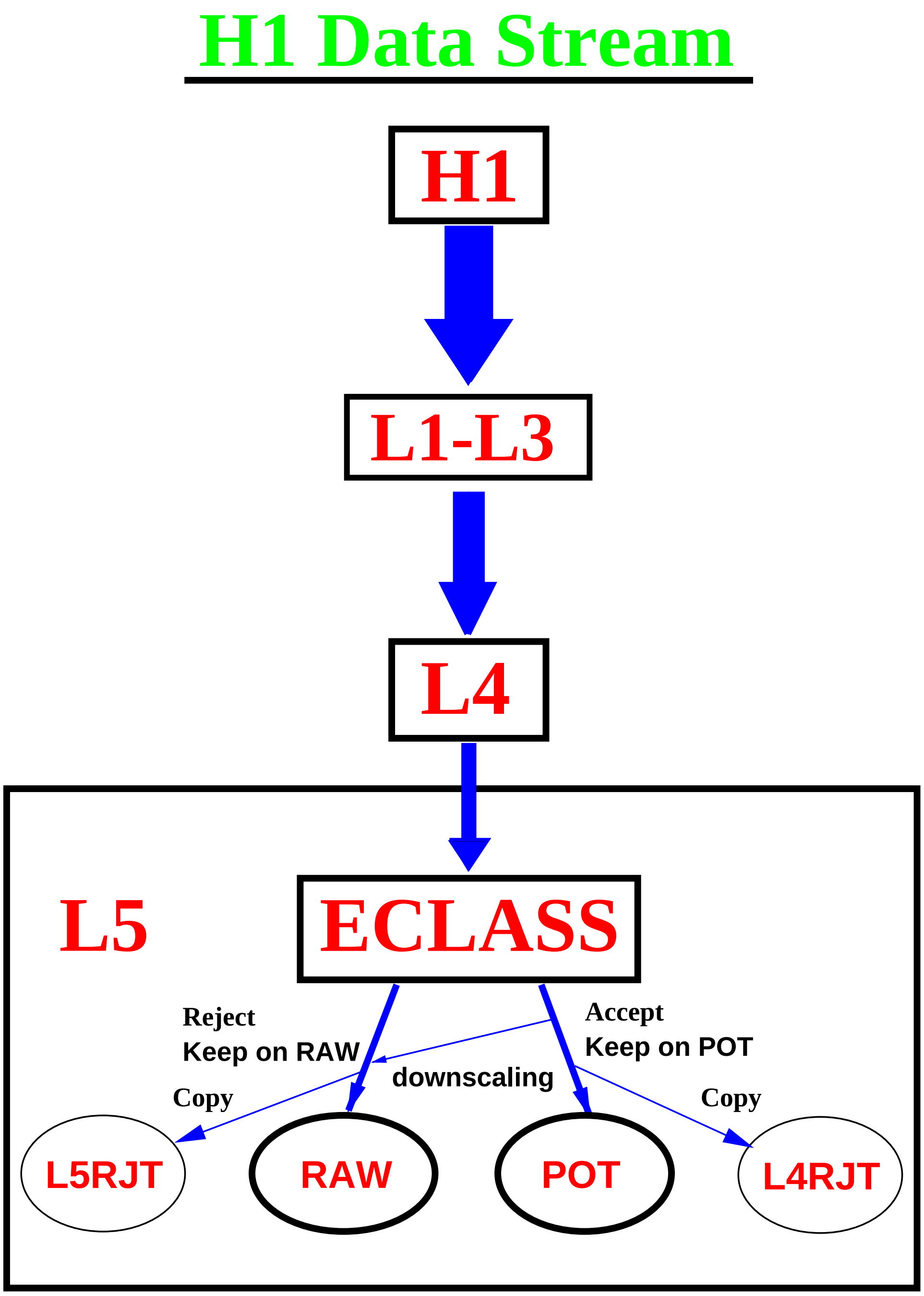


Figure 5.2: A principal diagram of the data stream in H1. Three levels of triggers operating with fast signals from nine detector branches and one software trigger evaluating the complete raw data are responsible for the rejection of background. A fifth level L5 works on fully reconstructed events and categorizes them into physics classes that are stored as POT's. To avoid overflowing them, the classes may be downscaled. Events accepted by L4 but falling in no physics class at all are kept as raw data only. Small samples of L4 and L5 rejected data are retained for monitoring purposes.

As mentioned above, it is evident for NC DIS that finally a clear candidate for the scattered electron should be found. In distinction to the event classification, stricter cuts have to be fulfilled now. In **H1** a program package for **PH**ysics **AN**alysis (H1PHAN [36]) is available containing several algorithms for that purpose. Two of them, QFSELH/M and QESCAT, are applied. Both exploit the properties that the scattered electron should be isolated from the jet of the struck parton due to p_t-balance and that it deposits its energy in form of a compact electromagnetic shower mainly in the ECAL part of the LAr or in the BEMC.

In the case of QFSELH/M, some additional requirements are imposed on the found candidate. A cone of radius 0.5 in azimuthal angle ϕ and pseudo-rapidity η is drawn around the candidate. At most 5% of its energy is allowed to be measured within this cone. Additionally, a matching track has to be found within 5° for high Q^2. For electrons in the BEMC, a hit in the BPC with a maximal distance of 5 cm from the cluster center projected onto the BPC (x, y)-plane has to be measured and the cluster radius may not exceed 5 cm.

For the purpose of the preselection, events are accepted if a candidate is determined either way. In the final analysis QFSELH/M is utilized for the low Q^2 sample and QESCAT for the high Q^2 events. A comparison revealed marginal differences between the two electron finders.

At last, a minimal energy of the candidate of 9 GeV is called for, and it is ensured that safe regions of the calorimeters are hit by cuts on the polar angle of $10° < \theta < 150°$ for the LAr and $157° < \theta < 173°$ for the BEMC respectively. Because the minimal Q^2 demanded later is $49\,\mathrm{GeV}^2$, a slightly smaller value of $45\,\mathrm{GeV}^2$, calculated according to eq. (3.11), has to be surpassed by the low Q^2 events.

5.5 Final Cut Scenario

Before the last selection is applied, an additional routine rejects residual events due to cosmic myons. Furthermore, there may be no photons detected in the PD.

In order to achieve a clean DIS sample, the final cuts have to take into account detector acceptances, efficiencies and resolutions as well as the diverse background sources described in section 5.1. The latter have already been suppressed to some extent in the selection procedure above. Additionally, possible constraints of the theoretical model to compare with have to be considered.

The final cuts are subdivided into two basic categories:

- **Phase space cuts:**
  Due to unavoidable limitations either of the experimental apparatus or the theoretical model, it is not possible to choose at will a phase space to investigate. Hence, common kinematic requirements have to be imposed on the data as well as on the theory.

- **Data quality cuts:**
  These are necessary to ensure that background is suppressed and the measurement is of good quality. Then, the data can be corrected for detector effects and may be compared to theoretical predictions.

First, the phase space cuts will be defined. For a quick overview see table 5.2.

Phase space cuts:

- **low Q^2: $49 < Q^2/\,\mathrm{GeV}^2 < 100$, high Q^2: $196 < Q^2/\,\mathrm{GeV}^2 < 10000$:**
  To examine the Q-dependence of the event shapes, all events are grouped into eight bins in Q: 7–8 GeV, 8–10 GeV, 14–16 GeV, 16–20 GeV, 20–30 GeV, 30–50 GeV, 50–70 GeV and 70–100 GeV. Note that there is a gap from 10 GeV to 14 GeV which corresponds to the excluded transition region between the BEMC and LAr calorimeters. The lower bound of 7 GeV is motivated by the fact that an energy of $Q/2 = 3.5\,\mathrm{GeV}$ should be available in the CH of the Breit frame. For $Q > 100\,\mathrm{GeV}$ there is not enough statistics at hand.
  In case of data, Q^2 is identified with Q_e^2 according to eq. (3.11).

- **$0.05 < y < 0.8$:**
  As explained in section 3.2, the electron method is not well suited for the reconstruction of y-values as low as 0.05. Hence, these events are rejected. At high y, radiative corrections due to an additional real photon emission of the scattered electron (s. section 6.1.1) are enormous and have to be avoided.
  For data, y is identified with y_e according to eq. (3.10) down to 0.15. Below that value, y_h from eq. (3.12) is taken instead.[3] Note that nevertheless $y_e > 0.05$ is asked for.

- **low Q^2: $E_{e'} > 14\,\mathrm{GeV}$, high Q^2: $E_{e'} > 11\,\mathrm{GeV}$:**
  Lower limits on the electron energy are imposed for several reasons: First, the trigger efficiency is above 99% [37, 13]. Moreover, there are numerous clusters of hadronic origin, e.g. pions, which are low energetic and may fake electrons, especially when the real one did not even hit the main detectors e.g. in photoproduction. At last, true leptons which lost a large part of their original energy due to radiative effects are excluded. Thereby, a good measurement of $E_{e'}$ should be reached, which is most important for the boost into the Breit frame.

- **low Q^2: $157° < \theta_{e'} < 173°$, high Q^2: $30° < \theta_{e'} < 150°$:**
  These cuts reflect the coverage in polar angle of the BEMC and LAr calorimeters, although the upper limit of 173° for the low Q^2 sample is redundant owing to $Q^2 > 49\,\mathrm{GeV}^2$. For the high Q^2 events, the forward region $\theta_{e'} \leq 30°$ with its high hadronic activity is left out to avoid misidentifications of the scattered electron.

[3]Concerning the boost into the Breit frame, it follows that $x = Q_e^2/(sy_h)$ for $0.05 < y_e < 0.15$!

Phase space cuts		
	low Q^2 sample (BEMC)	**high Q^2 sample (LAr)**
Cut 1	$49\,\mathrm{GeV}^2 < Q^2 < 100\,\mathrm{GeV}^2$	$196\,\mathrm{GeV}^2 < Q^2 < 10000\,\mathrm{GeV}^2$
Cut 2	$0.05 < y < 0.8$	
Cut 3	$E_{e'} > 14\,\mathrm{GeV}$	$E_{e'} > 11\,\mathrm{GeV}$
Cut 4	$157° < \theta_{e'} < 173°$	$30° < \theta_{e'} < 150°$
Cut 5	$20° < \theta_q$	
Cut 6*	$E^\star > Q/10$	

Table 5.2: Final phase space cuts. Note that the starred cut no. 6 is applied for τ_P, B_P, ρ_E, τ_C and C_P only.

- **$20° < \theta_q$:**
  θ_q, eq. (4.7), indicates the direction of the $-z^\star$-axis in the laboratory system. The requirement $\theta_q > 20°$ ensures a sufficient detector resolution in polar angle (s. the discussion in section 4.1.2). In the low Q^2 regime, this cut-off is automatically fulfilled owing to the previous selection in Q^2 and y; a small number of high Q^2 events, however, is discarded, cf. fig. 5.3.

- **τ_P, B_P, ρ_E, τ_C and C_P only: $E^\star > Q/10$:**
  As explained in section 4.2, this cut is essential to keep the event shapes τ_P, B_P, ρ_E, τ_C and C_P infrared safe. In this sense, it is part of their **definition**. Experimentally, it ensures a minimum of hadronic activity in the CH and suppresses events substantially influenced by noise and leakage out of the RH.

For an overview of the data quality cuts see table 5.3.

Data quality cuts:

- **τ_P, B_P, ρ_E, τ_C and C_P only: $N_{\mathrm{CH}} \geq 2\,;\ (F - F_{\min})\,,(F_{\max} - F) > 5 \cdot 10^{-5}$:**
  Asking for at least two objects in the CH, "unnatural" peaks at zero for ρ_E, τ_C and C_P in the low Q^2 region are removed. "Unnatural" means that they are caused either by leakage out of the remnant fragmentation region or by cutting off "regular" jets just at the border between the two hemispheres. In fact, B_P does exhibit a pronounced peak at 0.5 for the rejected events showing the hadrons/clusters to have polar angles marginally larger than the required 90° in the Breit frame. For the same reason, the extreme values $F_{\min}$ and $F_{\max}$ are excluded by very small cut-offs.

- $\boldsymbol{E_{\mathrm{forw}} > 0.5\,\mathrm{GeV}}$:
  Requiring a minimal energy deposition in the forward region defined by $4° < \theta < 15°$, diffractive events not described by usual pQCD calculations are discarded.

- $\boldsymbol{5.7° \leq \theta_{\mathrm{cl}} \leq 170°}$:
  This cut reflects the angular acceptance for clusters that are completely contained in the LAr or BEMC calorimeters.

- $\boldsymbol{30\,\mathrm{GeV} < (E - P_z) < 65\,\mathrm{GeV}}$:
  If all emerging particles of an ep reaction could be measured perfectly, then

$$(E - P_z) := \sum_i (E_i - p_{z_i}) \tag{5.5}$$

  would be $E_p - p_{z_p} + E_e - p_{z_e} = 2 \cdot E_e = 55\,\mathrm{GeV}$. Losses due to limited acceptance, e.g. around the beam pipe, transform this peak into a broad distribution with tails down to very low values. Restricting the range of $(E - P_z)$ effectively reduces the photoproduction background and the size of radiative corrections.

- **low** $\boldsymbol{Q^2}$: $\boldsymbol{P_t < 7.5\,\mathrm{GeV}}$, **high** $\boldsymbol{Q^2}$: $\boldsymbol{P_t < 15\,\mathrm{GeV}}$:
  The total transverse momentum

$$P_t := \sqrt{\left(\sum_i p_{x_i}\right)^2 + \left(\sum_i p_{y_i}\right)^2} \tag{5.6}$$

  of a NC event should normally be zero. To suppress remaining background from CC reactions or badly measured events, maximal P_t's of 7.5 GeV and 15 GeV are allowed in the low Q^2 and high Q^2 samples respectively.

- $\boldsymbol{|(E_{e'} - E_{da})/E_{da}| < 0.25}$:
  According to

$$E_{da} := \frac{Q^2_{da}}{4E_e} + E_e(1 - y_{da}), \tag{5.7}$$

  the energy of the scattered electron can be derived from angular information only. This fact is exploited e.g. for the energy calibration of the calorimeter. Asking for both values to be compatible within $\approx 3\sigma$ suppresses events strongly affected by QED radiation.

- $\boldsymbol{-35\,\mathrm{cm} < z_{\mathrm{v}} - \langle z_{\mathrm{v}} \rangle < 35\,\mathrm{cm}}$:
  Here, it is enforced that an interaction vertex could be determined which lies within $\pm 35\,\mathrm{cm} \approx 3\sigma$ around the average z-position for the corresponding run period. In 1994, $\langle z_{\mathrm{v}} \rangle \approx 3\,\mathrm{cm}$, and in 1995–1997, $\langle z_{\mathrm{v}} \rangle \approx -1\,\mathrm{cm}$.

Data quality cuts		
	low Q^2 sample (BEMC)	**high Q^2 sample (LAr)**
Cut 7*	$N_{\rm CH} \geq 2\,, \quad F - F_{\rm min} > 5 \cdot 10^{-5}\,, \quad F_{\rm max} - F > 5 \cdot 10^{-5}$	
Cut 8	$E_{\rm forw} > 0.5\,{\rm GeV}$	
Cut 9	$5.7^\circ \leq \theta_{\rm cl} \leq 170^\circ$	
Cut 10	$30\,{\rm GeV} < (E - P_z) < 65\,{\rm GeV}$	
Cut 11	$P_t < 7.5\,{\rm GeV}$	$P_t < 15\,{\rm GeV}$
Cut 12	$\lvert(E_{e'} - E_{da})/E_{da}\rvert < 0.25$	
Cut 13	$-35\,{\rm cm} < z_{\rm v} - \langle z_{\rm v}\rangle < 35\,{\rm cm}$	
Cut 14	$E_{\rm backw} < 10\,{\rm GeV}$	
Cut 15		$\lvert\phi_{\rm imp} - \phi_{\rm cr}\rvert \geq 2^\circ$ $\lvert z_{\rm imp} - z_{\rm cr}\rvert \geq 5\,{\rm cm}$

Table 5.3: Final data quality cuts. Note that the starred cut no. 7 is applied for τ_P, B_P, ρ_E, τ_C and C_P only. Depending on the theory to compare with, the cuts nos. 7 and 8 may have to be considered as phase space cuts, cf. the discussion in section 7.3.

- $\boldsymbol{E_{\rm backw} < 10\,{\rm GeV}}$:
  Due to the very limited capability of the BEMC calorimeter to measure hadronic energies, events with significant activity there are discarded. In addition, unidentified scattered electrons in the BEMC may lead to a rejection.

- **LAr only:** $\boldsymbol{\lvert\phi_{\rm imp} - \phi_{\rm cr}\rvert \geq 2^\circ}$,
  $\boldsymbol{\neg\,(-65\,{\rm cm} < z_{\rm imp} < -55\,{\rm cm}) \wedge \neg\,(15\,{\rm cm} < z_{\rm imp} < 25\,{\rm cm})}$:
  In order to ensure a reliable measurement of the scattered electron, partially inefficient regions such as cracks between calorimeter modules (ϕ-cracks) or wheels (z-cracks) have to be avoided. $\phi_{\rm imp}$ and $z_{\rm imp}$ denote the impact coordinates of the electron.

The only considerable background remaining hereafter stems from photoproduction. In the low Q^2 sample it is estimated to be less than 3%, for the high Q^2 region it is negligible [23]. Residual radiative effects are accounted for by the correction procedure described in ch. 7.

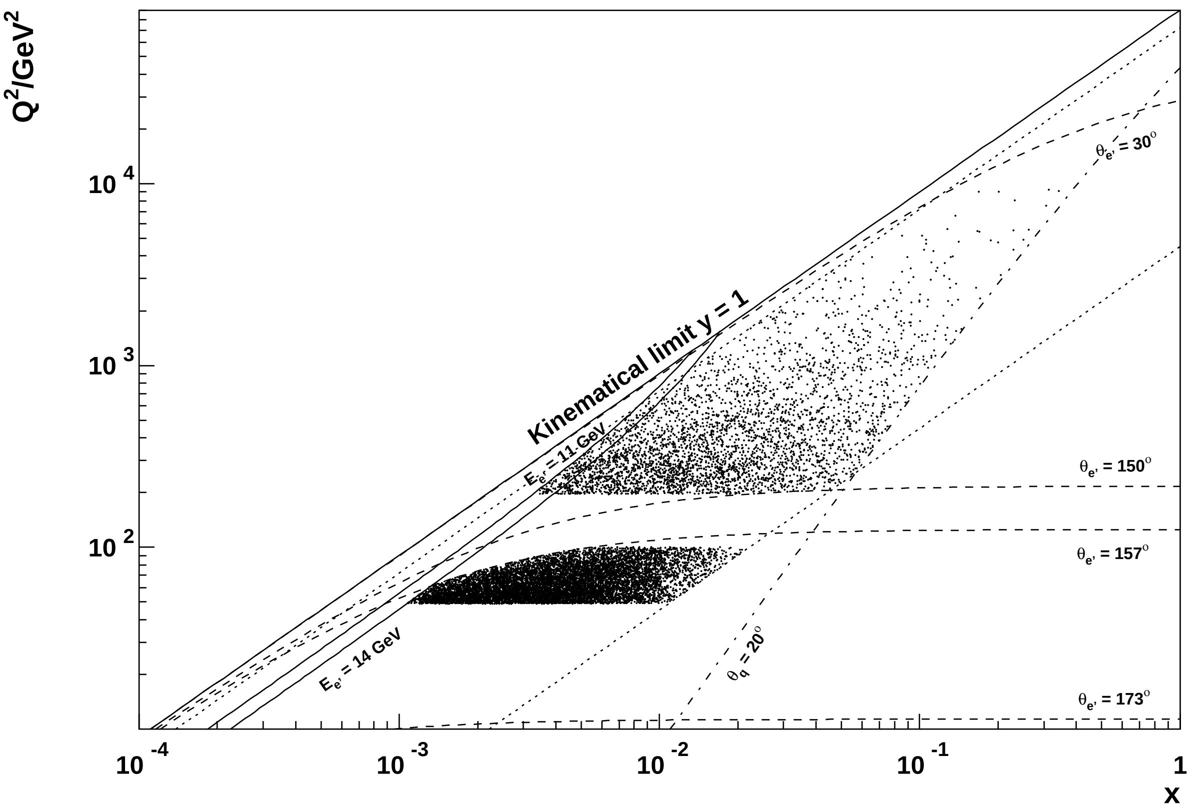


Figure 5.3: Distribution of finally selected events in the (x, Q^2)-plane. For clarity, 1994 data only are shown. The curves mark out the phase space cuts nos. 2–5 in y, $E_{e'}$, $\theta_{e'}$ and θ_q as indicated; the cuts in y correspond to the (unlabelled) dotted lines.

To illustrate the stability of the described selection procedure, the number of events accumulated in 1994 for the low Q^2 as well as the high Q^2 sample is plotted versus the integrated luminosity $\mathcal{L}_{int}$ in fig. 5.4. Also demonstrated is the constancy of the number of events gathered per $\mathcal{L}_{int}$ of $\approx 1\,\mathrm{pb}^{-1}$ for all four years contributing to the high Q^2 sample. The term "pre-final" refers to the production of data n-tuples where additionally to the preselection the cosmic filter, the anti-photon tag and the cuts nos. 8, 13 and 14 are in effect already. QFSELM/QESCAT is employed for the determination of the scattered electron. Statistics on the selection procedure can be looked up in table 5.4.

year	$\mathcal{L}_{int}/\mathrm{pb}^{-1}$	# preselected	# pre-final	# final
1994 high Q^2	3.20	12177	9467	3646
1995 high Q^2	4.29	16199	12817	5006
1996 high Q^2	8.49	30319	23924	9478
1997 high Q^2	22.26	100466	61711	24477
total high Q^2	38.24	159161	107919	42607
1994 low Q^2	3.18	35051	24690	9761

Table 5.4: Integrated luminosities and events gathered in the selection steps.

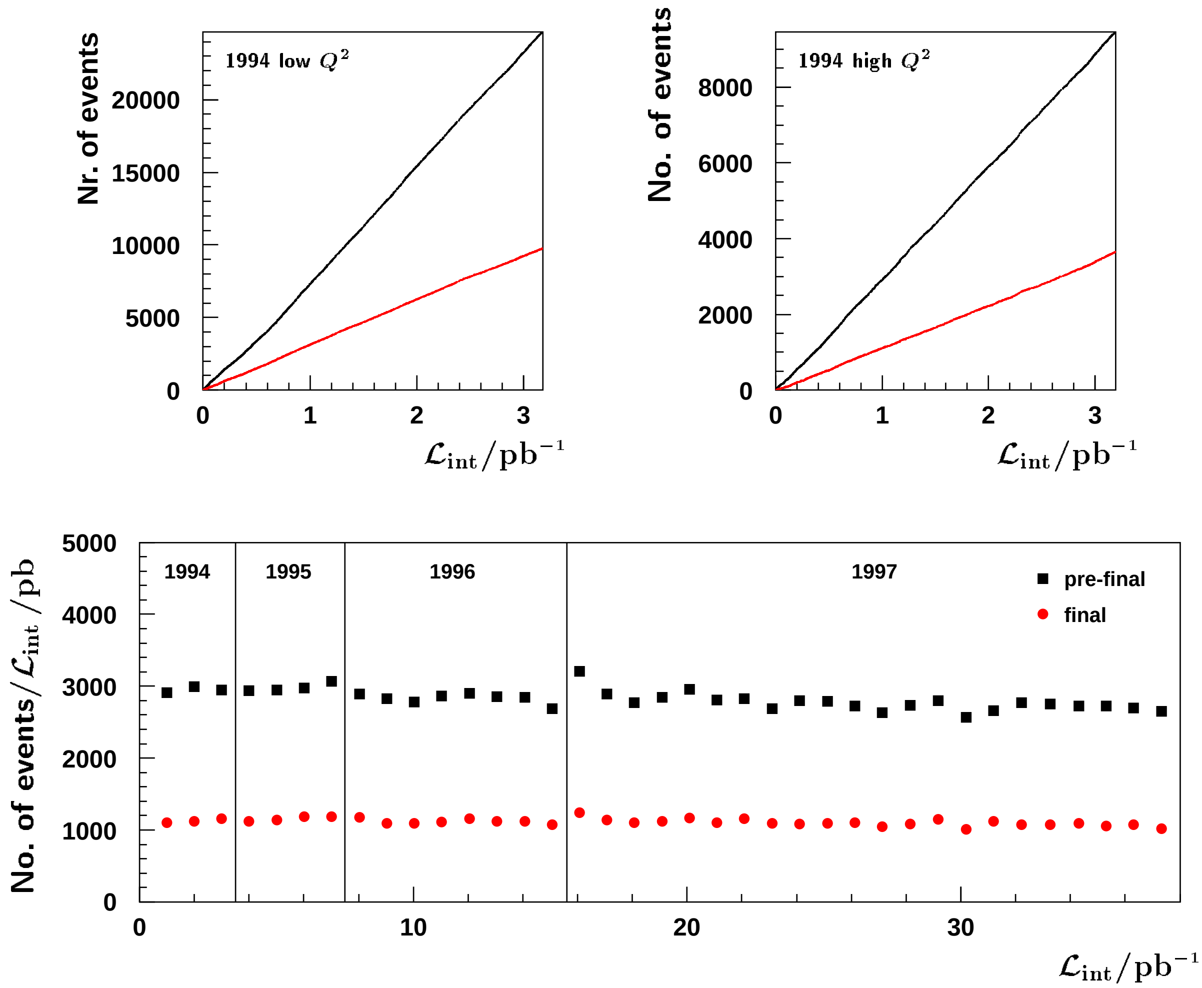


Figure 5.4: In the two plots on top, the accumulated events (upper curve: pre-final, lower curve: final) are drawn against $\mathcal{L}_{int}$ for the low and high Q^2 samples of 1994. A constant rate of incoming events is demonstrated. The lower plot presents the constancy of the number of events gathered per $\mathcal{L}_{int}$ of $\approx 1\,\mathrm{pb}^{-1}$ for all four years contributing to the high Q^2 sample. The statistical uncertainties are smaller than the marker sizes.

Chapter 6

Event Simulation

Measurements with the H1 detector essentially comprise clusters, i.e. energy depositions in the calorimeters, and tracks in the tracking devices that are caused by long-lived[1] particles, mainly hadrons. Yet, the objects dealt with in theoretical considerations of pQCD are partons or, equivalently, quarks, anti-quarks and gluons. Due to the complexity of the measuring apparatus and the underlying physical processes, a direct link from clusters and tracks backwards to hadrons or even partons can not be established. Nevertheless, information can be drawn from the selected data by comparing with model assumptions on a statistical basis. The first task to be carried out in this analysis chain consists in simulating the detector response to a given physics model. This simulation procedure is the subject of the next two sections.

6.1 Simulation

6.1.1 Parton Level

Starting with the calculation of the matrix element of the hard scattering to $\mathcal{O}(\alpha^2\alpha_s)$, complicated integrals arise. For the purpose of generating "events," they are solved by employing a *Monte Carlo integration* technique, s. e.g. [38]. Subsequently, the result has to be folded with pdfs describing the proton structure. For the purpose of easy access, the available sets of pdfs are compiled in the PDFLIB program library [39]. According to the probability distribution derived from the matrix element and the available phase space, a limited number of final state partons is "produced" within the framework of dedicated computer programs called **M**onte **C**arlo (MC) *generators*. For *ep* DIS two such programs, LEPTO [40] and HERWIG [41], are available.

To account for higher orders in a **L**eading **L**ogarithmic **A**pproximation (LLA), both offer an implementation of **P**arton **S**howers (PS) including coherence effects. These branching algorithms may be attached to incoming (**I**nitial **S**tate PS, ISPS) and outgoing partons (**F**inal **S**tate PS, FSPS) as long as their four-momentum squared t (*virtuality*) is above

[1] Here, long-lived means lifetimes $\tau > 10^{-8}$ s.

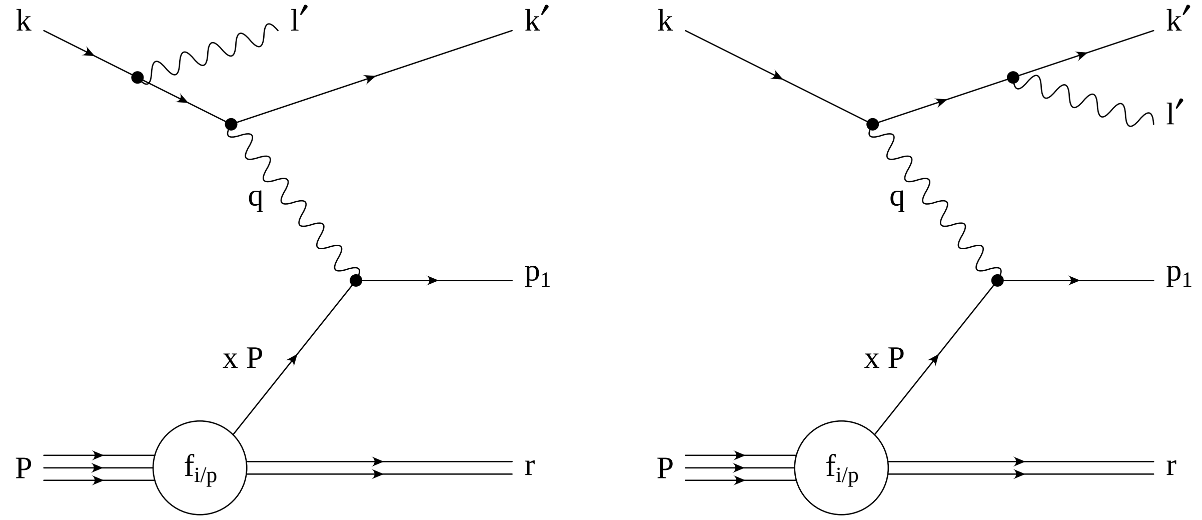


Figure 6.1: $\mathcal{O}(\alpha)$ real corrections to the QPM Feynman graph fig. 3.3: ISR (left) and FSR (right) on the lepton side.

some adjustable threshold, typically around $t_0 \approx 1$–$4\,\mathrm{GeV}^2$. Otherwise, the PS is terminated.

Alternatively, a **C**olour **D**ipole **M**odel (CDM) can be invoked to describe gluon radiation including the first emission in the QCDC reaction. ARIADNE [42] supplies an implementation of the CDM, but is not intended to be a stand-alone program. Instead, it provides an interface to LEPTO so that it may be used within its framework.

Apart from $\mathcal{O}(\alpha_s)$ QCD corrections to the Born cross section (3.21), also QED corrections of $\mathcal{O}(\alpha)$ may be sizable depending on the phase space. This is especially true for high y and low x. Following the singularities proportional to $1/(k \cdot l')$, $1/(k' \cdot l')$ and $1/(k - k' - l')^2$ that appear in the calculation of the real diagrams, they can instructively be labelled **I**nitial **S**tate **R**adiation (ISR), **F**inal **S**tate **R**adiation (FSR) and Compton contribution, where the latter plays only a minor role. Fig. 6.1 depicts the first two of them where the photons are radiated predominantly collinear to the incoming respectively outgoing lepton. Nevertheless, all three parts are defined in the whole phase space such that e.g. "ISR" photons may also be directed along the scattered lepton! In addition, this simple picture is only valid in LLA; beyond, the separation is not unique.

The event generator DJANGO [43] combines the abilities of LEPTO and HERACLES [44], which provides the QED radiative effects including virtual corrections due to 1-loop diagrams, into one software package offering the most complete description of ep DIS events available.

6.1.2 Hadron Level

Neglecting leptons and photons at this stage, we are left with partons of low virtualities $t \approx t_0$, where perturbation theory ceases to be applicable. Lacking better theoretical means, phenomenological models have to be invoked to perform the necessary fragmentation of partons into hadrons. Since the PS cut-off t_0 is an arbitrary parameter, the hadronization model should employ a similar scale from which to start in such a way that the dependence on t_0 largely cancels between the PS and the fragmentation process.

The two most popular models existing for this purpose are the *string* and the *cluster* fragmentation, both of which also consider colour coherence effects. The first one is implemented in JETSET [45], which is applied for this task by the MC generators of the last section except for HERWIG, which makes use of the second approach. Taking into account decays of unstable particles as well, the final outcome of this step consists of all particles traversing through a real detector.

6.1.3 Detector Simulation

So far the involved processes were of a basically theoretical nature and could be described by general purpose MC programs. Since now the detector response to the passing particles has to be reproduced, software specifically adapted to the measuring device is required. The software package employed by H1 is called H1SIM [46].

