A search for QCD-instantons in deep-inelastic ep scattering with the ZEUS detector at HERA

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

Sonja Hillert

aus Hamburg

Hamburg
2002

Gutachter der Dissertation: Prof. Dr. E. Lohrmann
Prof. Dr. R. Klanner

Gutachter der Disputation: Prof. Dr. E. Lohrmann
Prof. Dr. V. Blobel

Datum der Disputation: 20. Dezember 2002

Vorsitzender des Prüfungsausschusses: Prof. Dr. F.-W. Büßer

Vorsitzender des Promotionsausschusses: Prof. Dr. G. Huber

Dekan des Fachbereichs Physik: Prof. Dr. F.-W. Büßer

Abstract

A search for QCD-instanton induced events in deep-inelastic *ep* scattering (DIS) has been performed with the ZEUS detector at the HERA accelerator, using a neutral current sample of 1996/1997 data corresponding to an integrated luminosity of $38.6\,\mathrm{pb}^{-1}$. A kinematic range defined by cuts on the photon virtuality, $Q^2 > 120\,\mathrm{GeV}^2$, and on the Bjorken scaling variable, $x > 10^{-3}$, has been investigated. QCD-instanton induced events were modelled by the Monte Carlo generator QCDINS, predicting an instanton contribution to the inclusive sample of $\approx 0.5\%$. The background from standard DIS was described by a combination of Djangoh (Ariadne) with diffractive Rapgap and by Herwig, respectively. Different methods to enhance the fraction of instanton induced events were tried, Fisher's algorithm proving to be the best one. In an instanton-enhanced subsample, sphericity has been fitted by a combination of signal and background Monte Carlo, yielding an upper maximum likelihood 2σ limit on the fraction of instanton-induced events in the enhanced sample of 12% (8.1%) when Herwig (Djangoh with Rapgap) is assumed to describe the background correctly, compared to an expected fraction of 8.8%. However, there are indications that the simulation of standard DIS events is not reliable. Therefore, a most conservative 2σ limit of 206 instanton-induced events in the enhanced sample has been set, assuming the background to be zero, where 62 events are predicted by QCDINS.

Kurzfassung

Mit dem ZEUS-Detektor am HERA-Beschleuniger wurde eine Suche nach QCD-Instanton induzierten Ereignissen in tiefunelastischer *ep*-Streuung (DIS) durchgeführt. Der dabei verwendete Datensatz von Neutralstrom-Ereignissen aus den Jahren 1996/1997 entsprach einer integrierten Luminosität von $38.6\,\mathrm{pb}^{-1}$. Untersucht wurde der durch Schnitte auf die Photonvirtualität, $Q^2 > 120\,\mathrm{GeV}^2$, und auf die Bjorken-Skalenvariable, $x > 10^{-3}$, definierte kinematische Bereich. Instanton-induzierte Ereignisse wurden durch den Monte Carlo Generator QCDINS beschrieben, der einen Instanton-Beitrag von ca. 0.5% zum inklusiven Datensatz vorhersagt. Der Standard-DIS Hintergrund wurde zum einen durch eine Kombination von Djangoh(Ariadne) mit diffraktivem Rapgap, zum anderen durch Herwig beschrieben. Von verschiedenen Verfahren zur Anreicherung Instanton-induzierter Ereignisse erwies sich der Fisher-Algorithmus als am besten geeignet. In einem angereicherten Datensatz wurde eine Kombination von Signal- und Hintergrund-Monte Carlo an die Sphärizitätsverteilung angepasst. Die Anpassung ergab eine obere Maximum Likelihood 2σ-Grenze auf den Anteil Instanton-induzierter Ereignisse im angereicherten Datensatz von 12% (8.1%), wenn angenommen wird, daß Herwig (Djangoh mit Rapgap) die Daten korrekt beschreibt, verglichen mit einem vorhergesagten Anteil von 8.8%. Es gibt allerdings Hinweise, daß die Simulation von DIS-Ereignissen unzuverlässig ist. Daher wurde unter der Annahme, daß der DIS-Hintergrund keinen Beitrag liefert, eine konservative 2σ-Grenze von 206 Instanton-induzierten Ereignissen im angereicherten Datensatz gesetzt, für den QCDINS 62 Ereignisse vorhersagt.

Contents

List of Figures

Introduction

In the Standard Model, both the strong and the electroweak interaction are described by non-abelian gauge theories. In such theories, the ground state has a rich topological structure. Its simplest building blocks are non-perturbative fluctuations of the gauge fields [Bel75], called instantons ['t Hooft]. They can be interpreted as tunnelling processes of the gauge fields between topologically distinct types of vacuum states. The name instanton refers to the fact, that these fluctuations are localised in space and Euclidean time.

In quantum chromodynamics (QCD), the theory of strong interactions, a number of theoretical problems can be addressed by taking instanton effects into account, for a review see e.g. [Schä98]. For example, instanton theory predicts the non-conservation of chirality and provides an explanation for the fact that the η' mass is much larger than that of the η state.

Although instantons are theoretically required by the Standard Model, they have not yet been observed experimentally. While in electroweak interactions instantons might play a role only at cm energies of $\gg 10\,\mathrm{TeV}$ [Rin91, Gib95, Rin94, Rin02], in QCD instanton effects are expected to become sizable at much lower energies, because the strong coupling α_s is much larger than the equivalent parameter, α, in electroweak theory.

QCD instantons are predicted to have short distance implications [Bal93, Rin94, Moch97, Rin98]. In particular, they can induce characteristic events in neutral current (NC) deep-inelastic *ep* scattering (DIS) at HERA. The cross section of these events has been calculated by A. Ringwald, F. Schrempp and collaborators using instanton-perturbation theory [Moch97, Rin98]. Besides, the phenomenology of instanton-induced events at HERA has been studied [Rin94, Car99, Rin00]: such events are characterised by a high multiplicity final state of large transverse momentum, large sphericity, flavour democracy and maximal chirality violation.
The Monte Carlo (MC) generator QCDINS [Rin00] by A. Ringwald and F. Schrempp allows dedicated experimental searches for instanton-induced events to be performed. Results of such a search have recently been reported by the H1 collaboration [Ad02]. Upper limits for the cross section of 60 pb and 100 pb were determined, depending on the kinematic region considered.

In this thesis, a ZEUS analysis using a sample of 1996/1997 neutral current data corresponding to an integrated luminosity of $38\,\mathrm{pb}^{-1}$ is presented. As instanton-induced events are predicted to contribute only $\approx 0.5\%$ to the cross section of the inclusive sample used in this search, it is essential to find cuts that efficiently reject background events, thus enhancing the instanton contribution in the subsample. For that purpose, variables characterising the hadronic final state are studied. In one attempt, combinations of one- and two-dimensional cuts are used for restricting the sample. In another approach, first an optimal linear combination of the input variables is determined using the Fisher algorithm [Fish36, BlLo98]. The enhancement is then achieved by a cut on the resulting Fisher discriminant.

Upper limits on the fraction of instanton-induced events within instanton-enhanced samples are set using two different approaches: the fit approach extracts the information on the instanton contribution from the shape of characteristic variables. In the cut approach a limit is determined from the number of events passing a set of hard cuts.

Chapter 1

Theoretical Background

1.1 Electron-proton scattering

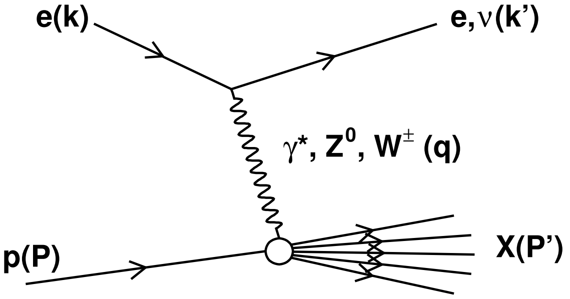


Figure 1.1: *A generic diagram of ep scattering. Four-momenta of the interacting particles are given in parentheses.*

In electron-proton scattering, a point-like particle, the electron[1], interacts via the electromagnetic or weak force with the proton, which has a complex substructure. The interaction can be described by the exchange of a photon, a Z^0 boson or a $W^\pm$, transferring a four-momentum q from the lepton to the proton as shown in Fig. 1.1.

Neutral current (NC) processes are mediated by a photon or Z^0, while $W^\pm$ exchange gives rise to charged current (CC) events. The contributions of NC and CC events to the total cross section depend on the virtuality $Q^2 = -q^2$ of the exchanged boson. While the NC cross section dominates at medium Q^2 due to the dominance of the electromagnetic over the weak contribution, at virtualities $Q^2 \approx 10^4\,\mathrm{GeV}^2$ the NC and CC cross sections are of similar size, see Fig. 1.2. This is evidence that the electromagnetic and the weak contributions can be described by a unified "electroweak" field theory, in which the couplings of the two contributions obtain similar values in that high Q^2

---

[1] In what follows, both positrons and electrons are referred to as electrons.

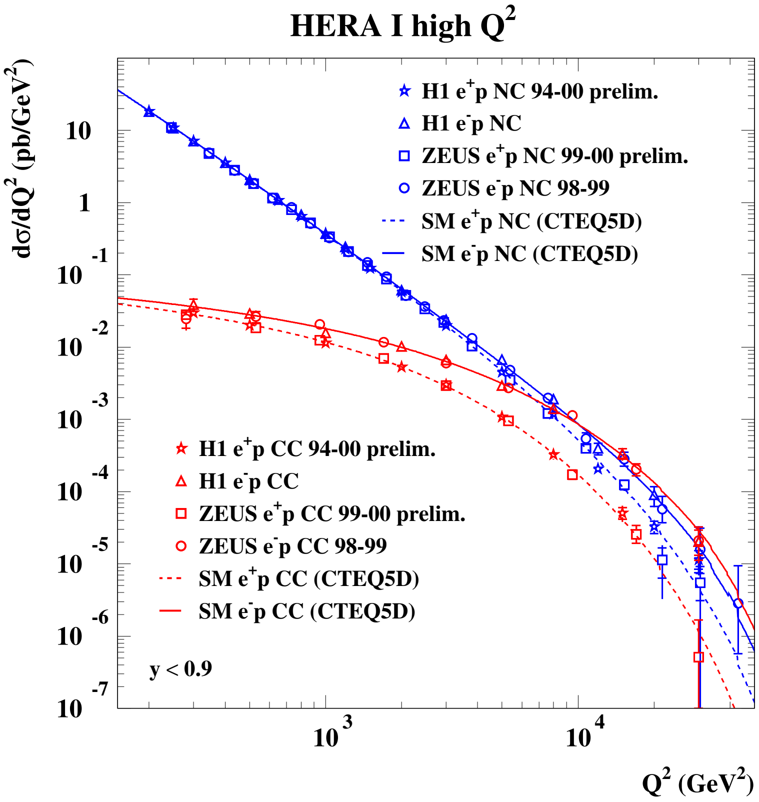


Figure 1.2: *Comparison of Neutral Current and Charged Current cross sections as a function of Q^2 measured in deep-inelastic ep scattering at HERA. From [Zha02].*

range.

For the analysis presented in this thesis, a NC deep-inelastic scattering (DIS) sample at Q^2 values, at which Z^0 exchange can be neglected, has been used. In that regime, the photon acts as a probe to study the substructure of the proton and the nature of the strong force giving rise to it. All processes observed in DIS, except for instanton-induced events, will be referred to as normal DIS (nDIS) events in what follows.

The Born cross section can be expressed in terms of the proton structure functions $F_i^{\rm NC}$, $i = 1, 2, 3$, and the kinematic variables x, y as

$$\left(\frac{d^2\sigma^{\rm NC}}{dx\; dQ^2}\right)_{\rm Born} = \frac{4\pi\alpha^2}{x\,Q^4}\left\{y^2 x F_1^{\rm NC}(x,Q^2) + (1-y)F_2^{\rm NC}(x,Q^2) \mp \left(y - \frac{y^2}{2}\right) xF_3^{\rm NC}(x,Q^2)\right\} .$$

Two structure functions, F_1 and F_2 are needed to describe the interaction of the photon with the proton. The combination

$$F_L(x,Q^2) = \left(1 + \frac{4M^2x}{Q^2}\right) F_2(x,Q^2) - 2xF_1(x,Q^2) \approx F_2 - 2xF_1 \quad ,$$

is the longitudinal structure function, where M is the proton mass. The function F_1 (F_L) corresponds to the absorption of transversely (longitudinally) polarised photons. The function xF_3^{NC} accounts for parity violation due to Z^0 exchange, hence contributing only at large Q^2.

The kinematic variables Q^2, x and y are defined as

$$Q^2 = -q^2 = -(k-k')^2 \quad , \qquad x = \frac{Q^2}{2P \cdot q} \quad , \qquad y = \frac{q \cdot P}{k \cdot P} \quad \text{(cf. Fig. 1.1)} \quad .$$

Because of the relation $Q^2 = x \cdot y \cdot s$, with s being the *ep* cm energy squared, the kinematics of the process is fully determined by any choice of two of the three variables x, y, Q^2.

The variables x and y are interpreted in the quark-parton model, which states that an *ep* scattering process can be approximated by the scattering of the electron off a quark inside the proton. In that model, the Bjorken variable x is the fraction of the proton momentum carried by the struck parton, while y represents the fraction of the positron energy transferred to the proton in the rest frame of the initial-state proton. The structure functions of the proton at large values of Q^2 have been precisely measured at HERA [StrFct].

1.2 Quantum Chromodynamics

The strong force is described by quantum chromodynamics (QCD), a renormalisable, non-abelian gauge field theory, see e.g. [Ell96] and the references therein. It is based on an $SU(3)$ symmetry between three types of colour charge carried by the quarks. However, because QCD is non-abelian, its gauge bosons, the gluons, also carry colour and can therefore interact with each other. The strength of the interaction is determined by the coupling α_s, which is not constant in a renormalisable gauge theory, but depends on the energy scale, usually chosen as Q.

At low energies, or equivalently, large distances, the coupling is large. Therefore, quarks and gluons cannot be observed as free particles, but are confined within colour-neutral composed objects, such as the proton. At those scales, interactions can only be described non-perturbatively. However, at large scales (short distances), α_s is small enough for the application of perturbation theory.

Physical observables $R(Q)$ can then be written in terms of α_s as

$$R(Q) = R_0(Q) + \alpha_s R_1(Q) + \alpha_s^2 R_2(Q) + \ldots \quad .$$

They should not depend on the renormalisation scale μ_R, at which ultraviolet divergences are subtracted, as μ_R can be chosen arbitrarily, so

$$\mu_R^2 \frac{d}{d\mu_R^2} R\left(\frac{Q^2}{\mu_R^2},\ \alpha_s\right) = 0 \quad .$$

If the renormalisation scale is chosen such that both $\alpha_s(Q)$ and $\alpha_s(\mu_R)$ are small enough, the energy scale dependence of α_s can be determined within the framework of perturbation theory. To lowest order it is given by [Ell96]

$$\alpha_s(Q^2) = \frac{\alpha_s\left(\mu_R^2\right)}{1+\alpha_s\left(\mu_R^2\right)bt} \quad , \; t = \ln\frac{Q^2}{\mu_R^2} \quad , \; b = \frac{1}{4\pi}\left(11-\frac{2}{3}n_f\right) \quad ,$$

where n_f is the number of quark flavours with masses below μ_R.
The value of α_s at a fixed scale, usually given by the mass of the Z boson, M_Z, can be determined in various ways, for an overview see e.g. [Beth00]. At HERA, α_s has been determined from jet physics [AlS], in particular from inclusive jet cross sections and dijet cross sections, as well as from measurements of the jet shape and the subjet multiplicity. Results are found to be in agreement with measurements from $p\bar{p}$ [Af02] and e^+e^- interactions, see [Beth00], and with the current world average [PDG00] of

$$\alpha_s(M_Z^2) = 0.1181 \pm 0.0020 \quad .$$

Both perturbative and non-perturbative effects have to be accounted for in the description of deep-inelastic *ep*-scattering. The factorisation theorem [Col85] states that the scattering process factorises with respect to these contributions, such that the hard scattering process, calculable in perturbative QCD, can be treated independently from non-perturbative effects. These are absorbed into universal parton distribution functions of the proton that have to be determined experimentally.
Three types of hard scattering processes contribute to the neutral current cross section up to $\mathcal{O}(\alpha_s)$: the quark-parton model process (QPM), the QCD Compton process (QCD-C) and the boson-gluon fusion (BGF), shown in Fig.1.3.

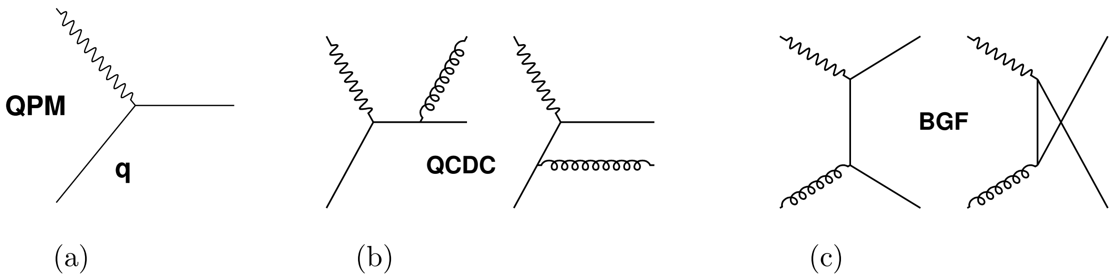


Figure 1.3: *Processes contributing to the NC cross section up to $\mathcal{O}(\alpha_s)$: a) Quark-Parton-Model, b) QCD-Compton, c) Boson-Gluon-Fusion.*

The structure functions $F_i(x, Q^2)$ of the proton, introduced in section 1.1, are related to the partonic structure functions $\hat{F}_i^a$, which describe the hard scattering process between the virtual photon and parton a (a being a gluon, quark or antiquark) via the convolution

$$F_i(x, Q^2) = \sum_a f_a(x, \mu_F) \otimes \hat{F}_i^a\left(x, \frac{Q}{\mu_F}, \alpha_s(\mu_R)\right) \quad ,$$

with the parton density functions $f_a(x, \mu_F)$. These expressions depend on the factorisation scale μ_F, which is commonly set equal to the renormalisation scale: $\mu_F = \mu_R = Q$. In addition to the factorisation scale dependence, f_a and $\hat{F}_i^a$ depend on the factorisation scheme, which determines, how much of the finite contribution to F_i is absorbed into the parton densities. Two schemes commonly used are the minimal subtraction ($\overline{MS}$) scheme and the DIS scheme.

The parton densities f_a are themselves scale dependent. However it is sufficient to experimentally determine them at an initial scale, as their evolution to different scales is predicted by QCD. That evolution is described by a set of $(2n_f - 1)$ equations referred to as the DGLAP equations, derived by Dokshitzer, Gribov, Lipatov, Altarelli and Parisi [DGLAP]. Using $t = \mu^2$, it can be written in the form

$$t\frac{\partial}{\partial t}\begin{pmatrix} q_i(x,t) \\ g(x,t) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi}\sum_{q_i,\bar{q}_j}\int_x^1 \frac{d\xi}{\xi}\times\begin{pmatrix} P_{q_iq_j}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{q_ig}\left(\frac{x}{\xi},\alpha_s(t)\right) \\ P_{gq_j}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{gg}\left(\frac{x}{\xi},\alpha_s(t)\right) \end{pmatrix}\begin{pmatrix} q_i(\xi,t) \\ g(\xi,t) \end{pmatrix} ,$$

where $q_i(x,t)$ and $g(x,t)$ are the (anti-)quark and gluon densities, respectively. The functions denoted by a capital P with quark/gluon indices are the splitting functions, each of which is calculable as a power series in α_s,

$$\begin{aligned} P_{q_iq_j}(z,\alpha_s) &= \delta_{ij}P^{(0)}_{qq}(z) + \tfrac{\alpha_s}{2\pi}P^{(1)}_{qq}(z) + \dots \\ P_{q_ig}(z,\alpha_s) &= \delta_{ij}P^{(0)}_{qg}(z) + \tfrac{\alpha_s}{2\pi}P^{(1)}_{qg}(z) + \dots \\ P_{gq_j}(z,\alpha_s) &= \delta_{ij}P^{(0)}_{gq}(z) + \tfrac{\alpha_s}{2\pi}P^{(1)}_{gq}(z) + \dots \\ P_{gg}(z,\alpha_s) &= \delta_{ij}P^{(0)}_{gg}(z) + \tfrac{\alpha_s}{2\pi}P^{(1)}_{gg}(z) + \dots \quad . \end{aligned}$$

They describe the probability of parton emission by quarks and gluons, depending on the types of partons involved, and on the distribution of the initial momentum to the resulting partons. As a consequence of charge conjugation invariance and $SU(n_f)$ flavour symmetry, the splitting functions are the same for quarks and antiquarks and independent of the quark flavour.

1.3 Instantons

Instantons are the simplest building blocks of the rich topological structure, which forms the ground state of QCD. Being a non-abelian gauge theory, QCD gives rise to topologically distinct vacuum states, which are separated from each other by finite energy barriers [Rin94]. However, the gauge fields may tunnel between the different states. These rapid tunnelling events, localised in space and Euclidean time, are described by instantons.

The presence of an instanton influences the hard subprocess, which forms the core of deep-inelastic *ep*-scattering events. As developed theoretically and phenomenologically by A. Ringwald and F. Schrempp [Rin94, Moch97, Rin98, Rin99, Rin00, Rin01], it results in a characteristic final state, that allows instanton-induced events to be distinguished from events stemming from normal DIS ("nDIS") processes. The Monte Carlo generator QCDINS [Rin00], describing instanton-induced events in the HERA regime, predicts a sizable rate of these events. To the experimentalist, HERA therefore offers a unique chance to discover instantons, which would be a confirmation of an essential non-perturbative Standard Model prediction.