It is responsible for tracking the MC particles through a virtual H1 detector and simulating the response signals in detail. The "events" produced that way look like real data. For comparison purposes it is desirable to have as many MC events as possible, at least about the same amount as data are available. However, the simulation step is very time consuming and takes several seconds of computing time per event on the computer systems at disposal. Therefore, the MC models to use have to be chosen carefully.

6.2 Reconstruction

6.2.1 Cluster Level

Because the "raw" information (real and simulated) in the form of wire hits, cell voltages, etc. is not very intuitive with respect to the physics of ep scattering, the data have to be refined. This *reconstruction* process is the task of another H1 software program, H1REC [34], which has to be identical for real and simulated events. As a result, it provides i.a. particle tracks and the calorimeter clusters extensively used in this study. For that reason, this stage is called "cluster level."

6.3 Comparison to Data

Once the simulation and reconstruction have been completed, the MC models must be confronted with data. Ideally, they should give a good account of both, "standard" distributions where selection cuts are applied (the energy spectrum of the scattered electron, $dn/dE_{e'}$, for example), as well as event shape distributions dn/dF that are of special interest here.

MC model	DJANGO 6.2	LEPTO 6.1	LEPTO 6.5	HERWIG 5.8
low Q^2 statistics	4446	4253	–	–
high Q^2 statistics	121483	8906	84561	12617

Table 6.1: Cluster level MC statistics.

In the previous publication [23], LEPTO 6.1 was shown to give the best description of the data. This is still true. However, it does not contain radiative corrections. Therefore, DJANGO 6.2 together with ARIADNE has been chosen as a replacement. In the following, the term "DJANGO 6.2" **always** refers to this combination!

For testing purposes a newer LEPTO version, LEPTO 6.5, and HERWIG 5.8 have been used. The basic differences to the old LEPTO version are changed default settings, a new cut-off scheme for the divergences of the matrix elements, an improved target remnant treatment and the introduction of soft colour interactions to facilitate the production of diffractive events within DIS samples. As will be seen, this has a considerable impact on the event shape distributions. Statistics of the employed MC files are given in table 6.1. In case of DJANGO 6.2 high Q^2, it should be noted that an extra very high Q^2 MC file was produced to increase the statistics in the Q-bins seven and eight. When combined with the other high Q^2 data sets, it would lead to unnatural steps in some of the distributions presented in the next section. Therefore, it was excluded there.

6.3.1 Standard Distributions

In figs. 6.3 and 6.4 the normalized differential distributions of $E_{e'}$, $\theta_{e'}$, $\ln(Q_e^2/\,\mathrm{GeV}^2)$, y_h, $(E - P_z)$ and θ_q are presented separately for the low Q^2 and high Q^2 sample. The comparison with DJANGO 6.2 as well as LEPTO 6.1 reveals a good overall agreement with the exception of low θ_q-values in case of low Q^2 events. Yet, the imposed cut-off no. 5 of $20° < \theta_q$ is well below the smallest occurring angle of $\theta_{q,\mathrm{min}} \approx 30°$, rendering the deviation harmless.

The dips in the high Q^2 $\theta_{e'}$-distribution stem from the rejection of events with the scattered lepton found in areas of the LAr where the energy measurement deteriorates (cut no. 15).

LEPTO 6.5 and HERWIG 5.8 are not shown here. Their agreement with data is similar.

6.3.2 Event Shape Distributions

The next step is to check the description of the event shape distributions. They are shown for four out of eight investigated bins in Q in figs. 6.5 and 6.6. Here, the similarity is not as satisfactory as before. DJANGO 6.2 systematically tends to overshoot the data in the low F-region which is compensated for by underestimating them for high values in F.

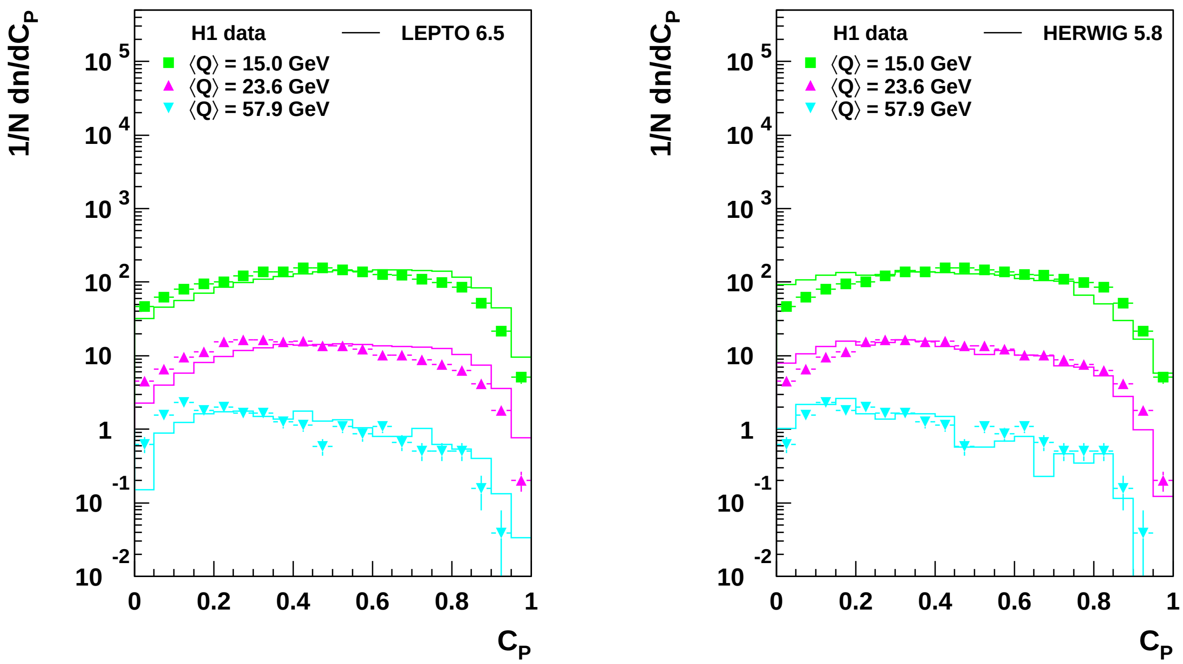


Figure 6.2: Normalized differential distributions of C_P on cluster level. H1 data (full symbols) are compared with LEPTO 6.5 (full lines) on the left-hand side and HERWIG 5.8 (full lines) on the right-hand side for four out of eight investigated bins in Q. The spectra for $\langle Q \rangle = 7.5\text{–}57.9\,\mathrm{GeV}$ are multiplied by factors of 10^n, $n = 0, 1, 2, 3$. The error bars represent statistical uncertainties only.

This is worst for τ_P and especially B_P, whose mean values come out too low in the MC. The uncorrected data means $\langle F \rangle$ are given in table 6.2.

Within the much lower statistics available for LEPTO 6.1, it seems to do a better job in reproducing the data distributions. Nevertheless, DJANGO 6.2 in combination with ARIADNE provides an acceptable account of the available data and, since it includes radiative corrections, is employed for the primary correction procedure discussed in the next chapter.

As an example of a rather bad performance, fig. 6.2 presents the normalized differential distributions of C_P for three bins of the high Q^2 sample in comparison with LEPTO 6.5 on the left-hand side and HERWIG 5.8 on the right-hand side. Severe deviations contrary to each other are observed.

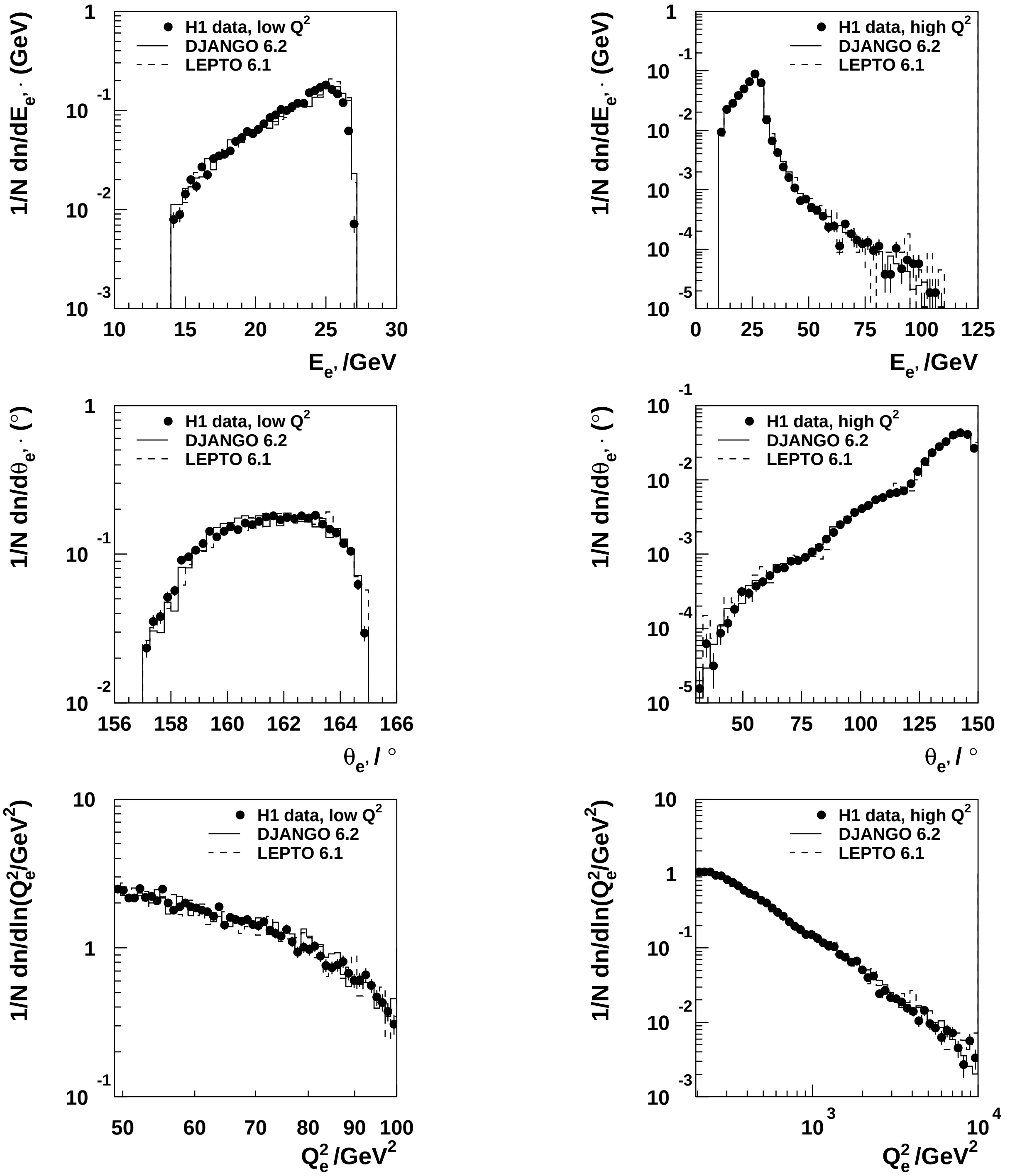


Figure 6.3: Normalized differential distributions of $E_{e'}$, $\theta_{e'}$ and $\ln(Q_e^2/\,\mathrm{GeV}^2)$ on cluster level for the low Q^2 (left) and high Q^2 (right) data sample (full symbols) in comparison with DJANGO 6.2 (full line) and LEPTO 6.1 (dashed line). The error bars represent statistical uncertainties only.

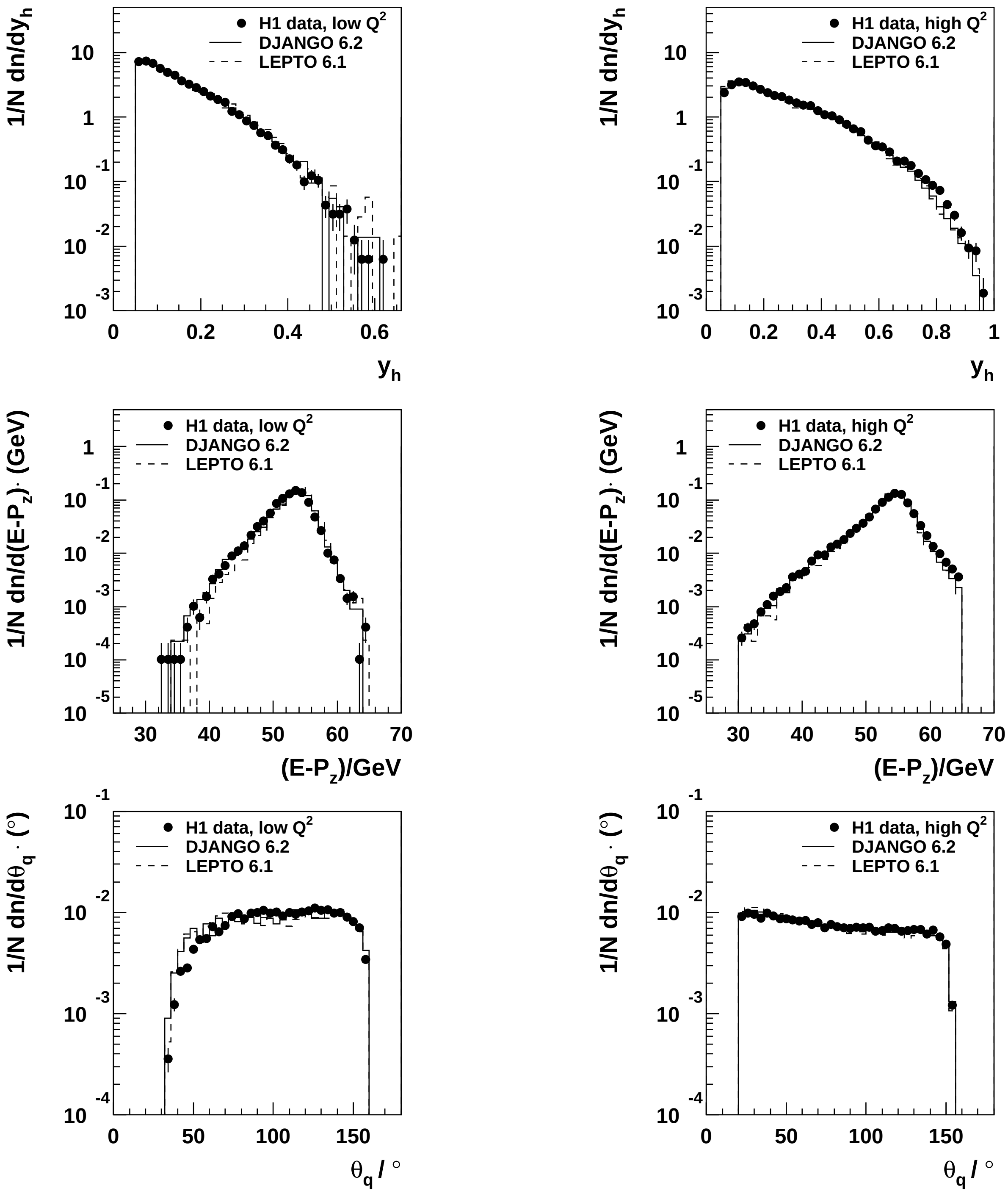


Figure 6.4: Normalized differential distributions of y_h, $(E - P_z)$ and θ_q on cluster level for the low Q^2 (left) and high Q^2 (right) data sample (full symbols) in comparison with DJANGO 6.2 (full line) and LEPTO 6.1 (dashed line). The error bars represent statistical uncertainties only.

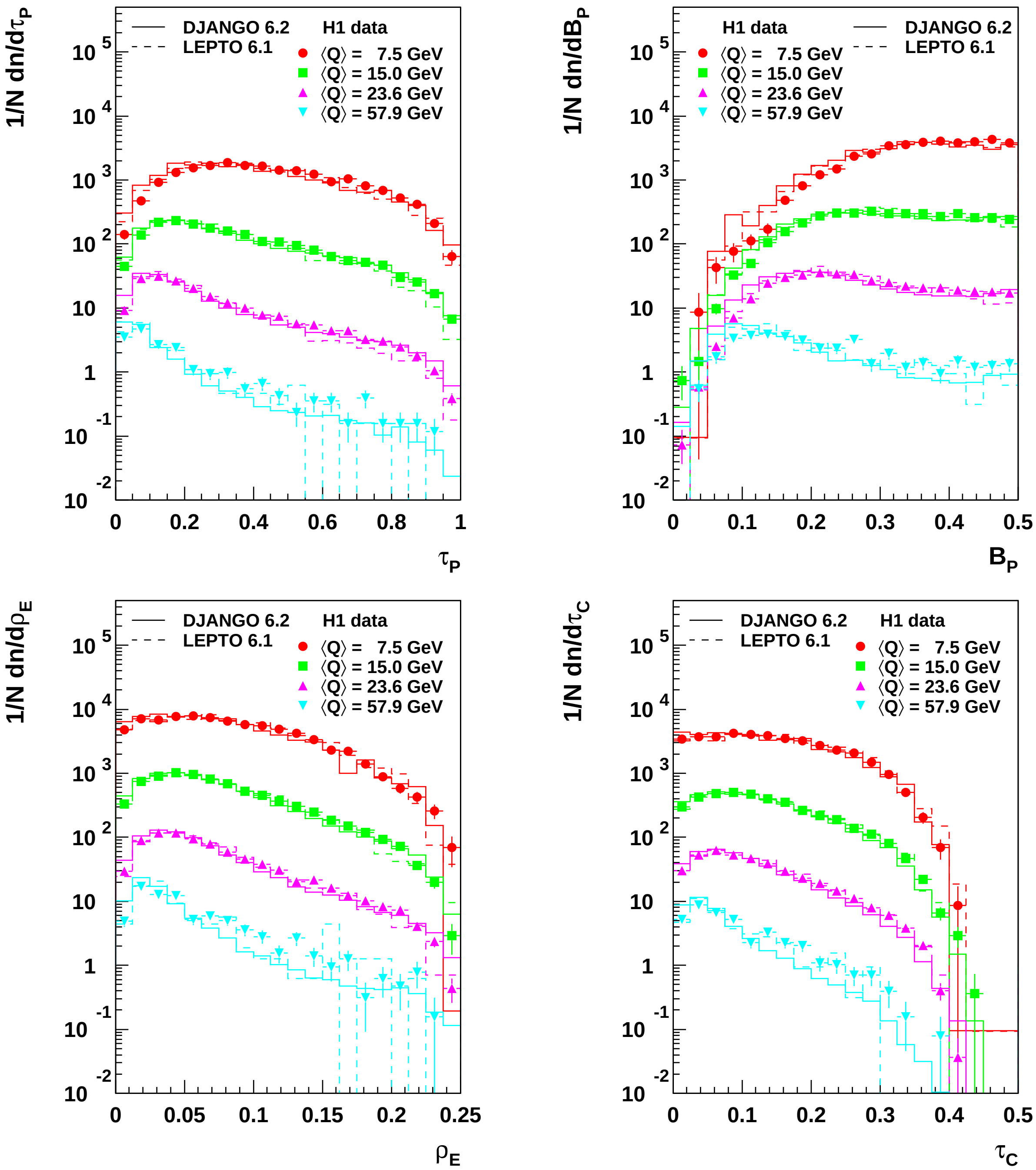


Figure 6.5: Normalized differential distributions of the event shapes τ_P, B_P, ρ_E and τ_C on cluster level. H1 data (full symbols) are compared with DJANGO 6.2 (full lines) and LEPTO 6.1 (dashed lines) for four out of eight investigated bins in Q. The spectra for $\langle Q \rangle = 7.5$–$57.9\,\mathrm{GeV}$ are multiplied by factors of 10^n, $n = 0, 1, 2, 3$. The error bars represent statistical uncertainties only.

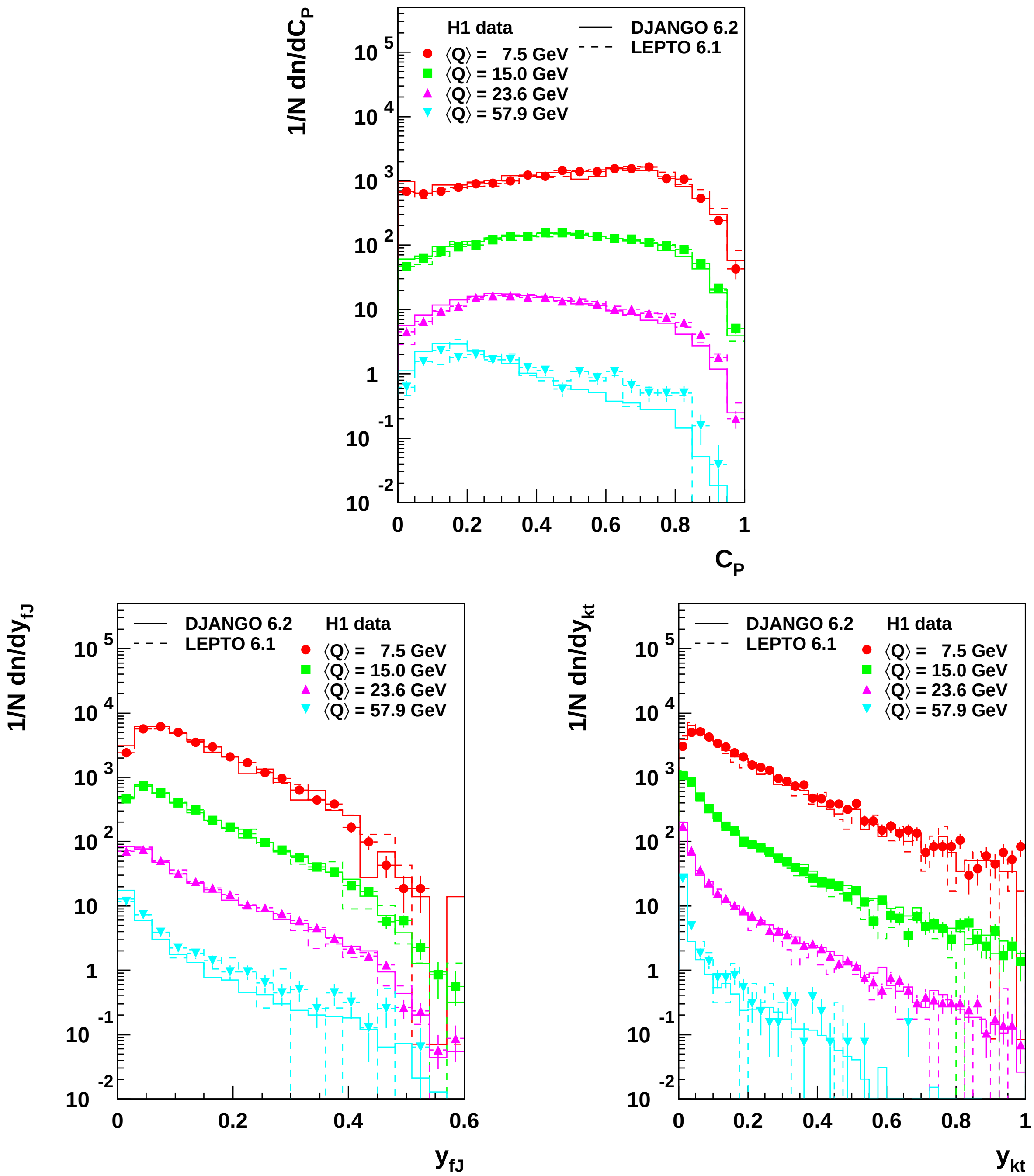


Figure 6.6: Normalized differential distributions of the event shapes C_P, y_{fJ} and y_{k_t} on cluster level. H1 data (full symbols) are compared with DJANGO 6.2 (full lines) and LEPTO 6.1 (dashed lines) for four out of eight investigated bins in Q. The spectra for $\langle Q \rangle = 7.5$–$57.9\,\text{GeV}$ are multiplied by factors of 10^n, $n = 0, 1, 2, 3$. The error bars represent statistical uncertainties only.

$\langle Q \rangle$ / GeV	$\langle \tau_P \rangle$	$\langle B_P \rangle$
7.46	0.4351 ± 0.0031	0.3651 ± 0.0013
8.74	0.4110 ± 0.0034	0.3543 ± 0.0014
14.97	0.3456 ± 0.0021	0.3186 ± 0.0010
17.75	0.3203 ± 0.0018	0.3039 ± 0.0009
23.62	0.2782 ± 0.0020	0.2790 ± 0.0010
36.72	0.2367 ± 0.0034	0.2502 ± 0.0019
57.93	0.2079 ± 0.0090	0.2271 ± 0.0053
81.32	0.1598 ± 0.0134	0.1891 ± 0.0082

$\langle Q \rangle$ / GeV	$\langle \rho_E \rangle$	$\langle \tau_C \rangle$	$\langle C_P \rangle$
7.46	0.0797 ± 0.0007	0.1401 ± 0.0013	0.4960 ± 0.0034
8.74	0.0805 ± 0.0008	0.1429 ± 0.0014	0.5048 ± 0.0038
14.97	0.0737 ± 0.0005	0.1297 ± 0.0008	0.4701 ± 0.0022
17.75	0.0725 ± 0.0004	0.1267 ± 0.0007	0.4603 ± 0.0019
23.62	0.0685 ± 0.0004	0.1179 ± 0.0008	0.4324 ± 0.0021
36.72	0.0626 ± 0.0008	0.1041 ± 0.0013	0.3847 ± 0.0036
57.93	0.0578 ± 0.0021	0.0950 ± 0.0033	0.3495 ± 0.0098
81.32	0.0482 ± 0.0030	0.0804 ± 0.0050	0.3041 ± 0.0147

$\langle Q \rangle$ / GeV	$\langle y_{fJ} \rangle$	$\langle y_{k_t} \rangle$
7.46	0.1273 ± 0.0012	0.1746 ± 0.0023
8.74	0.1274 ± 0.0013	0.1471 ± 0.0023
14.97	0.1131 ± 0.0009	0.1077 ± 0.0013
17.75	0.1115 ± 0.0008	0.0975 ± 0.0011
23.62	0.1047 ± 0.0009	0.0826 ± 0.0012
36.72	0.0911 ± 0.0015	0.0603 ± 0.0018
57.93	0.0874 ± 0.0044	0.0483 ± 0.0040
81.32	0.0681 ± 0.0060	0.0341 ± 0.0050

Table 6.2: Uncorrected mean values of the event shapes and their statistical uncertainty as a function of Q.

Chapter 7

Correction Procedure

After having established a sufficient description of data by MC models, the most important task remaining is disentangling the underlying physics from mere detector effects due to limited efficiencies and resolutions. Additionally, radiative corrections are taken into account.

In principle, it is feasible to do this in a one-step procedure. Compensating effects, however, may mislead to the conclusion that the corrections are small, even if indeed they are not. Moreover, one would like to identify the dominant influence. Another disadvantage is the complete neglect of migrations enforced by the lack of correlations.

For those reasons, the procedure applied here tries to differentiate between these contributions, and the unfolding of the data is correspondingly performed in two stages. The different correction methods available are explained in the next section.

7.1 Unfolding Methods

7.1.1 Factor Method

When only mean values do matter, the easiest thing to think of is to invoke one correction factor for each mean:

$$\langle F\rangle_{\mathrm{corr}} = c_{\mathrm{had/sim}} \cdot \langle F\rangle_{\mathrm{dat}} \,, \qquad c_{\mathrm{had/sim}} := \frac{\langle F\rangle^{\mathrm{MC}}_{\mathrm{had}}}{\langle F\rangle^{\mathrm{MC}}_{\mathrm{sim}}} \,. \tag{7.1}$$

7.1.2 Bin-to-bin Correction

By applying eq. (7.1) not to the mean values alone, but to each bin i of a differential distribution F_i,

$$F_{i,\mathrm{corr}} = c_{i,\mathrm{had/sim}} \cdot F_{i,\mathrm{dat}} \,, \qquad c_{i,\mathrm{had/sim}} := \frac{\frac{1}{N^{\mathrm{MC}}_{\mathrm{had}}} F^{\mathrm{MC}}_{i,\mathrm{had}}}{\frac{1}{N^{\mathrm{MC}}_{\mathrm{sim}}} F^{\mathrm{MC}}_{i,\mathrm{sim}}} \,, \tag{7.2}$$

the factor method can easily be extended to unfold complete distributions which again can be reevaluated to give a corrected mean $\langle F\rangle_{\mathrm{corr}}$.

As can be seen from eq. (7.2), however, this simple approach completely ignores the possibility of migrations from one bin i on hadron level to another bin j on cluster level. Depending on the quantity to deal with, this constitutes a severe disadvantage. A measure to improve the situation is to choose bin sizes in such a way that migrations are minimized.

7.1.3 Matrix Method

It is still better to take these migrations explicitly into consideration by employing a correction matrix C_{ij}:

$$F_{i,\mathrm{corr}} = \sum_j C_{ij} F_{j,\mathrm{dat}} \,. \tag{7.3}$$

The matrix C_{ij} can in principle be obtained from MC by inverting the transfer matrix T_{ij} which transforms the "true" values $F^{\mathrm{MC}}_{i,\mathrm{had}}$ into the observed ones

$$F^{\mathrm{MC}}_{i,\mathrm{sim}} = \sum_j T_{ij} F^{\mathrm{MC}}_{j,\mathrm{had}} \,. \tag{7.4}$$

In practice, however, the inversion leads to instabilities and oscillations unless extremely large MC statistics is available. Instead, we follow the strategy employed in [23] to define the correction matrix to be

$$C_{ij} := \frac{\rho_{ij}}{\sum\limits_k \rho_{kj}} \tag{7.5}$$

where ρ_{ij} represents a probability density derived from MC correlations as

$$F^{\mathrm{MC}}_{i,\mathrm{had}} = \sum_j \rho_{ij} F^{\mathrm{MC}}_{j,\mathrm{sim}} \,. \tag{7.6}$$

As a drawback of this scheme, there may remain some model dependence which can be overcome by an iterative procedure.

7.1.4 Bayesian Unfolding

The last method employs a program developed in [47]. It exploits Bayes' theorem on conditional probabilities to extract information from the observed distributions. To state it in a more general form, we label the fact of an event to have a "true" value F_{corr} allocated to bin i as "cause" C_i and the observed value F_{dat} in bin j as "effect" E_j. Since in general the binning and also the **domain of definition** may be different, we have a number of N_C causes and N_E effects. Denoting further the initial probability for cause C_i to happen with $P_0(C_i)$ and the conditional probability for the effect E_j to occur if C_i has already taken place with $P(E_j|C_i)$, Bayes' theorem can be written as

$$P(C_i|E_j) = \frac{P(E_j|C_i) \cdot P_0(C_i)}{\sum\limits_{k=1}^{N_C} P(E_j|C_k) \cdot P_0(C_k)} \,. \tag{7.7}$$

That is, given an observation of E_j, we get a probability that it was produced by cause C_i.

The result, however, depends on two ingredients: the conditional probabilities $P(E_j|C_i)$, which can be estimated from MC, and the initial probabilities $P_0(C_i)$, that are *a priori* unknown. At first, the latter have to be assumed by the observer according to his prejudices which may even result in a uniform distribution in the case of complete ignorance. But, given a number of experimental observations with frequencies $N(E) := \{N(E_1), \ldots, N(E_{N_E})\}$, Bayes' theorem can then be applied to gain information on $P_0(C_i)$. Based on these observed event numbers for the effects E_j, the expected number of events to be assigned to the different causes C_i can be inferred to be

$$\hat{N}(C_i) = \sum_{j=1}^{N_E} P(C_i|E_j) \cdot N(E_j)\,. \tag{7.8}$$

Ignoring the possibility of inefficiencies, i.e. the case when a cause does not produce one of the effects under consideration, the initial probability is estimated by

$$\hat{P}(C_i) := P(C_i|N(E)) = \frac{\hat{N}(C_i)}{\sum\limits_{i=1}^{N_C} \hat{N}(C_i)}\,. \tag{7.9}$$

If the initial distribution $P_0(C_i)$ is not consistent with the data, it will in general differ from the final one $\hat{P}(C_i)$. In an iterative procedure where the assumption on $P_0(C_i)$ is updated in each step, the observed events $N(E)$ can now be exploited to unfold the "true" distribution $\hat{N}(C)$.

7.2 Correction for Detector Effects

The first step consists in deriving a correction to hadron level from the simulated DJANGO 6.2 MC files. Here, special care has to be taken when specifying what is meant by "hadron level" since there may be radiative photons. But we are interested in detector effects only. Thus, the main influence of these photons has to be accounted for in a later step. This is accomplished by first enforcing a limitation in acceptance according to cut no. 9, thereby throwing away radiation collinear to the incoming lepton. Furthermore, with regard to the finite angular resolution of the LAr calorimeter, the scattered electron is redefined to be merged with radiative photons if they are closer in angle than 5°. This value is motivated by a study in [48]. As a consequence, the cuts nos. 10–12 may reject events with respect to this "radiative" hadron level.

Now one is free to draw a connection between the cluster and hadron level of the simulation and apply this to the data according to the four unfolding schemes described. Note that the selection criteria nos. 13–15 are imposed on the cluster level only such that they are accounted for in this first step.

Figs. 7.1 and 7.2 show the bin-wise detector corrections as derived from DJANGO 6.2 for four out of eight investigated bins in Q. Entries of zero (or correspondingly two, four or

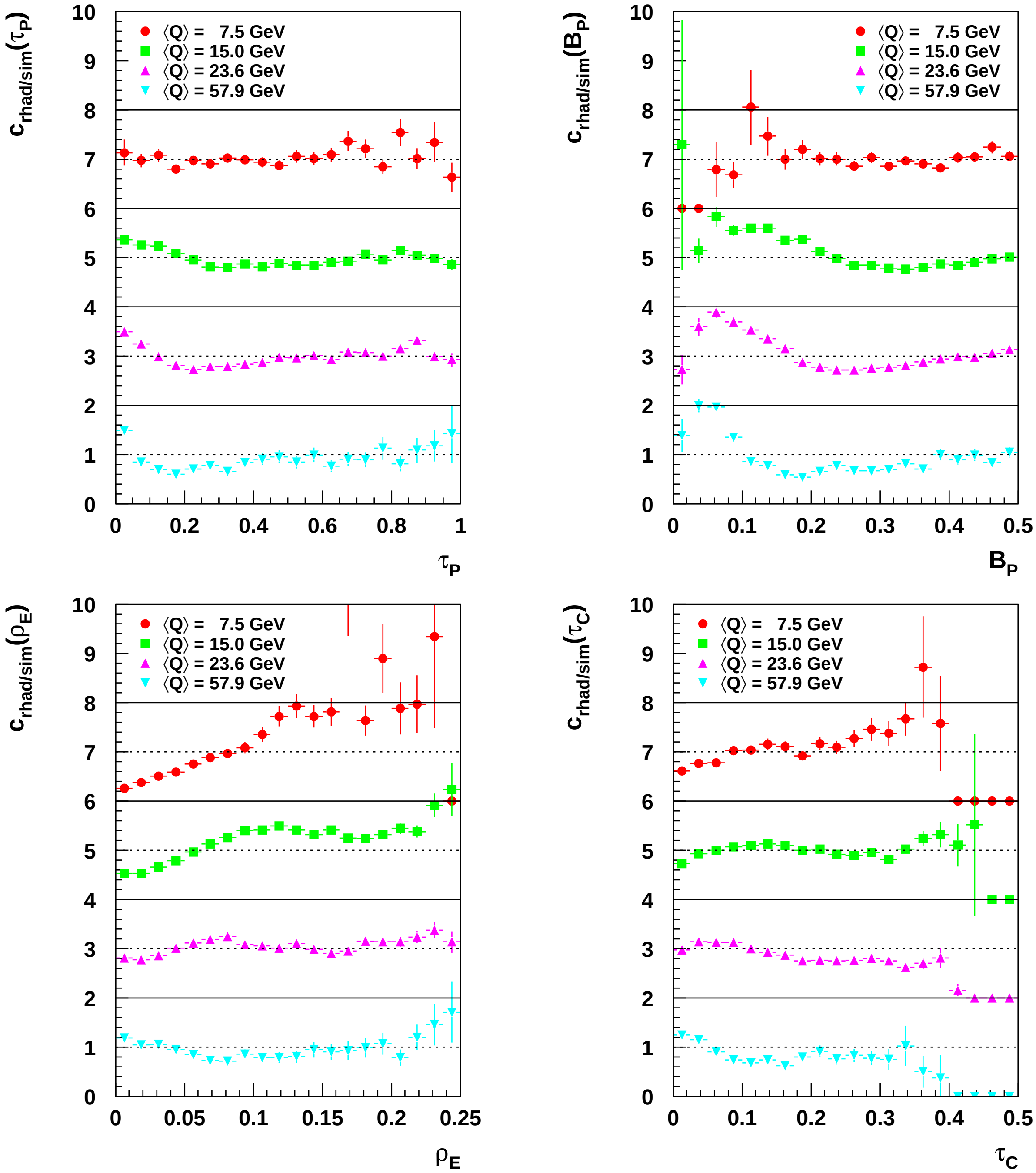


Figure 7.1: Bin-wise detector corrections for the event shapes τ_P, B_P, ρ_E and τ_C as derived from DJANGO 6.2 for four out of eight investigated bins in Q. The factors for $\langle Q \rangle = 7.5$–$57.9\,\mathrm{GeV}$ are shifted by offsets of $2 \cdot n$, $n = 0, 1, 2, 3$. The dashed lines delineate the corresponding position of unity. Entries exactly at zero (or 2, 4, 6) mean that no events were found on either level. The error bars represent the statistical uncertainty.

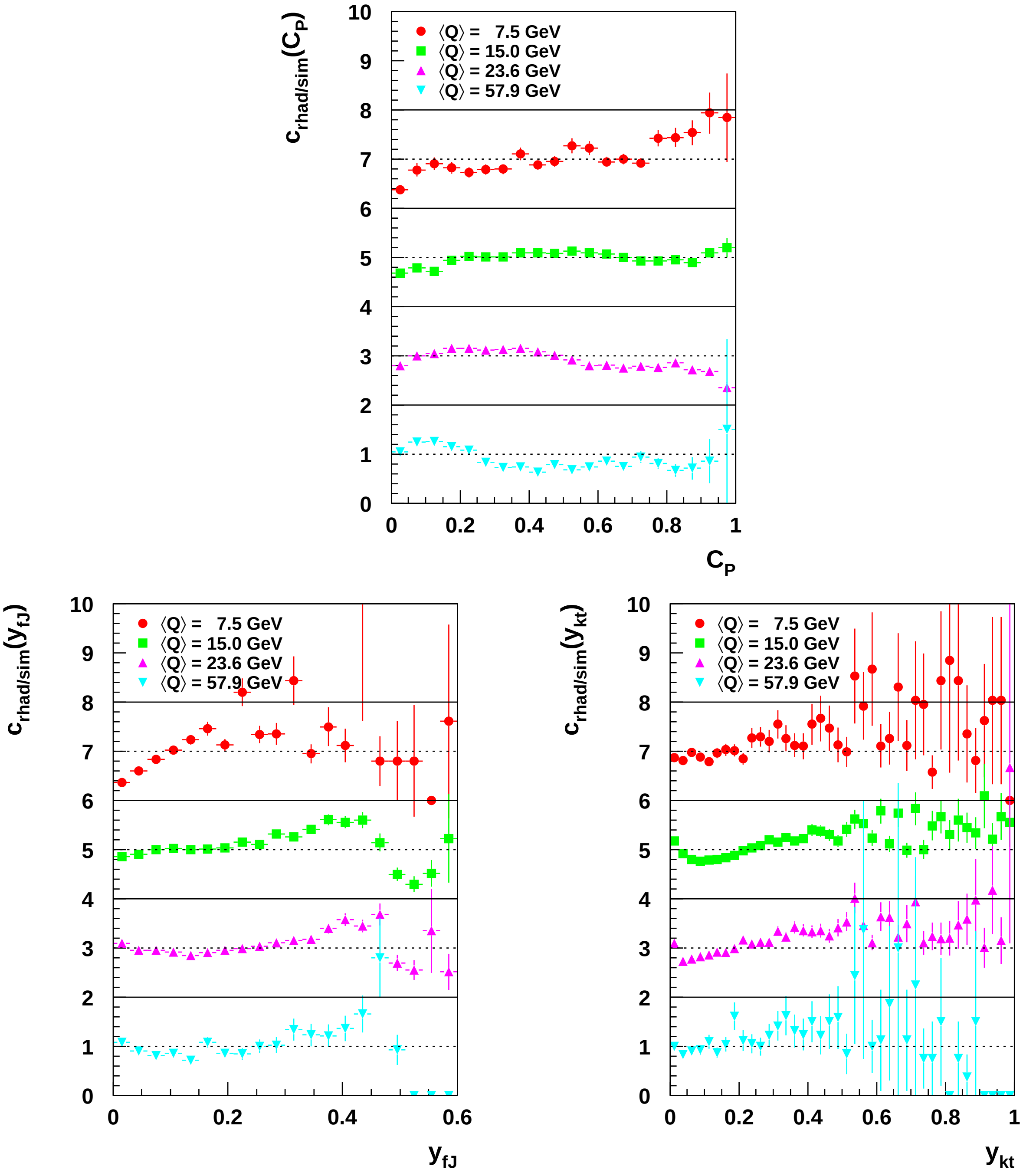


Figure 7.2: Bin-wise detector corrections for the event shapes C_P, y_{fJ} and y_{k_t} as derived from DJANGO 6.2 for four out of eight investigated bins in Q. The factors for $\langle Q \rangle = 7.5$–$57.9\,\mathrm{GeV}$ are shifted by offsets of $2 \cdot n$, $n = 0, 1, 2, 3$. The dashed lines delineate the corresponding position of unity. Entries exactly at zero (or 2, 4, 6) mean that no events were found on either level. The error bars represent the statistical uncertainty.

six) occurring at the edges mean that no events were found at either level. Concentrating on the statistically relevant points, it must be concluded that the corrections sometimes are larger than agreeable. Most often they are below 30%–50%, but may rise higher in some bins, especially for B_P and ρ_E.

Both, the matrix method as well as the Bayesian unfolding need the correlation matrices between hadron and cluster level as input. They are presented for each event shape in figs. 7.3–7.5 for two bins, one at low and one at high Q^2. Note that the box sizes are scaled **logarithmically** with the number of entries. From these we can learn mainly two things: First, correlations, which improve with rising Q, are clearly demonstrated, and second, with respect to the resolution, the bin sizes are too small for the bin-to-bin correction. Nevertheless, the binning is kept as a compromise, because the matrix and the Bayes unfolding account for migrations as well as resolution effects and the corrected means should not be affected by too coarse bin sizes.

With these ingredients given, two supplementary remarks are due. Concerning the matrix approach, the effect of iterations was revealed to be small, except for the fact that oscillations are gathered up. Hence, no iteration is performed here.

With respect to the Bayesian unfolding, it is recommended in [47] to avoid building up large fluctuations with an increasing number of iterations by smoothing a new "initial" distribution $\hat{P}(C_i)$ in the intermediate steps. This proved to be difficult for the event shapes and finally was not done. To start from uniformly distributed initial probabilities did work, but required some iterations before converging, and consequently fluctuations showed up. It could be circumvented by choosing the MC distribution for $P_0(C_i)$. Then, three iterations are fully sufficient.

7.3 Radiative Corrections

At this stage we have to perform the remaining correction from the previously defined radiative to a non-radiative hadron level. For that purpose a second kind of DJANGO 6.2 MC files is necessary. They have to be generated with exactly the same settings as the ones before, except for the radiative effects which have to be switched off. A simulation is not required. Because there are no more ISR or FSR photons, the acceptance cut no. 9 may be dropped leaving the cuts nos. 10–12 ineffective. In that way, the acceptance holes of the detector near the beam pipe are additionally compensated for. But due to our limited knowledge of that region, the extrapolation from MC possibly is not very precise. Notice that the cuts nos. 7 and 8 have been applied to **all** levels up to now and consequently they have not been accounted for by the unfolding procedure. Depending on the theory to compare with, e.g. the parton level of some MC, they may have to be considered as **phase space cuts**. Yet, they are not suitable in the framework of a pQCD calculation to NLO and hence are regarded as data quality cuts in this study.

For the factor and bin-to-bin method, the combined correction steps are fully equivalent to a one-step procedure. Yet, as a matter of fact, we now have to deal with two different MC files. That means we do not have any correlations at our disposal such that the unfolding schemes number three and four are not applicable. The bin-to-bin correction has to serve as the only available replacement.

Figs. 7.6 and 7.7 display the radiative correction factors for four out of eight investigated bins in Q. Except for τ_P and B_P where radiative effects are not negligible, they are very close to one. y_{k_t} is special. The additionally shown histogram with the full line usually near "one" corresponds to the correction factors where the acceptance cut no. 9 was kept! This demonstrates that it is not a radiative effect. y_{k_t} as applied here seems to be affected by the MC extrapolation, which is unfortunate and not understood at the moment. For all other event shapes including y_{fJ}, almost nothing changes whether cut no. 9 is active or not. In the case of τ_P up to C_P this could be expected since they only deal with the CH opposite to the remnant in the Breit frame.

7.4 Performance Check

A mandatory check of correction methods is to employ one MC to unfold another. As a first test, this was done for models without QED radiation. LEPTO 6.1 served as the MC, whereas LEPTO 6.5 and HERWIG 5.8 were treated like real data. Consequently, the unfolded means should predominantly lie close to the hadron level means of the corresponding MC (almost) regardless of what LEPTO 6.1 would prefer. To be able to discriminate between them, the models have to make sufficiently different predictions for the quantity under study. From fig. 6.2 it is clear that for C_P this is certainly true. The outcome is presented in the upper plots of fig. 7.8 proving all four correction schemes to work properly. Nevertheless, all other event shapes were tested, too, yielding similar results, although not as evident.

In a second test, LEPTO 6.1 events were smuggled in as data for DJANGO 6.2. Here, the additional complication arises that the latter does include radiative effects in contrast to the "data." As expected, the first correction step alone suffices since the radiative hadron level is designed to separate detector and radiative effects as much as possible. If the radiative factors are applied, too, then the sensible variables, i.e. τ_P and B_P, are even overdone with. This is demonstrated in the lower plots of fig. 7.8, where the "fully" corrected means of the right plot deviate considerably from the hadron level means of LEPTO 6.1 as opposed to the plot on the left-hand side. Using only DJANGO 6.2 events without radiative photons, the discrepancy disappeared.

Concluding, it can be stated that the unfolding works satisfactorily.

7.5 Final Means

To illustrate possible systematic effects due to the different unfolding schemes, all resulting corrected means are shown in figs. 7.9 and 7.10. The statistical uncertainties encompass data as well as MC statistics. For the factor method, they are obtained properly by error

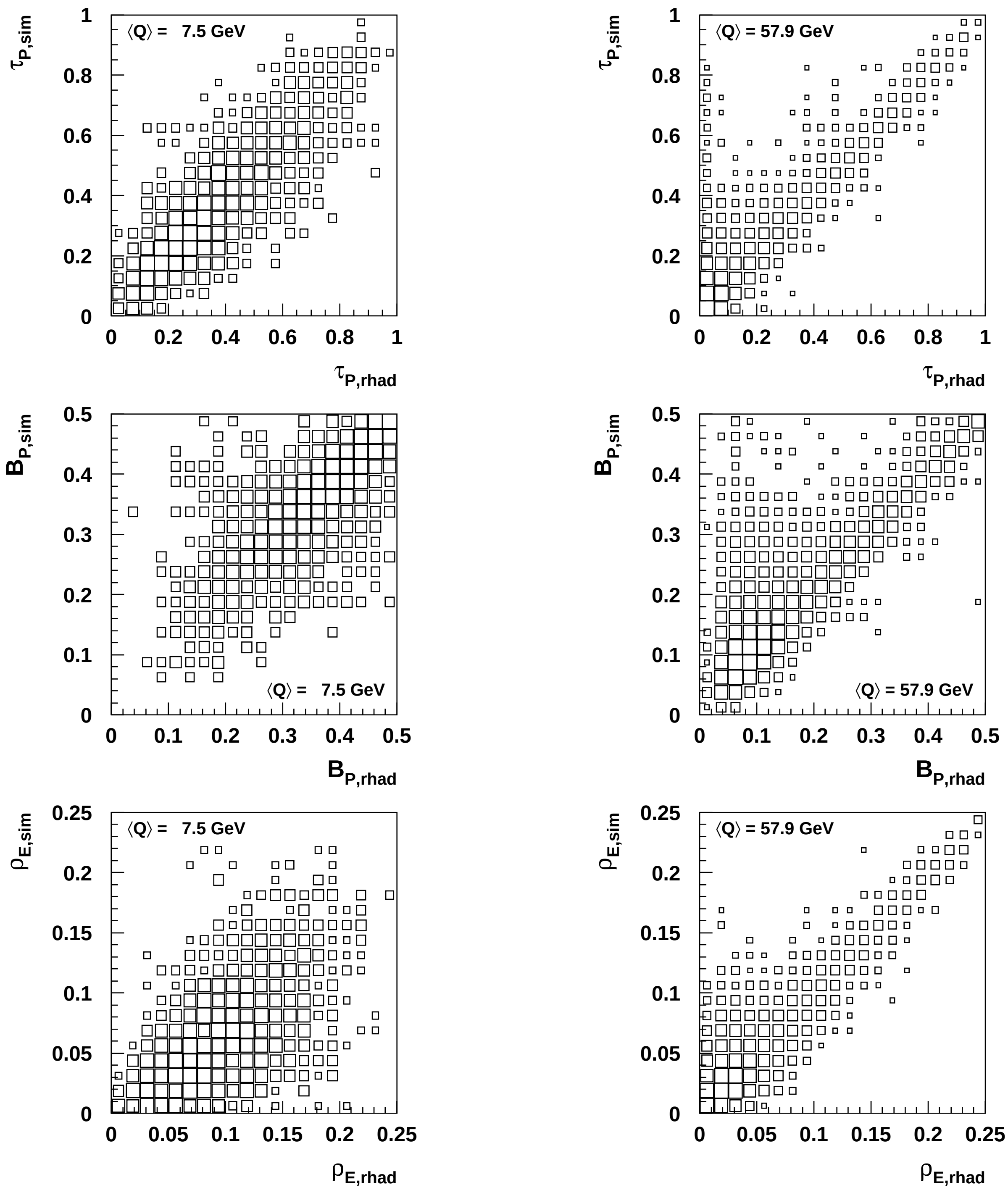


Figure 7.3: Correlation matrices for the event shapes τ_P, B_P and ρ_E as derived from DJANGO 6.2 for two out of eight investigated bins in Q. Note that the box sizes are scaled **logarithmically** with the number of entries.

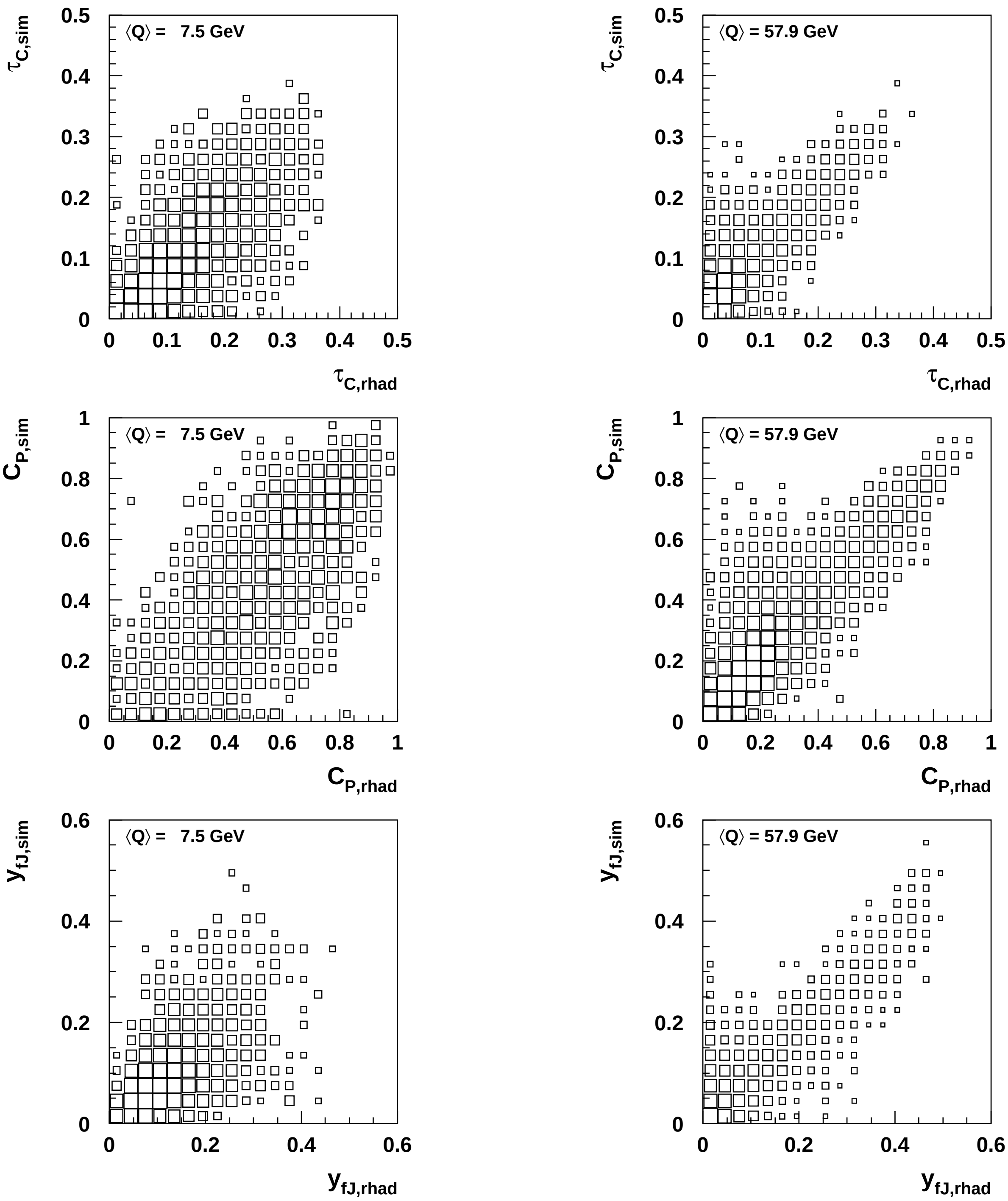


Figure 7.4: Correlation matrices for the event shapes τ_C, C_P and y_{fJ} as derived from DJANGO 6.2 for two out of eight investigated bins in Q. Note that the box sizes are scaled **logarithmically** with the number of entries.

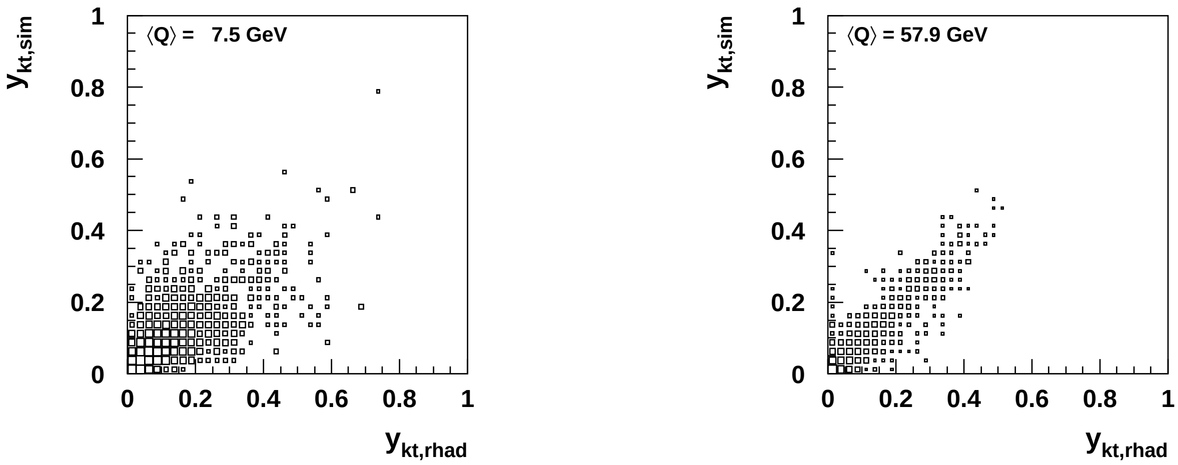


Figure 7.5: Correlation matrices for the event shape y_{k_t} as derived from DJANGO 6.2 for two out of eight investigated bins in Q. Note that the box sizes are scaled **logarithmically** with the number of entries.

propagation. In case of an evaluation from unfolded distributions, they are estimated by

$$\Delta \langle F \rangle = \sigma(\langle F \rangle) \cdot \left(\frac{1}{\sqrt{N_{\mathrm{dat}}}} + \frac{1}{\sqrt{N_{\mathrm{MC}}}} \right) , \tag{7.10}$$

where N_{dat}, N_{MC} are the appropriate event numbers and σ denotes the standard deviation. Note that $1/\sqrt{N_{\mathrm{dat}}}$ and $1/\sqrt{N_{\mathrm{MC}}}$ are not added conventionally in quadrature. This is motivated by the fact that the uncertainties derived this way are closer in value to those calculated for the factor method.

For the final results the differential distributions as well as the central values are taken from the most sophisticated approach, i.e. the Bayesian unfolding with subsequent bin-to-bin radiative correction. In case of the event shape spectra, no further error evaluation beyond statistical ones was attempted. For a comparison with pQCD calculations, their presentation is postponed to ch. 10, figs. 10.1 and 10.2.

Concerning the mean values, the other methods serve to judge systematic effects. The due asymmetric uncertainty is estimated to be half of the spread caused by the maximal deviation to larger respectively smaller values. The sum of both then corresponds to half the total spread which is somewhat smaller than $2/\sqrt{12} = 1/\sqrt{3}$ as appropriate for a uniform distribution.

A second source of systematic uncertainties are the electromagnetic and hadronic energy scales of the calorimeters. Due to recent improvements [13], the electron energies now can be measured to a precision of 1%, 2% and 3% up to z_{imp}-coordinates of 24 cm, 110 cm and the forward end of the LAr calorimeter respectively. Hadronic energies are known only to about 4%. In order to gain information on the influence of this uncertainty, the whole analysis is repeated scaling up and down separately for both scales. The discrepancies with respect to the normal central values are then attributed to two further asymmetric systematic uncertainties. All three, i.e. including the unfolding, added in quadrature, yield the total systematics given together with the final results in table 7.1.

Concerning individual contributions, the impact of the electron energy, which directly affects the boost into the Breit frame, is almost always largest for τ_P up to C_P, followed by unfolding effects. Realizing that hadronic systematics cancel between numerator and denominator for these event shapes, this is understandable. The situation is reversed for y_{fJ} and y_{k_t}. Here, no cancellation occurs and the systematic uncertainty due to hadronic energies dominates.

At last, figs. 7.11 and 7.12 present the final means including all uncertainties as well as the hadron level means of DJANGO 6.2 and LEPTO 6.1. Where available, already published results from [23] and [49] are shown for comparison.[1] τ_P and B_P are in good agreement, but they should not because radiative corrections leading to smaller values were previously neglected. Seemingly, this was compensated by an overestimation of the electron energy that happened before the new calibration was introduced into the analysis. Since ρ_E and τ_C are much less affected by QED radiation, the new data are shifted somewhat towards higher means.

[1] C_P is also given in [49], but the values are believed to be erroneous and therefore are not reproduced here.

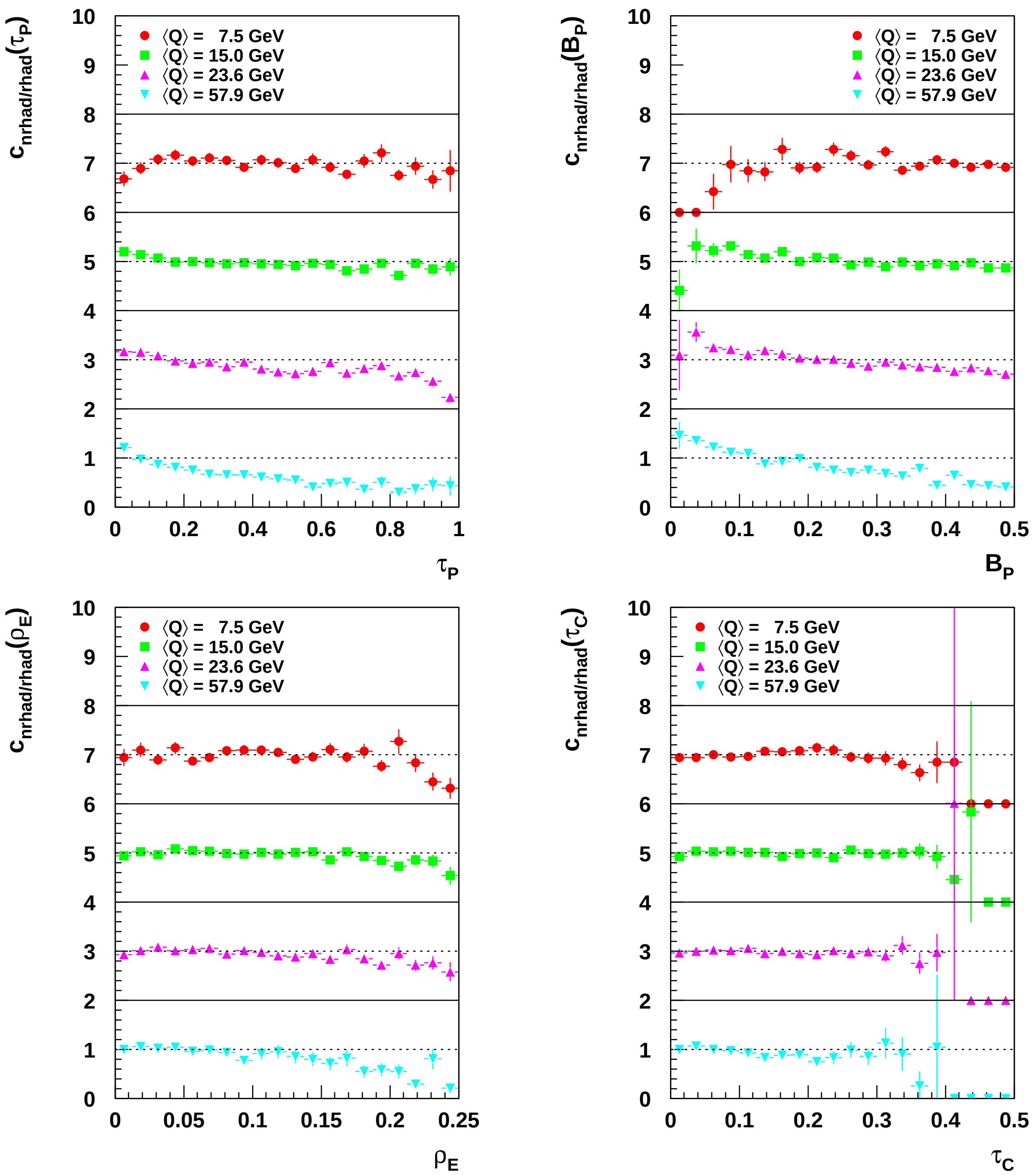


Figure 7.6: Bin-wise radiative corrections for the event shapes τ_P, B_P, ρ_E and τ_C as derived from DJANGO 6.2 for four out of eight investigated bins in Q. The factors for $\langle Q\rangle = 7.5$–$57.9\,\mathrm{GeV}$ are shifted by offsets of $2\cdot n$, $n = 0, 1, 2, 3$. The dashed lines delineate the corresponding position of unity. Entries exactly at zero (or 2, 4, 6) mean that no events were found on either level. The error bars represent the statistical uncertainty.