1.3.1 A quantum mechanical example

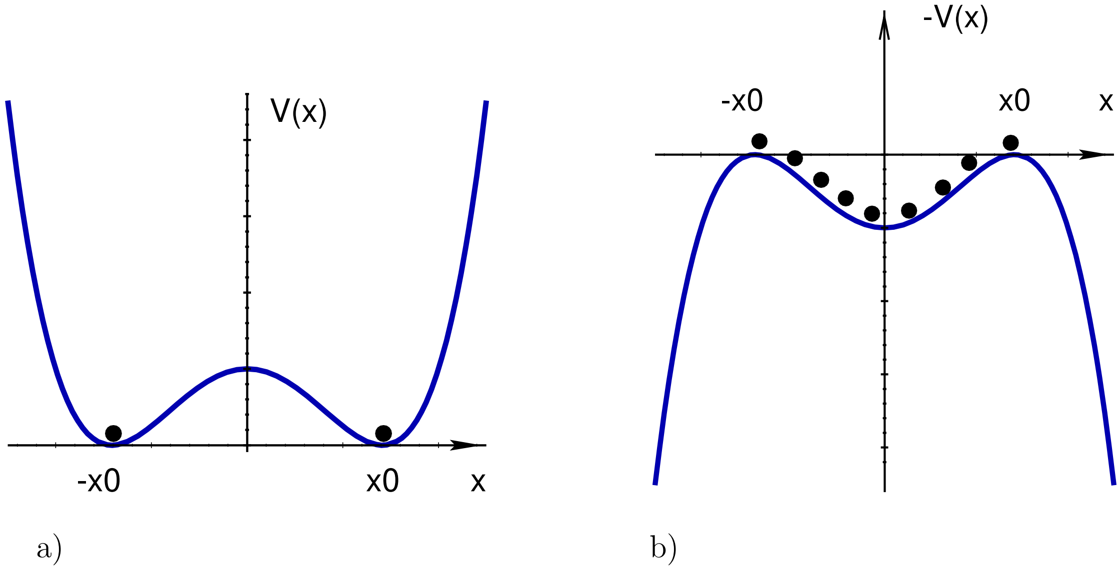


Figure 1.4: *The double well potential for (a) Minkowski time and (b) Euclidean time τ. In (a) the dots at $x = \pm x_0$ represent the classical solutions. In (b), the region $-x_0 < x < x_0$ becomes accessible. Figure from [Schr99].*

Some basic properties of instantons can be studied in the context of classical quantum mechanics. Consider the double-well potential with minima at $\pm x_0$, $V(x) = g(x^2 - x_0^2)^2$, illustrated in Fig. 1.4(a). Classically, the ground state has the two degenerate solutions $x = \pm x_0$. Quantum mechanically, tunnelling between these states can occur. Although being a quantum mechanical process, this can be described in a classical way. The classical equation of motion is given by

$$E = \frac{1}{2}\left(\frac{dx}{dt}\right)^2 + V(x) \quad .$$

By replacing time t by Euclidean time τ: $t \to -i\tau$, one obtains

$$-E = \frac{1}{2}\left(\frac{dx}{d\tau}\right)^2 - V(x)$$

instead. So the relative sign between the potential and the time derivative of x changes and the interval $-x_0 < x < x_0$ becomes accessible, cf. Fig. 1.4(b). In particular, for $E = 0$ one finds the solution

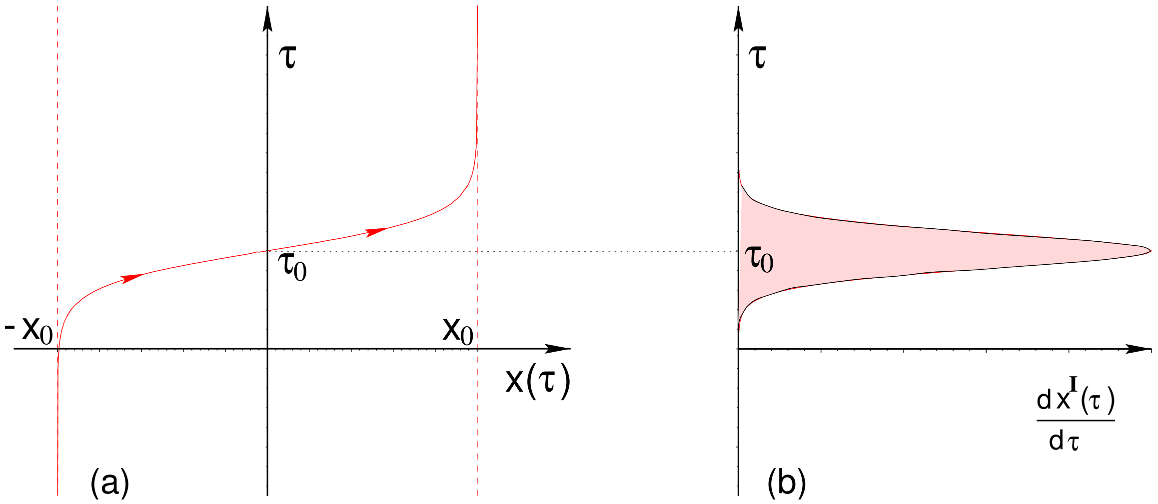


Figure 1.5: *(a): instanton solution for the double well potential, (b): the solution is located at τ_0. Figure adapted from [Schr99].*

$$x^{(\mathrm{I})}(\tau) = \pm x_0 \tanh\left(\sqrt{2g}x_0\,(\tau - \tau_0)\right) \quad ,$$

plotted in the $x - \tau-$plane in Fig. 1.5(a). This solution, the instanton $(+)$ or anti-instanton $(-)$ of the double-well potential, describes tunnelling between the classically degenerate ground states $\pm x_0$, hence lifting the degeneracy. Fig. 1.5(b) shows the time derivative of x as a function of Euclidean time τ. The distribution is sharply peaked at τ_0: the instanton solution is localised in (Euclidean) space-time. Tunnelling occurs very rapidly, almost instantaneously.

The double-well potential may also be treated using the Feynman path integral method. Because that approach can be directly transferred to the mathematically more subtle case of QCD, it will be outlined in the following. The path integral method allows the probability amplitude for the transition from initial state $x(t_i) = x_i$ to final state $x(t_f) = x_f$ to be calculated as an integral over all paths $[Dx]$ between these states, weighted with their action $S(x)$:

$$\langle x_f\,|\exp\left(-iH\,(t_f - t_i)\right)|\,x_i\rangle = \mathcal{N}\int_{x_i}^{x_f}[Dx]\exp\left(iS(x)\right) \quad ,$$

where H is the Hamiltonian and $\mathcal{N}$ a normalisation factor.
For Euclidean time this turns into

$$\langle x_f\,|\exp\left(H\,(\tau_f - \tau_i)\right)|\,x_i\rangle = \mathcal{N}\int_{x_i}^{x_f}[Dx]\exp\left(-S_E(x)\right) \quad ,$$

where the Euclidean action $S_E(x)$ is obtained from the one in Minkowski space as

$$S = \int dt L\left(x, \dot{x}\right) \Rightarrow iS_E \quad .$$

It depends on the Lagrangian, which for the double-well potential can be written as

$$L\left(x, \frac{dx}{dt}\right) = \frac{1}{2}\left(\frac{dx}{dt}\right)^2 - V(x) \quad .$$

That the action is finite shows that the instanton solution has a finite transition probability. For the path integral method, the classical (instanton) solutions correspond to the paths with minimal action in Euclidean space-time, which dominate the path integral. In Minkowski space-time one would instead have to integrate over various smaller contributions, which is unfeasible in practice. So using Euclidean instead of Minkowski space-time can be thought of as choosing a reference frame suited to the problem in the sense that it yields a mathematically concise description.

1.3.2 Instantons in QCD

QCD is a non-abelian gauge theory. The gauge field transforms as

$$\vec{A'} = U\left(\vec{x}\right)\vec{A}U^\dagger\left(\vec{x}\right) - iU\left(\vec{x}\right)\vec{\nabla}U^\dagger\left(\vec{x}\right) \quad ,$$

with a transformation $U\left(\vec{x}\right) \in SU(3)$. The vacuum is defined to be the state with minimum field energy. Other than in QED, where just one such state exists, there are infinitely many degenerate such states in QCD due to the non-abelian nature of QCD. From the transformation rule for the gauge field one can deduce that one obtains minimal field energy not only if $\vec{A} = 0$, but for all pure gauges

$$\vec{A}_{\text{vac}} = -iU\left(\vec{x}\right)\vec{\nabla}U^\dagger\left(\vec{x}\right) \quad .$$

Further, in $SU(3)$ one can distinguish between different types of gauge transformations $U\left(\vec{x}\right)$: some of these can be continuously deformed[2] into each other, others can not. All gauge transformations, which *can* be continuously deformed into each other, form a homotopy class. There are infinitely many of these classes. The different classes can be distinguished by their topological winding number

$$n = 0, \pm 1, \pm 2, \ldots \quad .$$

Not only are there different types of gauge transformations $U\left(\vec{x}\right)$, but also different types of vacua associated with the homotopy classes: The winding number of a gauge field A is given by the Chern-Simons number [Rin94]

$$N_{\text{CS}} = \frac{g^2}{16\pi^2}\int d^3x \epsilon_{\text{ijk}}\left(A_i^a \partial_j A_k^a - \frac{g}{3}\epsilon_{\text{abc}} A_i^a A_j^b A_k^c\right) \quad ,$$

---

[2] Rotation is an example of a continuous transformation, reflexion an example of a non-continuous one.

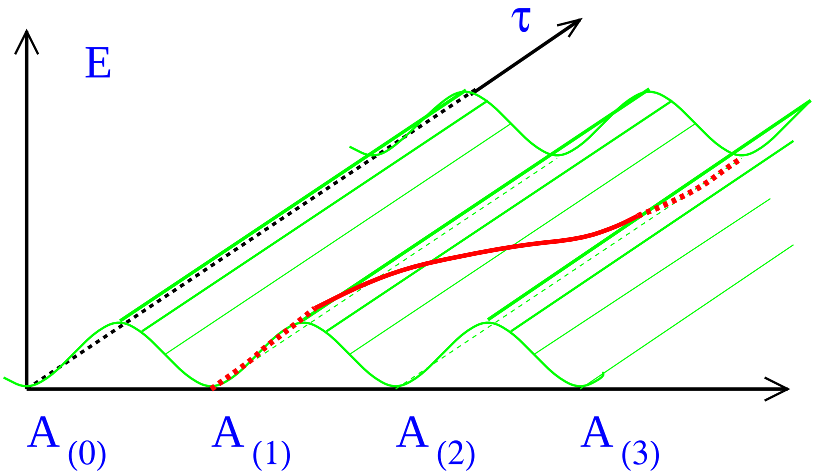


Figure 1.6: *The different QCD vacua, indicated by different pure gauges $A_{(\mathrm{n})}$, where n is the winding number, are separated by energy barriers. The degeneracy of the vacua is lifted by instantons, tunnelling solutions between "neighbouring" vacua ($\Delta N_{\mathrm{CS}} = 1$). Figure adapted from [For00].*

where ϵ_{ijk} is the totally antisymmetric tensor,

$$\epsilon_{\mathrm{ijk}} = \begin{cases} +1 & \text{for ijk an even permutation of 1, 2, 3} \\ -1 & \text{for ijk an odd permutation of 1, 2, 3} \\ 0 & \text{otherwise} \end{cases} .$$

Each type of vacuum comprises all pure gauges $\vec{A}_{\mathrm{vac}}$, which are obtained from the various gauge transformations $U^{(\mathrm{n})}(\vec{x})$ of one specific homotopy class (n) and is separated by energy barriers from those fields $\vec{A}_{\mathrm{vac}}$, which one obtains from a transformation $U^{(\mathrm{m})}(\vec{x})$ of any different class (m). Hence one arrives at the conclusion that the ground state of QCD should be infinitely degenerate. However, this is not the case. In the same way as in the simple quantum-mechanical example, the degeneracy between the states is lifted by tunnelling, i.e. instantons. In QCD, instantons describe tunnelling between neighbouring vacua, separated from each other by only one energy barrier as depicted in Fig. 1.6 and differing by only one unit of the topological winding number N_{CS}. The instanton is therefore assigned a topological charge $Q = \Delta N_{\mathrm{CS}} = +1$. (There are also anti-instantons with $Q = -1$.)

1.3.3 Instanton size and distance

The one-instanton solution of the classical Euclidean field equation in QCD can be obtained analytically [Bel75]. If the Lagrange density $\mathcal{L}\left[A_\mu^{(\mathrm{I})}(x)\right]$ is calculated with

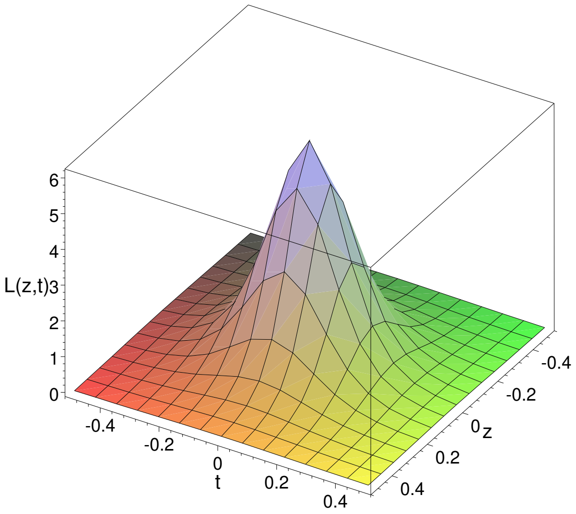


Figure 1.7: *Lagrange density of an instanton at* $x_0 = 0$*, from [Schr98].*

this known instanton solution, one finds a peaked behaviour

$$\mathcal{L}\left[A_\mu^{(\mathrm{I})}(x,\rho)\right] = \frac{12}{\pi\alpha_s}\frac{\rho^4}{\left(x^2+\rho^2\right)^4} \quad ,$$

as depicted in Fig. 1.7 . One should note that $\mathcal{L}$ only depends on the square of the coordinate x in Euclidean space-time. The instanton size ρ is a measure of the width of the Lagrange density distribution. From $\mathcal{L}$ one would find the Euclidean action S_E to be

$$S_E = \frac{2\pi|Q|}{\alpha_s} \quad ,$$

independent of ρ. It is determining the tunnelling amplitude

$$T = f\left(\alpha_s\right)\exp\left(-S_E\right)(1+\text{corrections})$$

and hence the cross sections

$$\sigma \propto \exp\left(-\frac{4\pi}{\alpha_s}\right) \quad .$$

One sees that the tunnelling is exponentially suppressed, as one might expect from quantum mechanics. However, due to the large QCD coupling α_s, the QCD instanton cross section is nevertheless sizable[3]. To make predictions of experimentally observable effects of QCD instantons in DIS one is not primarily interested in the exact form of the one-instanton solution, but has to take into account that Euclidean space is filled with many instantons of various sizes at some distance R from each other. This is illustrated in Fig. 1.8, which shows the Lagrange density as a function of one spatial

---

[3] In electroweak theory, which is also non-abelian and hence also has instanton solutions, the respective cross section is much more strongly suppressed than the one for QCD, as $\alpha \ll \alpha_s$.

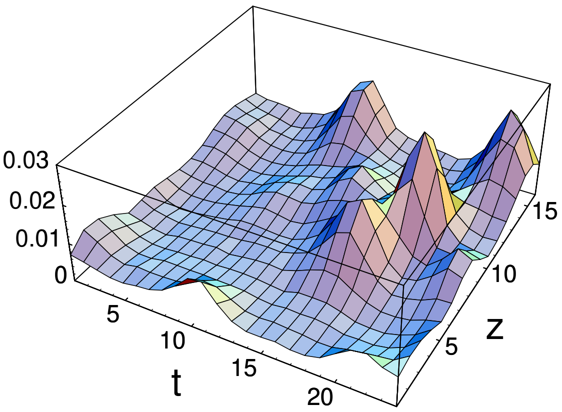
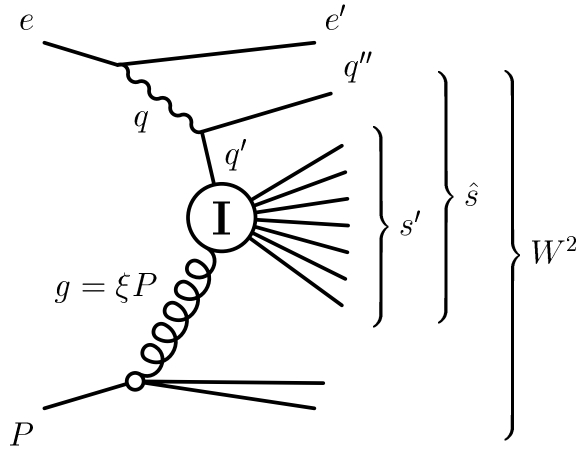


Figure 1.8: *Lagrange density of a slice of space time containing several instantons (size ρ, distance R). Figure from [Chu94].*

coordinate and Euclidean time obtained from a lattice calculation [Chu94].
For the kinematic range accessible in deep inelastic ep scattering at HERA, the predicted effects are induced by instantons of small size compared to their distance, and hence the dilute gas approximation is valid. It is then possible to calculate the cross section of the instanton-induced hard subprocess using instanton-perturbation theory. This is performed by the Monte Carlo generator QCDINS, cf. section 3.1.

1.3.4 Characteristics of instanton-induced events



DIS variables:
$Q^2 = -q^2 = -(e - e')^2$
$x = Q^2/(2P \cdot q)$
$W^2 = (q + P)^2 = Q^2(1 - x)/x$
$\hat{s} = (q + g)^2$
$\xi = x(1 + \hat{s}/Q^2)$

Variables of the instanton subprocess:
$Q'^2 = -q'^2$
$x' = Q'^2/(2g \cdot q')$
$s' = (q' + g)^2 = Q'^2(1 - x')/x'$

Figure 1.9: *Kinematics of instanton-induced events at HERA, from [Kuhl97]*

Fig. 1.9 shows the kinematics of instanton-induced events at HERA. The incoming lepton emits a photon, which in turn transforms into a quark-antiquark pair. One of these hadronises to form the current jet. The other one, denoted by q', fuses with a gluon from the proton. The phenomenological characteristics of instanton-induced events, as worked out in [Rin94], can be summarised as follows: In the hard subprocess exactly one $q\bar{q}$ pair of each of the n_f quark flavours, which are kinematically accessible, participates in the hard subprocess of each event, see Fig. 1.10. This feature is often termed “flavour democracy”. The number of gluons emitted per event follows a Poisson distribution with average value 3. The gluons plus $2n_f - 1$ (anti-)quarks give rise to a high multiplicity final state of high transverse energy. Particles stemming from the instanton are expected to be isotropically distributed in their centre of mass frame. Therefore, identification of events possibly caused by instantons requires a good separation of the instanton from the current jet and would profit from a good reconstruction of the instanton centre of mass frame.

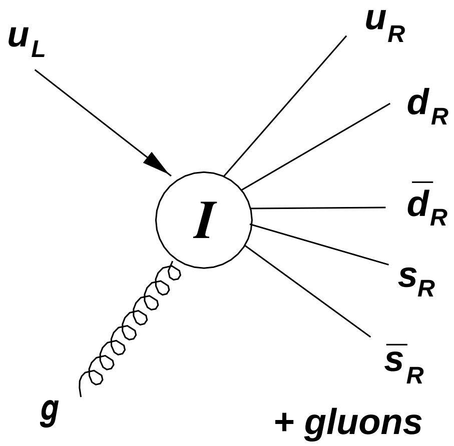


Figure 1.10: *The hard subprocess of instanton-induced events. Figure from [Ger98].*

The most prominent characteristic of instanton-induced events from a theoretical point of view, but also the one most difficult to observe experimentally, is chirality violation. In instanton-induced events, all quarks produced in the hard subprocess are emitted with the same handedness. This means that the axial charge Q_5, defined as the difference between the numbers of right handed and left handed fermions, $Q_5 = \#R - \#L$, changes by the maximal possible amount of $\Delta Q_5 = 2n_f Q$ in each event.

Chapter 2

The ZEUS detector at the HERA collider

2.1 The electron proton collider HERA

The Hadron-Elektron-Ring-Anlage (HERA) [HERA81] is the only accelerator world wide, in which electrons and protons are collided. The machine consists of two storage rings mounted on top of each other in a 6.3 km tunnel $15-25$ m underground, see Fig. 2.1.

The injection of the two beams is accomplished by a series of preaccelerators, some of which already existed when HERA was planned. Electrons are accelerated in a

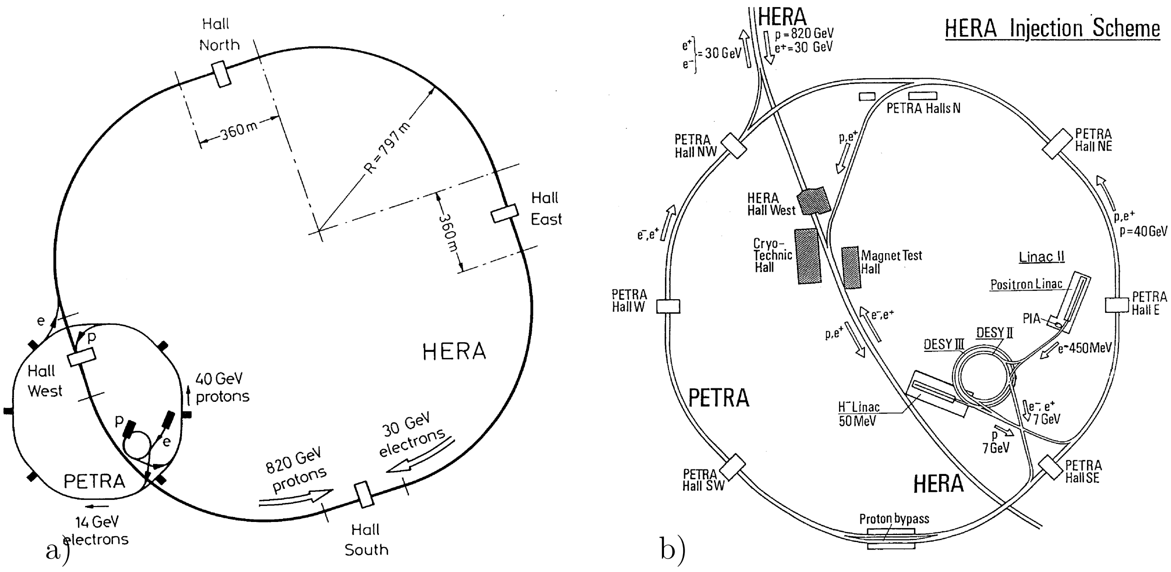


Figure 2.1: *(a): the HERA storage ring and (b): its preaccelerators.*

system of two linear accelerators, reaching energies of 220 MeV in the first (LINAC 1) and 450 MeV in the second (LINAC 2), respectively. Positrons are accumulated in the positron intensity accumulator (PIA), from which they are transferred to DESY II, where they are accelerated to an energy of 7 GeV. Up to 70 bunches, already with the HERA spacing of 96 ns, are then accelerated in PETRA II to 14 GeV, and finally injected into HERA.
The proton beam is obtained by accelerating H^- ions up to 50 MeV in a linear accelerator (LINAC). After removing the electrons, the protons are consecutively injected into DESY III and PETRA, where they are accelerated to 7.5 GeV and 40 GeV, respectively.
The final increase in energy in the HERA accelerator itself yields an electron beam of 27.5 GeV and a proton beam of 820 GeV (920 GeV since 1998), corresponding to a centre of mass energy of approximately 300 GeV. The magnetic field of 4.68 T, necessary to bend the high momentum proton beam in the four curved sections of the storage ring, is supplied by superconducting magnets, operating at 4.4 K. The lifetime of the proton (positron) beam is of the order of a few days (≈ 8 hours).
While up to 210 bunches of each particle type can be stored in HERA, the actual

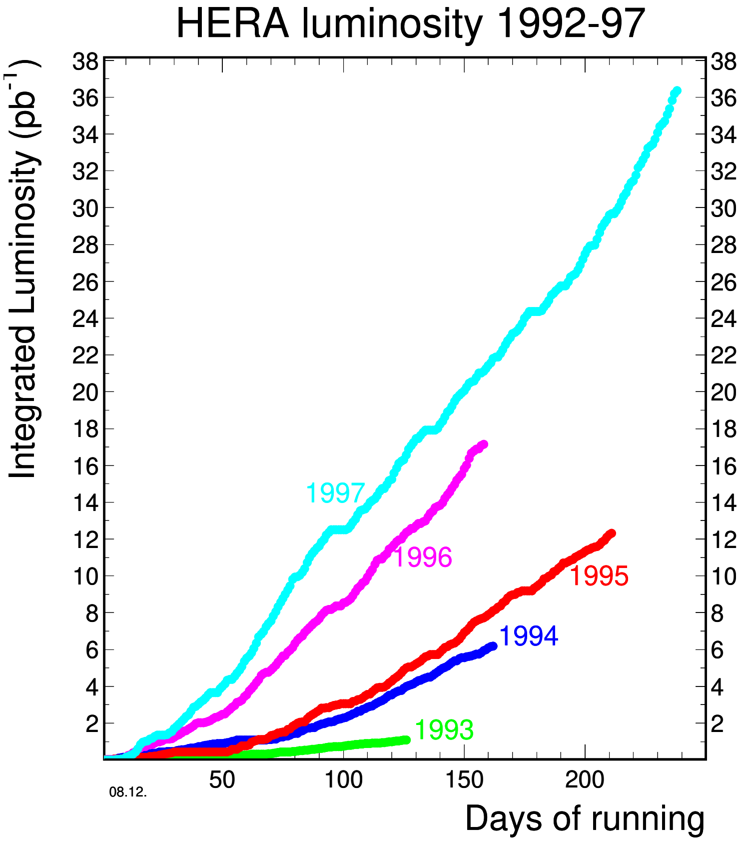


Figure 2.2: *Luminosity delivered by HERA in the years 1993 to 1997.*

HERA parameters	**Design Values**		**1996**		**1997**	
	e^-	p	e^+	p	e^+	p
Energy [GeV]	30	820	27.5	820	27.5	820
$\mathcal{L}_{\text{int}}$ [pb^{-1}]	35		17.6		36.35	
Current [mA]	58	163	23	60	30	75
Number of bunches	210		$177 + 28$	$177 + 3$	$176 + 20$	$176 + 5$

Table 2.1: *Some design parameters of the HERA accelerator and average values found during the '96/'97 data taking period. Both the number of colliding bunches and the number of non-colliding bunches are given. From [Kl01].*

number is lower most of the time, allowing for background studies using non-colliding bunches for beam-related background and empty bunches for background due to cosmic muons. Table 2.1 gives an overview over some HERA design parameters and the actual performance during the 1996/1997 data taking period. The integrated luminosity, $\mathcal{L}_{\text{int}} = \int \mathcal{L}\text{d}t$, has steadily been increased over the years as can be seen from Fig. 2.2.

Four experiments are located at the straight sections of the HERA storage ring: the two collider experiments H1 and ZEUS, and the fixed target experiments HERMES to study the spin structure of the nucleon and HERA-B. At HERMES, longitudinally polarised electrons are inelastically scattered by a polarised gas jet target. HERA-B uses a wire-target in the halo of the proton beam. All four experiments, as well as HERA itself, underwent a major upgrade in 2000 [HERA98] in order to increase the luminosity by a factor of five and to improve the detector performance.

2.2 The ZEUS detector

The ZEUS detector [ZEUS93] is an almost hermetic multipurpose detector designed to identify particles produced in *ep*-scattering events and to measure their momenta and energies. Since the detector was first commissioned in 1992 it has continuously been modified to improve its performance and to extend the range of observables measured. The ZEUS coordinate system is an orthogonal, right-handed system with the z-axis, the 'forward direction', being defined by the direction of the proton beam. The x-axis points towards the centre of the HERA-ring, the y-axis points upwards.
As the energies of electron and proton differ strongly, the centre of mass system of the reaction is moving with respect to the laboratory frame. The asymmetric detector design, shown in Fig. 2.3, accounts for this effect. In the data taking period considered, the innermost detector component surrounding the nominal interaction point[1] was the central track detector (CTD), described in section 2.2.1. Besides the CTD, the track-

---

[1] The vertex detector VXD had been removed in 1995 and was replaced by a silicon microvertex detector (MVD) in 2001.

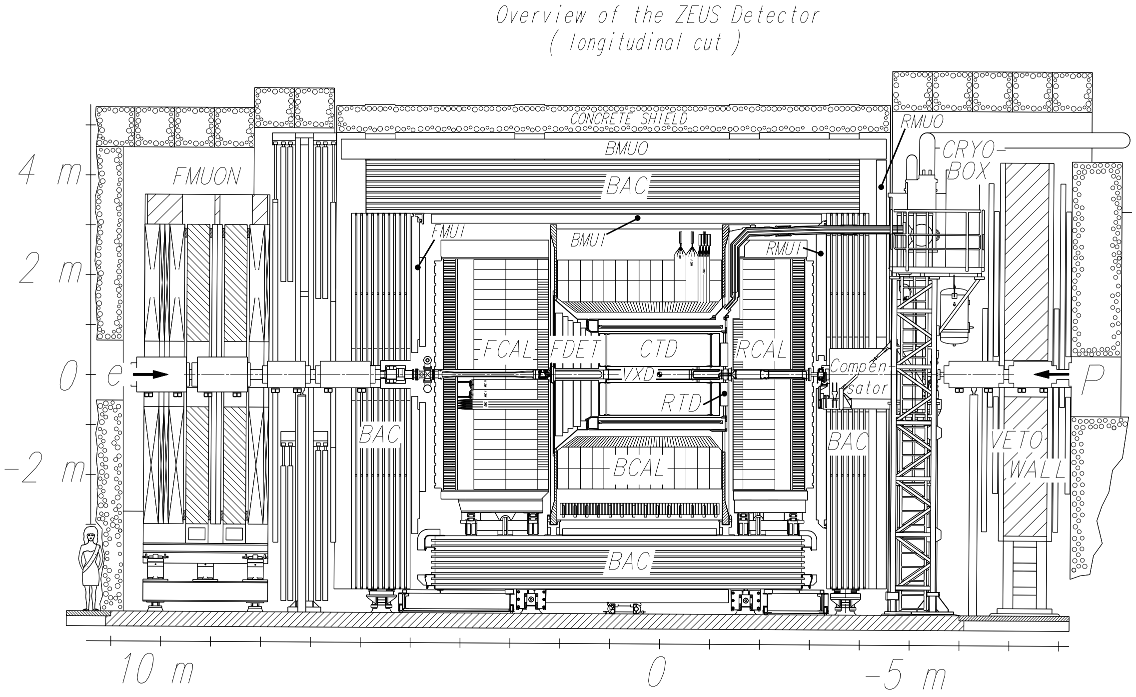


Figure 2.3: *A cross section through the ZEUS detector*

ing system comprises the FDET in the forward- and the RTD in the rear direction. While the RTD is a planar drift chamber, the forward tracking detector consists of three parts, FTD1 (the innermost) to FTD3. In the gaps between these, transition radiation detectors (TRD) are located. The tracking system is enclosed in a solenoid providing a magnetic field of 1.43 T. Momentum and charge of particles traversing the CTD can be determined from the curvature of their tracks in this magnetic field.

Energies of both charged and neutral particles are measured in the calorimeter (CAL), cf. section 2.2.2. Before particles enter the calorimeter, they pass the presampler (PRES), one layer of scintillator tiles, which allow correcting for energy losses that are caused by showering in dead material in front of the CAL. The hadron-electron-separator (HES) consists of silicon diodes, placed 2 cm below the surface of the forward and rear part of the calorimeter in order to improve the discrimination between electromagnetic and hadronic showers for particles of energies below 5 GeV.

The calorimeter is surrounded by a yoke of 7.3 cm thick iron plates, serving as return path for the magnetic flux of the solenoid. Besides, the iron-yoke is used as absorber for the backing calorimeter (BAC). In a set of proportional chambers, the BAC measures energies of showers which extend beyond the main calorimeter. Muons pass all the sub-detectors mentioned so far. They can be observed in the muon detectors FMUI, BMUI and RMUI on the inner side of the BAC and in the FMUON, BMUON and RMUON on its outer side.

Supplementary to the central detector described so far, the ZEUS experiment extends in both directions along the beam pipe, where several components have been added to increase the angular range of measurement. Between RTD and RCAL, the small angle rear tracking detector (SRTD) with a radius of $\approx 340\,\mathrm{mm}$ has been constructed around the beam pipe. In the RCAL beam hole, an electromagnetic beam pipe calorimeter (BPC) was installed in 1995, followed in 1997 by the beam pipe tracker (BPT), located in front of the BPC towards the main detector to improve the position resolution of the BPC.

In the forward direction, the forward plug calorimeter (FPC) was installed in 1998, extending the calorimeter coverage by one unit in pseudorapidity $\eta = -\ln \tan \theta/2$. The lead-scintillator strip detectors forming the proton remnant tagger (PRT) are located at $z = 5.1\,\mathrm{m}$ in the forward direction, covering a pseudorapidity range of $4.3 < \eta < 5.8$. The leading proton spectrometer (LPS) consists of six separate silicon strip detectors, placed into the gaps between HERA beam magnets at distances between 24 m and 90 m from the interaction point. Forward neutrons can be detected in the lead-scintillator calorimeter FNC at $z = 103.6\,\mathrm{m}$. Three small electromagnetic calorimeters, the 8 m-, 35 m- and 44 m taggers, serve to measure electrons close to the beam pipe. Also the luminosity monitor is located near the forward beam pipe, see section 2.2.3.

In the rear direction, at $z = -7.3\,\mathrm{m}$, two layers of scintillator hodoscopes surrounding an iron wall (VETO) are used to reject proton-beam related background at the trigger level. The trigger system, outlined in section 2.2.4, is distributed over the entire detector volume.

2.2.1 The Central Track Detector (CTD)

The central track detector (CTD) [CTD] is a cylindrical wire drift chamber. Signals are collected on sense wires, located at regular intervals, while additional wires provide gain and drift fields. The CTD active volume is filled with a gas mixture consisting of 89% argon, 5% CO_2 and 12% ethane. It has a length of 2030 mm with an inner radius of 190 mm and an outer radius of 785 mm, respectively, covering a range of $11.6° < \theta < 168.0°$ in the polar angle θ. This is equivalent to pseudorapidities η within $2.31 > \eta > -2.26$.

Charged particles traversing the CTD ionise the gas along their path, producing $10 - 20$ ions/ cm for minimum ionising particles. With a drift velocity of about $50\,\mu\mathrm{m/ns}$, the freed electrons move towards the sense wires, where a stronger field gives rise to an increase in the number of electrons by a factor of $\approx 10^4$, hence yielding a measurable signal. Sense wires are read out every 9.6 ns with 8 bit precision. The timing information is used to reconstruct the hit position, which for a single hit is ambiguous, as it may be located on either side of the wire. Only the track reconstruction, using information from many wires, allows one to distinguish between real and ghost hits.

Sense wires are arranged in 9 superlayers, each consisting of 8 layers. Within each superlayer, 8 sense wires form a cell, along with the wires necessary to provide the gain and drift fields. Altogether, the CTD contains 576 cells. One octant of the CTD is

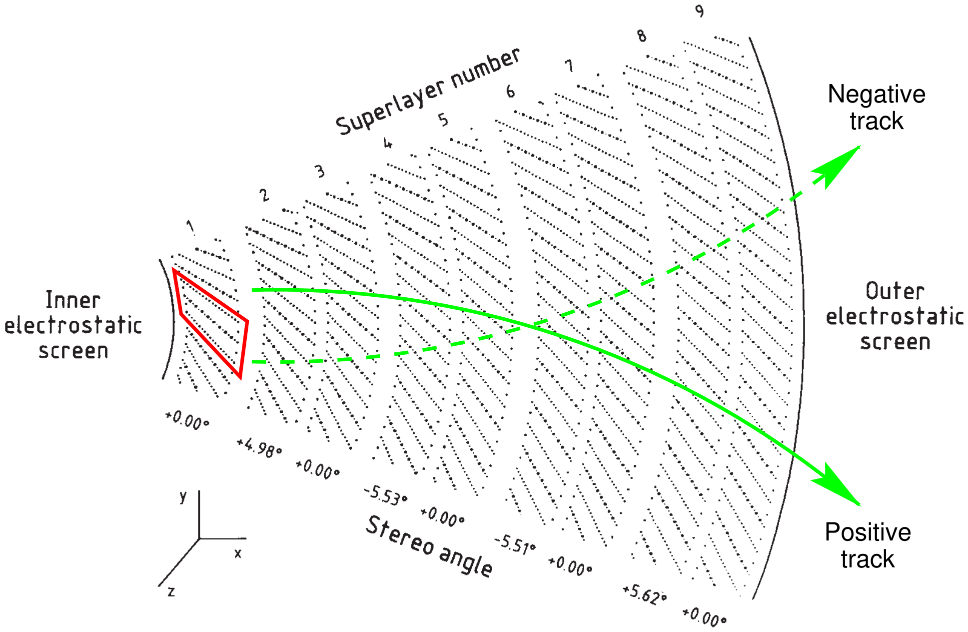


Figure 2.4: *An octant of the Central Track Detector (CTD). Examples of paths traversed by negative and positive tracks are shown, as well as the contour of a cell.*

shown in Fig. 2.4. Cells are inclined by 45° with respect to the radial direction. This ensures that straight tracks from the primary vertex have to cross the wire planes of the cells they pass under a sizeable angle. Therefore, hits are obtained in the vicinity of one or more sense wires per cell, and, for high momentum tracks, in more than one cell per superlayer, helping to disentangle the ambiguity between real and ghost hits. While wires run parallel to the CTD axis in the odd-numbered superlayers, the "axial" layers, they are inclined by $\pm 5°$ in the even-numbered, "stereo", layers. The z position resolution of single tracks obtained from the stereo layers is $1.0 - 1.4\,\mathrm{mm}$, yielding an improved vertex resolution of $\approx 2\,\mathrm{mm}$. However, this information is only available for the third level trigger (cf. section 2.2.4), after reconstruction. For a fast trigger decision at the first trigger level, a vertex position estimate with an accuracy of $\approx 44\,\mathrm{mm}$ is used. It is obtained from the "z-by-timing" system installed in the first three axial layers. In the (r, ϕ)-direction, the position resolution of the CTD is $120 - 130\,\mu\mathrm{m}$ for single tracks and $\approx 1\,\mathrm{mm}$ for the event vertex.
From a study of simulated tracks from the $K2\pi D^*$ channel [HWt99], the transverse momentum (p_t) resolution of tracks with a $p_t > 150\,\mathrm{MeV}$, stemming from the main vertex and traversing at least three superlayers, was found to be

$$\sigma(p_t)/p_t = 0.0058\; p_t \oplus 0.0065 \oplus 0.0014/p_t \quad , p_t \text{ in GeV} \quad .$$

The first term represents the intrinsic resolution of the CTD while the second and third terms account for multiple scattering of charged particles inside and in front of the CTD, respectively.

2.2.2 The Uranium Calorimeter (CAL)

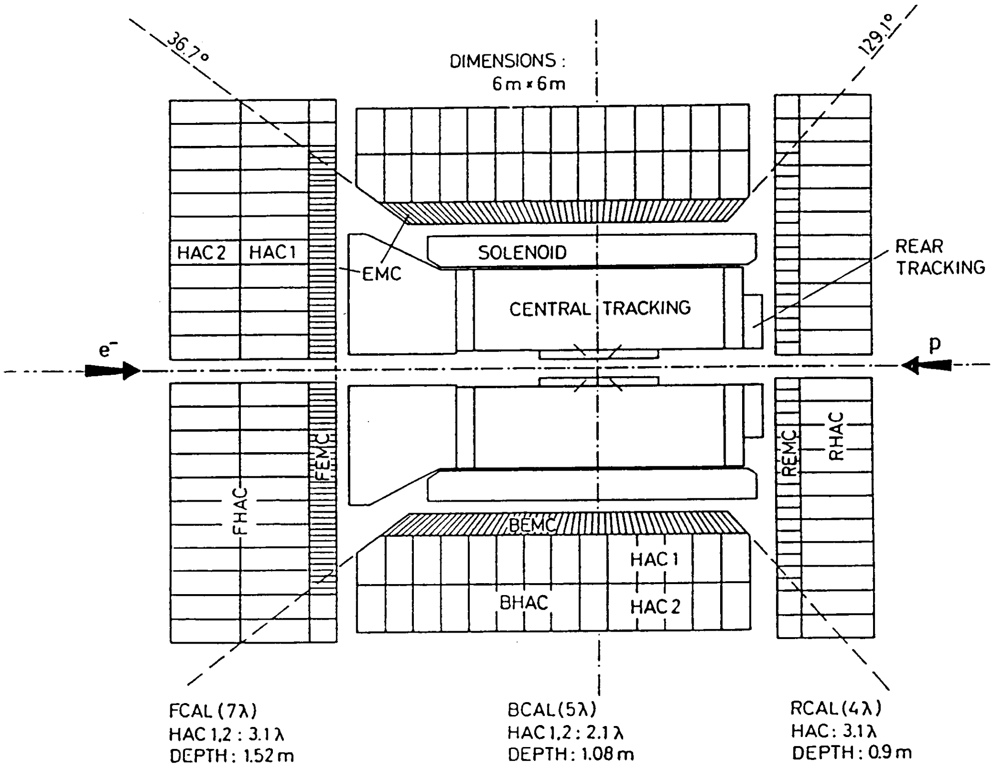


Figure 2.5: *Schematic view of the ZEUS Uranium Calorimeter (CAL). One can discern the three parts (F/B/RCAL) and their subdivision into EMC and HAC sections, see text.*

The CTD is surrounded by the calorimeter [CAL91], consisting of a forward (FCAL), a barrel (BCAL) and a rear (RCAL) part, as shown in Fig. 2.5. Each part is subdivided into an electromagnetic (EMC) and one (RCAL) or two (FCAL, BCAL) hadronic (HAC) sections. The BCAL-EMC is the only calorimeter section which is arranged projectively towards the interaction point, each wedge covering an azimuthal angle of 11.25°. The overall solid angle coverage of the calorimeter is 99.8% in the forward direction and 99.5% in the backward direction.

In the EMC sections, incident electrons and photons cause bremsstrahlung and pair production processes, which initiate electromagnetic showers. For electromagnetic energy deposits, the energy resolution determined in test beam measurements is

$$\frac{\sigma_E}{E} \simeq \frac{18\%}{\sqrt{E(\mathrm{GeV})}} \quad .$$

The development of hadronic showers in the HAC sections is more complex. While

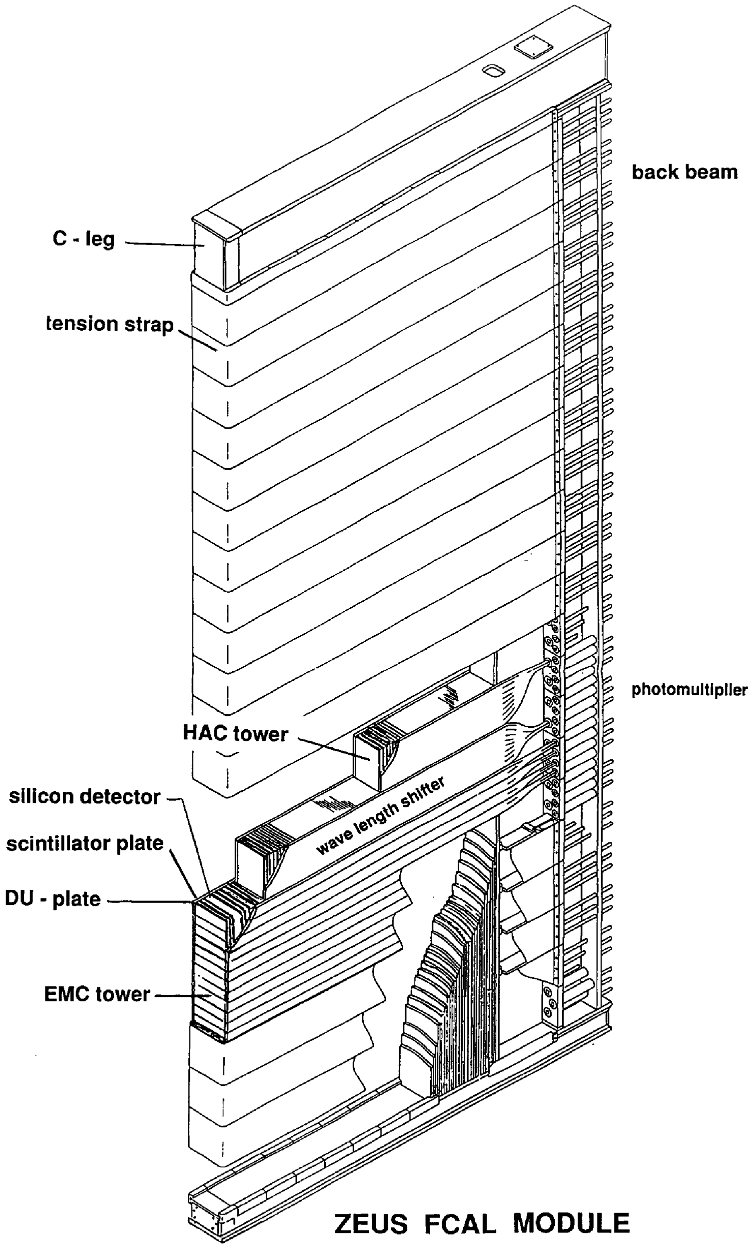


Figure 2.6: *A module of the forward calorimeter.*

starting mainly from inelastic scattering of hadrons off the nuclei in the absorber, hadronic showers propagate via around 300 different processes of similar cross section [Her90], which limits the precision of the hadronic energy measurement. Precision is further reduced by the fact that compared to the electromagnetic case a higher fraction of energy is lost without contributing to the signal, e.g. by intranuclear collisions and

neutral particles produced in the cascade.
The hadronic energy resolution of the ZEUS calorimeter of

$$\frac{\sigma_E}{E} \simeq \frac{35\%}{\sqrt{E(\mathrm{GeV})}} \quad ,$$

found in test beam measurements, is at the limit of what is feasible given the intrinsic restrictions. It is attained by using depleted uranium as absorber and by choosing the thicknesses of absorber and active material such as to achieve optimal compensation for the loss of signal from hadronic showers. Uranium plates of one radiation length $X_0 = 3.3\,\mathrm{mm}$ are sandwiched with 2.6 mm of organic scintillator to sample the energy deposition. The dark current stemming from the radioactive decay of the depleted uranium is used for calibration of each channel, performed once a day.
Besides the energy measurement, the calorimeter provides information on the position of the incident particles. The position resolution depends on the CAL granularity. EMC and HAC sections are built up of cells arranged into towers. Towers in turn are assembled to modules, as depicted in Fig. 2.6. EMC cell sizes vary from $5 \times 10\,\mathrm{cm}^2$ in the inner region of the RCAL to $20 \times 20\,\mathrm{cm}^2$ in the external EMC regions. HAC cells are typically sized $20 \times 20\,\mathrm{cm}^2$.
From each side of a cell, signals are transferred to a set of photomultipliers (PMTs) via wavelength shifters and optical fibres. A comparison of the two signals provides information on the horizontal impact position of a particle.
The time of incidence can be measured at a level of $\leq$1 ns for energies above 4.5 GeV. It is used by the trigger system to reduce background due to beam gas events.