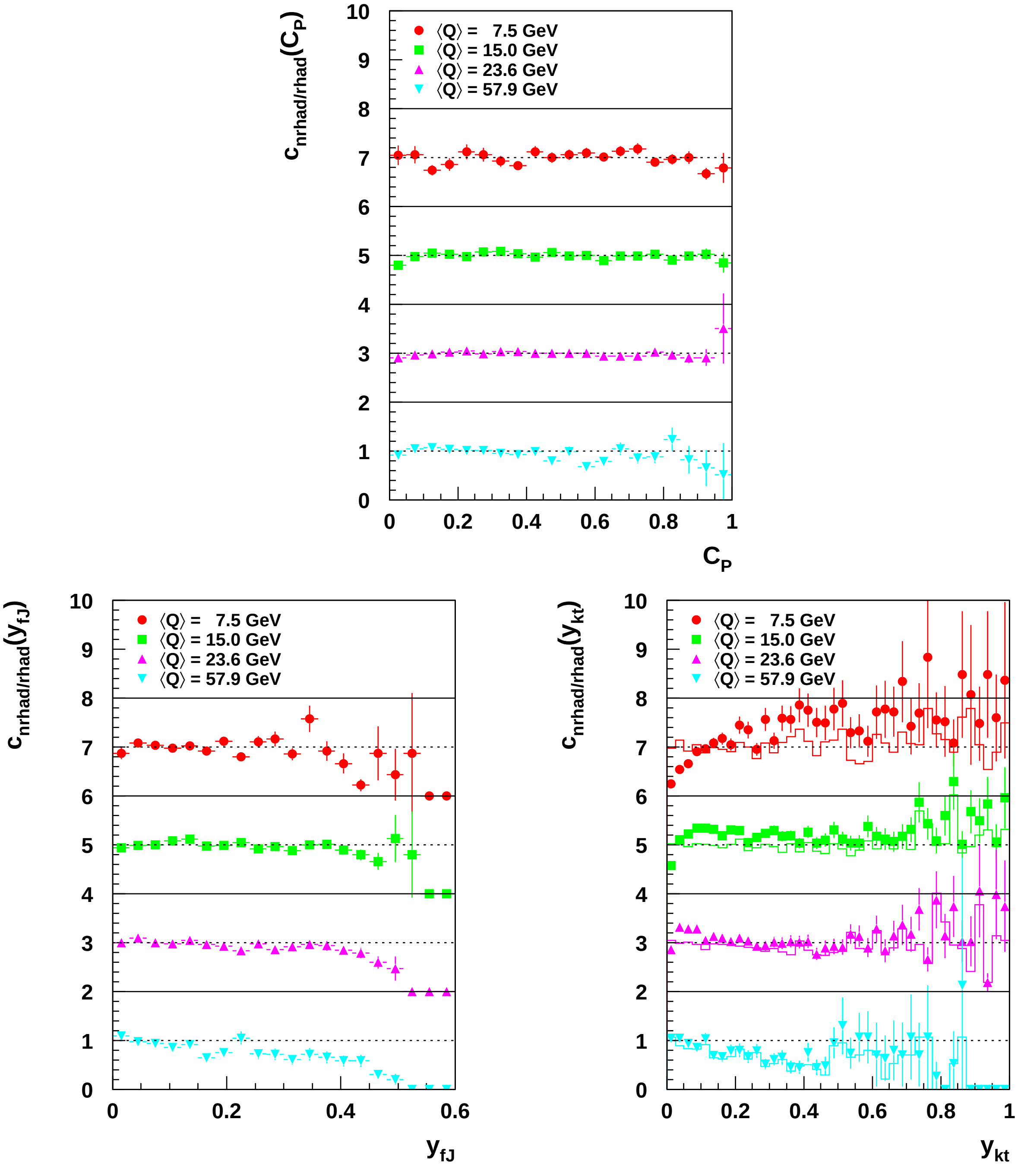


Figure 7.7: Bin-wise radiative corrections for the event shapes C_P, y_{fJ} and y_{k_t} as derived from DJANGO 6.2 for four out of eight investigated bins in Q. The factors for $\langle Q\rangle = 7.5$–$57.9\,\mathrm{GeV}$ are shifted by offsets of $2 \cdot n$, $n = 0, 1, 2, 3$. The dashed lines delineate the corresponding position of unity. Entries exactly at zero (or 2, 4, 6) mean that no events were found on either level. The error bars represent the statistical uncertainty. For y_{k_t} the additional histogram (full line) shows the correction factor without acceptance extrapolation.

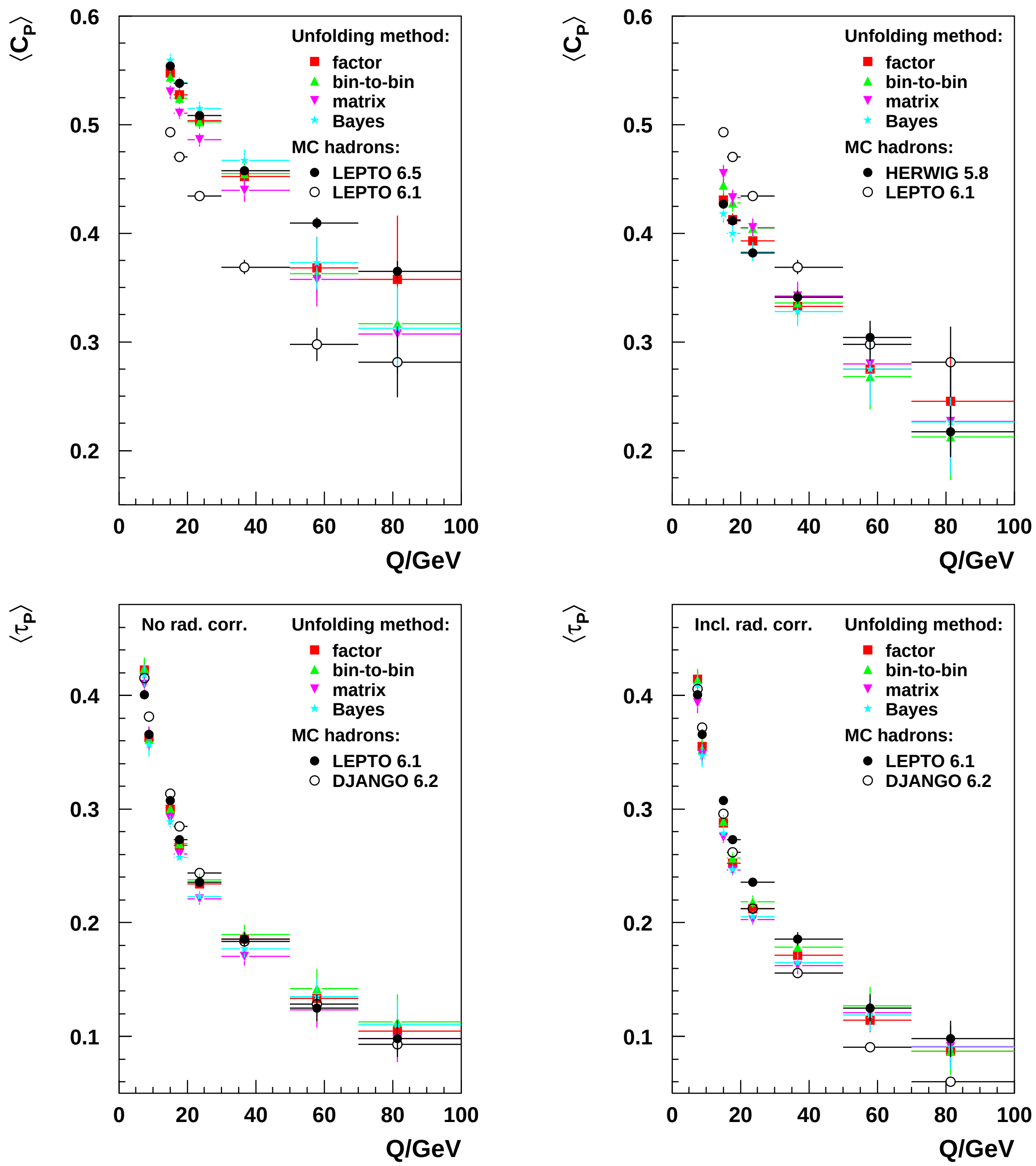


Figure 7.8: Top: Unfolded means of C_P as a function of Q for simulated high Q^2 LEPTO 6.5 (left) and HERWIG 5.8 (right) event samples. LEPTO 6.1 was employed as unfolding MC. Bottom: Unfolded means of τ_P as a function of Q for simulated LEPTO 6.1 event samples with (right) and without (left) radiative corrections derived from DJANGO 6.2. In addition, the corresponding hadron levels of the MC's are shown. The error bars represent statistical uncertainties only.

$\langle Q \rangle$ / GeV	$\langle \tau_P \rangle$	$\langle B_P \rangle$
7.46	$0.4402 \pm 0.0082 \,^{+0.0111}_{-0.0122}$	$0.3624 \pm 0.0034 \,^{+0.0046}_{-0.0038}$
8.74	$0.4017 \pm 0.0090 \,^{+0.0196}_{-0.0080}$	$0.3435 \pm 0.0040 \,^{+0.0076}_{-0.0037}$
14.97	$0.3052 \pm 0.0034 \,^{+0.0081}_{-0.0075}$	$0.2921 \pm 0.0017 \,^{+0.0044}_{-0.0033}$
17.75	$0.2762 \pm 0.0029 \,^{+0.0091}_{-0.0059}$	$0.2760 \pm 0.0016 \,^{+0.0054}_{-0.0034}$
23.62	$0.2279 \pm 0.0031 \,^{+0.0125}_{-0.0071}$	$0.2452 \pm 0.0018 \,^{+0.0078}_{-0.0042}$
36.72	$0.1814 \pm 0.0049 \,^{+0.0107}_{-0.0065}$	$0.2094 \pm 0.0031 \,^{+0.0083}_{-0.0053}$
57.93	$0.1330 \pm 0.0089 \,^{+0.0092}_{-0.0092}$	$0.1717 \pm 0.0062 \,^{+0.0109}_{-0.0096}$
81.32	$0.0984 \pm 0.0130 \,^{+0.0045}_{-0.0051}$	$0.1346 \pm 0.0088 \,^{+0.0117}_{-0.0041}$

$\langle Q \rangle$ / GeV	$\langle \rho_E \rangle$	$\langle \tau_C \rangle$	$\langle C_P \rangle$
7.46	$0.1115 \pm 0.0019 \,^{+0.0011}_{-0.0013}$	$0.1637 \pm 0.0032 \,^{+0.0030}_{-0.0029}$	$0.5601 \pm 0.0082 \,^{+0.0083}_{-0.0073}$
8.74	$0.1044 \pm 0.0021 \,^{+0.0017}_{-0.0000}$	$0.1600 \pm 0.0037 \,^{+0.0060}_{-0.0031}$	$0.5524 \pm 0.0094 \,^{+0.0074}_{-0.0054}$
14.97	$0.0872 \pm 0.0007 \,^{+0.0007}_{-0.0014}$	$0.1333 \pm 0.0013 \,^{+0.0013}_{-0.0019}$	$0.4824 \pm 0.0034 \,^{+0.0036}_{-0.0051}$
17.75	$0.0826 \pm 0.0007 \,^{+0.0013}_{-0.0015}$	$0.1263 \pm 0.0011 \,^{+0.0014}_{-0.0024}$	$0.4621 \pm 0.0030 \,^{+0.0045}_{-0.0069}$
23.62	$0.0714 \pm 0.0007 \,^{+0.0019}_{-0.0016}$	$0.1098 \pm 0.0012 \,^{+0.0020}_{-0.0026}$	$0.4112 \pm 0.0033 \,^{+0.0056}_{-0.0076}$
36.72	$0.0634 \pm 0.0012 \,^{+0.0013}_{-0.0013}$	$0.0985 \pm 0.0021 \,^{+0.0012}_{-0.0023}$	$0.3644 \pm 0.0058 \,^{+0.0029}_{-0.0058}$
57.93	$0.0518 \pm 0.0023 \,^{+0.0016}_{-0.0025}$	$0.0834 \pm 0.0040 \,^{+0.0015}_{-0.0046}$	$0.3127 \pm 0.0122 \,^{+0.0065}_{-0.0131}$
81.32	$0.0410 \pm 0.0034 \,^{+0.0022}_{-0.0016}$	$0.0663 \pm 0.0057 \,^{+0.0031}_{-0.0025}$	$0.2529 \pm 0.0173 \,^{+0.0160}_{-0.0065}$

$\langle Q \rangle$ / GeV	$\langle y_{fJ} \rangle$	$\langle y_{k_t} \rangle$
7.46	$0.1554 \pm 0.0031 \,^{+0.0052}_{-0.0057}$	$0.2814 \pm 0.0077 \,^{+0.0072}_{-0.0123}$
8.74	$0.1399 \pm 0.0034 \,^{+0.0056}_{-0.0048}$	$0.2196 \pm 0.0078 \,^{+0.0109}_{-0.0088}$
14.97	$0.1226 \pm 0.0014 \,^{+0.0053}_{-0.0047}$	$0.1364 \pm 0.0026 \,^{+0.0080}_{-0.0071}$
17.75	$0.1164 \pm 0.0014 \,^{+0.0047}_{-0.0047}$	$0.1156 \pm 0.0021 \,^{+0.0068}_{-0.0061}$
23.62	$0.1015 \pm 0.0015 \,^{+0.0044}_{-0.0042}$	$0.0943 \pm 0.0022 \,^{+0.0059}_{-0.0041}$
36.72	$0.0898 \pm 0.0026 \,^{+0.0037}_{-0.0041}$	$0.0665 \pm 0.0030 \,^{+0.0041}_{-0.0015}$
57.93	$0.0752 \pm 0.0050 \,^{+0.0039}_{-0.0056}$	$0.0451 \pm 0.0046 \,^{+0.0064}_{-0.0024}$
81.32	$0.0557 \pm 0.0063 \,^{+0.0045}_{-0.0056}$	$0.0295 \pm 0.0042 \,^{+0.0046}_{-0.0028}$

Table 7.1: Corrected mean values of the event shapes as a function of Q. The first error is statistical, the second systematic and comprises unfolding as well as energy scale uncertainties added quadratically.

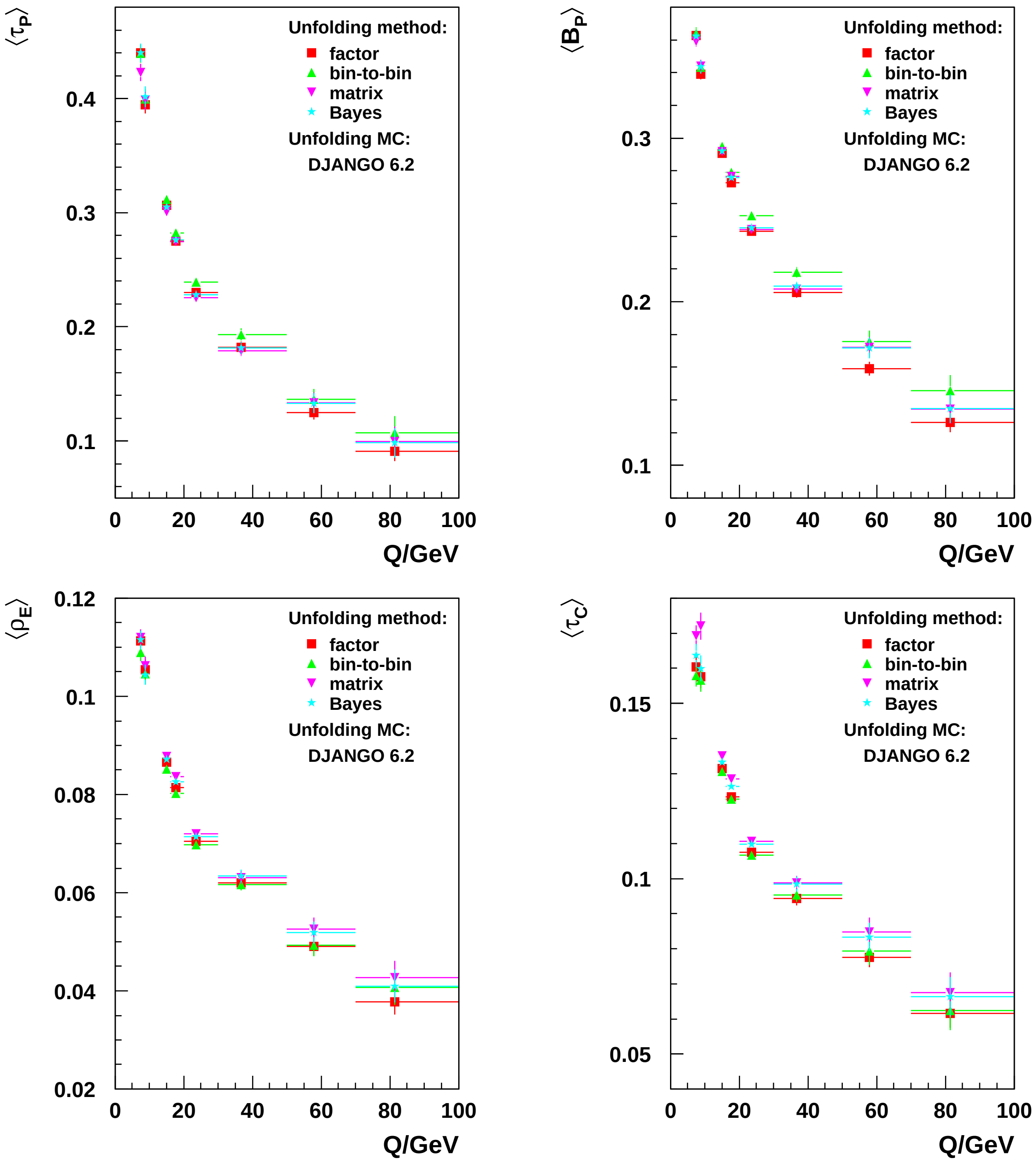


Figure 7.9: Unfolded means of τ_P, B_P, ρ_E and τ_C as a function of Q for the four described correction schemes. DJANGO 6.2 served as unfolding MC. Note that the step from radiative to non-radiative hadron level can be accomplished by the factor or bin-to-bin methods only. The latter was employed for this purpose in case of the matrix and Bayes procedures. The error bars represent statistical uncertainties only.

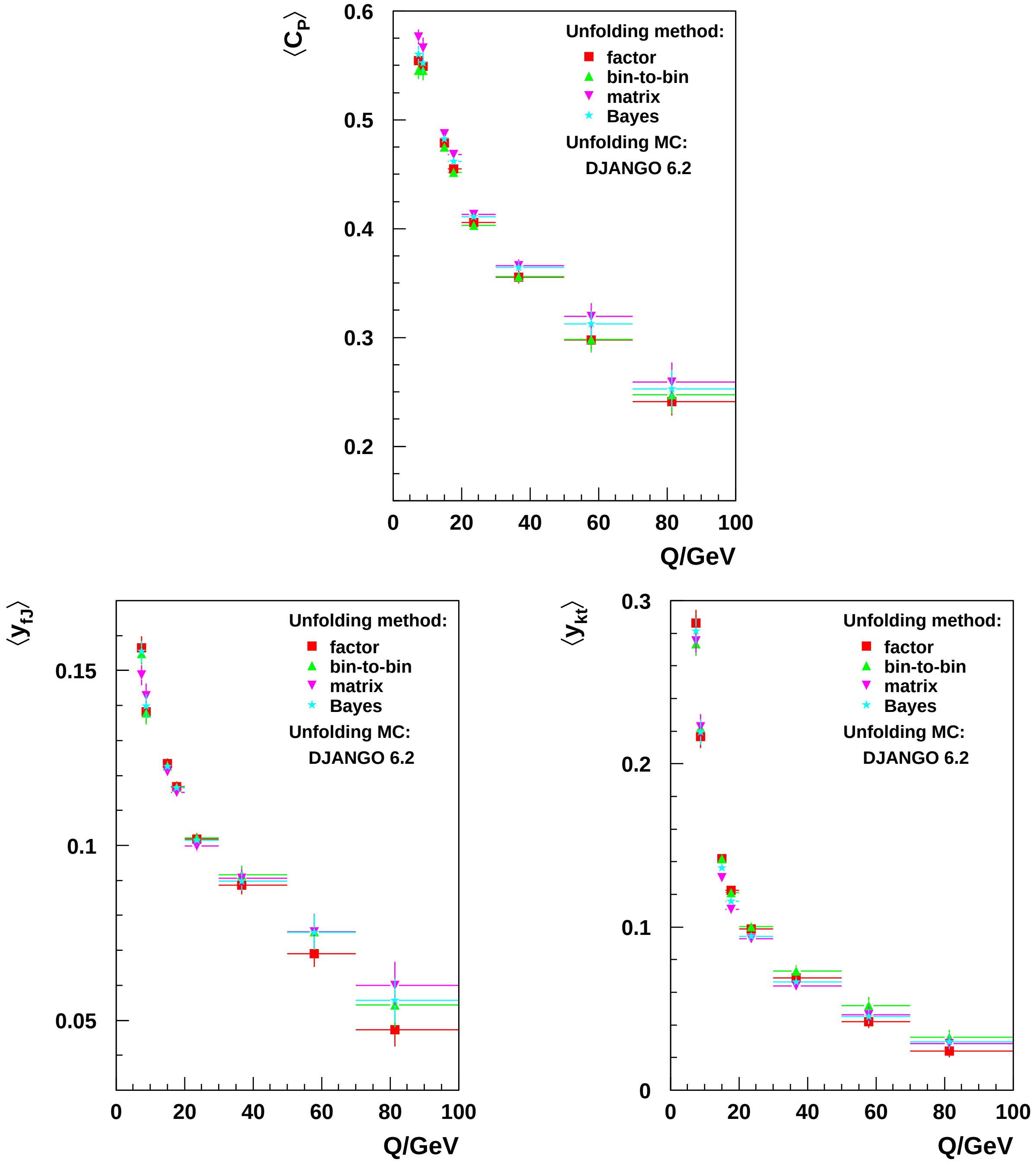


Figure 7.10: Unfolded means of C_P, y_{fJ} and y_{k_t} as a function of Q for the four described correction schemes. DJANGO 6.2 served as unfolding MC. Note that the step from radiative to non-radiative hadron level can be accomplished by the factor or bin-to-bin methods only. The latter was employed for this purpose in case of the matrix and Bayes procedures. The error bars represent statistical uncertainties only.

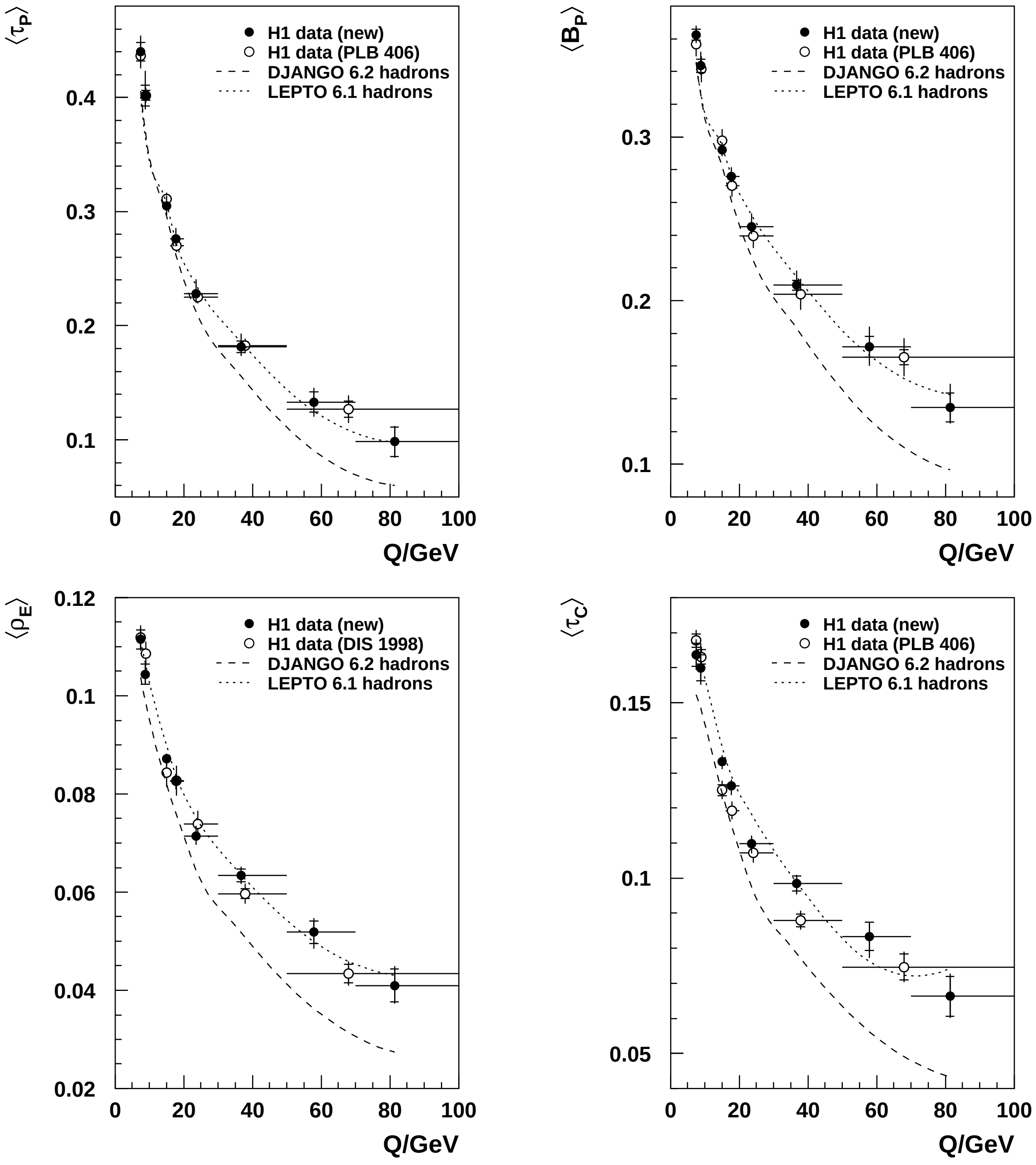


Figure 7.11: Corrected means (full symbols) of τ_P, B_P, ρ_E and τ_C as a function of Q. The inner error bars represent the statistical uncertainty, the outer ones the total statistic and systematic uncertainty. They are compared with the hadron level mean values of the DJANGO 6.2 (dashed) and LEPTO 6.1 (dotted) MC and with already published results by H1 (hollow symbols).

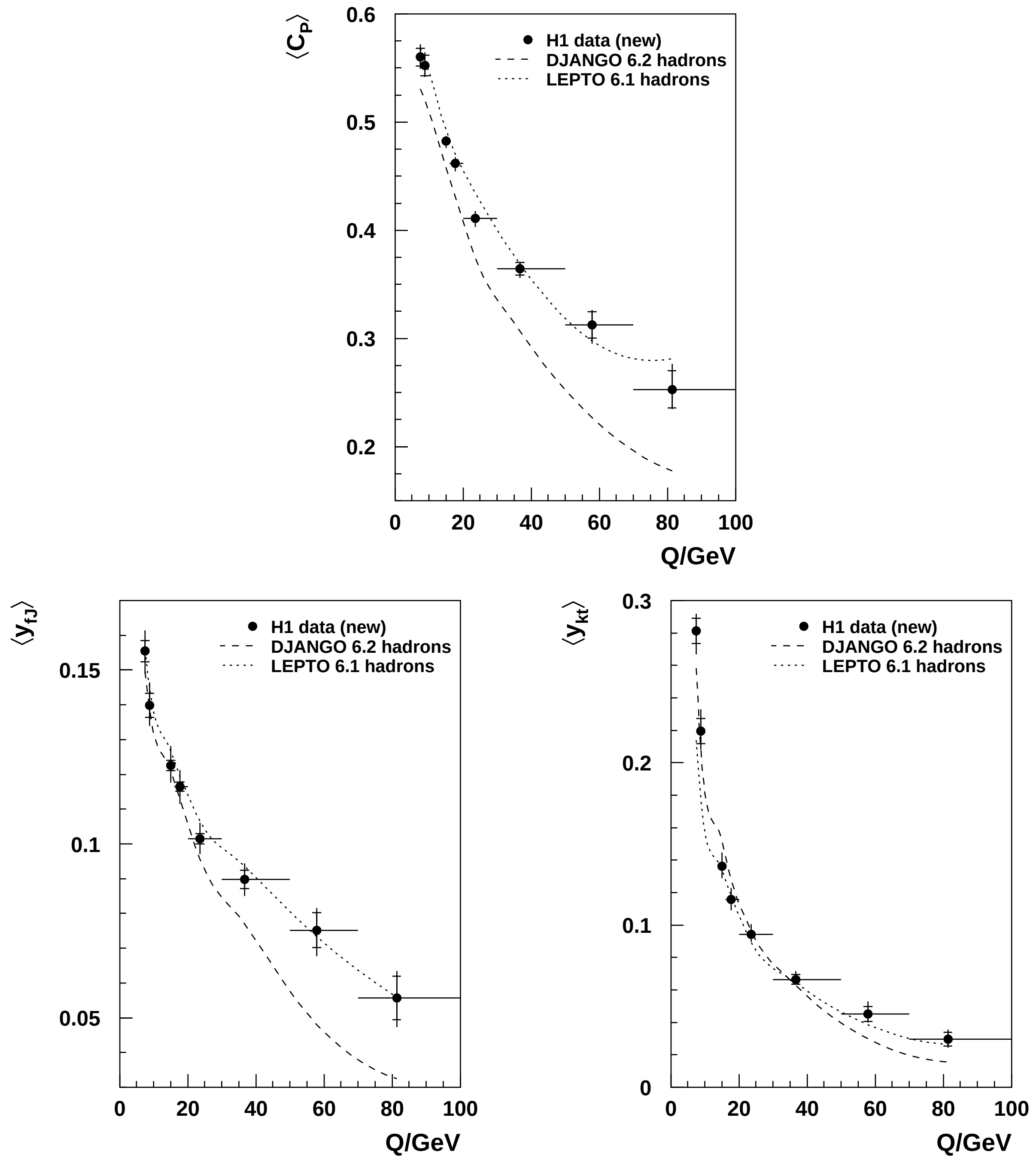


Figure 7.12: Corrected means (full symbols) of C_P, y_{fJ} and y_{k_t} as a function of Q. The inner error bars represent the statistical uncertainty, the outer ones the total statistic and systematic uncertainty. They are compared with the hadron level mean values of the DJANGO 6.2 (dashed) and LEPTO 6.1 (dotted) MC.

Chapter 8

NLO Integration Programs

According to the definition of the event shapes F in section 4.2, they will take on non-trivial values, i.e. $F \neq 0$, when derived to $\mathcal{O}(\alpha_s)$ at least. In terms of cross sections, the distributions dn/dF may be written to LO accuracy as

$$\frac{1}{\sigma_{\mathrm{tot}}}\frac{d\sigma}{dF} = C_1(F)\alpha_s\,. \tag{8.1}$$

The first moment or mean value, characterizing the distribution in F of an event sample in a simple way, now reads

$$\langle F\rangle := \frac{\int\limits_0^{F_{\max}} F\frac{d\sigma}{dF}dF}{\int\limits_0^{F_{\max}} \frac{d\sigma}{dF}dF} = \frac{1}{\sigma_{\mathrm{tot}}}\int\limits_0^{F_{\max}} F\frac{d\sigma}{dF}dF\,, \tag{8.2}$$

which can be expanded to LO in α_s as well:

$$\langle F\rangle = c_{1,F}\alpha_s\,. \tag{8.3}$$

However, theoretical arguments show that the achieved precision of LO calculations is not sufficient to extract much information on pQCD from experimental data. Especially the concept of a *running coupling constant* $\alpha_s(\mu_R)$, explained in the next section, requires the distributions and mean values to be known at least to NLO, i.e. $\mathcal{O}(\alpha_s^2)$.

8.1 The Running Coupling Constant

Following an approach from [50], one may assume that a dimensionless quantity $\langle F\rangle$ depends on a single energy scale Q which is much larger than any other parameter measured in units of energy such that quark masses may be set to zero.[1] Then, the conclusion is

[1] Note that this step requires a more careful consideration, cf. the ref. given above.

that $\langle F \rangle$ has a constant value independent of Q; there is no other possibility to remove the dimension of Q. However, this is not true in a renormalizable quantum field theory such as QCD.

Perturbative calculations operating with "naked" charges e_0 (or masses m_0) of point particles have to deal with ultraviolet divergences occurring in the integrations of loop momenta $d^4k/(2\pi)^4$. The procedure allowing us to absorb these divergences in physical charges e is called *renormalization*. Results concerning the (measured) physical charges are then all finite.

There exist several recipes how this can be done [51]. One of the most popular, also adopted in this analysis, is the *modified minimal subtraction* scheme $\overline{\mathrm{MS}}$, where the singularities are regulated by a reduction of the space-time dimensions to $n < 4$:

$$\frac{d^4k}{(2\pi)^4} \to \mu^{2\epsilon} \frac{d^{4-2\epsilon}k}{(2\pi)^{4-2\epsilon}} \,. \tag{8.4}$$

At an arbitrary scale called *renormalization scale* μ_R, the integrals are subdivided into finite and divergent parts for $\epsilon \to 0$. The singular terms (and some finite ones depending on the renormalization scheme) are then attributed to a renormalized charge e. However, the remaining finite terms also modify the charge e or equivalently the coupling constant $\alpha := e^2/4\pi$, which now explicitly varies with μ_R!