2.2.3 The Luminosity Monitor

A measurement of the time-integrated luminosity $\int \mathcal{L}\mathrm{d}t$ is required for calculating cross-sections from a given number of measured events. At ZEUS, the luminosity is determined from the rate of the Bethe-Heitler process $ep \rightarrow ep\gamma$, for which the cross-section is known to within 0.5% from QED calculations.
The rate of this process is determined by detecting its photon-component in a lead-scintillator calorimeter [And92], located at $z = -107$ m, as shown in Fig. 2.7. The energy resolution of the LUMI-γ detector was found to be $18\%/\sqrt{E(\mathrm{GeV})}$ in test beam measurements, but within the ZEUS experiment it is degraded to only $23\%/\sqrt{E(\mathrm{GeV})}$ by a carbon lead filter needed as protection against low energetic synchrotron radiation photons. The measurement of the photon rate is corrected for background due to bremsstrahlung of electrons with beam gas, using the measurements for empty proton bunches as a reference.
Another detector, LUMI-e, constructed similarly and placed at $z = -34$ m, was intended to measure the electron-component in coincidence with LUMI-γ. As the LUMI-γ detector alone yields an accurate luminosity value [Piot94] within 1.5%, the LUMI-e

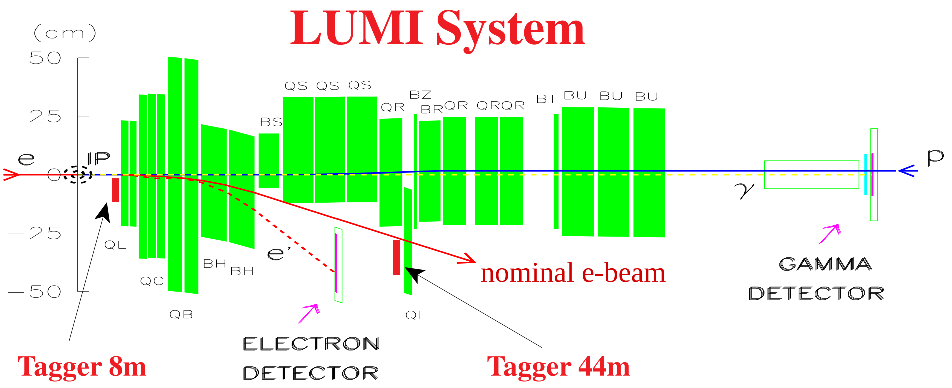


Figure 2.7: *The luminosity detectors and electron taggers along the beam-pipe.*

system is now used to detect outgoing electrons, scattered under very low angles in photoproduction events.

2.2.4 The Trigger and Data Acquisition System

The $2.5 \cdot 10^5$ channels of the ZEUS detector are read out for each bunch crossing, i.e. at 96 ns intervals. Compared to this 10 MHz readout frequency, interesting physics events occur only seldom, namely with a rate of a few Hz. The main source of background is formed by upstream collisions of protons with residual beam gas at a rate of $10 - 100$ kHz. Further sources of background are collisions of the electron beam with beam gas, beam halo interactions and cosmic ray events. A three level trigger system [YoSm], is employed to distinguish between background and potential physics events, reducing the rate by several orders of magnitude, see Fig. 2.8.

At the first level of the trigger system (FLT) each component has its own hardware trigger, located in the vicinity of the component to reduce latencies and to keep the noise low. Local trigger decisions, based on quantities such as energy sums, timing information and signals in veto counters, are transferred to the global first level trigger (GFLT) within 2 µs. The GFLT decision is taken within 4.4 µs, during which time the event is stored in a special pipeline buffer of a length of 58 bunch crossings [Her90]. It was designed for the calorimeter, but is also used by other components. The FLT output rate is ≈ 600 Hz.

Events accepted by the FLT are digitised and sent to the second level trigger (SLT), running in parallel on a network of transputers [Bot93]. Being a software trigger, the SLT can handle more complex event properties like vertex information, limited tracking and calorimeter timing. The global second level trigger (GSLT) combines the component decisions, yielding an output rate of the order of $50 - 100$ Hz.

For events that have passed the GSLT, part of the reconstruction code is run on the event builder [Beh93], yielding detailed tracking information and performing tagging of

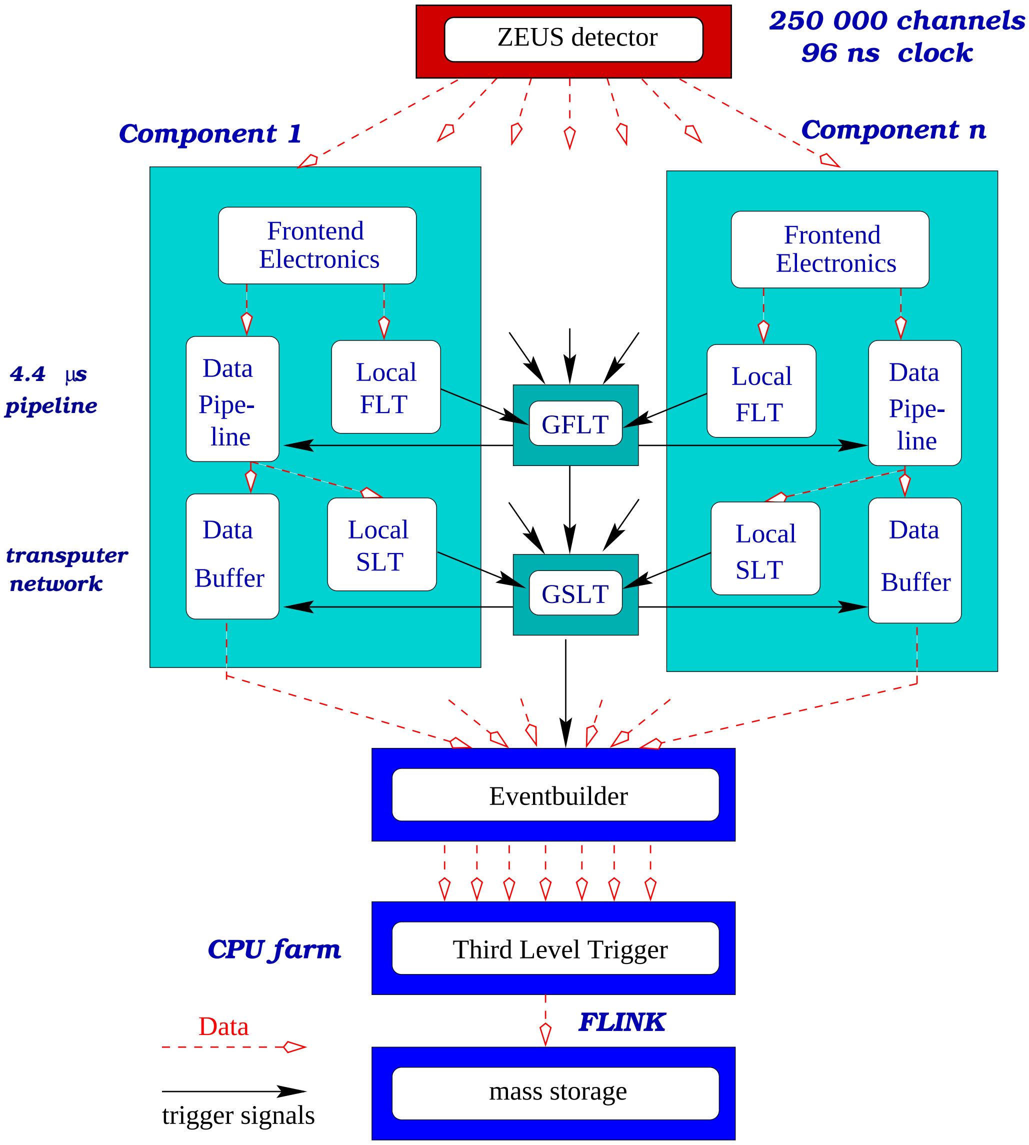


Figure 2.8: *The trigger and data acquisition system.*

the scattered DIS electron and jet finding. Based on that information, the third level trigger (TLT) decides, which events are finally sent via a fibre link (FLINK) connection to the DESY computer centre, where they are written to tape. The TLT output rate is $\approx 5 - 10\,\text{Hz}$. A typical event comprises $\approx$ 100kB information in the ADAMO format [FiPa94].
As some calibration constants are determined by averaging over each run, the full event reconstruction can only be performed offline.

Chapter 3

Monte Carlo Simulation

Statistical models, called Monte Carlo (MC) generators, are a widely-used tool to make predictions of the complex physics processes and event topologies encountered in high energy physics. The simulation is usually divided into one part modelling the hard subprocess and parton showering (fragmentation), which can be described perturbatively, and another part simulating the non-perturbative hadronisation phase. The results of both the first phase, the "parton level", and the second one, the "hadron level", are provided as tables of four-momenta by common MC programs. The hadron level results are used as input for a program which simulates the detector response, yielding a "detector level" description.

3.1 Simulation of instanton-induced events: QCDINS

Instanton-induced events are simulated using the MC generator QCDINS [Rin00] in its newest version 2.0. QCDINS simulates the hard subprocess in the background of an instanton. For the fragmentation and hadronisation phases of the event simulation, HERWIG is used in this analysis. The program QCDINS also provides the possibility to choose the JETSET hadronisation instead.
The simulation of the hard subprocess is accomplished using instanton-perturbation theory around the one-instanton solution with expansion parameter $\alpha_s(Q)$. Only the three lightest quark flavours are taken into account, as active quark flavours are required to have masses m_q that are small compared to the effective instanton-size, ρ_{eff}, that is $m_q \rho_{\text{eff}} \ll 1$. In the Bjorken limit of the instanton subprocess variables Q' and x', defined in Fig. 1.9, the cross section factorises as

$$\frac{\mathrm{d}\sigma^{(I)}_{\text{HERA}}}{\mathrm{d}x'\mathrm{d}Q'^2} \simeq \frac{\mathrm{d}\mathcal{L}^{(I)}_{q'g}}{\mathrm{d}x'\mathrm{d}Q'^2} \cdot \sigma^{(I)}_{q'g}(Q', x') \quad \text{for} \begin{cases} Q'^2 = -q'^2 > 0 & \text{large,} \\ 0 \leq x' = \frac{Q'^2}{2p \cdot q'} \leq 1 & \text{fixed .} \end{cases}$$

The term $\mathcal{L}^{(I)}_{q'g}$ gives the number of $q'g$ collisions per ep collision. It depends on the gluon density, the virtual photon flux and the flux of the virtual quark q' in the background of the instanton [Rin98]. The total cross-section $\sigma^{(I)}_{q'g}(Q',x')$ of the instanton-subprocess $q'g \xrightarrow{I} X$ forms the core of the calculation, describing the essential instanton dynamics [Schr02]. This quantity can be written in the form:

$$\sigma^{(I)}_{q'g} = \int d^4R \ \mathrm{e}^{\mathrm{i}(p+q')\cdot R} \int_0^\infty d\rho \int_0^\infty d\bar{\rho} \ \mathrm{e}^{-(\rho+\bar{\rho})Q'} D(\rho) D(\bar{\rho}) \int dU \ \mathrm{e}^{-\frac{4\pi}{\alpha_s}\Omega\left(U,\frac{R^2}{\rho\bar{\rho}},\frac{\bar{\rho}}{\rho}\right)} \{...\} \quad ,$$

where the dots in parentheses represent some additional contributions, e.g. from external partons. Instanton-induced processes, in which a quark is emitted by the proton are neglected, as they are suppressed by a factor α_s^2 with respect to the gluon-initiated process [Rin98].

The cross section $\sigma^{(I)}_{q'g}$ is obtained by integrating the (anti-) instanton sizes $(\bar{\rho})$ ρ, the instanton-anti-instanton distance four-vector R_μ and the relative $I\bar{I}$ colour orientation matrix U. The instanton size distribution $D(\rho)$ ['t Hooft, Ber79], and the instanton-anti-instanton interaction $\Omega\left(U,\frac{R^2}{\rho\bar{\rho}},\frac{\bar{\rho}}{\rho}\right)$ [Kho91, Ver91] , are known within instanton perturbation theory.

In general, a difficulty arises from the power-law behaviour of the instanton size distribution [Schr02],

$$D(\rho) \propto \rho^{\beta_0 - 5 + \mathcal{O}(\alpha_s)} \quad \text{with} \quad \beta_0 = \frac{11}{3}N_C - \frac{2}{3}n_f \ \ (N_C = 3) \quad .$$

It can give rise to an infrared divergence of the integrations over $I(\bar{I})$ sizes for large values of $\rho(\bar{\rho})$. However, deep-inelastic scattering is an exception from this general rule. In this particular case, $\sigma^{(I)}_{q'g}$ contains an exponential "form factor" $\exp^{-Q'(\rho+\bar{\rho})}$, [Moch97], which ensures convergence and small instantons at large enough Q', rendering the cross section calculable.

When the kinematic variables Q' and x' are lowered, $\sigma^{(I)}_{q'g}$ grows steeply. At some point, instanton perturbation theory ceases to be valid. The calculation implemented in the QCDINS generator has been compared to high quality lattice simulations of QCD. Cuts defining the fiducial region of the Monte Carlo have been derived [Rin99] from these lattice simulations:

$$Q' \geq {Q'}_{\mathrm{min}} = 30.8 \ \Lambda^{(n_f)}_{\overline{\mathrm{MS}}}$$

$$x' \geq {x'}_{\mathrm{min}} = 0.35 \quad .$$

The value for $n_f = 3$ flavours,

$$\Lambda^{(3)}_{\overline{\mathrm{MS}}} = 0.346 \ {}^{+0.031}_{-0.029} \ \mathrm{GeV} \quad ,$$

was found from a standard three-loop perturbative flavour reduction from the 1998 world average of the running coupling [PDG98],

$$\alpha_s(\mathrm{M_Z}) = 0.119 \pm 0.002 \quad \overset{\text{3 loop}}{\Longleftrightarrow} \quad \Lambda^{(5)}_{\overline{\mathrm{MS}}} = 0.219 \ {}^{+0.025}_{-0.023} \ \mathrm{GeV} \quad .$$

This results in a lower bound of ${Q'}^2 > 113\,\mathrm{GeV}^2$.

In order to restrict the QCDINS sample to the region, where the calculation is reliable, cuts on the generated x' and ${Q'}^2$ values were made when generating the QCDINS sample. No corresponding cuts could be made on nDIS MC and on the data, which therefore might contain a slightly larger contribution from instanton-induced events than predicted by the QCDINS sample used in the search. However, lattice calculations show a steep decrease of the instanton-contribution towards small $I\bar{I}$-separation, equivalent to $x' < 0.35$, suggesting that this region can be neglected [Rin00b, Rin99, Rin01].

In order to suppress non-planar contributions [Moch97], which are not taken account of in the calculation, a further cut on the photon virtuality is recommended and applied by default,

$$Q \geq Q_{\mathrm{min}} = {Q'}_{\mathrm{min}} \quad .$$

In the QCDINS sample used in the search, this cut was lowered to $Q^2 > 50\,\mathrm{GeV}^2$ when generating the sample. At the analysis stage, a cut was applied on the reconstructed photon virtuality, see chapter 4, $Q^2_{\mathrm{DA}} > 120\,\mathrm{GeV}^2$, which is above the QCDINS default value of $\approx 113\,\mathrm{GeV}^2$.

The uncertainty of the value of $\Lambda_{\overline{\mathrm{MS}}}$ gives rise to an uncertainty of the QCDINS predicted cross section of $\approx 30\%$ within the default cuts on x', ${Q'}^2$, Q^2 and additional cuts on the Bjorken variables $x \geq 10^{-3}$, $0.1 \leq y \leq 0.9$ [Rin00b].

3.2 Simulation of the normal DIS background

The precise description of the background is essential in a search for a signal expected to contribute only 0.5% to the predicted cross section. It is made difficult by the fact that the signal is located in an extreme region of phase space, in which it is not clear to what extent the assumptions, on which current models are based, are still valid. On the other hand, only in an extreme place in the tails of the background distributions there is any chance to observe such a small signal.

In this analysis, the MC generator DJANGOH [Cha94] was used as the main Monte Carlo to simulate the normal DIS background. To describe the hard subprocess, the version based on the Colour Dipole model (CDM) [Gus] as implemented in the program ARIADNE [Lön92] was chosen, because this model is known to yield overall the best description of most final state data [Kuhl97].

In the Colour Dipole model, QCD radiation is described in the form of parton showers evolving from independently radiating colour dipoles, as depicted in Fig. 3.1. In DIS, the model takes into account the spatial extension of the proton remnant. No distinction is made between initial and final state radiation, so there is no need of a special treatment of interference effects between them. Similarly, soft gluon interference is automatically described. In Ariadne, the boson-gluon-fusion process is added "by hand" applying a matrix-element approach.

To simulate the hadronisation phase, the Lund string fragmentation [Art] is applied as implemented in JETSET [Sjö94]. In this model, the colour field between $q\bar{q}$ pairs

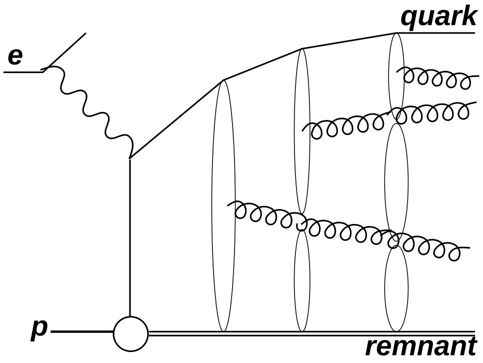


Figure 3.1: *Schematic view of the Colour Dipole model approach to simulate deep-inelastic ep scattering events. Figure from [Kuhl97].*

moving away from each other is described as a string of constant energy ($\approx 1\,\mathrm{GeV/fm}$) to take the self-interaction of gluons into account. Due to the constant energy density along the string, the total string energy increases with increasing $q\bar{q}$ distance, until it suffices to create a $q\bar{q}$ pair from the vacuum, thus splitting the string consecutively into smaller parts. The process comes to a halt when the energy of the original $q\bar{q}$ pair is exhausted.

Further aspects of DIS events that are modelled in Djangoh are the event kinematics and QED radiative corrections, which are both supplied by the program HERACLES [Kwi92]. In the determination of the kinematics, the CTEQ4 [Lai97] parton distribution functions were used.

To account for model uncertainties it is common practice not to rely exclusively on one specific Monte Carlo simulation, but to cross-check results with a different MC generator as part of the systematic studies. For this purpose, the generator HERWIG [Mar92] has been used in this analysis, because it is the model, to which the instanton MC is by default interfaced for the description of fragmentation and hadronisation.

HERWIG provides an exact calculation of the leading-order matrix element of the hard subprocess. Higher orders of perturbation theory are summed up to all orders using the DGLAP leading log approximation. A planar parton shower is generated, in which the energy fractions in each branching process are determined by the Altarelli-Parisi-splitting functions [DGLAP], and angular-ordering is taken into account. When the scale of the partons reaches a value of $\approx 1\,\mathrm{GeV}^2$, perturbation theory ceases to be applicable.

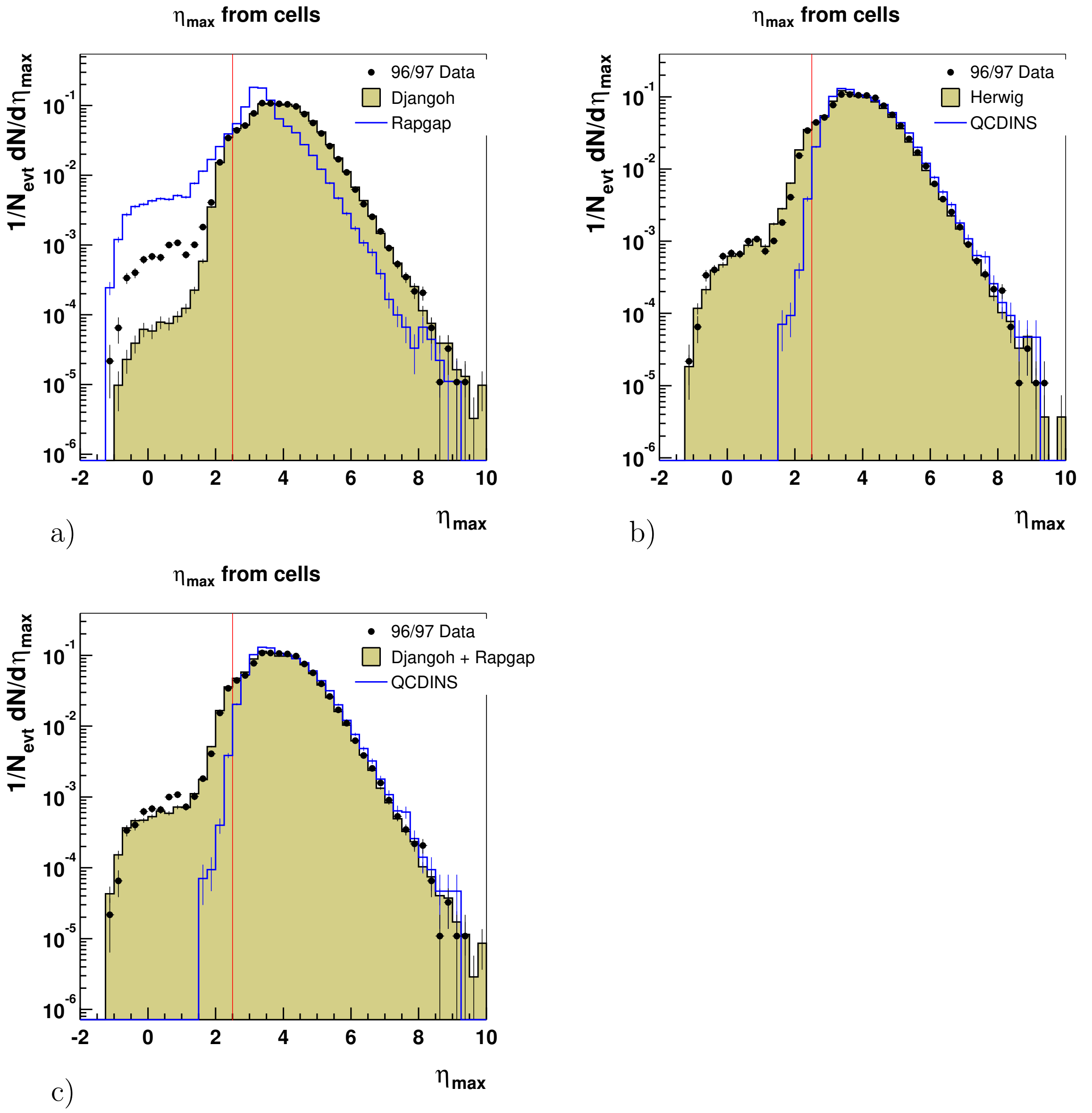
For the final part of the event simulation, HERWIG uses a cluster hadronisation model [Web84], based on the idea of preconfinement of colour [Am79]. Gluons are non-perturbatively split into $q\bar{q}$ pairs. Neighbouring quarks and antiquarks are then combined to colour-singlet clusters of low mass, which is typically two to three times the square root of the cut-off scale. While more massive clusters are first split into lighter ones, clusters of lower mass either decay isotropically into pairs of hadrons or turn into hadrons themselves by momentum exchange with a neighbouring cluster in order to adjust the mass. Finally, a decay of unstable hadrons into stable ones is simulated.

3.3 Contribution of diffractive events

In typical DIS events, a colour-field is stretched between the scattered quark and the proton remnant, which gives rise to a homogeneous rapidity distribution of final state hadrons. However, a sizable fraction of the events in the inclusive sample, so called-diffractive events, lack this colour-connection. Therefore, those events do more often than others contain a large rapidity gap between the proton remnant, which moves undetected along the beampipe, and the most forward energy deposit in the detector. The contribution of diffractive events is determined from a χ^2 fit of the distribution of $\eta_{\max}$, defined as lab frame pseudorapidity of the calorimeter energy deposit with lowest polar angle and an energy above 400 MeV. Fig. 3.2 shows the $\eta_{\max}$ distribution for the nDIS background Monte Carlos used in this analysis. While Herwig describes the distribution well by itself, a clear excess of the data over the Djangoh prediction is observed at low $\eta_{\max}$ values. In order to describe the diffractive contribution, a dedicated Monte Carlo, RAPGAP [Jun95], which is also based on Ariadne, is added to the Djangoh sample. The fit is restricted to the part of the distribution, which is sensitive to the diffractive contribution, $\eta_{\max} < 2.5$. An admixture of 12.0% diffractive MC has been found to give the best description of that region. The resulting $\eta_{\max}$ distribution is shown in Fig. 3.2 c).

3.4 Monte Carlo Tuning

Before investigating instanton-induced events, it was checked whether the Monte Carlos were able to describe the data. For this purpose, an inclusive sample was considered, as defined by cuts given in section 4.2. Within this sample, the predicted contribution from instanton-induced events, corresponding to only $\approx 0.5\%$ of the total cross section, is negligible. For the inclusive sample, distributions of various observables characterising the event kinematics and the hadronic final state were plotted. These control plots in general showed a good agreement between data and MC. However, Djangoh with Rapgap failed to describe the sphericity of the part of the hadronic final state, assigned to the instanton, defined in section 5.1. In the course of a systematic variation of some of the input parameters of that MC model ("tuning"), the adjustment of a parameter describing the width of the initial hadron p_T distribution was found to improve the description. For the analysis, this modified Monte Carlo sample has been used. The details of this study can be found in section 6.4.



a) b) c)

Figure 3.2: *Determination of the diffractive contribution from the* $\eta_{\max}$ *distribution in the range* $\eta_{\max} < 2.5$*: The* $\eta_{\max}$ *distribution of the data (dots) is compared to the one for Monte Carlo: (a) comparison to pure Djangoh (full histogram) and Rapgap, (b) comparison to Herwig (full histogram) and QCDINS, (c) comparison to a mixture of Djangoh and* 12.0% *Rapgap (full histogram) and QCDINS. All distributions are normalised to 1.*

3.5 Detector Simulation

The MC simulations described in the previous sections yield a description of the physics processes, which are expected to take place at the HERA accelerator. However, in order to compare the model predictions to the data, it is necessary to also simulate the interaction of the final state particles with the ZEUS detector. This task is performed by the program MOZART [Haa92], based on the CERN package GEANT [Br87]. A detailed description of material, geometry and relative position of the various detector components is built into the program, as well as information on the response of the components, obtained in several test beam measurements. The program keeps track of each particle, simulating e.g. energy loss, multiple scattering or decay, and generates the simulation of the according signals from the detector sub-components. That information is processed by the trigger simulation program ZGANA [Wai93], and, in case of a positive “trigger decision”, passed on to the event reconstruction package ZEPHYR [Glae91]. From that step onwards, data and MC events are technically treated equally.

Chapter 4

Event Selection

4.1 Event Reconstruction

4.1.1 Hadronic final state and electron reconstruction

Events are reconstructed from the data collected with the CTD and the calorimeter, cf. sections 2.2.1 and 2.2.2. By means of the program VCTRAK [Har97], run in CTD-only mode, trajectories in the CTD are recognised, track fits performed and the event vertex is determined. Track finding starts from a "seed" combination of hits in the

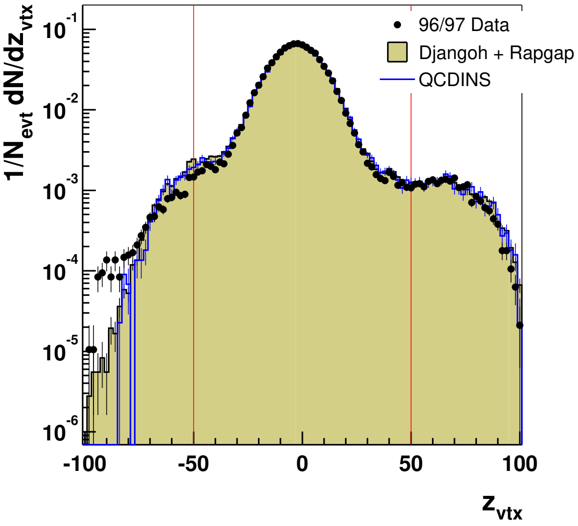


Figure 4.1: *Distribution of the z position of the event vertex, z_{Vtx}. All offline cuts except for the z_{Vtx} cut (indicated by the vertical lines) are applied, cf. section 4.2.*

outer part of the CTD and proceeds towards a very broad "virtual hit" added at the beam line. Regions of the CTD with known problems are not used at this stage. To improve the track measurement, various CTD hit corrections are performed. The final tracks are described by a five parameter helix model. The parameters are determined by "swimming" each track in the magnetic field, starting near the innermost measured hit and moving outward. In a further step, an extrapolation of the track towards the calorimeter is attempted. If possible, the entry point of the trajectory is determined. From the resulting tracks, the primary event vertex is fitted after trajectories, which are incompatible with the beamline or contribute too much to the χ^2 value, have been discarded. The distribution of the vertex z-position for the inclusive data sample used in this analysis is shown in Fig. 4.1.

For the reconstruction of energy flow objects ("efos" or "zufos"), only primary vertex-fitted tracks are used. They are required to traverse at least 4 CTD-superlayers and to have a transverse momentum, p_T, in the range $0.1\,\text{GeV} < p_T < 20\,\text{GeV}$, unless they cross more than seven superlayers, in which case the transverse momentum should be in the range $0.1\,\text{GeV} < p_T < 25\,\text{GeV}$.

zufo type	definition
0	unmatched track
1	1-to-1 match between track and CAL island
2	2 tracks match 1 island
3	3 tracks match 1 island
12	1 track matches 2 islands
22	2 tracks match 2 islands
37	1 track matches 2 islands, CAL energy and tracking angle used
41	1-to-1 match, CAL energy, tracking angle used
30	some track match, but only CAL information used
31	unmatched island

Table 4.1: *Definition of the different types of zufos, from [Tun01].*

In the reconstruction of energy deposits in the calorimeter, cells with low or zero signal or with high noise are masked. For the data, the noise due to Uranium decay is suppressed by a cut on the energy deposit in each EMC (HAC) cell at 60 MeV (110 MeV) for cells assigned to clusters and at 100 MeV (160 MeV) for isolated cells. Cells are clustered into cell-islands by linking each cell to its highest energy neighbour within the same EMC (HAC) section. In a second step, cell-islands from different sections are clustered into cone-islands depending on their angular separation. Cone-islands may comprise cell-islands from different calorimeter parts, e.g. one from the BCAL edge and one from the RCAL.

The combination of tracking and calorimeter information, see Fig. 4.2, yields the most complete reconstruction of the hadronic final state: the angular acceptance of the calorimeter exceeds that of the CTD, while the CTD reconstruction is more precise at the edges between the different calorimeter parts and at low track momenta.

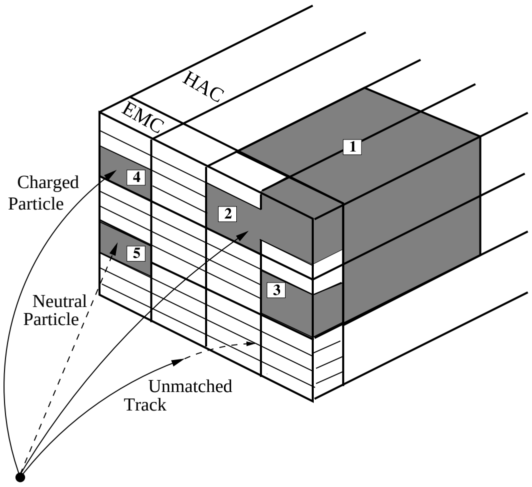


Figure 4.2: *Four-momenta of energy flow objects are calculated using tracking and calorimeter information. Schematic view taken from [Tun01].*

From CAL and CTD information, energy flow objects ("zufos") are formed using the zufosNT code [Br98, Tun01], including a backsplash correction. Tracks are assigned to cone islands on the basis of the distance of closest approach, DCA, of the track impact position on the calorimeter from the centre of the cone- island. A track is said to match an island, if the DCA is either below 20 cm or less than the island radius. Depending on the properties of islands and tracks considered and on their relative positions, ten different types of zufos can be distinguished, as specified in table 4.1. Type 0 zufos are assigned one, type 3 zufos three pion masses, respectively. Type 31 zufos are treated as massless particles, assuming these energy deposits to be caused by photons. The contributions of the different types to zufo multiplicity and scalar E_T in the hadronic centre of mass frame are shown in Fig. 4.3 for the inclusive MC samples obtained from the cuts given in section 4.2. While about 65% of the zufos are calculated from calorimeter information only, the zufo types depending on track information have on average a higher E_T. One can also see, that the relative contribution from tracks to the zufo multiplicity is higher in QCDINS events than in the nDIS MC.

From the list of zufos provided by zufosNT, the track and/or energy depositions assigned to the electron candidate and signals from calorimeter cells stemming from backsplash are removed to obtain the hadronic final state of the event.

The scattered electron is identified by the program SINISTRA [Ab95]. It is based on a neural network, which was trained using Monte Carlo events. In case the electron finder identifies more than one possible electron candidate, the decision which one of these is used, is made by means of the program FINDIS, run with option 5. That option chooses the electron candidate, which has been assigned the largest probability by the neural network. If there is another candidate of same probability, the one of higher energy is used. Depending on where the electron hits the CAL, it is required

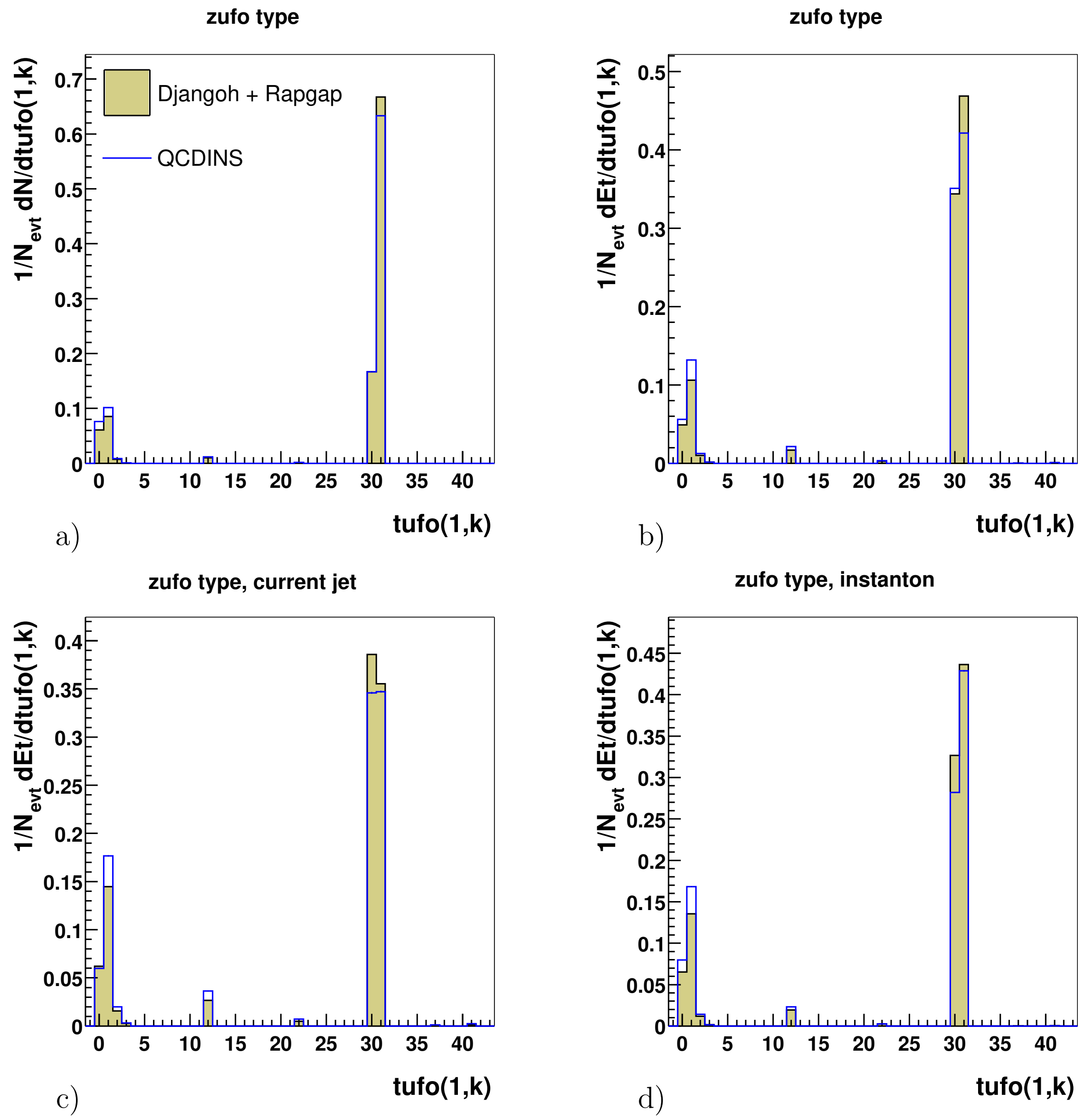


Figure 4.3: *Contributions of different zufo types to (a) zufo multiplicity and to the scalar* E_T *of (b) the hadronic final state, (c) the current jet part of the hadronic final state and (d) the instanton candidate, reconstructed as described in chapter 5.*

that a track points at the energy deposits in the calorimeter. Once the electron has been found, its four-momentum is calculated. In the high Q^2 region considered, the electron energy is given by the double angle method [Ben92]:

$$E_{\mathrm{e,DA}} = 2E_e \, \frac{\sin\gamma_h}{\sin\theta + \sin\gamma_h - \sin(\theta + \gamma_h)} \quad ,$$

where E_e is the electron beam energy, θ the electron polar angle and γ_{h} the hadronic

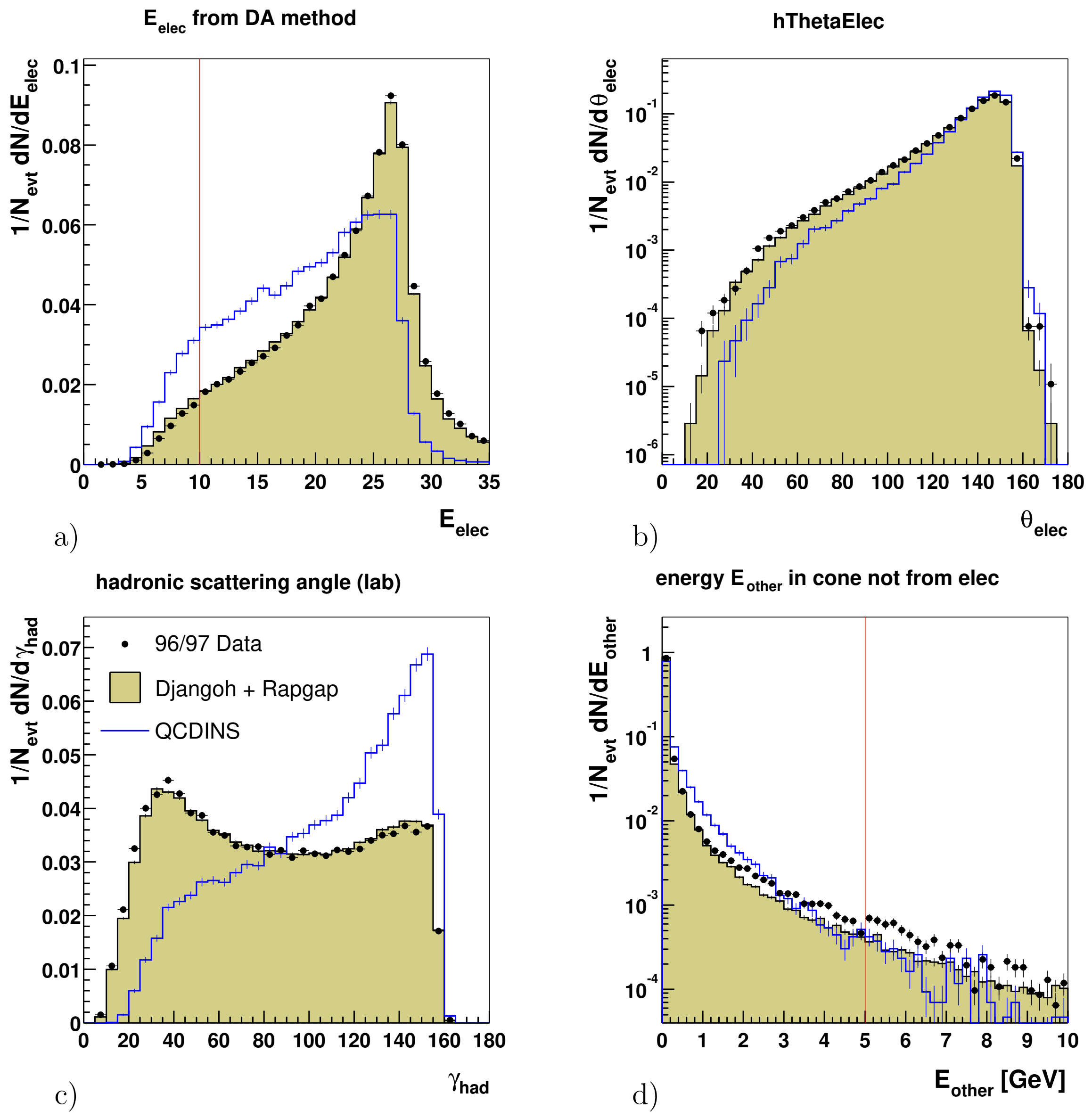


Figure 4.4: *Control plots for (a) electron energy* $E_{\mathrm{e,DA}}$*, (b) electron polar angle* θ*, (c) hadronic scattering angle* γ_{h} *and (d)* E_{cne}*, see section 4.2.*

scattering angle, both measured in the laboratory frame. The hadronic scattering angle is defined by

$$\cos\gamma_{\mathrm{h}} = \frac{p_{\mathrm{T,h}}^2 - \delta_{\mathrm{h}}^2}{p_{\mathrm{T,h}}^2 + \delta_{\mathrm{h}}^2} \quad .$$

In this definition,

$$p_{T,h}^2 = (\sum_j p_{x,j})^2 + (\sum_j p_{y,j})^2$$

is the missing transverse momentum of the hadronic final state and

$$\delta_{\mathrm{h}} = \sum_j E_j - p_{j,z}$$

is its $E - P_z$ value. For both $p^2_{T,h}$ and δ_{h}, the sums run over all zufos j of the hadronic final state.