Relating this to QCD, we have a strong coupling constant $\alpha_s := g^2/4\pi$ containing μ_R — a second mass scale that enters the game. $\langle F \rangle$ may therefore depend on the ratio Q^2/μ_R^2. But since μ_R is an arbitrary parameter of the renormalization procedure, it should not make an appearance in physical observables like $\langle F \rangle$. This can be expressed mathematically by

$$\mu_R^2 \frac{d}{d\mu_R^2} \langle F \rangle \left(Q^2/\mu_R^2, \alpha_s\right) = \left[\mu_R^2 \frac{\partial}{\partial \mu_R^2} + \mu_R^2 \frac{\partial \alpha_s}{\partial \mu_R^2} \frac{\partial}{\partial \alpha_s}\right] \langle F \rangle \left(Q^2/\mu_R^2, \alpha_s\right) = 0 \tag{8.5}$$

or, using the notation

$$t = \ln\left(\frac{Q^2}{\mu_R^2}\right) , \tag{8.6}$$

$$\beta(\alpha_s) = \mu_R^2 \frac{\partial \alpha_s}{\partial \mu_R^2} , \tag{8.7}$$

as

$$\left[-\frac{\partial}{\partial t} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s}\right] \langle F \rangle \left(e^t, \alpha_s\right) = 0 \,. \tag{8.8}$$

A solution to this partial differential equation can be given implicitly by defining the running coupling $\alpha_s(Q^2)$ as

$$t = \int_{\alpha_s}^{\alpha_s(Q^2)} \frac{dx}{\beta(x)} , \quad \alpha_s := \alpha_s(\mu_R^2) \tag{8.9}$$

such that $\langle F\rangle(1, \alpha_s(Q^2))$ fulfils eq. (8.8). The scale dependence of any $\langle F\rangle$ calculated perturbatively now follows from $\alpha_s(Q^2)$ provided eq. (8.9) could be solved. Since QCD is an asymptotically free theory, i.e. $\lim_{Q\to\infty} \alpha_s(Q^2) = 0$, this can also be achieved by expanding the β-function of the **R**enormalization **G**roup **E**quation (8.7) (RGE) in powers of α_s:

$$\frac{\beta(\alpha_s)}{4\pi} = -\sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi}\right)^{n+2} . \tag{8.10}$$

The first two coefficients defined that way are

$$\beta_0 = \frac{33 - 2N_f}{3} , \tag{8.11}$$

$$\beta_1 = \frac{306 - 38N_f}{3} , \tag{8.12}$$

where N_f is the number of active flavours. β_2 and further coefficients are scheme dependent.

Retaining only the first term, the 1-loop solution for $\alpha_s(Q)$ can be written as

$$\alpha_s(Q) = \frac{\alpha_s(\mu_R)}{1 + \frac{\beta_0}{2\pi}\alpha_s(\mu_R)\ln\left(\frac{Q}{\mu_R}\right)} . \tag{8.13}$$

Using this equation, α_s can be evaluated at any (sufficiently high) scale provided it is known at one point, say M_Z. Historically, a dimensional parameter Λ derived from

$$\ln\frac{Q^2}{\Lambda^2} = -\int_{\alpha_s(Q^2)}^{\infty} \frac{dx}{\beta(x)} \tag{8.14}$$

was determined instead of $\alpha_s(M_Z)$. The 1-loop formula then reads

$$\alpha_s(Q) = \frac{2\pi}{\beta_0 \ln\left(\frac{Q}{\Lambda}\right)} \tag{8.15}$$

and demonstrates that Λ indicates the scale at which the coupling would diverge if extrapolated outside the perturbative regime.

In principle, both approaches are equivalent and any measurement of $\alpha_s(M_Z)$ can be converted into a corresponding value for Λ and vice versa. The latter, however, has some disadvantages: It is not dimensionless, it depends on N_f and on the renormalization scheme[2] and — most important — there exist two slightly different definitions for Λ in the literature [5, 50].

[2]That is, it has to be labelled e.g. $\Lambda_{N_f,\overline{\mathrm{MS}}}$.

Except for the consistency check in section 9.2 employing $\Lambda_{5,\overline{\rm MS}}$, we therefore follow ref. [52] and apply the 2-loop equation according to

$$\alpha_s(Q) = \frac{\alpha_s(M_Z)}{1 + \alpha_s(M_Z) \cdot L^{(n)}\left(\frac{Q}{M_Z}\right)} \tag{8.16}$$

where

$$L^{(1)}\left(\frac{Q}{M_Z}\right) = \frac{\beta_0}{2\pi} \ln \frac{Q}{M_Z} \tag{8.17}$$

reproduces the 1-loop formula above and

$$L^{(2)}\left(\frac{Q}{M_Z}\right) = \left(\frac{\beta_0}{2\pi} + \frac{\beta_1}{8\pi^2}\alpha_s(M_Z)\right) \ln \frac{Q}{M_Z} \tag{8.18}$$

gives the 2-loop result.

Returning to $\langle F \rangle$, the perturbative expression to all orders

$$\langle F \rangle = \sum_{n=1}^{\infty} c_n(\mu_R)\alpha_s^n(\mu_R) \tag{8.19}$$

does not depend on μ_R. But we do not know all coefficients, and therefore any approximation by a truncated series does. To be more precise,

$$\mu_R^2 \frac{d}{d\mu_R^2} \sum_{n=1}^{N} c_n(\mu_R)\alpha_s^n(\mu_R) \propto \mathcal{O}(\alpha_s^{N+1}(\mu_R))\,. \tag{8.20}$$

According to current knowledge, the dependence on μ_R is much reduced if more terms are inserted. Hence, the theoretical uncertainty usually associated with the choice of scale, whose variation reflects the unknown higher orders, is diminished. Up to now, we merely considered c_1, which itself does not contain μ_R, and the complete μ_R-dependence stems from $\alpha_s(\mu_R)$ and is rather large. It is reduced with the inclusion of $c_2(\mu_R)\alpha_s^2(\mu_R)$ with

$$c_2(Q) = c_2(1) + \frac{\beta_0}{2\pi} \ln \frac{\mu_R}{Q} c_1\,, \tag{8.21}$$

which can be calculated by the three NLO integration programs presented in the next sections.

8.2 NLO Integration Techniques

As mentioned in section 3.4 on the cross section to $\mathcal{O}(\alpha_s)$, again infrared and collinear divergences are present in calculations to $\mathcal{O}(\alpha_s^2)$ which cancel between real and virtual diagrams. Examples of real corrections to the QCDC and BGF processes with now three final state partons are given in fig. 8.1.

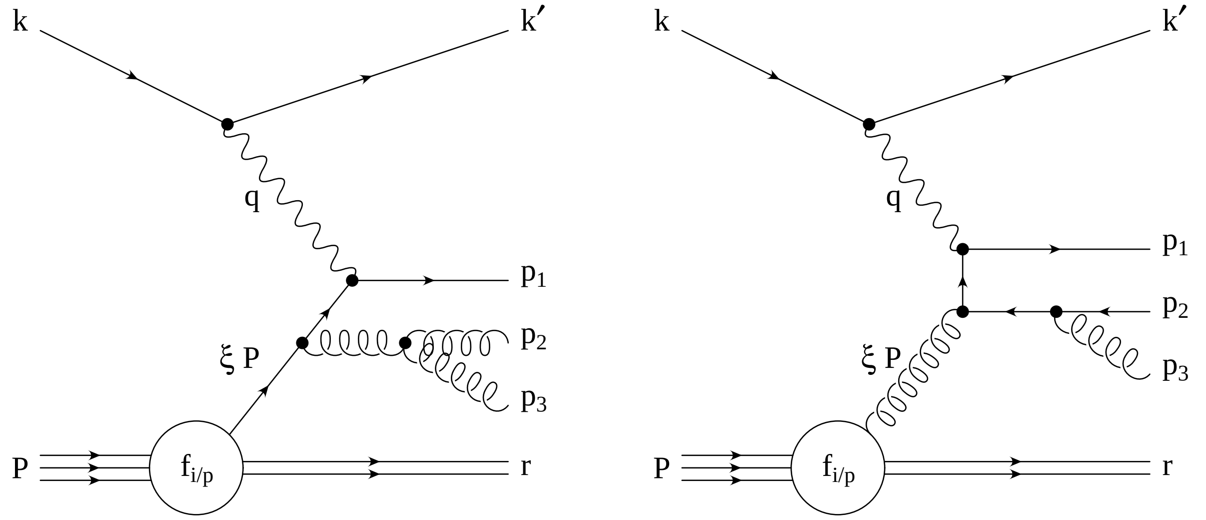


Figure 8.1: Examples of real corrections to QCDC (left) and BGF (right) Feynman graphs.

In order to eliminate the occurring singularities and obtain numerically stable results, the necessary cancellations were performed analytically for a certain jet algorithm in the programs DISJET [53] and PROJET [54]. However, these programs lack the flexibility to evaluate NLO corrections for arbitrary infrared-safe observables and they contain approximations that are inappropriate for certain phase space regions. Therefore, more elaborate methods have been developed that ensure the cancellations and evaluate numerical results for arbitrary observables. Basically, two such schemes are known: the *phase space slicing* method [55] and the *subtraction* method [28].

In the phase space slicing method, the separation of singular from non-singular phase space regions is done by surrounding partons with a small cone. Real emissions of partons outside the cone defined by $s_{ij} = (p_i + p_j)^2 > s_{\min}$, where $s_{\min} \approx 0.1\,\mathrm{GeV}^2$, are computed exactly. Inside the cone, soft and collinear approximations are used. A drawback of this procedure is the possibility of a residual dependence on the technical parameter $s_{\min}$. For any calculation it has to be established that a plateau has been reached for small enough cut-offs. In addition, this scheme prevents the integration of the event shape distributions $F d\sigma/dF$ to be carried out over the whole phase space from zero to $F_{\max}$. Instead, a lower bound $F_{\mathrm{cut}} > 0$ has to be imposed and $s_{\min}$ has to be adjusted accordingly. Moreover, the introduction of *crossing functions* [56], that have to be evaluated for every parton density parameterization, is required.

In case of the subtraction method, the singularities are cancelled point-by-point in phase space. A cut-off like $s_{\min}$ is not formally required by this scheme, although a tiny value of $s_{\min} \approx 10^{-8}$–$10^{-10}\,\mathrm{GeV}^2$ is applied to set the number of significant digits. Hence, the event shape distributions $F d\sigma/dF$ can be integrated down to zero [57].

The first method is employed in the NLO integration program MEPJET [58], whereas the subtraction procedure is used in DISENT [59] and DISASTER++ [60]. All three programs, however, do not contain the full set of diagrams to $\mathcal{O}(\alpha_s^2)$. Additional 2-

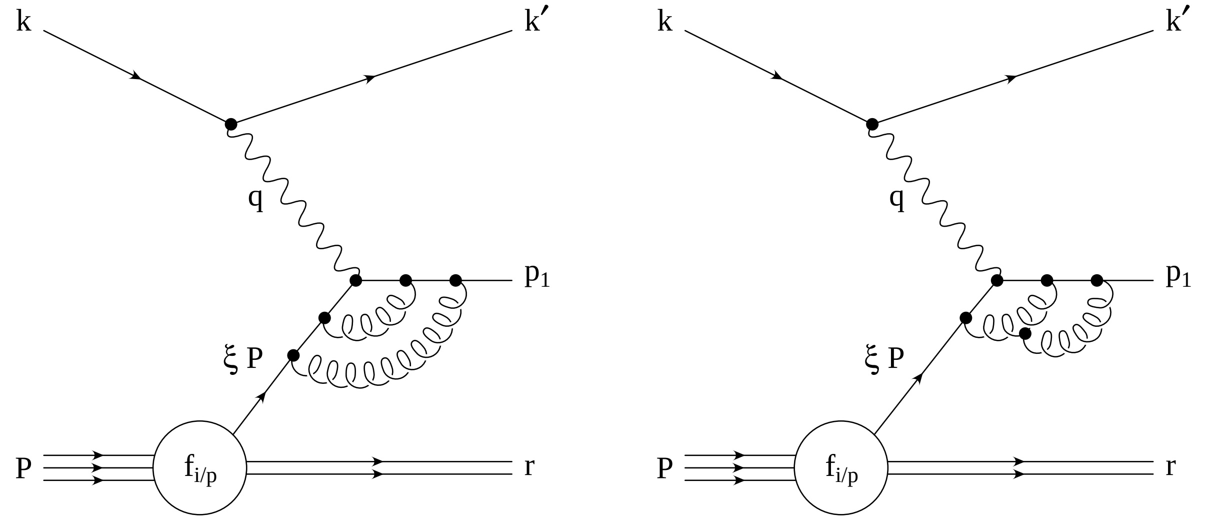


Figure 8.2: Examples of 2-loop virtual correction Feynman graphs.

loop virtual corrections from graphs like the ones in fig. 8.2 appear and are difficult to handle. Yet, they contribute as interference term to the Born process only and therefore do not affect the event shape distributions except for the bin at zero. The integral $\int_0^{F_{\max}} F\frac{d\sigma}{dF}dF$ still contains integrable divergences, but they are not harmful for DISENT or DISASTER++ [61]. A calculation of the total cross section $\int_0^{F_{\max}} \frac{d\sigma}{dF}dF$, needed for the normalization, to $\mathcal{O}(\alpha_s^2)$ (i.e. NNLO), however, is not possible. A Taylor expansion of eq. (8.2)

$$\begin{aligned}
\langle F\rangle &= \frac{\int\limits_0^{F_{\max}} F\frac{d\sigma}{dF}dF}{\int\limits_0^{F_{\max}} \frac{d\sigma}{dF}dF} && (8.22)\\
&\approx \frac{a_1\alpha_s^1 + a_2\alpha_s^2}{b_0\alpha_s^0 + b_1\alpha_s^1 + b_2\alpha_s^2} && (8.23)\\
&= \frac{a_1}{b_0}\alpha_s^1 + \frac{a_2 - \frac{a_1 b_1}{b_0}}{b_0}\alpha_s^2 + \mathcal{O}(\alpha_s^3) && (8.24)
\end{aligned}$$

shows that, owing to the property of F to be zero in the QPM limit and hence $a_0 = 0$, $\sigma_{\rm tot}$ needs only be known to $\mathcal{O}(\alpha_s)$ for $\langle F\rangle$ to be still correct to $\mathcal{O}(\alpha_s^2)$.

8.3 First Comparisons of NLO Programs

In view of the complexity of the calculations and their implementations into program code, it is mandatory to compare the results where possible. A first comparison of MEPJET 1.4 and DISENT 0.0 was performed in [25]. Due to the phase space slicing method applied

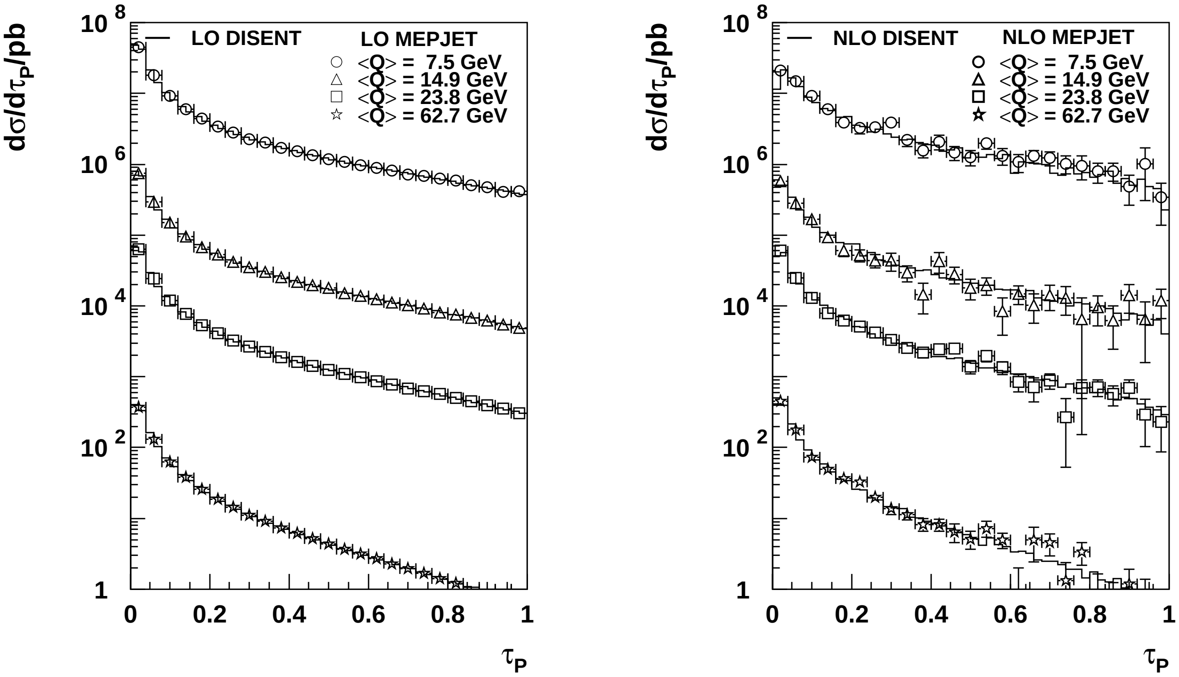


Figure 8.3: Differential distributions of τ_P for four out of seven investigated bins in Q. LO (left) and NLO (right) pQCD calculations of MEPJET 1.4 (hollow symbols) and DISENT 0.0 (full lines) are compared. The spectra for $\langle Q \rangle = 7.5$–$62.7\,\mathrm{GeV}$ are multiplied by factors of 10^n, $n = 0, 1, 2, 3$.

in MEPJET, the differential distributions for $F > F_{\mathrm{cut}} > 0$ only could be considered. In order to keep F_{cut} small, a cut-off of $s_{\min} = 0.01\ \mathrm{GeV}^2$ instead of $0.1\ \mathrm{GeV}^2$ was used. Nevertheless, the left-most bins extending down to zero should not be taken seriously, even if they accidentally do agree. Fig. 8.3 shows a perfect agreement of the LO spectra and, within the calculational precision of 3–5%, compatible results for the NLO distributions.

With the advent of DISASTER++, a check of all three programs against each other was carried out and some discrepancies have been revealed [60]. Partially, these are understood and meanwhile an improved version DISENT 0.1 has been released. In event shape spectra like fig. 8.3, effects are hardly visible, but the mean values have increased considerably in the low Q^2 region compared to the calculations that entered in [23]. The worst case was the jet broadening B_P, whose mean value at $Q \approx 7.5\,\mathrm{GeV}$ rose by about 14%.

The most recent program versions are MEPJET 2.2, DISENT 0.1 and DISASTER++ 1.0.1. In this analysis only DISENT will be used, but more elaborate comparisons are under way in the DESY workshop on *Monte Carlo Generators for HERA Physics*.

Chapter 9

DISENT Results

DISENT, the NLO integration program mainly employed here, is structured similarly to a MC event generator. A main routine generates configurations of parton momenta in the allowed phase space region — for our purposes defined by the cuts nos. 1–6 of section 5.5 —, stores them in an event record and calculates the weight according to the matrix elements. Consequently, a user routine may analyze these events with respect to arbitrary infrared and collinear safe observables. For physical cross sections, the weights have to be multiplied by the strong coupling constant raised to the proper power and by weights extracted from the pdfs of the proton. These are not inherent to DISENT but must be gathered from somewhere else, e.g. the PDFLIB [39]. As standard options we use the set MRSA'-115 [62] and $N_f = 5$ as number of active flavours. Note that DISENT currently neglects quark masses as well as electroweak corrections proportional to $Q^2/(Q^2 + M_Z^2)$ and $(Q^2/(Q^2 + M_Z^2))^2$. Fortunately, only the highest Q-bin investigated with $\langle Q \rangle = 81.32\,\mathrm{GeV}$ may be affected noticeably [13]. In the near future, DISENT may be extended to include the electroweak terms [63].

In distinction to MC generators, however, several different but correlated event records, in total 14, are produced for **one** DISENT "event." They correspond to contributions of real and virtual diagrams plus counter terms that ensure the necessary cancellation of divergences. These contributions have coupled weights that are positive as well as negative. For the error evaluation, it is mandatory that these weights are allowed to counterbalance each other before finally e.g. being histogrammed. This may be done by creating an intermediate array or histogram, where only the contributions of one event are entered. All bins not empty at the end are then transferred to the final array with their respective weights. A more detailed program description can be found in [64].

9.1 Evaluation of the Statistical Uncertainty

Given the framework of the DISENT program, the calculation of differential distributions $d\sigma/dF$ and mean values $\langle F \rangle$ is straightforward. Positive and negative weights of the different contributions are summed up. Yet, the evaluation of the statistical uncertainties, where the technique of intermediate histograms is presupposed, is more involved.

First, we want to consider cross sections that can be written as a sum of weights w_i, given by DISENT, which are already normalized to the number of events N:

$$\sigma = \sum_{i=1}^{N} w_i \,. \tag{9.1}$$

By reintroducing unnormalized weights v_i, one obtains

$$\sigma = \frac{1}{N} \sum_{i=1}^{N} v_i = \langle v \rangle \,, \tag{9.2}$$

demonstrating that σ is computed analogously to a mean value. Its statistical uncertainty can be estimated by

$$\begin{aligned}
\Delta\sigma &= \sqrt{\frac{1}{N-1}\left(\langle v^2 \rangle - \langle v \rangle^2\right)} \\
&= \sqrt{\frac{1}{N-1}\left(\frac{1}{N}\sum_{i=1}^{N} v_i^2 - \frac{1}{N^2}\left(\sum_{i=1}^{N} v_i\right)^2\right)} \qquad (9.3) \\
&= \sqrt{\frac{1}{N-1}\left(N\sum_{i=1}^{N} w_i^2 - \left(\sum_{i=1}^{N} w_i\right)^2\right)} \,.
\end{aligned}$$

With $N \approx N-1$, this finally gives

$$\Delta\sigma = \sqrt{\sum_{i=1}^{N} w_i^2 - \frac{\left(\sum\limits_{i=1}^{N} w_i\right)^2}{N}} \,, \tag{9.4}$$

which in the limit of very large N becomes

$$\Delta\sigma = \sqrt{\sum_{i=1}^{N} w_i^2} \,. \tag{9.5}$$

For our purposes the second term is not negligible and formula (9.4) has to be applied. Turning to the event shape means

$$\langle F \rangle = \frac{\sum\limits_{i=1}^{N} F_i w_i}{\sum\limits_{i=1}^{N} w_i'} \tag{9.6}$$

one could be misled to the assumption that a similar uncertainty as in eq. (9.4) could be derived here. Yet, formula (9.6) emphasizes that the weights w_i and $w_i^{'}$ are not identical because the numerator is evaluated to $\mathcal{O}(\alpha_s^2)$ and the cross section in the denominator to $\mathcal{O}(\alpha_s)$ only. Instead, the uncertainties have to be estimated separately, and afterwards the error of the quotient can be obtained by the usual formulae. The final result reads:

$$\Delta\langle F\rangle = \langle F\rangle \sqrt{\frac{\sum\limits_{i=1}^{N}(F_i w_i)^2}{\left(\sum\limits_{i=1}^{N} F_i w_i\right)^2} + \frac{\sum\limits_{i=1}^{N} w_i^{'2}}{\left(\sum\limits_{i=1}^{N} w_i^{'}\right)^2} - \frac{2}{N}}\,. \tag{9.7}$$

As a cross-check, the same DISENT 0.0 calculation has been repeated about a hundred times with different seeds of the random number generator for 0.1, 0.2, 0.5 and 1.0 million generated parton events each. Fig. 9.1 shows on the left-hand side histograms of the obtained cross sections σ_{tot} and mean values $\langle\tau_P\rangle$ together with the fit of a Gaussian distribution for half a million events. It describes the distribution of the results well and its standard deviation is compatible with the deduced uncertainty $\mathrm{RMS}_{\mathrm{calc}}$. The right-hand side presents the dependence of the uncertainties for both, the statistical and the computational derivation, on the number of events produced. It nicely exhibits errors shrinking proportional to $1/\sqrt{N}$. The computation underestimates only slightly the statistically achieved variance.

The error bars in differential distributions are evaluated according to eq. (9.5).

9.2 x -Dependence

$\alpha_s(M_Z)$ is the fundamental parameter of QCD that has to be determined by experiment. An assumption on it, however, is already contained in our cross section and mean value calculations. Therefore, they cannot directly be used to extract information from our data. In [25] and [23] this problem was avoided by fitting eq. (8.19) up to $\mathcal{O}(\alpha_s^2)$ to the obtained mean values of each event shape. The resulting coefficients $c_{1,F}$ and $c_{2,F}$ are subsequently employed for an $\alpha_s(M_Z)$-dependent parameterization of the event shape means.

In contrast to e^+e^- physics, where $c_{1,F}$ and $c_{2,F}$ truly are pure numbers (s. e.g. [65]), this is not the case in *ep* DIS because in addition to the matrix elements x-dependent pdfs enter their evaluation. In general, the coefficients will therefore be **functions of x** which was taken into account in an approximate way only by the above ansatz. Fig. 9.2 presents as examples for a strong — the "worst" case being y_{k_t} — and a weak variation of the mean values with x curves of $\langle\tau_P\rangle$ and $\langle C_P\rangle$ versus Q for four different bins in x. Note that in this investigation the two highest Q-bins were merged.

A solution to this problem is the direct calculation of $c_{1,F}$ and $c_{2,F}$ for every Q-bin with its specific x-range separately. This method corresponds to an approximation of functions $c_{n,F}(x)$ by step functions instead of one suitably chosen number.

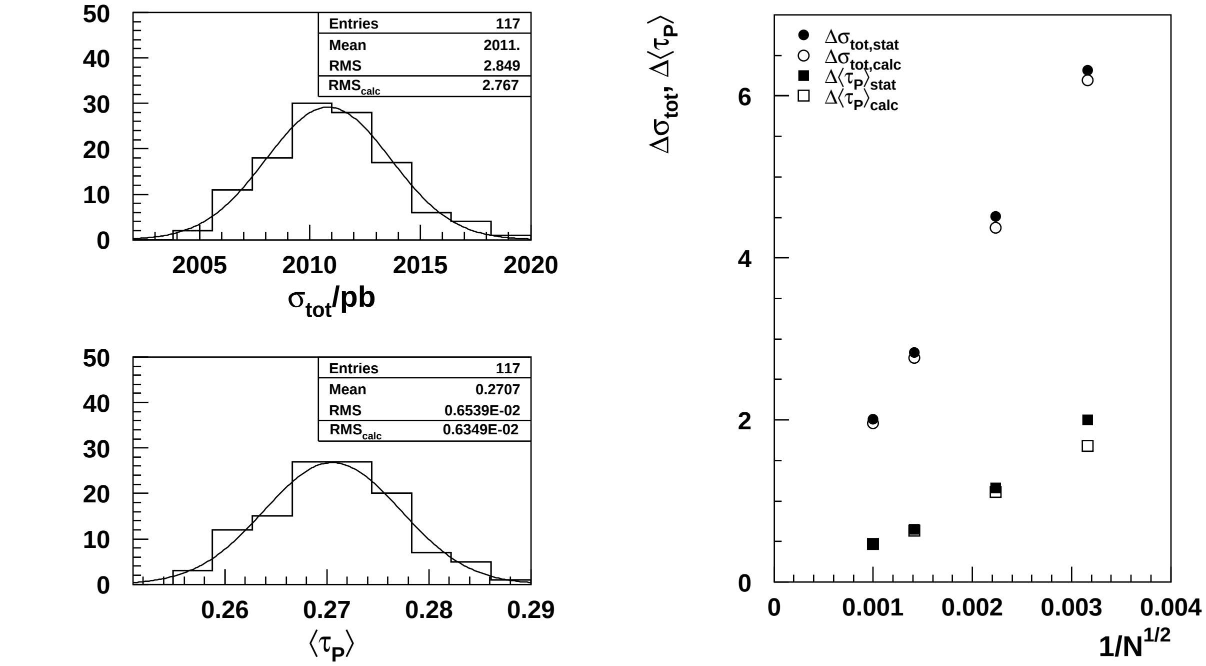


Figure 9.1: Left: Histograms of $\sigma_{\rm tot}$ (top) and $\langle\tau_P\rangle$ (bottom) plus Gaussian fit resulting from 117 DISENT 0.0 calculations with half a million events each. $\mathrm{RMS}_{\rm calc}$ denotes the on average estimated standard deviation. Right: Dependence on $1/\sqrt{N}$ of the statistical (full symbols) and computational (hollow symbols) uncertainty estimates for $\sigma_{\rm tot}$ and $\langle\tau_P\rangle$.

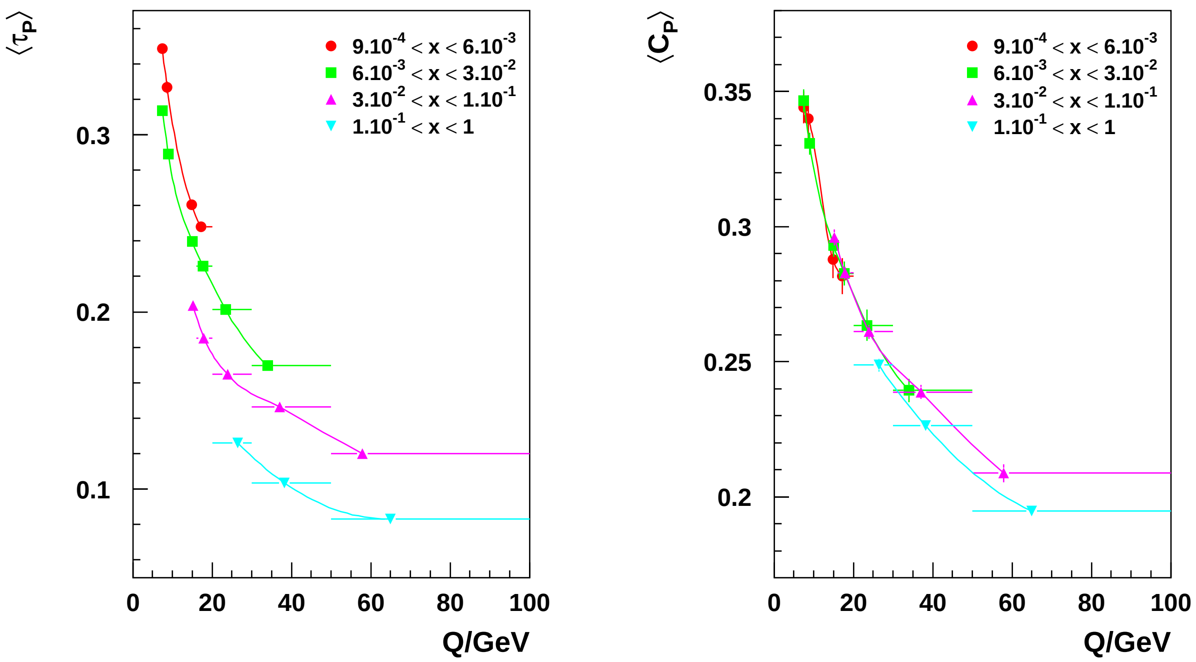
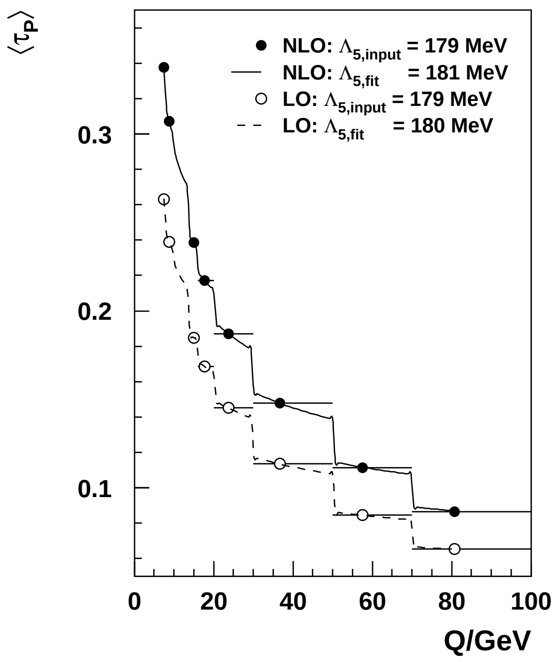


Figure 9.2: Mean values of τ_P (left) and C_P (right) versus Q in four different bins of x.