To calculate the electron polar and azimuthal angle, preferably the tracking information is used, as it is more precise. However, this is only done, if the polar angle in the CAL is above 0.3, i.e. within the CTD acceptance region, and if the track traverses more than three CTD superlayers. Distributions of the electron energy $E_{\mathrm{e,DA}}$, of its polar angle and of the hadronic scattering angle are shown in Fig. 4.4.

4.1.2 Current jet identification

In each event, part of the hadronic final state stems from the hadronisation of the current quark. This current jet needs to be identified in order to calculate the virtuality Q' of the quark entering the hard subprocess as shown in Fig. 1.9. In instanton-induced events it is necessary to separate the current jet from the part of the hadronic final state assigned to the instanton.

For finding the current jet, all zufos are boosted to the hadronic centre of mass system (hcms), defined as system, in which $P + q = 0$, where P is the four-momentum of the incoming proton, as shown in Fig. 1.9, and q the one of the virtual photon. The direction, in which the photon moves, is chosen as positive z-direction. The advantage of using the hcms for the jet-finding is that the part of the hadronic final state, which is of interest, is well separated from the proton remnant in that system.

For the jet-finding, the k_T-cluster algorithm [Ca93] in its longitudinally invariant form [Ell93] is run on all zufos in the hcms. This recursive algorithm merges input four-momenta to jets on the basis of distances $d_{i,j}$ calculated as follows: From differences in pseudorapidity η and azimuthal angle ϕ between the input four-momenta, the quantity $\Delta R_{i,j}$ is computed, defined as

$$\Delta R_{i,j} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad ,$$

These values are weighted with transverse energy E_T to give

$$d_{i,j} = \min{}^2(E_{T,i},\, E_{T,j}) \cdot (\Delta R_{i,j})^2 \quad .$$

Besides, the distance $d_{i,r}$ from the proton remnant is computed for each four-momentum:

$$d_{i,r} = E^2_{T,i} \cdot R^2_0 \text{ with } R_0 = 1.0 \quad .$$

The smallest of these values is determined,

$$d_{\min} = \min\{d_{i,j},\, d_{i,r}\} \quad .$$

If the minimum distance $d_{\min}$ is a distance between input four-momenta, they are merged, i.e. they are treated as a new input particle k in the next run of the algorithm. Their four-momenta are added to yield the four-momentum of the merged particle. If the minimum distance $d_{\min}$ is found to be the distance of four-momentum i from the beam, the recursion is stopped for that particle, which is added to the list of output-jets. At the end of the program run, all zufos have been assigned to "jets"[1].

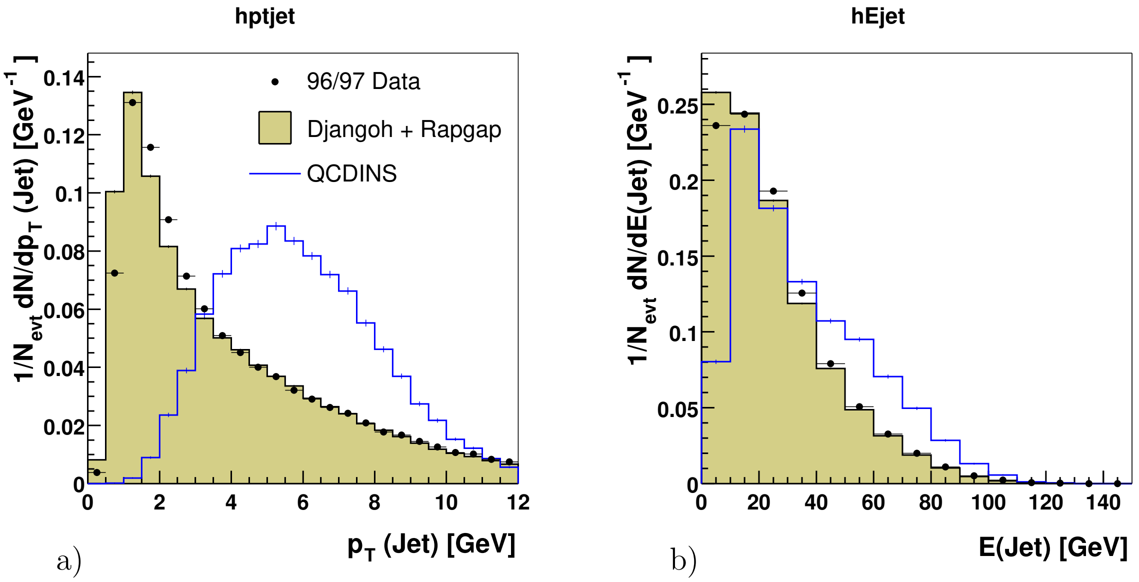


Figure 4.5: *(a): transverse momentum and (b): energy of the current jet in the hadronic centre of mass frame for the inclusive sample.*

Monte Carlo studies at parton level [Ger98] have shown that in the QCDINS enriched region at high transverse jet momenta, the current quark on average has a larger hcms pseudorapidity than the partons assigned to the instanton. Therefore, the current jet in this analysis is reconstructed as follows: The output list of jets from the k_T cluster algorithm is sought for jets with a pseudorapidity $\eta_{\text{jet}}^{\text{hcm}}$ exceeding the E_T-weighted mean pseudorapidity of all zufos in the event:

$$\eta_{\text{jet}}^{\text{hcm}} > \frac{\sum_{\text{zufos}} E_T \eta}{\sum_{\text{zufos}} E_T} \quad .$$

Of the jets fulfilling that requirement, the one with the highest transverse energy is chosen as current jet.
Fig. 4.5 shows the transverse momentum p_T^{Jet} and the energy of the current jet in the hadronic centre of mass frame. One can see that for both variables in the range of values, in which the QCDINS distribution is above the nDIS distribution when both

---

[1] Some of these "jets" are so low energetic, that they would not be considered as jets in e.g. jet multiplicity studies.

are normalised to area 1, the data are well described by nDIS, while discrepancies only occur at lower values. A good description of the region, where the QCDINS maximum is located, is particularly important for p_T^{Jet}, which is used as a cut variable to enhance the instanton contribution.

4.1.3 Reconstruction of the kinematic variables

In this analysis, x and Q^2 define the kinematic region investigated. They are reconstructed [Ben92] using the double angle method (by default) and the electron method (as a systematic check). With the energy E_p of the proton beam and further notation as introduced in section 4.1.1, the double-angle variables are defined as follows:

$$Q^2_{\text{DA}} = 4E_e^2 \; \frac{\sin\gamma_h(1+\cos\theta)}{\sin\theta + \sin\gamma_h - \sin(\theta+\gamma_h)}$$

$$x_{\text{DA}} = \frac{E_e}{E_p} \; \frac{\sin\gamma_h + \sin\theta + \sin(\gamma_h+\theta)}{\sin\gamma_h + \sin\theta - \sin(\gamma_h+\theta)} \quad .$$

The method is referred to as "double-angle method" as both the electron polar angle θ and the hadronic scattering angle γ_{h} are used.
The electron method variables are given by the expressions

$$\begin{aligned} Q^2_{\text{el}} &= 2E_eE'(1+\cos\theta) = 4E_eE'\cos^2\frac{\theta}{2} \quad , \\ x_{\text{el}} &= \frac{E_e}{E_p} \cdot \frac{E'(1+\cos\theta)}{2E_e - E'(1-\cos\theta)} = \frac{Q^2_{\text{el}}}{sy_{\text{el}}} \quad , \\ y_{\text{el}} &= 1 - \frac{E'}{2E_e}(1-\cos\theta) = 1 - \frac{E'}{E_e}\sin^2\frac{\theta}{2} \quad . \end{aligned}$$

According to the Jacquet-Blondel method [Jacq79], y is calculated from the hadronic final state:

$$y_{\text{JB}} = \delta_{\text{h}}/(2E_e) \quad .$$

In the analysis, a cleaning cut on that variable is applied, cf. section 4.2. Distributions of Q^2_{DA}, $-\log_{10} x_{\text{DA}}$, y_{el} and y_{JB} can be found in Fig. 4.6 . Regions of phase space, in which the two y variables are poorly described, are removed from the samples. The x_{DA} and Q^2_{DA} description is good in the entire phase space considered.

As explained in section 3.1, the QCDINS Monte Carlo is valid only for high enough virtualities Q'^2 of the (anti-) quark entering the hard subprocess. Different methods to reconstruct Q'^2 have been investigated. As a result of these studies, the double angle method from [Siev00] has been chosen, as no better method could be found. In that method, Q'^2 is calculated according to

$$Q'^2_{\text{DA}} = Q^2_{\text{DA}} + \frac{W^2_{\text{DA}} - Q^2_{\text{DA}}}{W_{\text{DA}}} \cdot \sum_{i\in\text{Jet}} E_i - \frac{W^2_{\text{DA}} + Q^2_{\text{DA}}}{W_{\text{DA}}} \cdot \sum_{i\in\text{Jet}} p_{z,i} - M^2_{\text{Jet}} \quad ,$$

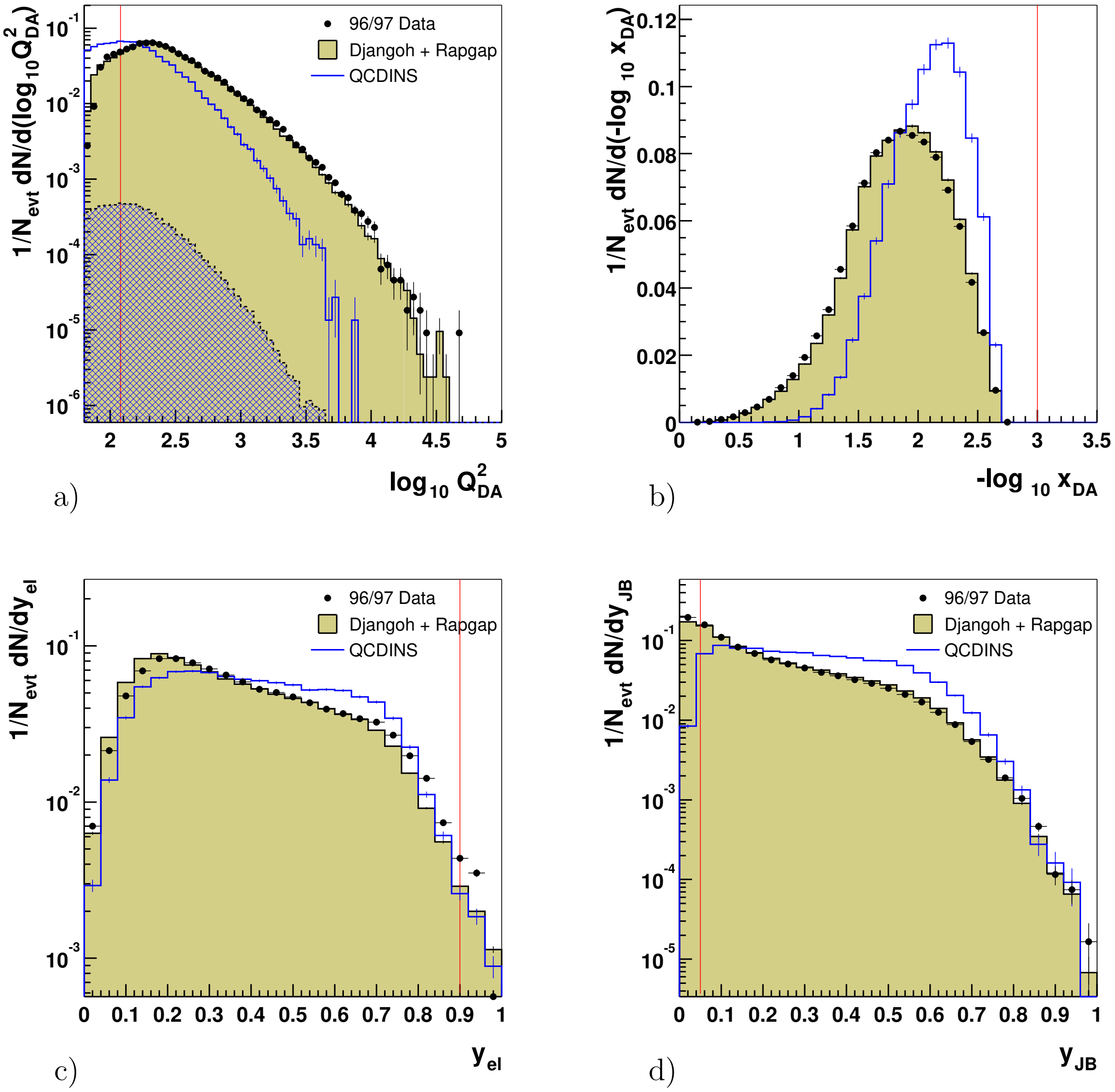


Figure 4.6: *Distributions of (a):* $Q^2_{\rm DA}$*, (b):* $-\log_{10} x_{\rm DA}$*, (c):* $y_{\rm el}$ *and (d):* $y_{\rm JB}$*. All offline cuts of section 4.2 are applied, except for the cut on the variable plotted, which is indicated by the vertical line. In (a) the QCDINS MC distribution is shown both normalised to area 1 (line) and normalised to the predicted instanton fraction in the inclusive sample (hatched).*

with

$$W_{\rm DA} = \sqrt{Q^2_{\rm DA}(1 - x_{\rm DA})/x_{\rm DA} + m_p^2}\ ,$$

the jet mass $M_{\rm Jet}$ and the proton mass m_p. The sums run over all zufos inside a cone of radius $R = 0.7$ in η, ϕ around the axis of the current jet. Fig. 4.7 (a) shows that the

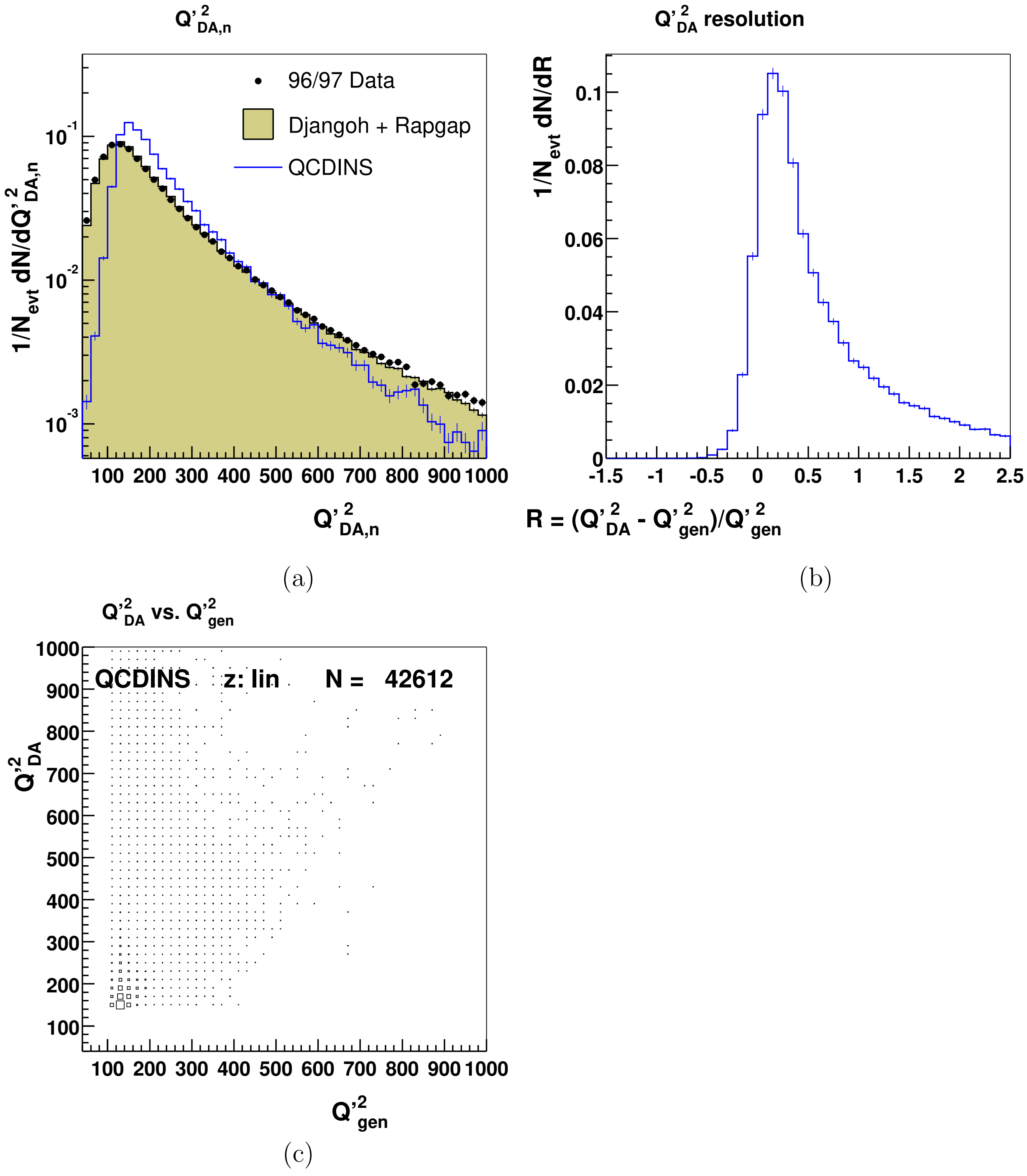
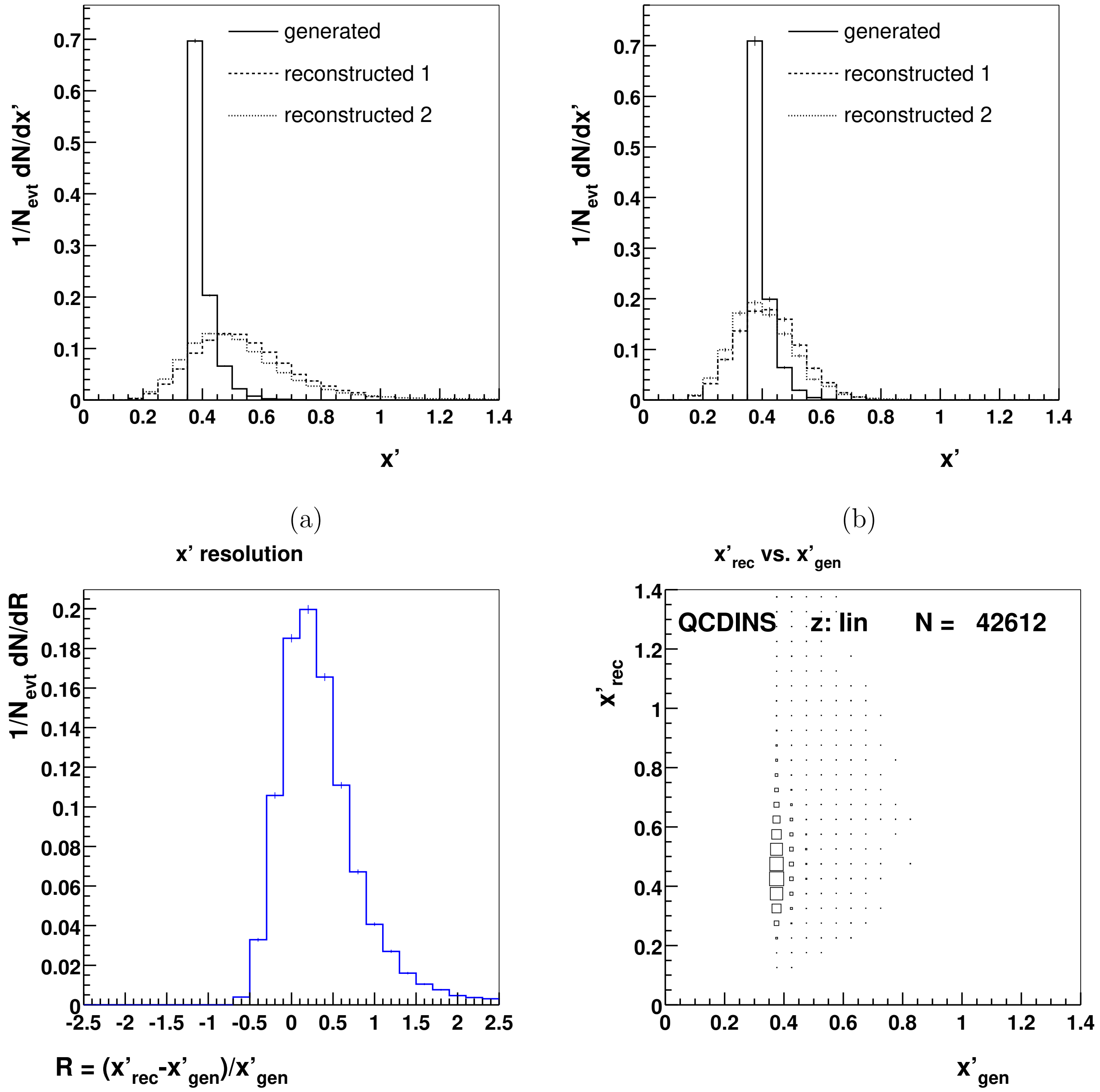


Figure 4.7: *Reconstruction of* Q'^2*: (a): distribution of* Q'^2_{DA}*; all offline cuts given in section 4.2 are applied, except for the* Q'^2_{DA} *cut. Comparisons of the reconstructed value,* Q'^2_{DA}*, to the QCDINS generated one,* Q'^2_{gen}*, are shown in (b):* $(Q'^2_{\mathrm{DA}} - Q'^2_{\mathrm{gen}})/Q'^2_{\mathrm{DA}}$ *and (c):* Q'^2_{DA} *vs.* Q'^2_{gen}*.*



(a) (b)

(c) (d)

Figure 4.8: *Reconstruction of x': values $x'_{\rm rec}$ obtained from two reconstruction methods are compared to the generated distribution $x'_{\rm gen}$, (a): for inclusive sample, cf. section 4.2, (b): for a sample with one-dimensional instanton enhancing cuts on $p_T^{\rm Jet}$, $Q'^{\,2}_{\rm DA}$ and $N_{\rm eft}$, see section 5.2.1. For the second reconstruction method, (c) shows $(x'_{\rm rec} - x'_{\rm gen})/x'_{\rm gen}$ and (d): $x'_{\rm rec}$ vs $x'_{\rm gen}$.*

variable is well described. However, a comparison of the reconstructed values of ${Q'}^2_{\mathrm{DA}}$ to the QCDINS MC generated ones, ${Q'}^2_{\mathrm{gen}}$, see Fig. 4.7 (b) and (c), shows that Q'^2 is badly reconstructed, some reconstructed values being much too large. This would also be the case for ${Q'}^2_{\mathrm{gen}}$ values below the minimum value ${Q'}^2_{\mathrm{min}} = 113\,\mathrm{GeV}^2$ that defines the fiducial region of the QCDINS MC, were they not removed when generating the QCDINS sample. Unfortunately, the maximum of the QCDINS MC ${Q'}^2$ distribution is found at low ${Q'}^2$ values, which poses a problem that cannot fully be solved: From Fig. 4.7 (c) one can see that in order to remove the events with real Q'^2 below ${Q'}^2_{\mathrm{min}}$ completely, one would have to choose the ${Q'}^2_{\mathrm{DA}}$ cut at so large a value, that all QCDINS events would be rejected. The chosen cut value of $140\,\mathrm{GeV}^2$, applied in Fig. 4.7 (c), seems to be a reasonable compromise, rejecting most events outside the fiducial region while still keeping large enough a sample.