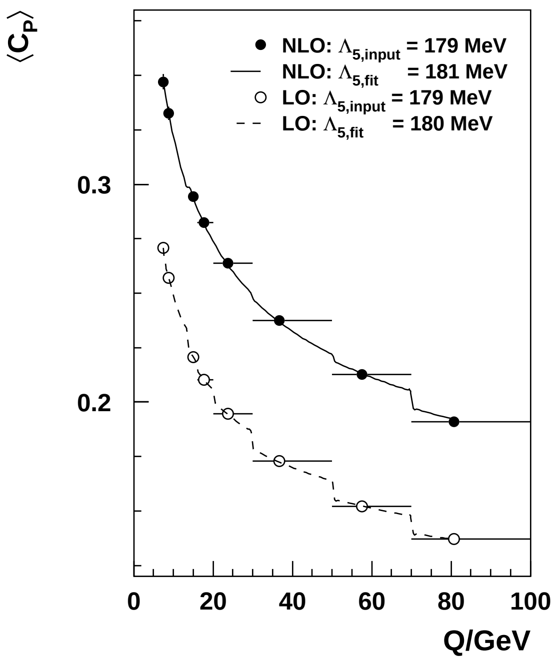


Figure 9.3: Fits (dashed and full lines) of $\Lambda_{5,\overline{\mathrm{MS}}}$ for the LO (hollow symbols) and NLO (full symbols) means of τ_P (left) and C_P (right). The steps express the x-dependence of the perturbative coefficients $c_{1,F}$ and $c_{2,F}$.

As a consistency check, the perturbative mean values at LO and NLO are fit with the coefficients derived in the same computation. The outcome for $\Lambda_{5,\overline{\mathrm{MS}}}$ should be identical to the initial choice, that is $\Lambda_{5,\overline{\mathrm{MS}}} = 179$ MeV for MRSA'-115. The achieved fits for $\langle \tau_P \rangle$ and $\langle C_P \rangle$ are displayed in fig. 9.3. Again, the much steeper steps for $\langle \tau_P \rangle$ indicate its stronger x-dependence. The $\Lambda_{5,\overline{\mathrm{MS}}}$ values extracted turn out to be very well compatible with the expectation for all event shapes except for the jet broadening B_P in NLO! This may be a hint that there is a problem in the DISENT 0.1 calculations at low Q^2.

9.3 Final Perturbative Coefficients

The final results obtained with DISENT 0.1 are compiled in the tables 9.1–9.8. The first one contains the mean Q- and x-values to NLO for the eight bins of section 5.5 as well as the LO and NLO cross sections. The latter are differentiated by indices "1H" and "2H" to indicate whether the CH alone (1H, energy cut no. 6) or both hemispheres (2H, no energy cut) are allowed to contribute to the hadronic final state. To LO there is no difference because the only parton available has $E^\star = Q/2$ and runs into the CH by definition.

Tables 9.2–9.8 comprise the mean values of all event shapes for our standard pdf set MRSA'-115 and the perturbative coefficients $c_{1,F}$ and $c_{2,F}$ needed for the $\alpha_s(M_Z)$-dependent parameterization of the means.

$\langle Q \rangle$ / GeV	$\langle x \rangle$	σ_{LO}/pb	$\sigma_{\mathrm{NLO,1H}}/pb$	$\sigma_{\mathrm{NLO,2H}}/pb$
7.46	$4.58 \cdot 10^{-3}$	2519.9 ± 0.3	2232.5 ± 0.4	2402.6 ± 0.4
8.80	$7.21 \cdot 10^{-3}$	2357.5 ± 0.4	2111.9 ± 0.5	2245.8 ± 0.5
14.95	$1.60 \cdot 10^{-2}$	534.4 ± 0.1	484.4 ± 0.1	503.8 ± 0.1
17.73	$2.16 \cdot 10^{-2}$	584.7 ± 0.1	533.4 ± 0.1	551.1 ± 0.1
23.75	$3.16 \cdot 10^{-2}$	465.8 ± 0.1	428.4 ± 0.1	438.9 ± 0.1
36.69	$5.80 \cdot 10^{-2}$	156.2 ± 0.1	145.9 ± 0.1	147.9 ± 0.1
57.61	$1.09 \cdot 10^{-1}$	24.79 ± 0.02	23.63 ± 0.02	23.75 ± 0.02
80.76	$1.74 \cdot 10^{-1}$	7.502 ± 0.006	7.291 ± 0.006	7.305 ± 0.006

Table 9.1: Mean Q- and x-values as well as the LO and NLO cross sections as calculated by DISENT 0.1. The indices "1H" and "2H" differentiate between cross sections with and without the minimal energy cut in the CH.

$\langle Q \rangle$ / GeV	$\langle \tau_P \rangle_{\mathrm{LO}}$	$\langle \tau_P \rangle_{\mathrm{NLO}}$	c_{1,τ_P}	c_{2,τ_P}
7.46	0.2633 ± 0.0001	0.3377 ± 0.0014	1.4559 ± 0.0006	2.250 ± 0.041
8.80	0.2391 ± 0.0001	0.3070 ± 0.0012	1.3735 ± 0.0008	2.226 ± 0.038
14.95	0.1847 ± 0.0001	0.2387 ± 0.0009	1.1895 ± 0.0008	2.222 ± 0.038
17.73	0.1684 ± 0.0001	0.2173 ± 0.0008	1.1217 ± 0.0007	2.159 ± 0.037
23.75	0.1453 ± 0.0002	0.1871 ± 0.0013	1.0209 ± 0.0011	2.065 ± 0.062
36.69	0.1138 ± 0.0001	0.1480 ± 0.0007	0.8622 ± 0.0011	1.978 ± 0.041
57.61	0.0846 ± 0.0002	0.1113 ± 0.0008	0.6909 ± 0.0014	1.811 ± 0.055
80.76	0.0654 ± 0.0002	0.0864 ± 0.0006	0.5622 ± 0.0022	1.604 ± 0.043

Table 9.2: Mean values and perturbative coefficients for τ_P as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle B_P \rangle_{\mathrm{LO}}$	$\langle B_P \rangle_{\mathrm{NLO}}$	c_{1,B_P}	c_{2,B_P}
7.46	0.3923 ± 0.0001	0.2524 ± 0.0018	2.1696 ± 0.0007	-3.444 ± 0.049
8.80	0.3604 ± 0.0002	0.2510 ± 0.0018	2.0699 ± 0.0009	-2.894 ± 0.052
14.95	0.2879 ± 0.0001	0.2309 ± 0.0014	1.8545 ± 0.0009	-1.847 ± 0.051
17.73	0.2659 ± 0.0001	0.2215 ± 0.0012	1.7713 ± 0.0008	-1.513 ± 0.047
23.75	0.2350 ± 0.0002	0.2063 ± 0.0013	1.6515 ± 0.0013	-1.034 ± 0.056
36.69	0.1921 ± 0.0002	0.1795 ± 0.0009	1.4565 ± 0.0015	-0.437 ± 0.047
57.61	0.1521 ± 0.0002	0.1528 ± 0.0007	1.2431 ± 0.0021	0.245 ± 0.042
80.76	0.1254 ± 0.0002	0.1309 ± 0.0005	1.0798 ± 0.0023	0.571 ± 0.034

Table 9.3: Mean values and perturbative coefficients for B_P as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle \rho_E \rangle_{\mathrm{LO}}$	$\langle \rho_E \rangle_{\mathrm{NLO}}$	c_{1,ρ_E}	c_{2,ρ_E}
7.46	0.0464 ± 0.0001	0.0538 ± 0.0006	0.2564 ± 0.0002	0.243 ± 0.017
8.80	0.0437 ± 0.0001	0.0522 ± 0.0005	0.2509 ± 0.0002	0.293 ± 0.014
14.95	0.0367 ± 0.0001	0.0464 ± 0.0004	0.2363 ± 0.0002	0.405 ± 0.015
17.73	0.0347 ± 0.0001	0.0445 ± 0.0003	0.2309 ± 0.0003	0.434 ± 0.014
23.75	0.0316 ± 0.0001	0.0406 ± 0.0009	0.2222 ± 0.0004	0.448 ± 0.040
36.69	0.0274 ± 0.0001	0.0367 ± 0.0003	0.2076 ± 0.0004	0.537 ± 0.019
57.61	0.0233 ± 0.0001	0.0321 ± 0.0004	0.1907 ± 0.0007	0.589 ± 0.028
80.76	0.0204 ± 0.0001	0.0279 ± 0.0002	0.1761 ± 0.0009	0.569 ± 0.017

Table 9.4: Mean values and perturbative coefficients for ρ_E as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle \tau_C \rangle_{\mathrm{LO}}$	$\langle \tau_C \rangle_{\mathrm{NLO}}$	c_{1,τ_C}	c_{2,τ_C}
7.46	0.0701 ± 0.0001	0.0907 ± 0.0010	0.3876 ± 0.0003	0.620 ± 0.028
8.80	0.0664 ± 0.0001	0.0872 ± 0.0008	0.3816 ± 0.0004	0.675 ± 0.025
14.95	0.0569 ± 0.0001	0.0770 ± 0.0007	0.3666 ± 0.0004	0.814 ± 0.026
17.73	0.0542 ± 0.0001	0.0738 ± 0.0006	0.3607 ± 0.0004	0.854 ± 0.025
23.75	0.0500 ± 0.0001	0.0688 ± 0.0006	0.3515 ± 0.0006	0.911 ± 0.030
36.69	0.0443 ± 0.0001	0.0618 ± 0.0005	0.3357 ± 0.0006	0.999 ± 0.030
57.61	0.0389 ± 0.0001	0.0550 ± 0.0005	0.3179 ± 0.0007	1.077 ± 0.034
80.76	0.0348 ± 0.0001	0.0490 ± 0.0004	0.2991 ± 0.0009	1.073 ± 0.028

Table 9.5: Mean values and perturbative coefficients for τ_C as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle C_P \rangle_{\mathrm{LO}}$	$\langle C_P \rangle_{\mathrm{NLO}}$	c_{1,C_P}	c_{2,C_P}
7.46	0.2708 ± 0.0002	0.3470 ± 0.0035	1.4977 ± 0.0010	2.304 ± 0.096
8.80	0.2570 ± 0.0002	0.3325 ± 0.0031	1.4765 ± 0.0013	2.472 ± 0.091
14.95	0.2207 ± 0.0002	0.2945 ± 0.0024	1.4220 ± 0.0015	3.000 ± 0.094
17.73	0.2103 ± 0.0002	0.2826 ± 0.0021	1.4008 ± 0.0015	3.151 ± 0.088
23.75	0.1945 ± 0.0003	0.2638 ± 0.0024	1.3679 ± 0.0022	3.371 ± 0.112
36.69	0.1728 ± 0.0003	0.2374 ± 0.0019	1.3109 ± 0.0024	3.689 ± 0.107
57.61	0.1522 ± 0.0004	0.2126 ± 0.0019	1.2451 ± 0.0033	4.049 ± 0.124
80.76	0.1371 ± 0.0004	0.1909 ± 0.0013	1.1796 ± 0.0044	4.063 ± 0.095

Table 9.6: Mean values and perturbative coefficients for C_P as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle y_{fJ} \rangle_{\mathrm{LO}}$	$\langle y_{fJ} \rangle_{\mathrm{NLO}}$	$c_{1,y_{fJ}}$	$c_{2,y_{fJ}}$
7.46	0.1694 ± 0.0001	0.1575 ± 0.0007	0.9369 ± 0.0004	-0.197 ± 0.020
8.80	0.1553 ± 0.0001	0.1493 ± 0.0007	0.8918 ± 0.0005	-0.043 ± 0.021
14.95	0.1219 ± 0.0001	0.1253 ± 0.0006	0.7852 ± 0.0005	0.259 ± 0.021
17.73	0.1119 ± 0.0001	0.1172 ± 0.0005	0.7452 ± 0.0005	0.341 ± 0.020
23.75	0.0976 ± 0.0001	0.1054 ± 0.0005	0.6859 ± 0.0007	0.473 ± 0.024
36.69	0.0779 ± 0.0001	0.0875 ± 0.0004	0.5905 ± 0.0007	0.618 ± 0.019
57.61	0.0594 ± 0.0001	0.0697 ± 0.0003	0.4854 ± 0.0007	0.739 ± 0.018
80.76	0.0470 ± 0.0001	0.0564 ± 0.0002	0.4052 ± 0.0009	0.739 ± 0.014

Table 9.7: Mean values and perturbative coefficients for y_{fJ} as derived from DISENT 0.1.

$\langle Q \rangle$ / GeV	$\langle y_{k_t} \rangle_{\mathrm{LO}}$	$\langle y_{k_t} \rangle_{\mathrm{NLO}}$	$c_{1,y_{k_t}}$	$c_{2,y_{k_t}}$
7.46	0.2138 ± 0.0001	0.2942 ± 0.0004	1.1825 ± 0.0004	2.529 ± 0.016
8.80	0.1803 ± 0.0001	0.2366 ± 0.0003	1.0348 ± 0.0005	1.929 ± 0.016
14.95	0.1214 ± 0.0001	0.1514 ± 0.0002	0.7818 ± 0.0004	1.299 ± 0.013
17.73	0.1045 ± 0.0001	0.1288 ± 0.0002	0.6960 ± 0.0004	1.123 ± 0.011
23.75	0.0830 ± 0.0001	0.1008 ± 0.0002	0.5830 ± 0.0005	0.920 ± 0.013
36.69	0.0562 ± 0.0001	0.0676 ± 0.0001	0.4256 ± 0.0004	0.686 ± 0.009
57.61	0.0349 ± 0.0001	0.0423 ± 0.0001	0.2851 ± 0.0007	0.515 ± 0.008
80.76	0.0235 ± 0.0001	0.0287 ± 0.0001	0.2015 ± 0.0009	0.405 ± 0.005

Table 9.8: Mean values and perturbative coefficients for y_{k_t} as derived from DISENT 0.1.

Chapter 10

Power Corrections

To get a first impression of the extent to which pQCD is able to approximate the data, figs. 10.1 and 10.2 present the normalized differential distributions of the seven investigated event shapes, corrected to hadron level, in comparison with DISENT 0.1 NLO calculations. Note that the first data bin was left out in both normalizations. This is necessary since according to section 8.2 it is not possible to go down to $F = 0$ in $\mathcal{O}(\alpha_s^2)$. Starting with bin one, the NLO histogram could alternatively be normalized to $\sigma_{\rm NLO}$ yielding a slight mismatch by a factor of $\sigma_{\rm NLO}/\sigma_{\rm NNLO}$. Except for a small shift in height complicating a comparison of shapes, the conclusions would not change.

Obviously, there are large discrepancies between data and the NLO predictions. Common to all event shapes, however, is the clear tendency that the description of data improves considerably with rising Q; to what degree depends on the variable under consideration. y_{fJ} and y_{k_t} exhibit a very fast convergence hinting at small hadronization corrections, whereas the prediction for C_P approaches the data slowly.

Turning back to the mean values $\langle F \rangle$, we can, except for the y shapes, expect them to be substantially larger than predicted by pQCD with decreasing differences for increasing Q. Indeed, physical quantities receive, in addition to pQCD, non-perturbative contributions which basically behave power-like, i.e. $\propto 1/Q^p$, hence called *power corrections*. The cross section $e^+e^- \to$ hadrons, $\sigma_{\rm had}$, for example exhibits a $1/Q^4$-term which is also understandable on theoretical grounds [66]. Less inclusive variables like event shapes are not as well-behaved as $\sigma_{\rm had}$ and reveal $1/Q$-corrections that are sizable even at $Q = M_Z$ [67]. They can be made plausible in the framework of a longitudinal phase space or *tube model* [68], which essentially is the simplest version of the string fragmentation implemented in JETSET. One possibility to compare with pQCD is based on hadronization models like the one implemented in JETSET by unfolding the data in a third correction step to some kind of (**ambiguous**) parton level.

We do not follow this recipe. Instead, we will call upon explicit formulae trying to parameterize these effects. They are presented in sections 10.1 and 10.3 for the tube model and a new approach respectively. The main goal of this work is to test these models.

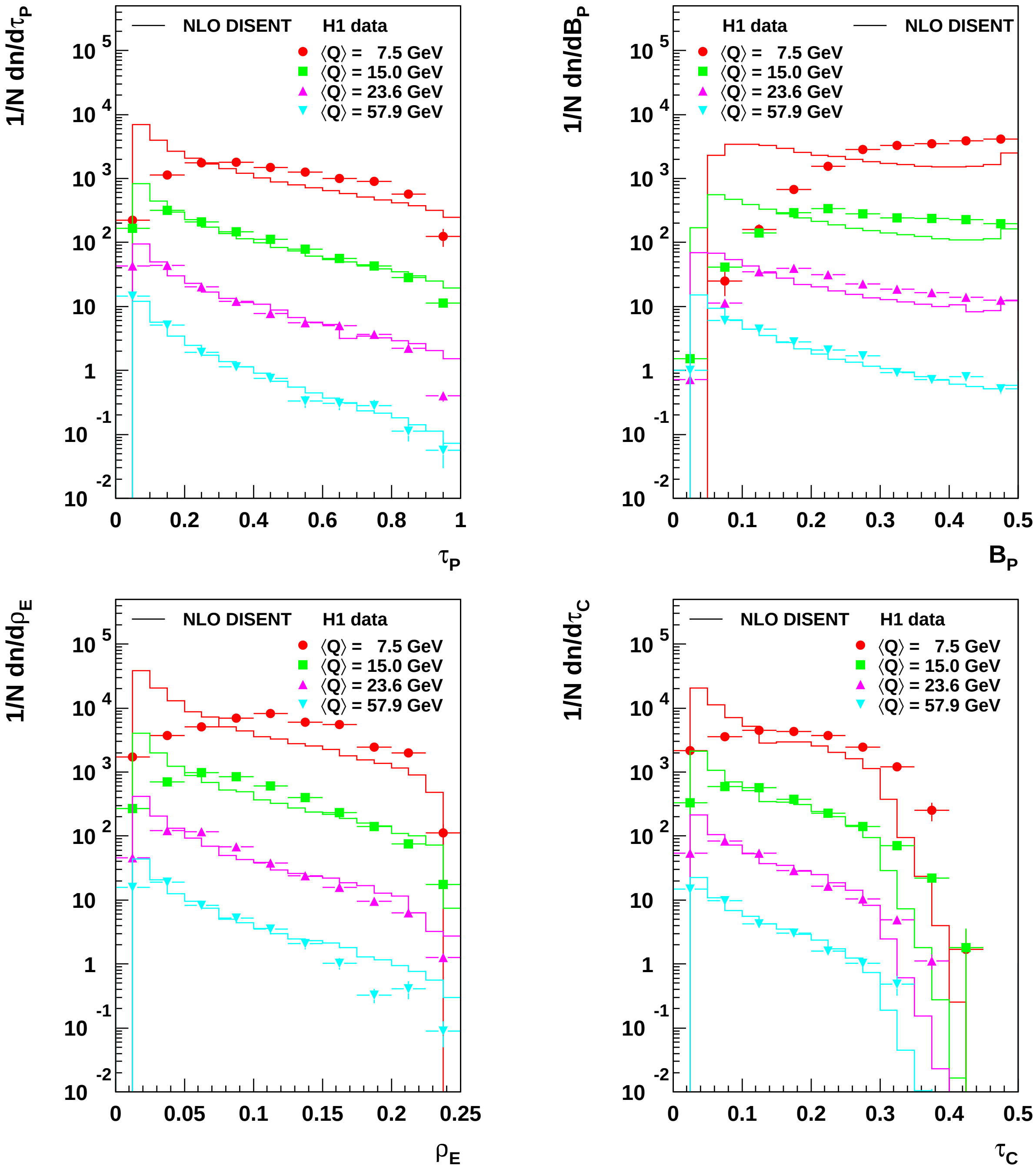


Figure 10.1: Normalized differential distributions of the event shapes τ_P, B_P, ρ_E and τ_C corrected to hadron level. H1 data (full symbols) are compared with DISENT 0.1 NLO calculations (full lines) for four out of eight investigated bins in Q. The spectra for $\langle Q\rangle = 7.5$–$57.9\,\mathrm{GeV}$ are multiplied by factors of 10^n, $n = 0, 1, 2, 3$. The error bars represent statistical uncertainties only. Note that the first data bin was left out in both normalizations.

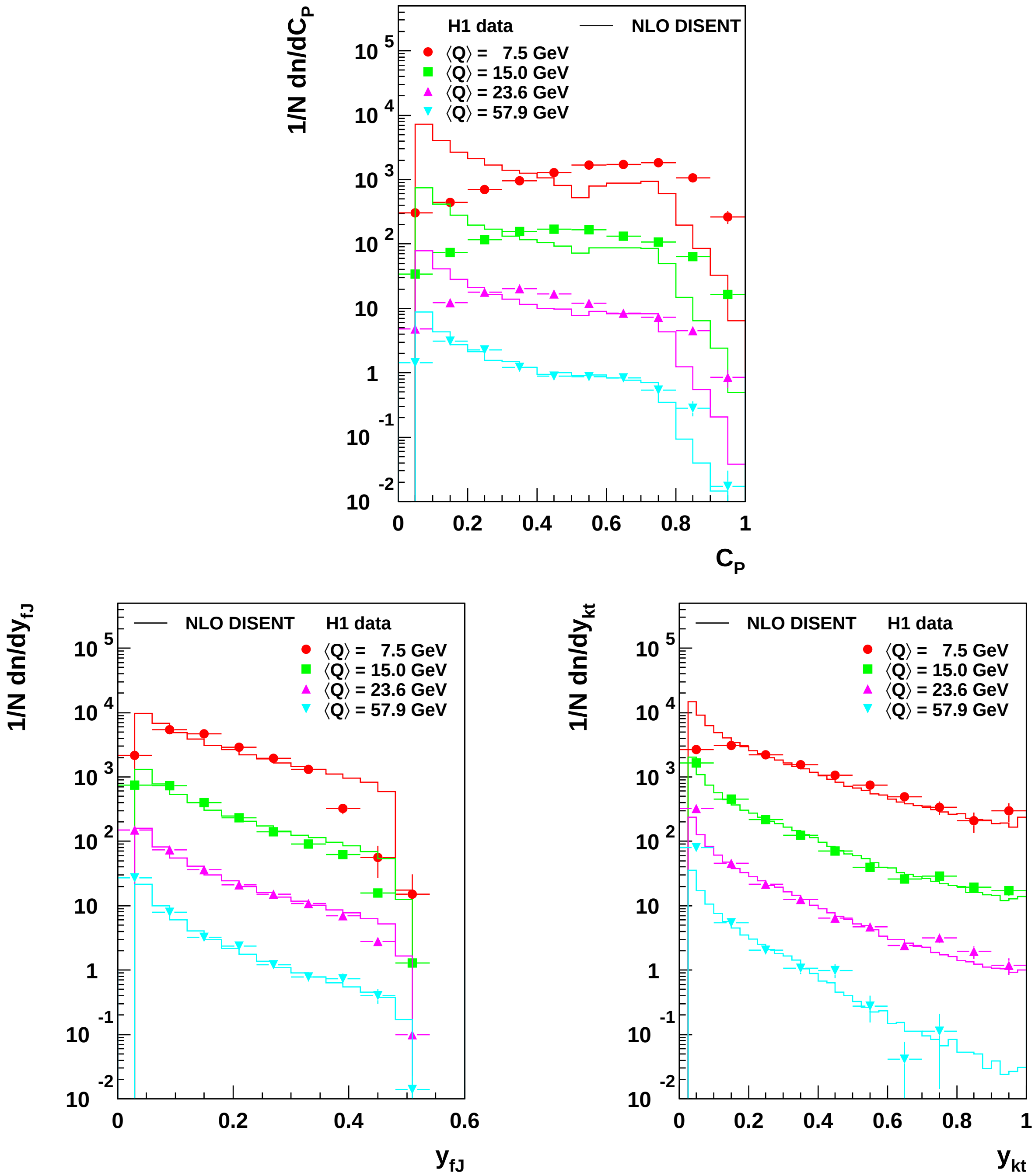


Figure 10.2: Normalized differential distributions of the event shapes C_P, y_{fJ} and y_{k_t} corrected to hadron level. H1 data (full symbols) are compared with DISENT 0.1 NLO calculations (full lines) for four out of eight investigated bins in Q. The spectra for $\langle Q \rangle$ = 7.5–57.9 GeV are multiplied by factors of 10^n, $n = 0, 1, 2, 3$. The error bars represent statistical uncertainties only. Note that the first data bin was left out in both normalizations.

10.1 The Tube Model

Measurements of the momentum spectra of secondary particles produced by a primary parton of energy $E_0 = Q/2$ approximately obey rather simple scaling laws [3]. Characterizing a particular secondary by E, p_t, p_l and m with respect to the direction of the original parton, the distribution in p_t, $\rho(p_t)$, decreases exponentially almost independently of E_0 and p_l, whereas the distribution in p_l rises proportional to E_0. Making use of the rapidity variable y_R (s. eq. (5.3)),

$$y_R := \frac{1}{2} \ln \frac{E + p_l}{E - p_l} = \ln \frac{E + p_l}{\sqrt{p_t^2 + m^2}}\,, \tag{10.1}$$

with the property $dy_R = dp_l/E$, this can be expressed in the form of

$$\Phi(y_R, p_t) = B \cdot \rho(p_t)\,. \tag{10.2}$$

B is the (constant) number of hadrons per unity in rapidity and Φ the distribution function of light hadrons in the produced jet corresponding to a tube in (y_R, p_t)-space. Following refs. [50, 69], the energy E_j and longitudinal momentum $P_{l,j}$ of a jet of length $y_{R,\max}$ in rapidity can be estimated by:

$$E_j = \int_0^{y_{R,\max}} \int_0^{\infty} E \cdot B\rho(p_t) dp_t dy_R \approx \lambda \sinh y_{R,\max} \approx p_{l,\max} \approx E_0\,, \tag{10.3}$$

$$P_{l,j} = \int_0^{y_{R,\max}} \int_0^{\infty} p_{l,j} \cdot B\rho(p_t) dp_t dy_R \approx \lambda(\cosh y_{R,\max} - 1) \approx E_0 - \lambda\,. \tag{10.4}$$

Recall that $E = m_t \cosh y_R$ and $p_l = m_t \sinh y_R$ with the transverse mass

$$m_t := \sqrt{p_t^2 + m^2} \approx p_t \tag{10.5}$$

for light hadrons. λ is proportional to the average transverse momentum:

$$\lambda := B \int_0^{\infty} p_t \rho(p_t) dp_t\,. \tag{10.6}$$

The power corrections to mean $\tau = 1-$thrust and jet mass now become

$$\langle\tau\rangle^{\text{pow}} = 1 - \tfrac{P_{l,j}}{E_j} \approx 1 - \tfrac{E_0 - \lambda}{E_0} = \frac{2\lambda}{Q}\,, \tag{10.7}$$

$$\langle\rho\rangle^{\text{pow}} \approx \tfrac{E_j^2 - P_{l,j}^2}{Q^2} \approx \tfrac{E_0^2 - (E_0 - \lambda)^2}{Q^2} \approx \frac{\lambda}{Q}\,. \tag{10.8}$$

With $B \approx 2$ and $\langle p_t \rangle \approx 0.3\,\text{GeV}$ [69], we can estimate $\lambda \approx 0.6\,\text{GeV}$. Testing this simple ansatz with event shapes in ep DIS is the topic of the next section.

10.2 $1/Q^p$ -Fits

The input for all fits performed here and in section 10.4 consists of the corrected data means of table 7.1 and the perturbative coefficients given in tables 9.2–9.8 together with their statistical uncertainties. Systematic uncertainties are not taken into account; they will be analyzed in section 10.5.

The standard fit procedure employs the method of least squares (s. e.g. [70]) to determine the best estimates for the parameters of the supplied model. Technically, this is accomplished by using the program MINUIT [71] which additionally provides the χ^2-value and, where appropriate, the maximally encountered correlation κ. In the tables collecting the results, the reduced χ^2-values, i.e. χ^2 divided by the number of **d**egrees **o**f **f**reedom (dof), are entered. For one- and two-parameter fits we have seven respectively six dofs yielding a maximal $\chi^2/\mathrm{dof} \approx 2$ at a confidence level of 5%.

The first test carried out is to allow a variation of $\alpha_s(M_Z)$ in the perturbative expression

$$\langle F\rangle^{\mathrm{pert}} = c_{1,F}(x)\,\alpha_s(\mu_R) + \left(c_{2,F}(x) + \frac{\beta_0}{2\pi}\ln\frac{\mu_R}{Q}c_{1,F}(x)\right)\alpha_s^2(\mu_R) \tag{10.9}$$

only, where, for the time being, μ_R is identified with Q. As can be seen from table 10.1, this works surprisingly well except for B_P. Seemingly, the decrease of $\langle F\rangle$ with Q can be accounted for by the logarithmic Q-dependence of α_s as long as $\alpha_s(M_Z)$ may be adapted to the corresponding hadronization correction. Compared to the world average $\alpha_s(M_Z) = 0.119$ [5], we obtain high results for τ_P and B_P, very high ones for ρ_E, τ_C, C_P and lower values for y_{fJ}, y_{k_t}. Taking this as an indication for the impact of hadronization, it can be concluded that our event shapes are affected by medium, large and small (**negative**) corrections respectively.

Sometimes, the ratio of NLO to LO results, the K-factor, is considered to signal big effects of higher orders. However, although evaluating them for τ_P and ρ_E in Q-bin one to be $K_{\tau_P} = 1.28$ and $K_{\rho_E} = 1.16$, we get ratios of the data means with respect to NLO of 1.30 and 2.07 calling this argument into question.

Writing

$$\langle F\rangle = \langle F\rangle^{\mathrm{pert}} + \langle F\rangle^{\mathrm{pow}} \tag{10.10}$$

according to [6], we extend our fit to include terms of the form

$$\langle F\rangle^{\mathrm{pow}} = \frac{\lambda}{Q} \tag{10.11}$$

and

$$\langle F\rangle^{\mathrm{pow}} = \frac{\mu}{Q^2}\,. \tag{10.12}$$

At first, $\alpha_s(M_Z)$ will be kept fixed to 0.119 yielding estimates for λ and μ as given in table 10.1. With the exception of B_P, the χ^2-values worsen for both formulae, even dramatically in the case of μ/Q^2. Simple $1/Q^2$-corrections are therefore ruled out, but notably y_{k_t} exhibits the best behaviour of all. Also, pure $1/Q$-terms do not work very well, yet, at least the order of magnitude for $\lambda \approx 0.5$–$1.0\,\mathrm{GeV}$ is confirmed for τ_P and ρ_E,

$\langle F \rangle$	$\alpha_s(M_Z)$	χ^2/dof	$\lambda/\,\mathrm{GeV}$	χ^2/dof	$\mu/\,\mathrm{GeV}^2$	χ^2/dof
$\langle \tau_P \rangle$	0.1311 ± 0.0005	0.5	0.705 ± 0.028	1.4	7.17 ± 0.32	22.8
$\langle B_P \rangle$	0.1341 ± 0.0005	24.0	0.549 ± 0.016	0.7	5.22 ± 0.16	28.0
$\langle \rho_E \rangle$	0.1601 ± 0.0006	0.6	0.539 ± 0.007	38.4	5.17 ± 0.08	282.5
$\langle \tau_C \rangle$	0.1520 ± 0.0005	1.5	0.734 ± 0.011	29.2	7.12 ± 0.13	212.2
$\langle C_P \rangle$	0.1481 ± 0.0004	5.4	2.325 ± 0.032	66.2	20.82 ± 0.36	326.6
$\langle y_{fJ} \rangle$	0.1110 ± 0.0005	1.7	-0.185 ± 0.012	2.0	-1.74 ± 0.13	10.5
$\langle y_{k_t} \rangle$	0.1082 ± 0.0007	1.8	-0.360 ± 0.021	1.4	-4.21 ± 0.27	9.3

Table 10.1: Results of one-parameter fits according to eqs. (10.9)–(10.12) for the event shape means. Uncertainties are statistical only.

although a relative factor of two is not found between them. According to our previous expectations, C_P receives a larger λ, whereas for y_{fJ} and y_{k_t} negative values are assumed.