Similarly as for the standard kinematic variables, a different reconstruction method, namely the photon method from [Siev00], has been used for ${Q'}^2$ as part of the systematic checks. For that method, ${Q'}^2$ is reconstructed as

$$Q'^2_\gamma = Q^2_\gamma + 2 \cdot q_\gamma p_{\mathrm{Jet}} - M^2_{\mathrm{Jet}}$$

with the four-momenta of the photon, q_γ, and of the jet, p_{Jet}, respectively, and $Q^2_\gamma = -q^2_\gamma$.

In order to reconstruct x', two methods have been tried, the second one yielding a slightly better resolution. The first method relies only on the virtuality ${Q'}^2$ and on the mass of the instanton region, M_{Inst}, as defined in section 5.1:

$${x'}_{\mathrm{rec}} = \frac{{Q'}^2}{{Q'}^2 + M^2_{\mathrm{Inst}}} \quad .$$

The second method is to calculate x' according to

$${x'}_{\mathrm{rec}} = \frac{{Q'}^2}{\hat{s} + {Q'}^2 - 2x_{\mathrm{DA}}\frac{\hat{s}+Q^2_{\mathrm{DA}}}{Q^2_{\mathrm{DA}}} \cdot P\, p_{\mathrm{Jet}}} \quad ,$$

with the notation of Fig. 1.9.

Distributions of the generated and of the reconstructed x' values for the inclusive sample defined in section 4.2 and for a sample with some of the instanton enhancing cuts given in section 5.2.1 applied, are shown in Fig. 4.8 (a) and (b), respectively. The distributions of both the reconstructed variables are broader than the generated distributions, with the peak being shifted to larger values for the inclusive sample. Within the instanton enhancing cuts, the reconstruction is improved compared to the inclusive sample, with the maximum being in the same bin as for the generated distribution. However, the fraction of reconstructed values below the value 0.35, defining the QCDINS fiducial region, is larger for the enhanced sample than it is for the inclusive one.

Fig. 4.8 (c) shows the resolution

$$R = \frac{{x'}_{\mathrm{rec}} - {x'}_{\mathrm{gen}}}{{x'}_{\mathrm{gen}}}$$

for the better of the reconstruction methods, i.e. the second one. In Fig. 4.8 (d), reconstructed values are plotted versus generated ones, showing that the reconstructed value is hardly correlated with the generated value.

In summary, neither ${Q'}^2$ nor x' can be well reconstructed. The QCDINS sample has been generated with default cuts on the generated ${Q'}^2$ and x' applied. For ${Q'}^2$, a cut on the reconstructed variable at $140\,\mathrm{GeV}^2$ is included in the offline cuts defining the inclusive sample, see section 4.2. No cut is made on the reconstructed x'.

4.2 Selection of neutral current DIS events

A '96/'97 data sample corresponding to an integrated luminosity of $38\,\mathrm{pb}^{-1}$ is used in the search. A high Q^2 NC DIS sample is selected using bit DIS03 of the third level trigger (TLT), corresponding to the requirement of an electron with an energy greater than 4 GeV at a position outside a radius of 25 cm on the calorimeter. DST bit 12, requiring $Q^2 > 80\,\mathrm{GeV}^2$ is used when the ntuple is written, which is done using the program orange [Dus02]. At the same stage, a SINISTRA electron with a radius cut of 36 cm at the calorimeter is required. The SINISTRA probability for the candidate actually being the scattered electron, must exceed the value 0.9. Further cuts are applied offline as listed below, defining what will be referred to as inclusive sample in what follows:

- **kinematic cuts**:
  - ○ $Q^2_{\mathrm{DA}} > 120\,\mathrm{GeV}^2$;
  - ○ $x_{\mathrm{DA}} > 10^{-3}$;
  - ○ $y_{\mathrm{JB}} > 0.05$;
- **vertex cut**:
  - ○ z position of the vertex, $|z_{\mathrm{Vtx}}| < 50\,\mathrm{cm}$
- **cuts to ensure the quality of the electron reconstruction**:
  - ○ $E_{\mathrm{e,DA}} > 10\,\mathrm{GeV}$;
  - ○ $E_{\mathrm{cne}} < 5\,\mathrm{GeV}$;
    E_{cne} is the energy not stemming from the electron, which is found inside a cone around the track of the electron candidate. This cone is defined by having the (η, ϕ)-radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.8$ at the impact point of the electron track on the calorimeter surface;
- **suppression of photoproduction events**:
  - ○ $y_{\mathrm{el}} < 0.90$;
  - ○ $35\,\mathrm{GeV} < E - P_z < 65\,\mathrm{GeV}$,

where $E - P_z$ is obtained from the sum of the values found for the electron and for the hadronic final state, reconstructed using zufos;

- ○ DCA $< 10\,\mathrm{cm}$:
  The distance of closest approach, DCA, is calculated between the electron track and the cone island, which is closest to the tracks' impact point on the calorimeter surface;

- **suppression of cosmics**:
  - ○ $\not{p}_t/\sqrt{E_{\mathrm{t,s}}} < 3\sqrt{\mathrm{GeV}}$;
  - ○ $\not{p}_t/E_{\mathrm{t,s}} < 0.7$;
    The missing transverse momentum, $\not{p}_t = \sqrt{\left(\sum p_x\right)^2 + \left(\sum p_y\right)^2}$ is compared to the scalar sum of the transverse energy, $E_{\mathrm{t,s}} = \sum \left(p_x^2 + p_y^2\right)$.
    Both quantities are calculated from calorimeter cells;

- **restriction to the fiducial region of the QCDINS MC**:
  - ○ $Q'^2_{\mathrm{DA}} > 140\,\mathrm{GeV}^2$.

	DATA	QCDINS	DJANGOH + RAPGAP	HERWIG
luminosity $[\mathrm{pb}^{-1}]$	38.6	2737	132.8	136.5
#(events) generated	—	161285	1839897	1579045
Q^2_{DA}, x_{DA}	235761	905	238850	210656
y_{el}	232481	891	236727	210656
$\lvert z_{\mathrm{Vtx}}\rvert$	219218	845	224212	208894
$\not{p}_t/\sqrt{E_{\mathrm{t,s}}}$	218697	845	224127	197680
$\not{p}_t/E_{\mathrm{t,s}}$	218390	845	223904	197535
y_{JB}	152792	828	163928	197364
$E_{\mathrm{e,DA}}$	144923	737	154057	140847
$E - P_z$	144545	736	153885	132084
DCA	140276	724	151941	131976
E_{cne}	139442	722	151315	130356
Q'^2_{DA}	92116	601	101159	76997

Table 4.2: *Luminosities and numbers of events obtained for the available samples when consecutively applying the offline cuts. Except for the number of generated events, numbers of events are normalised to data luminosity. The last line corresponds to the inclusive sample used in the analysis.*

The luminosities of the available data and MC samples and the numbers of events remaining after consecutively applying the offline selection cuts for the inclusive NC DIS sample, are listed in table 4.2. Except for the number of generated events, numbers of events in the Monte Carlo samples are normalised to data luminosity. One can see, that the number of events in the inclusive sample from Djangoh with Rapgap exceeds the number of data events by ≈ 10%, while for Herwig, the number is below the number of data events by ≈ 15%.

The discrepancies between Djangoh with Rapgap and Herwig for the inclusive sample have several reasons: The cross sections of Djangoh with Rapgap and Herwig within the kinematic cuts on Q^2_{DA} and x_{DA} differ by about 13%, as Herwig does not explicitly contain radiative corrections and diffractive events, which are included in Djangoh with Rapgap.

When consecutively applying the offline selection cuts, differences between the nDIS Monte Carlos increase, reaching a level of 16% before the Q'^2 cut is applied. The fraction of events removed by the Q'^2 cut, which is specific to this analysis, is larger for Herwig than for both Djangoh with Rapgap and the data. Differences between Djangoh with Rapgap and the data mainly arise from the cuts on y_{JB} and on the distance of closest approach, DCA, while within the cuts on Q^2_{DA} and x_{DA} Djangoh with Rapgap agrees with the data to better than 1.5%.

In order not to be biased by initial discrepancies between the nDIS Monte Carlo and the data when restricting the inclusive sample to an instanton-enhanced subsample, numbers of nDIS Monte Carlo events have been normalised to the number of data events in the inclusive sample, where the predicted QCDINS contribution is negligible (≈ 0.5%):

$$N_E = N_{\mathrm{E,raw}} \cdot D_{\mathrm{O,raw}} / N_{\mathrm{O,raw}} \quad ,$$

with

$D_{\mathrm{O,raw}}$:	raw number of data events in inclusive sample;
$N_{\mathrm{O,raw}}$:	raw number of nDIS events in inclusive sample;
$N_{\mathrm{E,raw}}$	raw number of nDIS events in instanton-enhanced sample;
N_E	normalised number of nDIS events in instanton-enhanced sample .

Chapter 5

Instanton enhancement

5.1 Reconstruction of the instanton

Based on the characteristics of instanton-induced events, described in chapter 3, observables are defined, which have different distributions for the signal than for the nDIS background. Most of these variables are calculated from a subset of hadronic final state zufos, which is assigned to the instanton. These zufos are referred to as "instanton part of the hadronic final state" or "instanton candidate". Although these terms are meaningful only for instanton-induced events, they will also be used for background events, referring to the part of the event, which is found by applying the same reconstruction algorithm used for instanton-induced events.

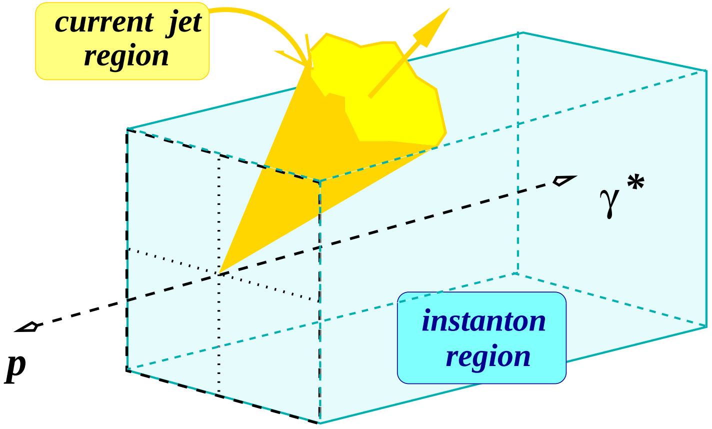


Figure 5.1: *Assignment of certain regions of the hadronic final state to current jet and instanton.*

Physically, the notion of an instanton part of the event is more clearly defined at parton level than at hadron level. At parton level, it comprises the $2n_f - 1$ quarks and the gluons stemming from the photon-gluon fusion process in the background of an instanton. However, in the process of fragmentation, these partons are connected by a colour flow to both the current quark and the proton remnant. Therefore, hadrons "stemming from the instanton" might also "stem from the current quark". The aim in identifying an instanton candidate at detector level is therefore to reconstruct as well as possible the information contained in the parton level instanton candidate. This

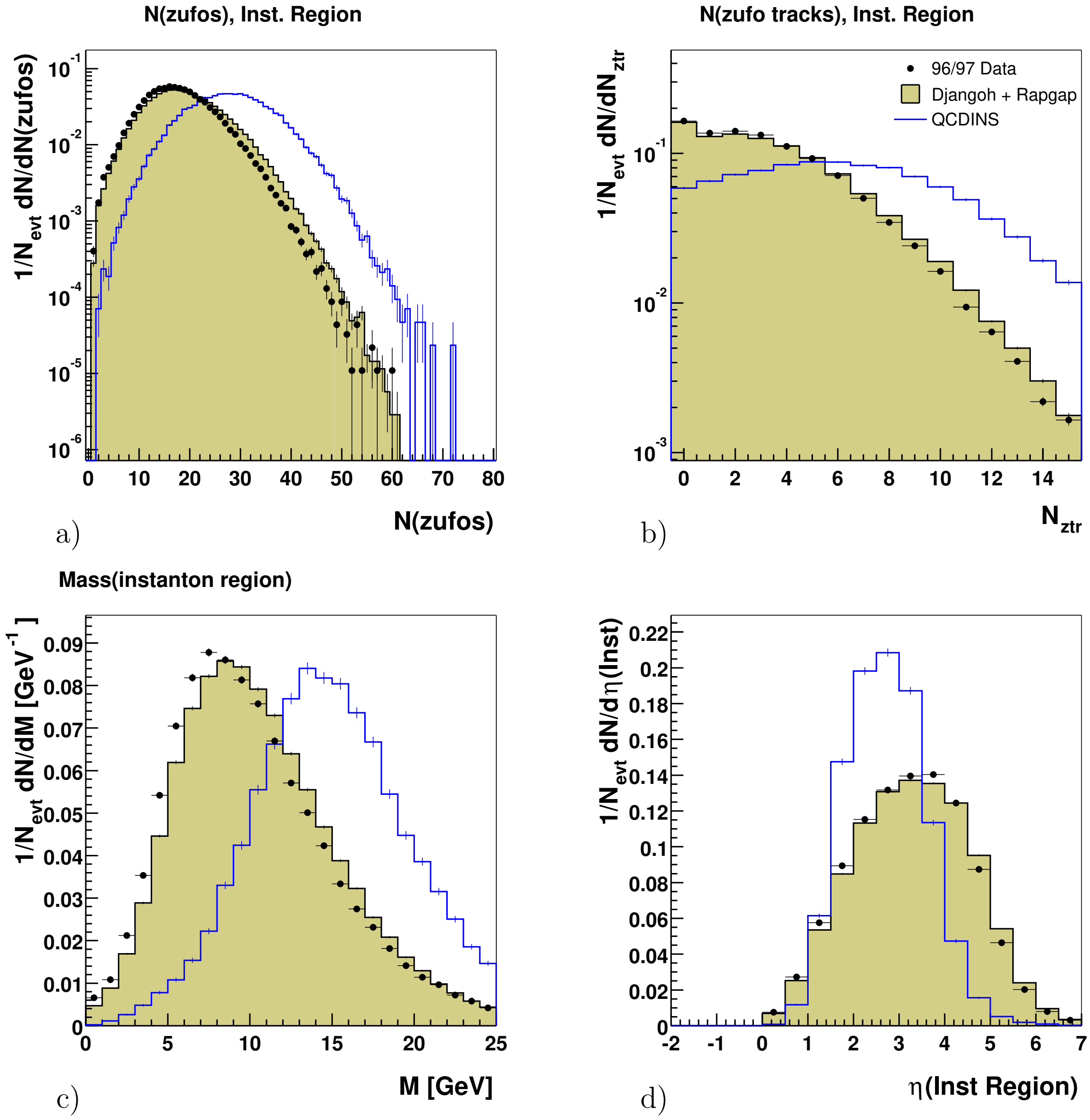


Figure 5.2: *Control plots for the instanton part of the hadronic final state for the inclusive sample: (a): zufo multiplicity* $N_{\rm efo}$, *(b): zufo track multiplicity* $N_{\rm eft}$, *(c): mass* $M_{\rm Inst}$ *of instanton candidate, (d): pseudorapidity* $\eta_{\rm Inst}$ *of instanton candidate in hcm frame.*

task is more difficult than the usual reconstruction of the current jet, as this information comprises four-vectors of several partons and is therefore more complex than the information associated with the current quark.
An additional caveat is that very often part of the event is lost in the beam-pipe. This affects especially particles that would contribute to the instanton part of the event, which on average is found further forward in the detector than the current jet.

The reconstruction method used in this analysis first requires the current jet to be reconstructed in the way described in chapter 4. Zufos, which have not been assigned to the current jet and which move away from the proton remnant in the hcms, $p_z^{\text{hcms}} > 0$, are assigned to the instanton, see Fig. 5.1.

Once the instanton part of the hadronic final state has been identified, two multiplicities are determined: the multiplicity of zufos (or efos), assigned to the instanton candidate, N_{efo}, and the multiplicity of zufo tracks (or "energy flow tracks"), N_{eft}. This variable counts the number of tracks contained in zufos, which are assigned to the instanton part of the event. For zufos of types 0, 1, 12, 37 and 41 (cf. table 4.1), N_{eft} is increased by 1, for types 2 and 22 by 2 and for type 3 by 3. Distributions of N_{efo} and N_{eft} are shown in Fig. 5.2. The QCDINS distribution peaks at higher multiplicities than the nDIS MC and the data. At high multiplicities, the normal DIS description overshoots the data, at low multiplicities, the MC prediction is below the data points. This might be due to the fact that the chosen method to reconstruct the instanton tends to assign many low momentum zufos to the instanton. The failure of the nDIS Monte Carlos to correctly describe the multiplicities of such zufos is a known problem.

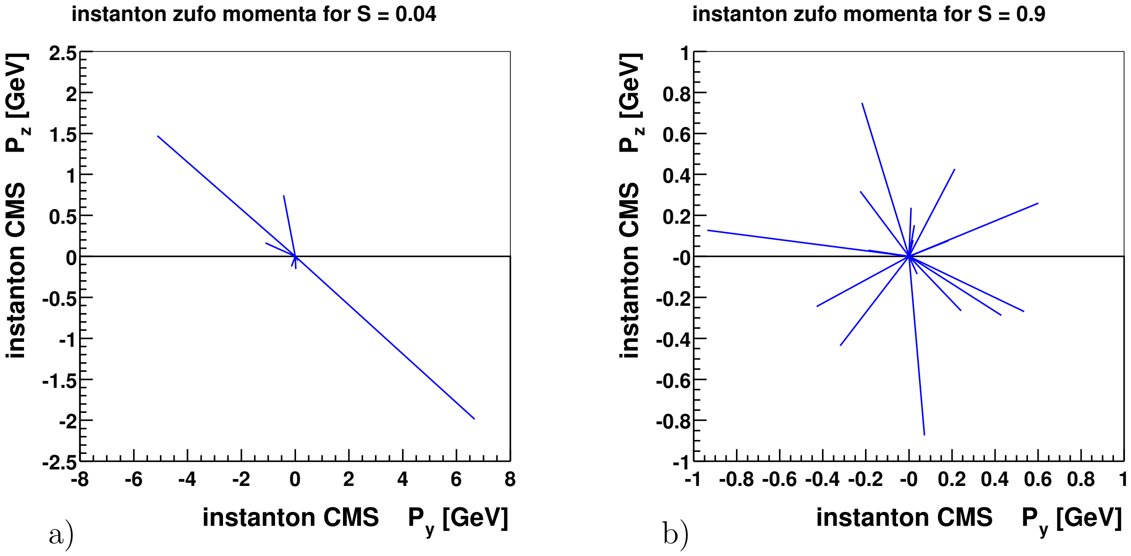
The four-momenta of all zufos in the instanton part of the hadronic final state are added, yielding the four-momentum of the instanton candidate. Its mass M_{Inst} and pseudorapidity η_{Inst} in the hadronic centre of mass frame are shown in Fig. 5.2. While the pseudorapidity is fairly well described, the nDIS distribution of M_{Inst} is clearly shifted towards higher values compared to the data.