In summary, this suggests that logarithmic as well as power-like Q-dependences play a role. The idea to fit both, $\alpha_s(M_Z)$ and λ, however, does not lead to the desired outcome as shown in table 10.2. Due to the extreme anti-correlation found between them, there is a tendency to minimize the power contribution at the cost of a very high $\alpha_s(M_Z)$. Merely B_P and the y event shapes produce correlated but reasonable numbers. Triggered by the x dependence of $\langle F \rangle^{\mathrm{pert}}$, one could assume that $\langle F \rangle^{\mathrm{pow}}$ should be a function of x as well. Yet, for corrections with power $p = 1$ this is not expected [72]. For an improvement a more sophisticated model is needed.

At last, an explicit check on the power p was performed by fitting

$$\langle F \rangle^{\mathrm{pow}} = \frac{\nu}{Q^p} \tag{10.13}$$

keeping again $\alpha_s(M_Z)$ fixed. The results for p lie between 0.5 and 1.0, but once more ν and p are strongly correlated.

10.3 The Model of Dokshitzer, Webber et al.

For shorthand notation the model introduced in the following is labelled after Dokshitzer and Webber, the authors of ref. [6], where the essential formula (10.15) needed here is first presented. Note that besides these initiating authors many others contributed to the currently very active field of power corrections in QCD, s. e.g. [73] and refs. therein.

As demonstrated above, pQCD applied in the form of NLO integration programs is not sufficient to describe our data. Ignoring the fact that the calculation of still higher orders will become forbidding due to the tremendous amount of Feynman diagrams to compute, it may not even be a good idea, since the QCD perturbation series is not necessarily expected to converge! One known source of singularities are diagrams with a chain of fermion loops

$\langle F\rangle$	$\alpha_s(M_Z)$	$\lambda/$ GeV	χ^2/dof	$\kappa/\%$	$\alpha_s(M_Z)$	$\mu/$ GeV^2	χ^2/dof	$\kappa/\%$
$\langle\tau_P\rangle$	0.1318	-0.044	0.6	-99	0.1314	-0.23	0.6	-89
$\langle B_P\rangle$	0.1195	0.532	0.8	-92	0.1279	3.07	1.7	-68
$\langle\rho_E\rangle$	0.1652	-0.089	0.5	-99	0.1614	-0.27	0.4	-87
$\langle\tau_C\rangle$	0.1573	-0.153	1.3	-99	0.1535	-0.57	1.1	-86
$\langle C_P\rangle$	0.1609	-1.393	2.0	-99	0.1519	-4.84	1.0	-85
$\langle y_{fJ}\rangle$	0.1142	-0.076	1.6	-97	0.1119	-0.28	1.7	-82
$\langle y_{k_t}\rangle$	0.1216	-0.453	1.6	-99	0.1050	1.33	1.6	-93

Table 10.2: Results of two-parameter fits according to eqs. (10.9)–(10.12) for the event shape means.

inserted into a gluon line. These *renormalon* chains lead to a factorial divergence of the coefficients in the series expansion (8.19). At first sight, we are at a loss here. But on the contrary, one can even gain more insight into non-perturbative effects of QCD by studying these divergences [66, 73]. In a manner of speaking, one approaches the non-perturbative regime from the perturbative side. Employing this renormalon approach [74], one ends up with ambiguities of the series expansion of the general form $(\Lambda/\mu_R)^p$ modulo logarithms $\ln^q(\mu_R/\Lambda)$, yet the exact result for an observable has to be single-valued, of course. Hence, power-like terms have to be added to regain definiteness.

Practically, this scheme can be exploited to give the leading corrections in terms of powers p and q for an observable, but the absolute normalization is unknown. Introducing the notion of a universal infrared-finite effective coupling $\alpha_{\text{eff}}(\mu_R)$ responsible for soft gluon emissions, the coefficients for different observables to the same power p can be related at the expense of a new but universal non-perturbative parameter $\overline{\alpha}_{p-1}(\mu_I)$. Above an *infrared matching* scale μ_I with

$$\Lambda \ll \mu_I \approx 2\,\text{GeV} \ll Q\,, \tag{10.14}$$

this effective coupling has to coincide with the normal perturbative one $\alpha_s(\mu_R)$. For event shapes in e^+e^- annihilation, this concept was first applied in [6] and extended in [24]:

$$\langle F\rangle^{\text{pow}} = \frac{\hat{a}_F}{p}\left(\frac{\mu_I}{\mu_R}\right)^p \ln^q\left(\frac{\mu_R}{\mu_I}\right)\left[\overline{\alpha}_{p-1}(\mu_I) - \alpha_s(\mu_R) - \frac{\beta_0}{2\pi}\left(\ln\frac{\mu_R}{\mu_I} + \frac{K}{\beta_0} + \frac{1}{p}\right)\alpha_s^2(\mu_R)\right]. \tag{10.15}$$

Here, $\hat{a}_F$ is an F-dependent but calculable coefficient and

$$K = \frac{67}{6} - \frac{\pi^2}{2} - \frac{5}{9}N_f \tag{10.16}$$

accounts for differences between the $\overline{\text{MS}}$ and the CMW renormalization scheme [75, 76]:

$$\alpha_{s,\text{CMW}} = \alpha_{s,\overline{\text{MS}}} + \frac{K}{2\pi}\alpha^2_{s,\overline{\text{MS}}}\,. \tag{10.17}$$

The subtractions proportional to α_s and α_s^2 should be evaluated for the same number of flavours, here $N_f = 5$, as in the perturbative part eq. (10.9) and serve to avoid double counting.

In terms of the effective coupling, $\overline{\alpha}_{p-1}(\mu_I)$ can be written as the moment

$$\overline{\alpha}_{p-1}(\mu_I) := \frac{p}{\mu_I^p} \int_0^{\mu_I} k^{p-1} \alpha_{\text{eff}}(k) dk \,, \tag{10.18}$$

i.e. $\overline{\alpha}_0(\mu_I)$ essentially corresponds to an average effective coupling at low scales and is deduced from e^+e^- data on thrust to be ≈ 0.5 for $\mu_I = 2\,\text{GeV}$ [6].

For our event shapes, $p = 1$ except for y_{k_t} where $p = 2$ is expected [24, 21]. In that case, other mechanisms can give contributions to the same order $1/Q^2$, preventing a straightforward calculation of $\hat{a}_{y_{k_t}}$. The other coefficients can be gathered from [24, 21] when performing the transition

$$\hat{a}_F \to \frac{16}{3\pi} a_F \tag{10.19}$$

to normalize to the $\hat{a}_F$-value for τ in e^+e^- physics [6].

In the form of the *dispersive* approach [77], this technique was applied to event shape distributions [78] and other subjects as well [79, 80]. In the meantime, however, it was discovered that some ambiguities unfortunately remain in deriving a_F-values. Refining the investigations to two-loop level [81, 82], the problem could be resolved and universality was reconstituted provided an additional factor, the *Milan factor*

$$\mathcal{M} = 1 + \frac{7.311 - 0.052 N_f}{\beta_0} \tag{10.20}$$

is applied. Recently, it was confirmed to hold also for DIS event shapes [83]. Due to a mismatch in definitions compared to (10.15), however, we have to multiply by $2/\pi \cdot \mathcal{M} \approx 1.143$ where, in general, $N_f = 3$ should be used [84].

Putting everything together, we get for $\langle F \rangle^{\text{pow}}$

$$\langle F \rangle^{\text{pow}} = a_F \frac{32}{3\pi^2} \frac{\mathcal{M}}{p} \left(\frac{\mu_I}{\mu_R} \right)^p \left[\overline{\alpha}_{p-1}(\mu_I) - \alpha_s(\mu_R) - \frac{\beta_0}{2\pi} \left(\ln \frac{\mu_R}{\mu_I} + \frac{K}{\beta_0} + \frac{1}{p} \right) \alpha_s^2(\mu_R) \right] \tag{10.21}$$

with the coefficients given in table 10.3.

Note that there is no logarithm with power q any more. Such a term with $q = 1$ was predicted for B_P only. In the previous publications [25, 23], this could not be supported and the theoretical calculations for it have been reexamined [76], leading to the much more complicated factor

$$a'_{B_P} = \frac{\pi}{2\sqrt{\frac{8}{3}\alpha_{s,\text{CMW}}(e^{-3/4}Q)}} + \frac{3}{4} - \frac{\beta_0}{16} - 0.6137 + \mathcal{O}(1) \,. \tag{10.22}$$

$\langle F\rangle$	a_F	p	$\overline{\alpha}_{p-1}(\mu_I = 2\,\mathrm{GeV})$	$\alpha_s(M_Z)$	χ^2/dof	$\kappa/\%$
$\langle\tau_P\rangle$	1	1	$0.480 \pm 0.028\ ^{+0.048}_{-0.062}$	$0.1174 \pm 0.0030\ ^{+0.0097}_{-0.0081}$	0.5	−97
$\langle B_P\rangle$	$1/2 \cdot a'_{B_P}$	1	$0.491 \pm 0.005\ ^{+0.032}_{-0.036}$	$0.1106 \pm 0.0012\ ^{+0.0060}_{-0.0057}$	0.7	−58
$\langle\rho_E\rangle$	1/2	1	$0.561 \pm 0.004\ ^{+0.051}_{-0.058}$	$0.1347 \pm 0.0015\ ^{+0.0111}_{-0.0100}$	1.2	+7
$\langle\tau_C\rangle$	1	1	$0.475 \pm 0.003\ ^{+0.044}_{-0.048}$	$0.1284 \pm 0.0014\ ^{+0.0100}_{-0.0092}$	1.3	+19
$\langle C_P\rangle$	$3\pi/2$	1	$0.425 \pm 0.002\ ^{+0.033}_{-0.039}$	$0.1273 \pm 0.0009\ ^{+0.0104}_{-0.0093}$	0.9	+63
$\langle B_P\rangle$	1/2	1	$0.602 \pm 0.014\ ^{+0.029}_{-0.042}$	$0.1164 \pm 0.0011\ ^{+0.0055}_{-0.0049}$	0.7	−84
$\langle y_{fJ}\rangle$	1	1	0.258 ± 0.004	0.1044 ± 0.0017	1.9	−61
$\langle y_{fJ}\rangle$	$-1/4^*$	1	0.460 ± 0.079	0.1183 ± 0.0027	1.5	+99
$\langle y_{k_t}\rangle$	$1^\dagger$	1	0.431 ± 0.034	0.0462 ± 0.0129	2.6	−99
$\langle y_{k_t}\rangle$	$1^\dagger$	2	0.586 ± 0.182	0.1047 ± 0.0019	1.6	+93
$\langle y_{k_t}\rangle$	-30^*	2	0.326 ± 0.034	0.1206 ± 0.0040	0.3	+99

Table 10.3: Results of two-parameter fits à la Dokshitzer, Webber et al., eqs. (10.9) and (10.21), for the event shape means. In case of B_P, the outcome without the additional factor a'_{B_P}, eq. 10.22, is given as well. The starred coefficients of the y variables are derived from a **circular** fit procedure described in section 10.4.3, whereas the coefficients marked with †'s are guesstimated. Where reasonable, statistical **and** total systematic uncertainties are presented.

10.4 Fits à la Dokshitzer, Webber et al.

10.4.1 Two-Parameter Fits

Two-parameter fits according to eqs. (10.9) and (10.21) are performed for each event shape separately. The results are compiled in table 10.3, where for comparison the fit of $\langle B_P\rangle$ without the new factor a'_{B_P} is given as well. For a discussion of the quoted total systematic uncertainties s. section 10.5.

Omitting at first the y variables, one immediately recognizes reduced correlations as compared to the tube model — with the exception of τ_P — and reasonable χ^2/dof-values. The five obtained $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$'s scatter around the expectation of ≈ 0.5 by about 20% and within errors are compatible with it. The uncertainties are dominated by systematics of 5%–10% being at least twice as large as the statistical ones.

Turning to $\alpha_s(M_Z)$, one observes a two-fold ambiguity. The event shapes employing $z^\star$ as event axis prefer low couplings, whereas the other three lead to $\alpha_s(M_Z) \approx 0.130$. Considering the errors of 1%–3% statistically and 5%–9% systematically, they are consistent with each other. Yet, it does evoke some suspicion.

Figs. 10.3 and 10.4 show the fit results in comparison with the corrected data as well as the LO and NLO predictions of DISENT 0.1 with respect to each corresponding $\alpha_s(M_Z)$. The NLO band represents a variation of the renormalization scale of $1/2 \leq \mu_R^2/\,\mathrm{GeV}^2 \leq 2$,

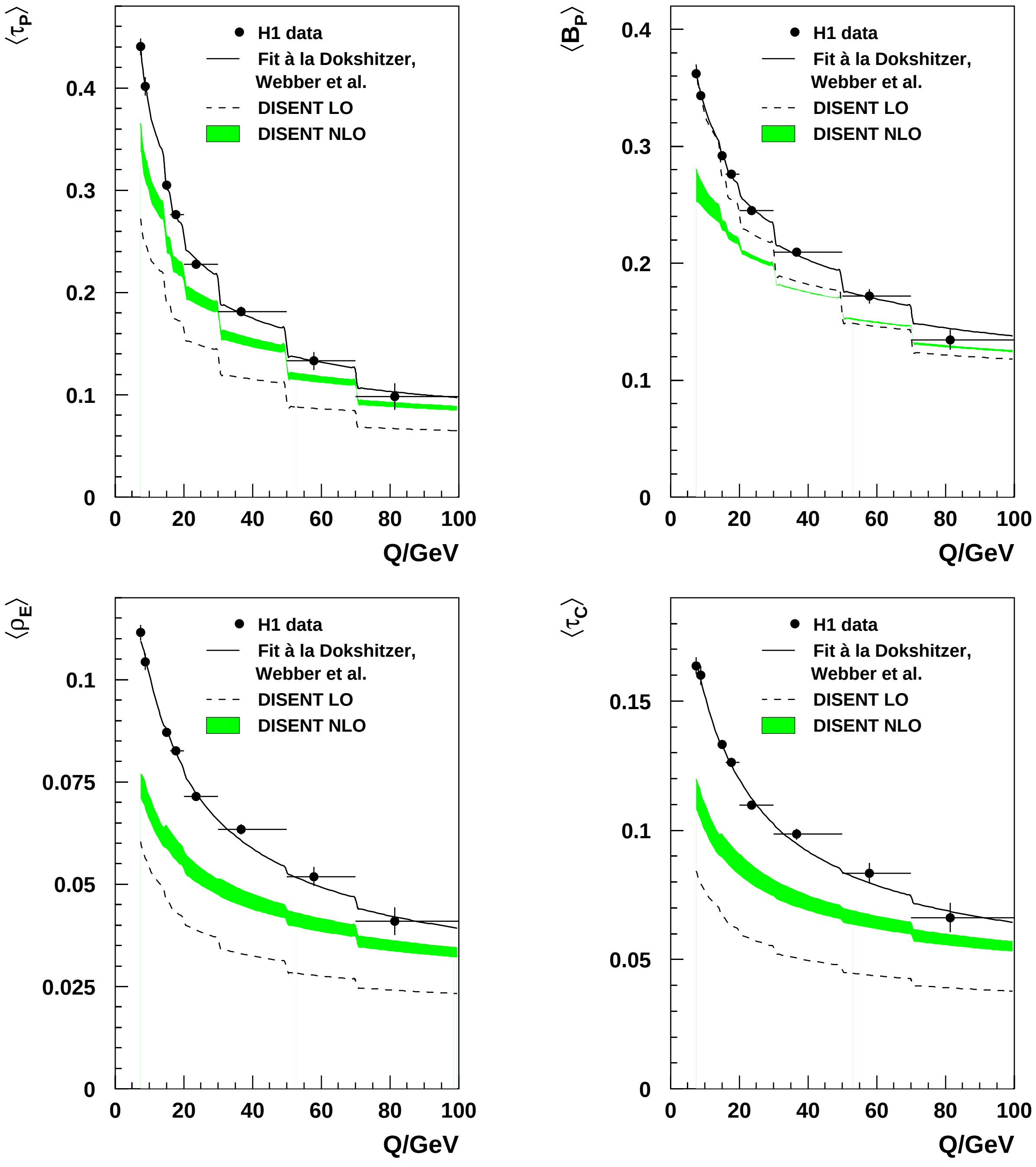


Figure 10.3: Corrected means (full symbols) of τ_P, B_P, ρ_E and τ_C as a function of Q. The error bars represent statistical uncertainties only. The full line corresponds to a power correction fit according to eqs. (10.9) and (10.21). For comparison the LO (dashed) and NLO (band) predictions of DISENT 0.1 are shown. The band reflects a variation of the renormalization scale of $1/2 \leq \mu_R^2/\,\mathrm{GeV}^2 \leq 2$.

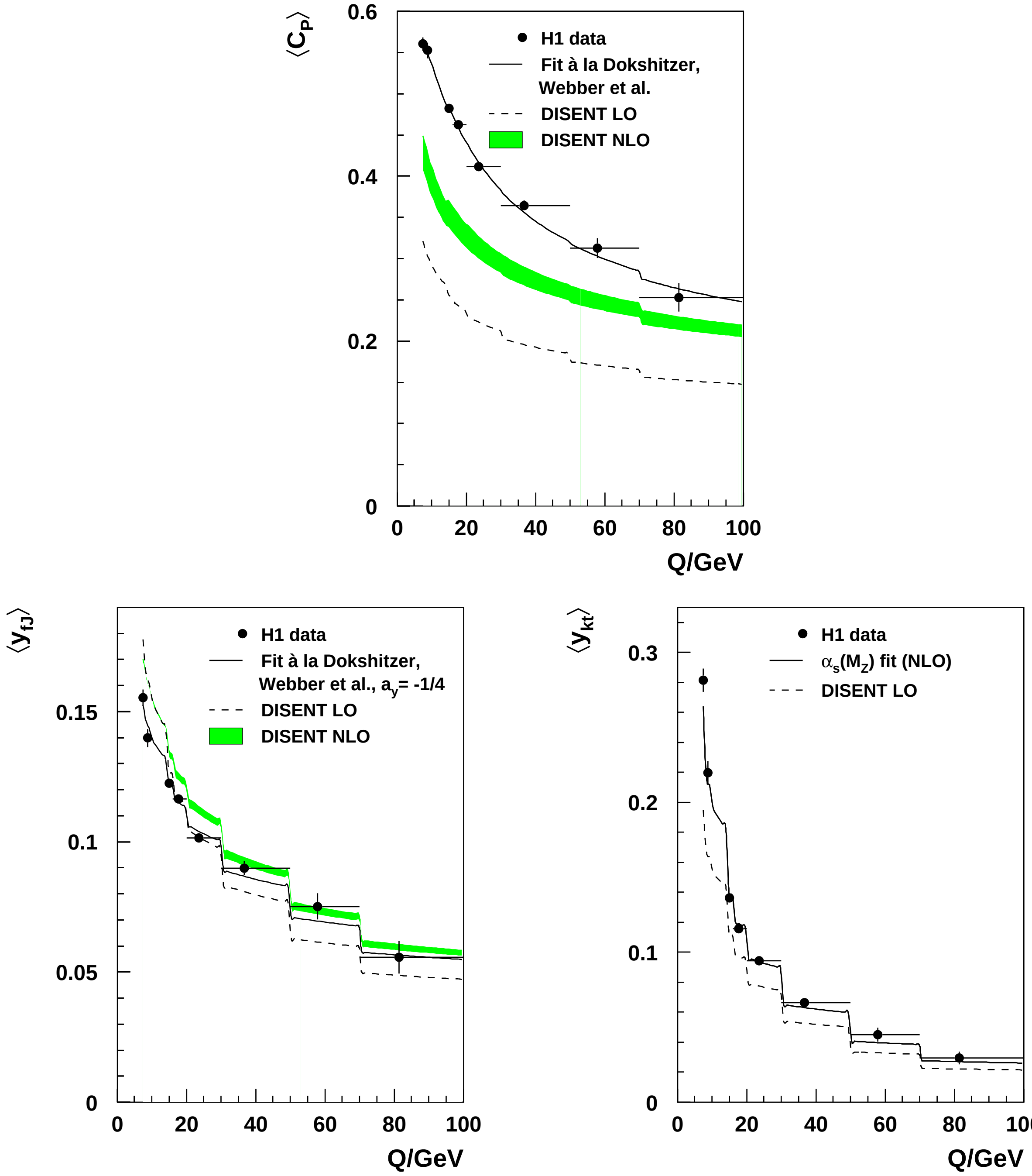


Figure 10.4: Corrected means (full symbols) of C_P, y_{fJ} and y_{k_t} as a function of Q. The error bars represent statistical uncertainties only. The full line corresponds to a power correction fit according to eqs. (10.9) and (10.21) except for y_{k_t} where no power contribution was assumed. For comparison the LO (dashed) and, if appropriate, NLO (band) predictions of DISENT 0.1 are shown. The band reflects a variation of the renormalization scale of $1/2 \leq \mu_R^2/\,\mathrm{GeV}^2 \leq 2$.

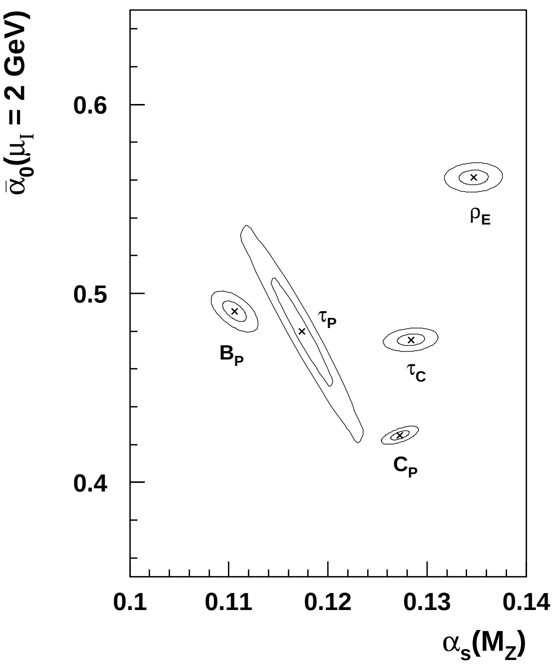


Figure 10.5: Results of the two-parameter fits à la Dokshitzer, Webber et al. for the means of τ_P, B_P, ρ_E, τ_C and C_P in the form of 1σ- and 2σ-contours in the $(\alpha_s, \overline{\alpha}_0)$-plane for statistical uncertainties only.

s. section 10.5 for details. Each fit taken separately, the data are nicely described. Putting all five in terms of $(\alpha_s, \overline{\alpha}_0)$-pairs into the one fig. 10.5, the discussed discrepancies become obvious. But recall again that systematic uncertainties are treated individually and therefore only statistical uncertainties are included!

With y_{fJ}, one obtains a good fit but rather low numbers for $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ and $\alpha_s(M_Z)$, leading to the conclusion that the coefficient $a_{y_{fJ}}$ given in the proceedings [24] may not be appropriate. Referring back to section 10.2, y_{fJ} does exhibit smaller hadronization corrections than thrust. The curve in fig. 10.4 was produced with $a_{y_{fJ}} = -1/4$ derived in section 10.4.3.

Concerning y_{k_t}, the coefficient $a_{y_{k_t}}$ is not known and the $1^\dagger$ in table 10.3 is pure guesswork. The fit for $p = 1$ works not very well supporting the conjecture of a $1/Q^2$-term instead of $1/Q$. In fact, both results obtained with $p = 2$ look more reasonable. Nevertheless, fig. 10.4 contains the outcome according to eq. (10.9) without any power correction.

10.4.2 Combined Fits

Based on QCD, the universality of α_s is on firm theoretical grounds. Not so, however, $\overline{\alpha}_0$ since here additional assumptions are made. It was therefore checked if fits allowing for different $\overline{\alpha}_0$'s are possible while enforcing one $\alpha_s(M_Z)$. Correlations between the

Comb. fits:	τ_P, B_P	ρ_E, τ_C, C_P	τ_P, ρ_E, τ_C, C_P	τ_P, B_P, ρ_E, τ_C, C_P
$\overline{\alpha}_0(\tau_P)$	0.532 ± 0.012	—	0.374 ± 0.010	0.414 ± 0.009
$\overline{\alpha}_0(B_P)$	0.488 ± 0.005	—	—	0.443 ± 0.006
$\overline{\alpha}_0(\rho_E)$	—	0.561 ± 0.004	0.561 ± 0.004	0.562 ± 0.004
$\overline{\alpha}_0(\tau_C)$	—	0.476 ± 0.003	0.476 ± 0.003	0.474 ± 0.003
$\overline{\alpha}_0(C_P)$	—	0.428 ± 0.002	0.427 ± 0.002	0.421 ± 0.002
$\alpha_s(M_Z)$	0.1115 ± 0.0011	0.1291 ± 0.0007	0.1285 ± 0.0007	0.1245 ± 0.0006
χ^2/dof	$11.6/13$	$38.4/20$	$55.5/27$	$222/34$

Table 10.4: Results of combined fits à la Dokshitzer, Webber et al. allowing for a separate $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ for each event shape, but enforcing one $\alpha_s(M_Z)$. Several combinations are tried; uncertainties are statistical only.

observables were ignored for that purpose. According to the systematic behaviour found previously concerning the $\alpha_s(M_Z)$-results, several combinations are tried. The outcome is collected in table 10.4.

The first two arrangements reflect the separation in τ_P, B_P on the one hand, and ρ_E, τ_C, C_P on the other. As expected, the numbers show resemblance to what we have obtained before. Yet, including step by step first τ_P and then B_P into the ρ_E, τ_C, C_P triple, the χ^2/dof-values worsen considerably.[1] Both times the fit results are dominated by the latter, which in addition to be more numerous, also have the smaller statistical uncertainties. $\alpha_s(M_Z)$ is forced to remain high, and thereby $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ for τ_P and B_P is shifted.

One concludes that combined fits with one $\alpha_s(M_Z)$ are possible, but notwithstanding the freedom to choose an individual $\overline{\alpha}_0$ for each event shape, they suffer from systematic effects. In order to determine $\alpha_s(M_Z)$, it would be interesting to know what their origin is and what biases are implied.

10.4.3 a_F-Fits

Motivated by the problems encountered with the a_F-coefficients for the y event shapes, one can try to do a three-parameter fit of $\overline{\alpha}_0$, $\alpha_s(M_Z)$ and a_F. However, this set of parameters is strongly correlated and with eight data points only it does not properly converge for τ_P, B_P, ρ_E, τ_C or C_P. The y variables are somewhat of an exception here, giving $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.58 \pm 0.20$, $\alpha_s(M_Z) = 0.134 \pm 0.015$ and $a_{y_{fJ}} = -0.71 \pm 0.18$ at $\chi^2/\mathrm{dof} = 1.6$ for y_{fJ}, and $\overline{\alpha}_1(\mu_I = 2\,\mathrm{GeV}) = 0.35 \pm 0.05$, $\alpha_s(M_Z) = 0.123 \pm 0.006$ and $a_{y_{k_t}} = -33.2 \pm 3.0$ at $\chi^2/\mathrm{dof} = 0.23$ for y_{k_t} respectively. Despite looking quite reasonable, one nevertheless has to exercise caution because of the very strong correlations.

[1] Note that with increasing dofs from 6 for the two-parameter fits to 34 in the last combination, the limit for the 5% confidence level falls below 1.5.

F	a_F	χ^2/dof	F	a_F	χ^2/dof	F	a_F	χ^2/dof
τ_P	0.84 ± 0.03	0.4	ρ_E	0.65 ± 0.01	11.2	y_{fJ}	-0.23 ± 0.01	1.3
B_P	0.41 ± 0.01	6.0	τ_C	0.88 ± 0.01	9.0	y_{k_t}	-0.41 ± 0.02	1.8
			C_P	2.85 ± 0.04	27.7			

Table 10.5: Results for fits of the coefficients a_F while presupposing the validity of eq. 10.21 with $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.5$ and $\alpha_s(M_Z) = 0.119$. Uncertainties are statistical only.

Being constrained that way to less than three parameters, we start from the other side by deriving a_F while presupposing the validity of eq. 10.21 with $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.5$ and $\alpha_s(M_Z) = 0.119$. Table 10.5 compiles the determined coefficients, which, although some rather large χ^2/dof's occur, are of a similar magnitude as the theoretical ones. Notably, y_{fJ} gives $a_F \approx -1/4$ instead of 1!

As a consistency check, this new value was reinserted to produce the two-parameter fit in fig. 10.4 with $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.46$ and $\alpha_s(M_Z) = 0.118$, s. again table 10.3. Performing the same, i.e. $p = 1$, circular procedure with y_{k_t}, one gets $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.34$ and $\alpha_s(M_Z) = 0.112$ which looks much better than the entry in table 10.3, but still is not satisfactory. Since $\overline{\alpha}_1$ is basically unknown, we must determine it in addition to a_F in case of $p = 2$. Refitting then with $a_F = -30$, one obtains $\overline{\alpha}_1(\mu_I = 2\,\mathrm{GeV}) = 0.33$ and $\alpha_s(M_Z) = 0.121$, s. table 10.3. Both refits, however, exhibited strong correlations between $\overline{\alpha}_0$ respectively $\overline{\alpha}_1$ and $\alpha_s(M_Z)$.

In a last test, all coefficients were multiplied by factors of 2 respectively 1/2 to study the effect on the $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$, $\alpha_s(M_Z)$ results. Having fig. 10.5 in mind, the outcome corresponds approximately to shifts along the main diagonal with large a_F's inducing smaller $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$'s and $\alpha_s(M_Z)$'s.

10.5 Evaluation of Systematic Uncertainties

The general procedure followed here is to repeat the fits for a variation in every prominent origin of systematic effects. The obtained discrepancy compared to the standard result is attributed to a corresponding uncertainty. In case of two deviations in one direction for the same primary source, e.g. an up- and downwards modification of an energy scale, only the larger one is considered for the evaluation of the total uncertainty. The latter is derived from all contributions by adding them up quadratically. An exception is the unfolding, whose influence is estimated in the same way as explained in section 7.5 on the final results of the data means.

Altogether, the following studies have been performed to estimate systematic effects:

- Experimental uncertainties:
    1. Usage of four unfolding procedures
    2. Variation of the electromagnetic energy scale of the calorimeters by ±1%, 2%, 3% depending on z_{imp}
    3. Variation of the hadronic energy scale of the LAr calorimeter by ±4%
- Theoretical uncertainties:
    1. Variation of the renormalization scale $\mu_R^2 = Q^2$ by factors of 2 and 1/2
    2. Variation of the factorization scale $\mu_F^2 = Q^2$ by factors of 4 and 1/4
    3. Variation of the infrared matching scale μ_I by $\pm 1/2\,\text{GeV}$
    4. Selection of the MRSA'-105 and MRSA'-130 pdfs [62] with lower respectively higher intrinsic $\alpha_s(M_Z)$ than in the standard pdfs
    5. Usage of the completely different pdf set CTEQ4A2 [85] with similar $\alpha_s(M_Z)$

The experimental sources have already been described in section 7.5, but concerning the theoretical ones, some remarks are in order. The renormalization scale μ_R as well as the factorization scale μ_F (s. [50] for details) are arbitrary in the sense that in a complete theory the calculations are not allowed to vary with any specific choice. Yet, in reality we have only an approximative theory at our disposal yielding residual dependences due to the neglected higher orders. To avoid the appearance of large logarithms in the computations, it is recommended to identify them with a process relevant scale which for our purposes is always Q. To estimate the effect of omitted higher orders, it is **conventional** to vary them by an **arbitrary** factor of 4 concerning the squares μ_R^2, μ_F^2. In the case of μ_R^2, we had to reduce this factor to 2 because of the condition (10.14): $\Lambda \ll \mu_I \approx 2\,\text{GeV} \ll Q$. Taking this into account, varied results can be achieved quickly via eq. (8.21). With respect to μ_F, however, the complete calculations have to be redone.

The change of the infrared matching scale μ_I by $\pm 1/2$ follows from the values used in the original proposal [6]. Note that this contributes only to the uncertainty for $\alpha_s(M_Z)$! $\overline{\alpha}_0$ explicitly depends on μ_I.

The next point accounts for the fact that implicitly $\alpha_s(M_Z)$ has already been used in deriving the pdfs from data which in turn may bias our computations. The same is true for the choice of parameterization of the pdfs. For those reasons, three alternative sets, two with different assumptions on $\alpha_s(M_Z)$ and one with approximately the same $\alpha_s(M_Z)$ but another parameterization, have been selected for a reevaluation with DISENT 0.1.

All uncertainties described are presented graphically in figs. 10.6–10.8, separately for each of the five event shapes where the coefficients a_F are known. Without such a prediction for y_{fJ} and y_{k_t}, a study of systematic effects is omitted.

Inspecting the figs. 10.6–10.8 again, it stands out that τ_P and B_P behave differently compared to ρ_E, τ_C and C_P. For the latter, the systematic uncertainties are clearly

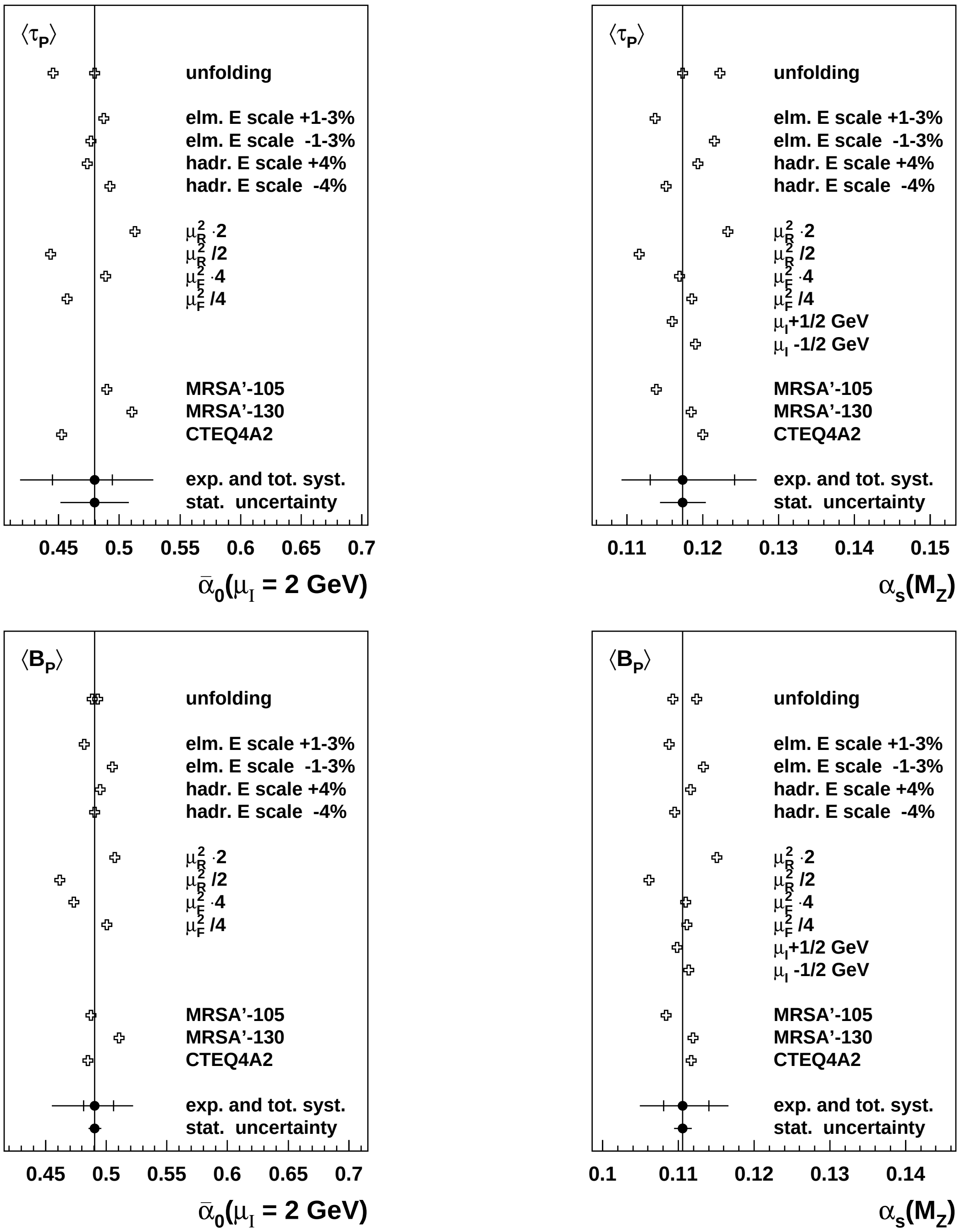


Figure 10.6: Systematic uncertainties of $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ (left) and $\alpha_s(M_Z)$ (right) for τ_P (top) and B_P (bottom).

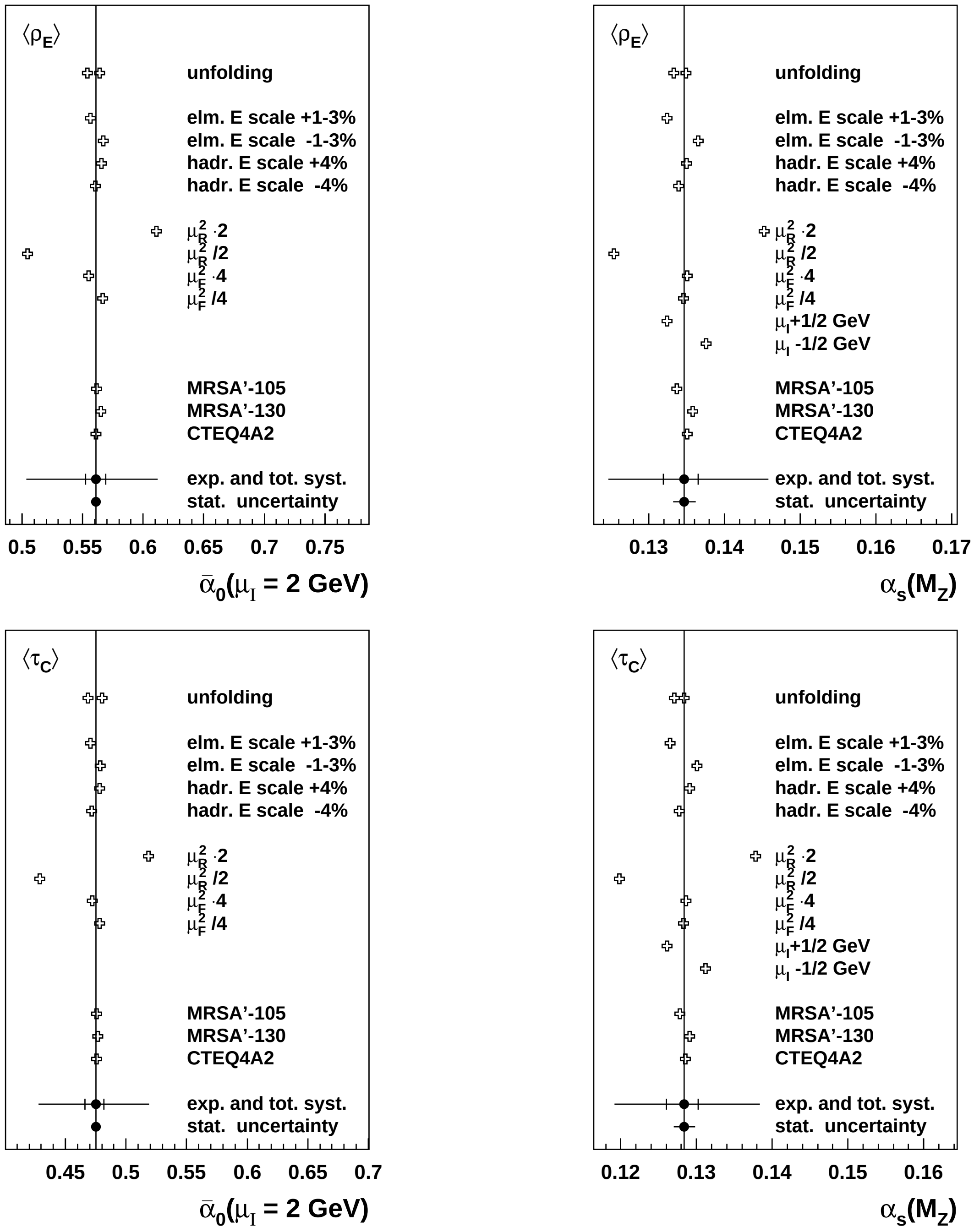


Figure 10.7: Systematic uncertainties of $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ (left) and $\alpha_s(M_Z)$ (right) for ρ_E (top) and τ_C (bottom).

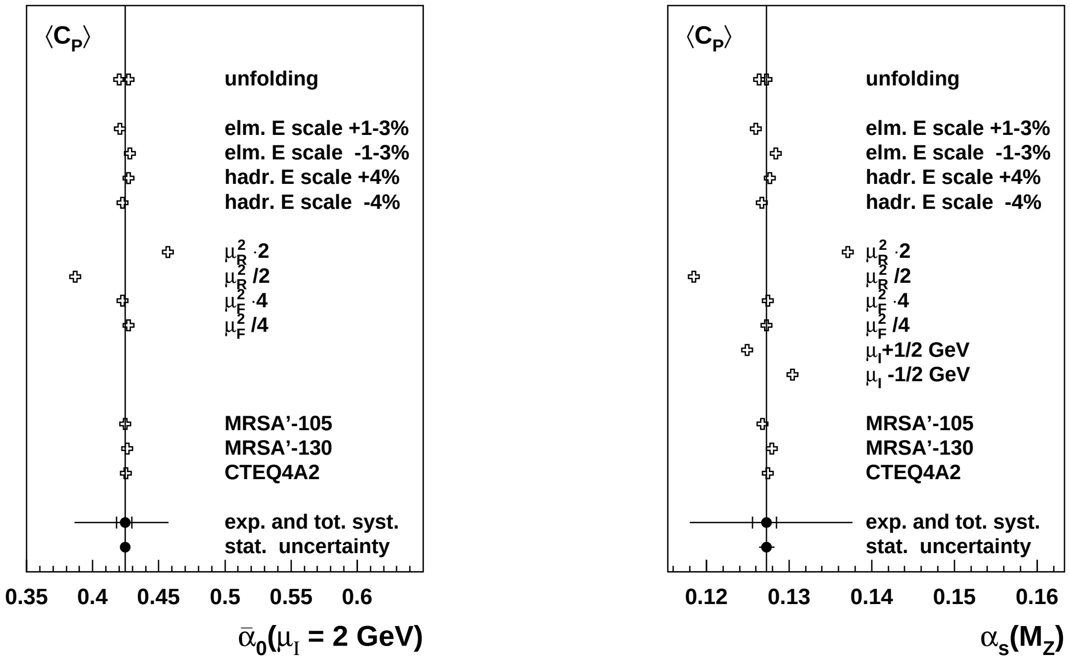


Figure 10.8: Systematic uncertainties of $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV})$ (left) and $\alpha_s(M_Z)$ (right) for C_P.

dominated by the variation of the renormalization scale. In case of τ_P and B_P, this is much less pronounced. The larger influence of experimental uncertainties probably can be related to the explicit reference to the boson axis implied in the definitions of τ_P and B_P.

10.6 Cross-Checks

Concerning systematic effects some additional cross-checks were performed:

1. In order to ensure a reliable measurement of the scattered electron, partially inefficient regions between calorimeter modules (ϕ-cracks) or wheels (z-cracks) are avoided. This selectively diminishes the contribution to certain phase space regions and is corrected for by the procedure described in section 7.2. Nevertheless, it was checked that the actual influence is negligible, even without unfolding.

2. The derivation of the power correction coefficients does not account for phase space constraints such as the cuts nos. 3 and 4 imposed on the scattered electron. Experimentally, they are absolutely necessary, but for testing purposes the NLO calculations were repeated without these cuts. Except for $\langle y_{k_t} \rangle$, which increased by about 8% in the low Q^2 region, all other mean values changed by less than 2%.

Chapter 11

Summary and Outlook

The first measurement of event shapes in ep DIS [25, 23], covering a range in Q from 7 GeV up to 100 GeV in a single experiment, has been substantially improved and extended to include the additional variables ρ_E, C_P, y_{fJ} and y_{k_t} as well as new data taken in 1997. The available statistics in the high Q^2 sample was more than doubled, facilitating a bisection of the highest Q-bin. The improvements encompass i.a. a more precise calibration of the energy of the scattered electron, an unfolding of the data distributions applying a Bayesian approach and a full account of radiative QED corrections. Taking into consideration the predominant sources of systematic influences, i.e. the accuracy of the electromagnetic and hadronic energy scale of the calorimeters and the unfolding procedure, corrected mean values of the event shape spectra together with estimates on their statistical and systematic uncertainty could be derived.

All event shape means exhibit a strong Q-dependence. They decrease with rising Q, i.e. the energy flow becomes more collimated. For the event shapes defined in the current hemisphere alone, it is interesting to compare with corresponding ones in e^+e^- collisions. This was already done in [23] for $\langle\tau_C\rangle$ and $\langle\tau_P\rangle$ as well as $\langle\rho_Q\rangle$. Despite the fact that the QCD dynamics in ep DIS is different,[1] a large similarity not only in the Q-dependence, but also in absolute value to their e^+e^- counterparts could be established. When consulting reanalyzed data of the JADE Collaboration [87], this still holds for the C parameter, but not for the jet broadening. B_P is about a factor two larger than the wide as well as the total jet broadening.

Employing calculations in pQCD to NLO accuracy, the influence of parton fragmentation and hadronization could be studied. As an improvement, the approximative treatment of the x-dependence of the perturbative coefficients applied in [23] has been replaced by a more precise bin-wise evaluation. A direct comparison with data revealed that the investigated event shapes can roughly be categorized into three classes:

[1] There are neither BGF diagrams in e^+e^- physics, nor is there a predefined event axis. In addition, it is impossible to have an empty hemisphere.

1. The event shapes τ_P and B_P employing the $z^\star$-axis as event axis are sensitive to QED radiation. They show hadronization corrections of medium size.

2. ρ_E, τ_C and C_P are strongly affected by the non-perturbative transition to hadrons. Their by far largest systematic uncertainty is due to the variation of the renormalization scale μ_R.

3. Both event shapes based on jet algorithms, y_{fJ} and y_{k_t}, exhibit small (even negative) hadronization corrections.

Fixing $\alpha_s(M_Z)$ to 0.119, the current world average [5], simple power corrections λ/Q and μ/Q^2 to parameterize the discrepancies between data and theory were tested. For λ, reasonable results of 0.7 GeV and 0.5 GeV could be obtained for τ_P and B_P, whereas y_{fJ} and y_{k_t} yielded negative values of -0.2 GeV and -0.4 GeV respectively. For the other event shapes the fit delivered poor χ^2/dof's. This is generally the case for μ/Q^2-terms; except for maybe y_{fJ} and y_{k_t}, they are ruled out. Fitting $\alpha_s(M_Z)$ in addition to λ or μ did not help either due to very large correlations.

Invoking the approach for power corrections initiated by Dokshitzer and Webber, the situation improves significantly. For τ_P, B_P, ρ_E, τ_C and C_P, where predictions for the coefficients a_F are available, correlations were reduced — although not as far as desirable — and the fits converged nicely. Note that for B_P the new coefficient a'_{B_P} replacing the formerly predicted logarithmic term had to be included. Within 20%, all determined $\overline{\alpha}_0$'s are compatible with a universal value of $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) \approx 0.5$. Very similar results are achieved with this model when applied to e^+e^- data [87, 88, 89] except for the wide jet broadening.

Concerning $\alpha_s(M_Z)$, an inconvenient spread ranging from 0.111–0.135 is observed. Although all five values turned out to be consistent with each other, a simple averaging seems to be inappropriate since the outcomes for $\alpha_s(M_Z)$ again reflect two of the categories defined above. τ_P and B_P mark the lower bound of the range in $\alpha_s(M_Z)$ as opposed to ρ_E, τ_C and C_P at the upper end. Combining several of these event shapes, the fits did work, but their quality severely declined, even allowing an individual $\overline{\alpha}_0$ for each.[2] As long as this systematic behaviour is not better understood, a precise determination of $\alpha_s(M_Z)$ is not feasible that way. Derived from the spread above, it can be stated at best that $\alpha_s(M_Z) = 0.123 \pm 0.012$.

Returning to the y event shapes, the conjectured coefficient of $a_{y_{fJ}} = 1$ has been excluded. Derived from data keeping $\overline{\alpha}_0(\mu_I = 2\,\mathrm{GeV}) = 0.5$ and $\alpha_s(M_Z) = 0.119$ fixed, a value of about $-1/4$ can be predicted. Without any coefficient given for y_{k_t}, it can only be checked for the power p of the $1/Q^p$ term. Due to the smallness of the observed hadronization corrections in comparison with experimental uncertainties, however, it is very difficult to extract them from data. In fact, the performed investigations hint at powers $p > 1$, but they are not conclusive enough to claim more.

[2] Note that correlations were ignored.

The achieved measurements can be complimented in several ways:

- Instead of looking for the Q-dependence alone, the data could additionally be divided into bins in x. This is done by the ZEUS Collaboration [86].
- In e^+e^- annihilation the model of Dokshitzer and Webber et al. has been extended to differential distributions [78, 76]. Although no predictions for ep DIS exist, it is straightforward to just look what comes out, since the unfolded event shape spectra are at hand. Actually, a first check was already performed [49] and revealed inconsistencies even between fits of the mean values and distributions of one shape variable. Resummed calculations which are not available for ep DIS may be helpful.
- By further subdividing a distribution into intervals as done by the DELPHI Collaboration, it is possible to examine where specific power-like contributions occur [89, 90].
- Alternatively, higher moments of already investigated event shapes can be studied. First steps into that direction were taken by the OPAL Collaboration [91].
- Another obvious extension is the usage of new variables, possibly with small hadronization corrections.

Especially concerning the last point, however, the question arises of what to do with them. For y_{k_t} it was already noticed that with the given experimental uncertainties power contributions can not be reliably extracted. Yet, just in this case it could be interesting to perform direct fits of $\alpha_s(M_Z)$ with the distributions at high Q^2. Remaining hadronization corrections must then be estimated by the usual procedure of correcting to the parton level of MC models. At the same time, this could be done for the event shapes of this analysis, such that a comparison of the fit results in $\alpha_s(M_Z)$ could provide insight into systematic effects in one (or both) of these methods.

At last, it should be mentioned that the *per se* sharp angular cut-off of 90° in the Breit frame is neither theoretically nor experimentally unproblematic, (s. sections 4.2 and 5.5). It has been suggested to replace it by a kind of weighting procedure to ensure a smooth transition. On the other hand, one could try to separate the remnant by employing e.g. the k_t-algorithm and treat then all remaining objects as input for the derivation of event shapes.

List of Figures

List of Tables

Bibliography

[1] C. Berger, *Teilchenphysik, Eine Einführung*, Springer-Verlag, Berlin Heidelberg (1992).

[2] F. Halzen, A.D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, Wiley, Singapur (1984).

[3] D.H. Perkins, *Introduction to High Energy Physics*, Addison Wesley, Wokingham (1987).

[4] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Benjamin/Cummings, Menlo Park (1983).

[5] Particle Data Group, C. Caso et al., *Review of Particle Physics*, *Eur. Phys. J.* **C 3** (1998) 1.

[6] Yu.L. Dokshitzer, B.R. Webber, *Phys. Lett.* **B 352** (1995) 451.

[7] F. Eisele, in Proceedings *XVIII. Physics in Collision*, Frascati, Italy, 1998.

[8] H1 Collaboration, I. Abt et al., *Nucl. Instr. and Meth.* **A 386** (1997) 310.

[9] ZEUS Collaboration, *The ZEUS Detector, Status Report 1993*, DESY, Hamburg (1993).

[10] H1 Collaboration, I. Abt et al., *Nucl. Instr. and Meth.* **A 386** (1997) 348.

[11] H1 Collaboration, *Technical Proposal to Upgrade the Backward Scattering Region of the H1 Detector*, DESY-PRC-93-02.

[12] H1 SPACAL Group, T. Nicholls et al., *Nucl. Instr. and Meth.* **A 374** (1996) 149.

[13] H1 Collaboration, Conference contribution abstract 533 to *International Conference on High Energy Physics*, Vancouver, Canada, 1998.

[14] F. Jacquet, A. Blondel, in Proceedings *Study of an ep Facility for Europe*, Hamburg, Germany, 1979.

[15] S. Bentvelsen, J. Engelen, P. Kooijman, in Proceedings *Physics at HERA*, Hamburg, Germany, 1991.

[16] U. Bassler, G. Bernardi, *Nucl. Instr. and Meth.* **A 361** (1995) 197.

[17] H1 Collaboration, I. Abt et al., *Phys. Lett.* **B 321** (1994) 161;
H1 Collaboration, S. Aid et al., *Phys. Lett.* **B 354** (1995) 494;
H1 Collaboration, S. Aid et al., *Nucl. Phys.* **B 470** (1996) 3.

[18] H1 Collaboration, C. Adloff et al., *Phys. Lett.* **B 393** (1997) 452.

[19] J.G. Körner, E. Mirkes, G.A. Schuler, *Int. J. of Mod. Phys.* **A 4** (1989) 1781.

[20] K.H. Streng, T.F. Walsh, P.M. Zerwas, *Z. Phys.* **C 2** (1979) 237.

[21] M. Dasgupta, B.R. Webber, *Eur. Phys. J.* **C 1** (1998) 539;
M. Dasgupta, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Chicago, USA, 1997.

[22] M.H. Seymour, D. Graudenz, private communication.

[23] H1 Collaboration, C. Adloff et al., *Phys. Lett.* **B 406** (1997) 256.

[24] B.R. Webber, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Paris, France, 1995.

[25] K. Rabbertz, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Chicago, USA, 1997;
K. Rabbertz, in Proceedings *Workshop on Low x Physics*, Madrid, Spain, 1997.

[26] S. Catani, G. Turnock, B.R. Webber, *Phys. Lett.* **B 295** (1992) 269.

[27] T. Chandrahoman, L. Clavelli, *Nucl. Phys.* **B 184** (1981) 365.

[28] R.K. Ellis, D.A. Ross, A.E. Terrano, *Nucl. Phys.* **B 178** (1981) 421.

[29] G. Sterman, S. Weinberg, *Phys. Rev. Lett.* **39** (1977) 1436.

[30] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock, B.R. Webber, *Phys. Lett.* **B 269** (1991) 432.

[31] S. Catani, Yu.L. Dokshitzer, B.R. Webber, *Phys. Lett.* **B 285** (1992) 291.

[32] JADE Collaboration, W. Bartel et al., *Z. Phys.* **C 33** (1986) 23.

[33] B.R. Webber, *J. Phys.* **G 19** (1993) 1567.

[34] internal H1 Software Package.

[35] V. Blobel et al., *F-PACK, F-package for input/output*, internal H1 Software Package.

[36] H. Albrecht et al., *a guide to PHAN, an H1 physics analysis package*, internal H1 Software Package.

[37] H1 BEMC GROUP, J. Bán et al., *Nucl. Instr. and Meth.* **A 372** (1996) 399.

[38] G.P. Lepage, *J. Comp. Phys.* **27** (1978) 192;
G.P. Lepage, *VEGAS, An adaptive multidimensional Integration Program*, CLNS-80/447, Cornell University, 1980.

[39] H. Plothow-Besch, *Comp. Phys. Comm.* **75** (1993) 396;
H. Plothow-Besch, *Int. J. of Mod. Phys.* **A 10** (1995) 2901.

[40] G. Ingelman, A. Edin, J. Rathsman, *Comp. Phys. Comm.* **101** (1997) 108.

[41] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, L. Stanco, *Comp. Phys. Comm.* **67** (1992) 465.

[42] L. Lönnblad, *Comp. Phys. Comm.* **71** (1992) 15.

[43] K. Charcula, G.A. Schuler, H. Spiesberger, *Comp. Phys. Comm.* **81** (1994) 381.

[44] A. Kwiatkowski, H.-J. Möhring, H. Spiesberger, *Comp. Phys. Comm.* **69** (1992) 155.

[45] T. Sjöstrand, *Comp. Phys. Comm.* **39** (1986) 347;
T. Sjöstrand, M. Bengtsson, *Comp. Phys. Comm.* **43** (1987) 367.

[46] J. Meyer, ed., et al., *Guide to Simulation program H1SIM*, internal H1 Software Package, H1 Software Note No. 3 (1991).

[47] G. D'Agostini, *Nucl. Instr. and Meth.* **A 362** (1995) 487.

[48] M. Wobisch, *Untersuchung von QED-Prozessen mit harter Photonabstrahlung in tiefinelastischer Elektron-Proton-Streuung im H1-Experiment bei HERA*, PITHA 95/38, Diplomarbeit, RWTH Aachen (1995).

[49] H.-U. Martyn, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Brussels, Belgium, 1998;
H1 Collaboration, Conference contribution abstract 530 to *International Conference on High Energy Physics*, Vancouver, Canada, 1998.

[50] R.K. Ellis, W.J. Stirling, B.R. Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge (1996).

[51] J.C. Collins, *Renormalization*, Cambridge University Press, Cambridge (1984).

[52] W.T. Giele, E.W.N. Glover, J. Yu, *Phys. Rev.* **D 53** (1996) 120.

[53] T. Brodkorb, E. Mirkes, hep-ph/9404287.

[54] D. Graudenz, *Comp. Phys. Comm.* **92** (1995) 65.

[55] K. Fabricius, G. Kramer, G. Schierholz, I. Schmitt, *Z. Phys.* **C 11** (1982) 315.

[56] W.T. Giele, E.W.N. Glover, D.A. Kosower, *Nucl. Phys.* **B 403** (1993) 633.

[57] M.H. Seymour, private communication.

[58] E. Mirkes, D. Zeppenfeld, *Phys. Lett.* **B 380** (1996) 205.

[59] S. Catani, M.H. Seymour, *Nucl. Phys.* **B 485** (1997) 291.

[60] D. Graudenz, hep-ph/9710244.

[61] M.H. Seymour, D. Graudenz, private communication.

[62] A.D. Martin, W.J. Stirling, R.G. Roberts, *Phys. Lett.* **B 356** (1995) 89.

[63] M.H. Seymour, private communication.

[64] S. Catani, M.H. Seymour, *The DISENT Program*, DISENT 0.1 program manual, unpublished (1997).

[65] G. Altarelli, ed., et al., *Z physics at LEP 1*, CERN 89-08 (Vol. 1-3) (1989).

[66] A.H. Mueller, in Proceedings *QCD — 20 Years Later*, Aachen, Germany, 1992.

[67] B.R. Webber, in Proceedings *International Conference on High Energy Physics*, Glasgow, Scotland, 1994.

[68] R.P. Feynman, *Photon Hadron Interactions*, W.A. Benjamin, New York (1972).

[69] M. Kuhlen, *QCD and the Hadronic Final State in Deep Inelastic Scattering at HERA*, MPI-PhE-97-33, hep-ph/9712505, Habilitation, Universität Hamburg (1997).

[70] A.G. Frodesen, O. Skjeggestad, H. Tøfte, *probability and statistics in particle physics*, Universitetsforlaget, Oslo (1979).

[71] F. James, *MINUIT — Function Minimization and Error Analysis*, CERN Program Library entry D506, Geneva (1994).

[72] M. Dasgupta, private communication.

[73] M. Beneke, hep-ph/9807443.

[74] B.R. Webber, *Phys. Lett.* **B 339** (1994) 148.

[75] S. Catani, G. Marchesini, B.R. Webber, *Nucl. Phys.* **B 349** (1991) 635.

[76] Yu.L. Dokshitzer, G. Marchesini, G.P. Salam, hep-ph/9812487.

[77] Yu.L. Dokshitzer, G. Marchesini, B.R. Webber, *Nucl. Phys.* **B 469** (1996) 93.

[78] Yu.L. Dokshitzer, B.R. Webber, *Phys. Lett.* **B 404** (1997) 321.

[79] M. Dasgupta, B.R. Webber, *Phys. Lett.* **B 382** (1996) 273.

[80] M. Dasgupta, B.R. Webber, *Nucl. Phys.* **B 484** (1997) 247.

[81] Yu.L. Dokshitzer, A. Lucenti, G. Marchesini, G.P. Salam, **JHEP** 05 (1998) 003.

[82] G.P. Salam, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Brussels, Belgium, 1998.

[83] M. Dasgupta, B.R. Webber, hep-ph/9809247.

[84] G.P. Salam, private communication.

[85] H.L. Lai et al., *Phys. Rev.* **D 55** (1997) 1280.

[86] R.G. Waugh, in Proceedings *Workshop on Deep Inelastic Scattering and QCD*, Brussels, Belgium, 1998;
ZEUS Collaboration, Conference contribution abstract 808 to *International Conference on High Energy Physics*, Vancouver, Canada, 1998.

[87] JADE Collaboration, P.A. Movilla Fernández et al., PITHA 98/21,

[88] D. Wicke, *Nucl. Phys. Proc. Suppl.* **64** (1998) 27.

[89] DELPHI Collaboration, P. Abreu et al., *Z. Phys.* **C 73** (1997) 229. hep-ex/9807007.

[90] D. Wicke, in Proceedings *33rd Rencontres de Moriond: QCD and High Energy Hadronic Interactions*, Les Arcs, France, 1998.

[91] OPAL Collaboration, Conference contribution abstract 178 to *International Europhysics Conference on High Energy Physics*, Jerusalem, Israel, 1997.

Curriculum Vitae

Geburtsdatum:	13. Januar 1967
Geburtsort:	Mönchengladbach
Familienstand:	ledig
Staatsangehörigkeit:	deutsch
Eltern:	Matthias Rabbertz, Weber Gertrud Rabbertz, geb. Hermanns, Bankangestellte
Geschwister:	Beate Reitler Gerda Boms

Schulbildung:	01.08.73 - 06.07.77	Gemeinschaftsgrundschule Wegberg
	01.08.77 - 12.06.86	Städtisches Maximilian-Kolbe-Gymnasium Wegberg Abschluß: Allgemeine Hochschulreife
Wehrdienst:	01.10.86 - 31.12.87	Grundausbildung in Pinneberg Tätigkeit als Sprechfunker in Geilenkirchen
Studium:	01.10.87 - 08.26.94	Rheinisch-Westfälische Technische Hochschule Aachen, Mathematisch-Naturwissenschaftliche Fakultät Studienfach: Physik Fachrichtung: Elementarteilchenphysik Abschluß: Diplom-Physiker
	01.11.89 - 15.10.92	Tätigkeit als studentische Hilfskraft am Institut für Mathematik, Prof. Dr. G. Hellwig
	seit 01.07.94	Wissenschaftlicher Angestellter am I. Physikalischen Institut der Rheinisch-Westfälischen Technischen Hochschule Aachen, Beginn des Promotionsvorhabens bei Prof. Dr. Ch. Berger
	seit 01.01.95	Mitglied der H1-Kollaboration am DESY, Hamburg

Acknowledgements

First of all, many thanks to my advisor, Prof. Dr. Christoph Berger, who put me on track of this exciting new topic, and to my collaborator Dr. Hans-Ulrich Martyn. Their constant and patient support were invaluable to me. Both had, at some point or another, to suffer from my stubbornness.

I am thankful to Prof. Dr. Siegfried Bethke for accepting the task of being my second referee.

An experiment as complicated as H1 can not be successfully operated without the dedicated and knowledgeable work of many people. I wish to thank the members of the H1 Collaboration, who contributed to this analysis in manifold ways, for their hospitality and the fascinating opportunity of researching in an international community.

For enlightening discussions and answering many questions, I am indebted to M. Dasgupta, Yu.L. Dokshitzer, D. Graudenz, E. Mirkes, G.P. Salam, M.H. Seymour and B.R. Webber.

On the important field of social interactions, I enjoyed, and still do so, very much the friendly and inspiring work climate of the Aachen groups, which sometimes led to less work and more climate ..., thank you:
Simone Baer-Lang, Sascha Caron, Carlo Duprel, Thomas Hadig, Martin Hampel, Heiko Itterbeck, Claus Keuker, Torsten Köhler, Carsten Krauss, Boris Leissner, Peer-Oliver Meyer, Christian Niedzballa, Konrad Rosenbauer, Wilhelm Rottkirchen, Jürgen Scheins, Lars Sonnenschein, Markus Wobisch and ... all the people I might have forgotten.

Special thanks to Thomas Hadig, Martin Hampel, Thorsten Wengler and Markus Wobisch for proofreading and lots of helpful suggestions.

Last but not least, I am heartily grateful to my family and especially my parents for their generous support and their confidence in me. It is very easy to take comforts for granted.