Two variables are used to describe the isotropy of the set of zufos assigned to the instanton. Both have values in the interval $[0, 1]$, low values corresponding to a jet like, large values to a more isotropic distribution as shown in Fig. 5.3. These variables are the circularity C of the instanton in the hadronic centre of mass frame with respect to the photon-proton axis and the sphericity S of the instanton part of the event in its centre of mass frame.
The circularity describes, how isotropically a set of four-momenta is distributed over a plane perpendicular to a given axis. The normalised two-dimensional momentum tensor,

$$M_{\alpha\beta}^{(2\text{D})} = \frac{\sum_j p_{j,\alpha} p_{j,\beta}}{\sum_j \left(p_{j,1}^2 + p_{j,2}^2\right)} \quad \text{with } \alpha, \beta = 1,\ 2\ ,$$

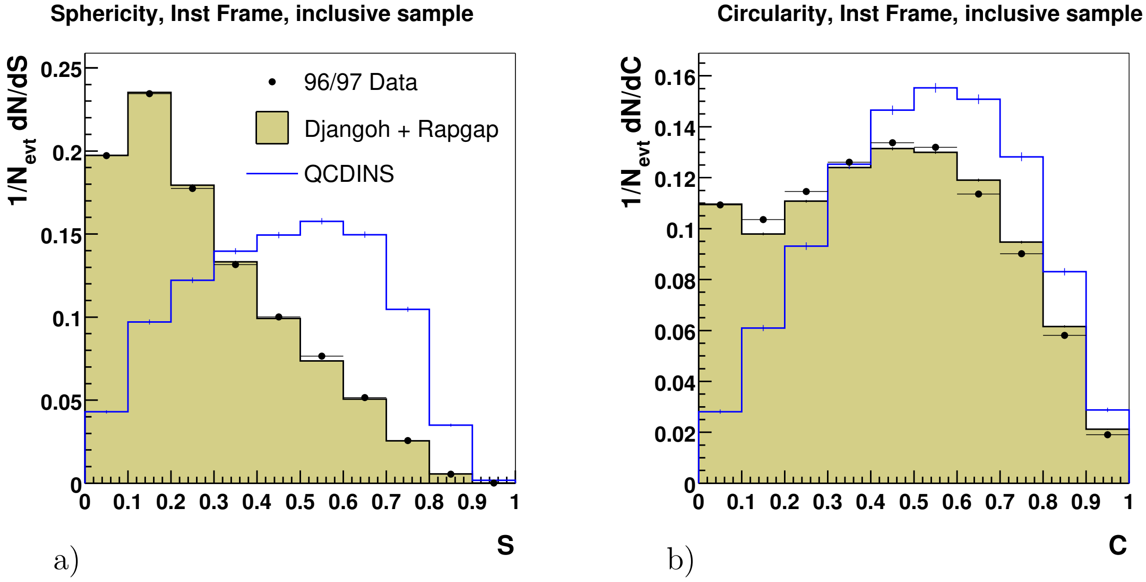
is computed from the hcms-four-momenta of all zufos assigned to the instanton. From its eigenvalues λ_1, λ_2, with $\lambda_1 > \lambda_2$, the circularity is obtained:

$$C = 2(1 - \lambda_2) \quad .$$



a) b)

Figure 5.3: *Projections of instanton four-momenta in instanton CMS to two-dimensions for (a): a low sphericity event and (b): an event with large sphericity.*



a) b)

Figure 5.4: *Control plots for the instanton part of the hadronic final state for the inclusive sample: (a): sphericity S in instanton CMS, (b): circularity C in hcms.*

The sphericity is a measure of how isotropically a collection of four-momenta is distributed in three dimensions. Similarly as in the circularity calculation, a normalised

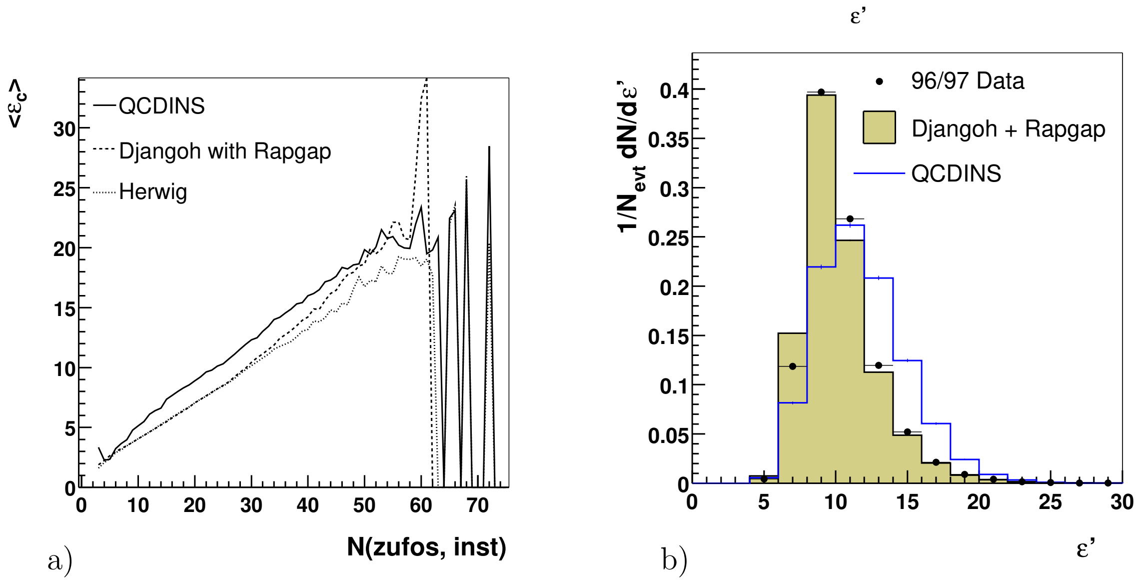


Figure 5.5: *Control plots for the instanton part of the hadronic final state for the inclusive sample: (a) dependence of* $< \epsilon_c >$ *on instanton zufo multiplicity, see text, (b)* ϵ'*.*

momentum tensor is calculated from the zufo four-momenta assigned to the instanton, but now in three dimensions and after having boosted the zufo momenta to their common rest frame:

$$M^{(3\mathrm{D})}_{\alpha\beta} = \frac{\sum_j p_{j,\alpha} p_{j,\beta}}{\sum_j \left(p^2_{j,1} + p^2_{j,2} + p^2_{j,3}\right)} \quad \text{with } \alpha, \beta = 1,\, 2,\, 3 \quad .$$

From the eigenvalues of $M^{(3\mathrm{D})}_{\alpha\beta}$, $0 \leq Q_1 \leq Q_2 \leq Q_3$, the sphericity is found according to the definition

$$S = \frac{3}{2} \left(Q_1 + Q_2\right) \quad .$$

Distributions of S and C for the inclusive sample are shown in Fig. 5.4 (a) and (b). Within these cuts, in which the instanton-induced contribution is expected to be negligible, the sphericity data distribution is well described by the nDIS Monte Carlo, while for circularity small discrepancies can be seen in some of the bins. The plots also show clearly, that the sphericity distribution is better suited for distinguishing between nDIS and QCDINS than is circularity, i.e. sphericity has a larger separation power.

Finally, the variable ϵ' is calculated, which is a measure of the average distance of zufos assigned to the instanton part of the event in η in the hcms, corrected for multiplicity dependence. It is based on work by A. Barakbaev and E. Boos [Bar02], and is obtained from the following algorithm:
All zufos assigned to the instanton part of the event are sorted with respect to η in the hcms, yielding N_{efo} values $\eta_1,\, \eta_2,\, \ldots,\, \eta_{N_{\mathrm{efo}}}$.

The interval $[\eta_i, \eta_{i+k}]$, $i + k \leq N_{\mathrm{efo}}$, contains k zufos. For each i, k, the value

$$\epsilon_{i,k} = \frac{k}{\eta_{i+k} - \eta_i}$$

is calculated. These values are averaged over all zufos $i < N_{\mathrm{efo}} - k$ of the event, yielding $N_{\mathrm{efo}} - 1$ values

$$\epsilon_k = \sum_{i=1}^{i=N_{\mathrm{efo}}-k} \frac{\epsilon_{i,k}}{N_{\mathrm{efo}} - k}$$

per event.

As the range of η values accessible is restricted by the detector acceptance, one would expect zufos to be closer in η in high multiplicity events and further apart for low multiplicities. To account for that effect, the variable

$$\epsilon_c \equiv \epsilon_k \quad \text{with} \quad k = \begin{cases} N_{\mathrm{efo}}/2 & \text{for } N_{\mathrm{efo}} \text{ even,} \\ (N_{\mathrm{efo}} + 1)/2 & \text{for } N_{\mathrm{efo}} \text{ odd} \end{cases}$$

for the central value of k in the event is chosen as observable.

Its mean value, $< \epsilon_c >$, averaged over events of same multiplicity, is plotted as a function of the multiplicity in Fig. 5.5 (a). The plot shows that $< \epsilon_c >$ increases with increasing multiplicity, so the corresponding η distance decreases and has been over-corrected by choosing ϵ_c. This might be due to the fact that the instanton part of the event does not always take up the full η range available, but only part of it. Which fraction of the full range that part corresponds to might also depend on the multiplicity.

To remove the residual multiplicity dependence, the slope b of the function $< \epsilon_c > (N_{\mathrm{efo}})$ in the range $20 \leq N_{\mathrm{efo}} \leq 40$ is determined from a first order polynomial fit with fit parameters a, b,

$$< \epsilon_c > = a + b \cdot N_{\mathrm{efo}}, \quad 20 \leq N_{\mathrm{efo}} \leq 40 \quad .$$

Fit parameters are averaged over the inclusive Herwig and Djangoh with Rapgap samples, yielding $b = 0.339$. The dependence is subtracted from ϵ_c to give the observable

$$\epsilon' = \epsilon_c - b \cdot (N_{\mathrm{efo}} - 30) \quad .$$

Distributions of ϵ' are shown in Fig. 5.5 (b).

5.2 Restriction to instanton enhanced subsample

In order to enhance the instanton contribution, cuts were applied on the variables p_T^{Jet}, Q'^2_{DA}, γ_{h}, N_{efo}, N_{eft}, M_{Inst}, η_{Inst}, C, S and ϵ' defined above. Different cut methods were tried out and compared, which will be described in detail in the following. The enhancement methods differ in the extent, to which correlations between observables are taken into account.

For each enhancing method, two samples have been considered, one with harder cuts used for the cut approach for setting a limit on the instanton contribution, and one with relaxed cuts used for the fit approach. Relaxing the cuts for the fit approach was necessary in order to retain enough statistics in the enhanced sample for a fit of the instanton sphericity distribution.
Different cut methods can be compared to each other by comparing how many normal DIS events they reject for a fixed percentage of QCDINS events from the inclusive sample that are kept in the enhanced sample. From the numbers

N_O : number of nDIS events in the inclusive sample
N_E : number of nDIS events in the enhanced sample
I_O : number of instanton events in the inclusive sample
I_E : number of instanton events in the enhanced sample ,

the efficiencies r_I of QCDINS and r_N of nDIS MC are calculated according to

$$r_I = \frac{I_E}{I_O} \quad , \qquad r_N = \frac{N_E}{N_O} \quad .$$

The separation power,

$$P_s = \frac{r_I}{r_N} = \frac{I_E/I_O}{N_E/N_O} \quad ,$$

plotted as a function of the QCDINS efficiency, can be used to determine, which method to enhance the instanton fraction works best.

5.2.1 Combination of one-dimensional cuts

In a first attempt to explore the separation power of the variables, correlations were widely ignored. Cut values were fixed on the basis of distributions for the inclusive sample. Only in a second step, each distribution was plotted with cuts on all other observables applied to ensure that the cut on the variable under consideration did not seem to be in an unreasonable position.

For the fit approach, the following cuts were applied:

- $p_T^{\text{Jet}} > 4\,\text{GeV}$
- $Q'^2_{\text{DA}} < 250\,\text{GeV}^2$
- $N_{\text{eft}} > 7$
- $C > 0.2$
- $S > 0.2$

The sample resulting from these cuts has an instanton efficiency of $r_I = 17\%$ and an nDIS efficiency of $r_N = 1.6\%$ for Djangoh with Rapgap and $r_N = 0.91\%$ for Herwig. The predicted fraction of instanton-induced events in the enhanced sample, as defined in section 6.1, is $f_{\mathrm{I,th}} = 8.8\%$. Fits of S and C distributions within these cuts will be presented in section 6.2.

For the cut approach, the same cut values as chosen for the fit approach were used for the variables p_T^{Jet}, Q'^2_{DA} and N_{eft}. Additionally, it was required that

- $C > 0.5$
- $S > 0.5$.

The resulting numbers of events are listed in table 6.2. A comparison of this enhancement method to the more advanced method presented in the next subsection is shown in Fig. 5.12.

5.2.2 Combination of two-dimensional cuts

Combining cuts on two-dimensional distributions accounts for correlations between variables to some extent. The choice of cuts was made in two steps: Firstly, efficient combinations of variables were chosen from a set of all possible combinations of 9 variables (the above mentioned ones except for Q'^2_{DA}). In a second step, cuts on four two-dimensional distributions were set. Each cut was allowed to vary to some extent and all combinations of the varied cuts were tested for their separation power. In each case, a one-dimensional cut on Q'^2_{DA} as in section 5.2.1 was applied in addition.

Step 1: All combinations of the following variables were considered as candidates for distributions on which to cut: sphericity S, circularity C, $\log_{10} p_T^{\mathrm{Jet}}$, N_{efo}, N_{eft}, γ_{h}, M_{Inst}, η_{Inst} and ϵ' as defined at the beginning of this chapter.
For each variable combination, a histogram of 10×10 bins was written out for each MC sample. The separation power P_s, as defined at the beginning of this section, was calculated for one specific cut. This test cut is defined by keeping the events from all bins of the two-dimensional histogram, which

- contain at least 25 events for each nDIS Monte Carlo[1] and
- would yield a separation power $p_i \equiv P_s(\text{bin i})$ above the value 2, if the cut would select the events of the respective bin only, $p_i > p_{i,\min} = 2$.

When choosing variable combinations for the definition of two-dimensional cuts it was required that the QCDINS efficiency, r_I, of the sample obtained from the test cut was at least 45% in order to keep enough events when combining several two-dimensional cuts. Besides, the efficiencies of the two nDIS Monte Carlos, r_{DjaRgp} and r_{Hwg}, defined

---

[1] It has been ensured that no region of large separation power has been rejected by this requirement.

analogously to r_N, see beginning of this section, were required not to differ by more than 15%,

$$0.85 \leq \frac{r_{\mathrm{DjaRgp}}}{r_{\mathrm{Hwg}}} \leq 1.15 \quad .$$

The minimum QCDINS efficiency of 45% is the stronger of these requirements. Variable combinations, for which too few QCDINS events are kept, often contain the observables η_{Inst}, γ_{h}, C and N_{eft}, while variable combinations, which yield too large a discrepancy between the nDIS Monte Carlos, include N_{eft} and $\log_{10} p_T^{\mathrm{Jet}}$.

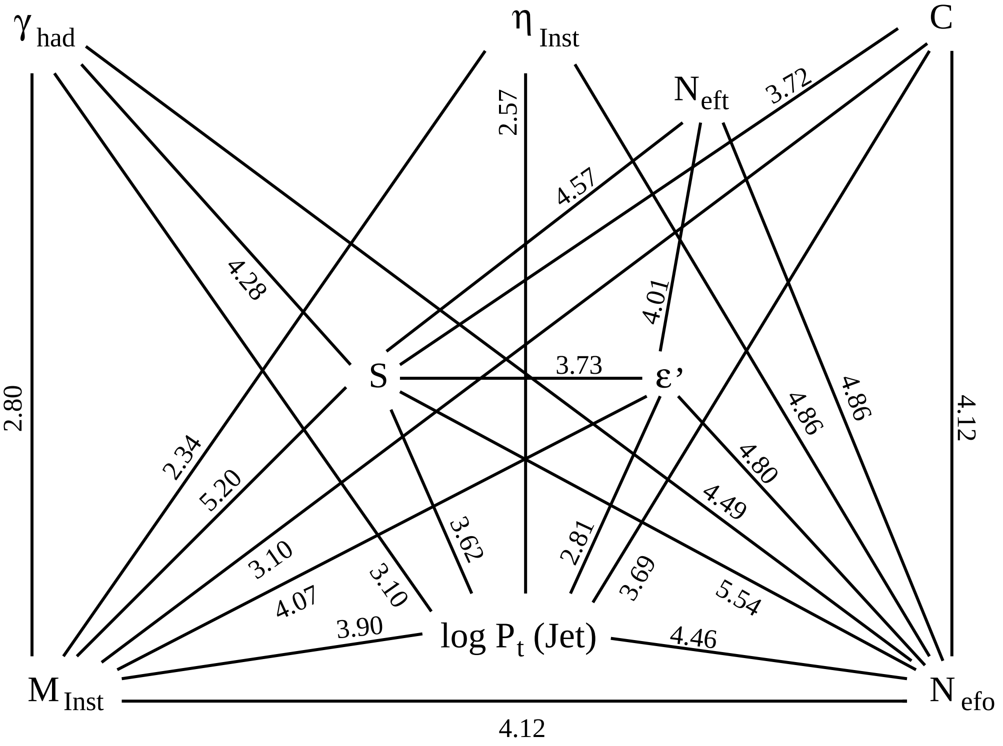


Figure 5.6: *The mean separation power* $< P_s >$*, obtained from the test cut on combinations of variables, averaged over distributions with* 8×8*,* 10×10 *and* 12×12 *bins, can be used as a figure of merit for the distribution, see text.*

The distributions which pass both requirements are listed in tables 1 to 3 in the Appendix, which also give the results one obtains when using 8×8 and 12×12 bins instead of 10×10. As expected, the change of bin size does not have a large influence on the set of distributions which pass the requirements. This proves, that distributions can be characterised by whether they pass the above conditions or not, and hence that a set of distributions for two-dimensional cuts may be selected on this basis.

The mean separation power, $< P_s >$, averaged over the three distributions with different bin sizes, can be used as a figure of merit of the distribution. These numbers are shown in the overview graph in Fig. 5.6, which summarises the results of this study.

The distributions, which were finally chosen for defining the two-dimensional cuts, are

- sphericity of the instanton region, S, vs. $\log_{10} p_T^{\mathrm{Jet}}$ of the current jet;
- circularity of the instanton region, C, vs. $\log_{10} p_T^{\mathrm{Jet}}$;
- S vs. zufo multiplicity in the instanton region, N_{efo};
- multiplicity of zufo tracks in the instanton region, N_{eft}, vs. N_{efo}.

These distributions are marked with an asterisk in tables 1 to 3 in the Appendix. They are shown in Fig. 5.10 for QCDINS and in Fig. 5.11 for Djangoh with Rapgap.

	S vs. $\log_{10} p_T^{\mathrm{Jet}}$	C vs. $\log_{10} p_T^{\mathrm{Jet}}$	S vs. N_{efo}	N_{eft} vs. N_{efo}
cut “A”	1.4	1.2	1.2	1.4
cut “B”	1.8	1.4	1.6	2.0
cut “C”	2.4	1.8	2.4	2.4

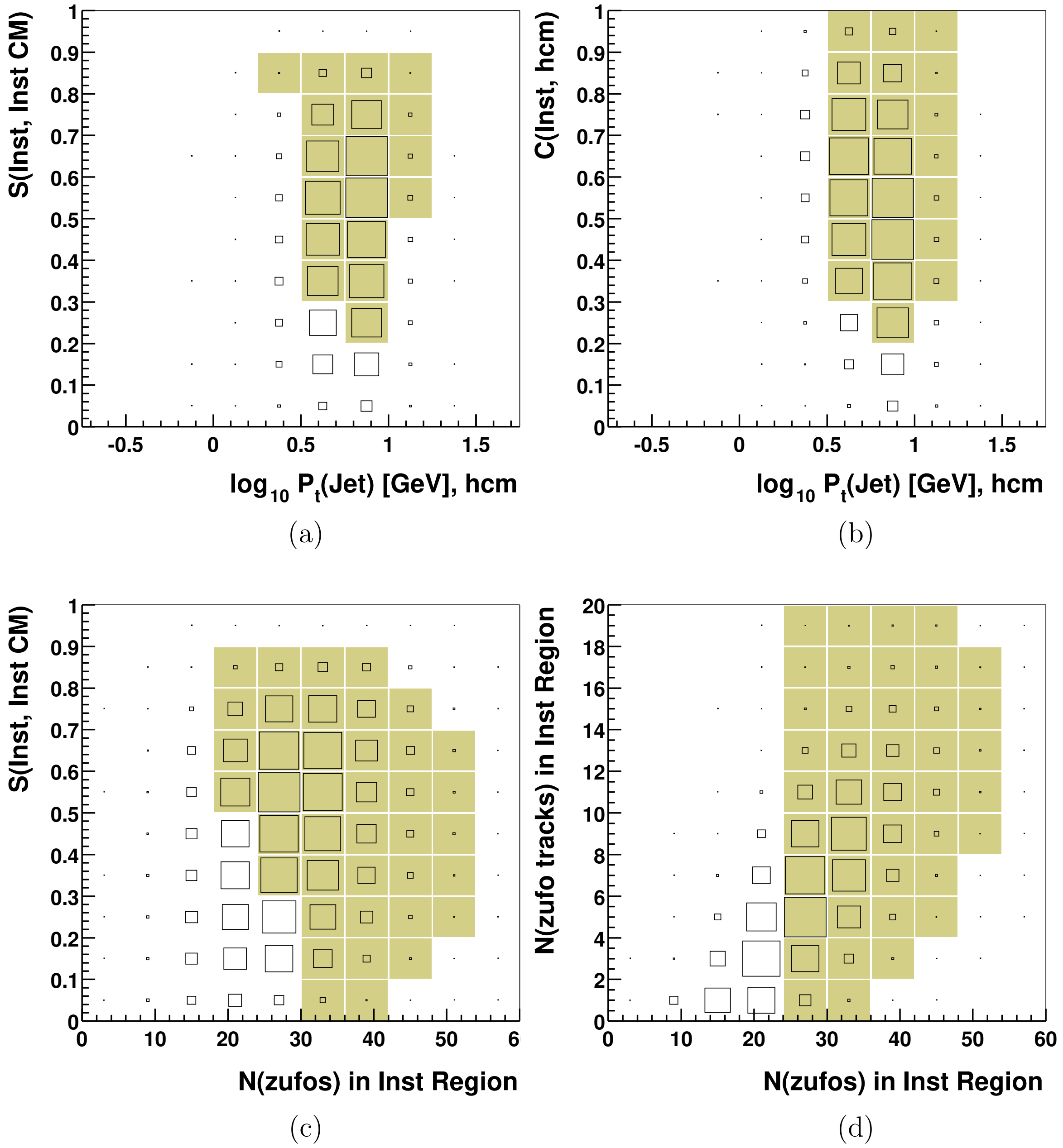
Table 5.1: *Minimum separation power $p_{i,\mathrm{min}}$ required in each bin to be included in bin-based cut on two-dimensional distributions.*

Step 2: Using the same set of four two-dimensional distributions, marked with an asterisk in tables 1 to 3, two different ways to cut have been tested.

In the first approach, called “bin based cuts” in what follows, the separation power of the two-dimensional bins from step 1 has again been used. For each of the four distributions, three cuts were defined labelled “A”, “B” and “C”, “C” being the most restrictive one. The part of a distribution selected by a cut is given by all histogram bins with more than 25 entries in the nDIS Monte Carlos and $p_i > p_{i,\mathrm{min}}$, where different values for $p_{i,\mathrm{min}}$ are used for cuts “A”, “B”, “C” and for the different distributions, as given in table 5.1 . The input distributions for the QCDINS MC are shown in Fig. 5.7 to 5.9, where the shaded bins indicate the area selected by the cuts. For each of the 81 combinations of three possible cuts on four distributions, the QCDINS efficiency r_I and the separation power P_s have been determined. The performance of this enhancement method can be judged from a plot of P_s vs. r_I, see Fig. 5.12. In the plot it has been ensured that from point to point P_s for Djangoh/Rapgap increases for decreasing QCDINS efficiency r_I, i.e. only the best of the 81 combinations are shown.

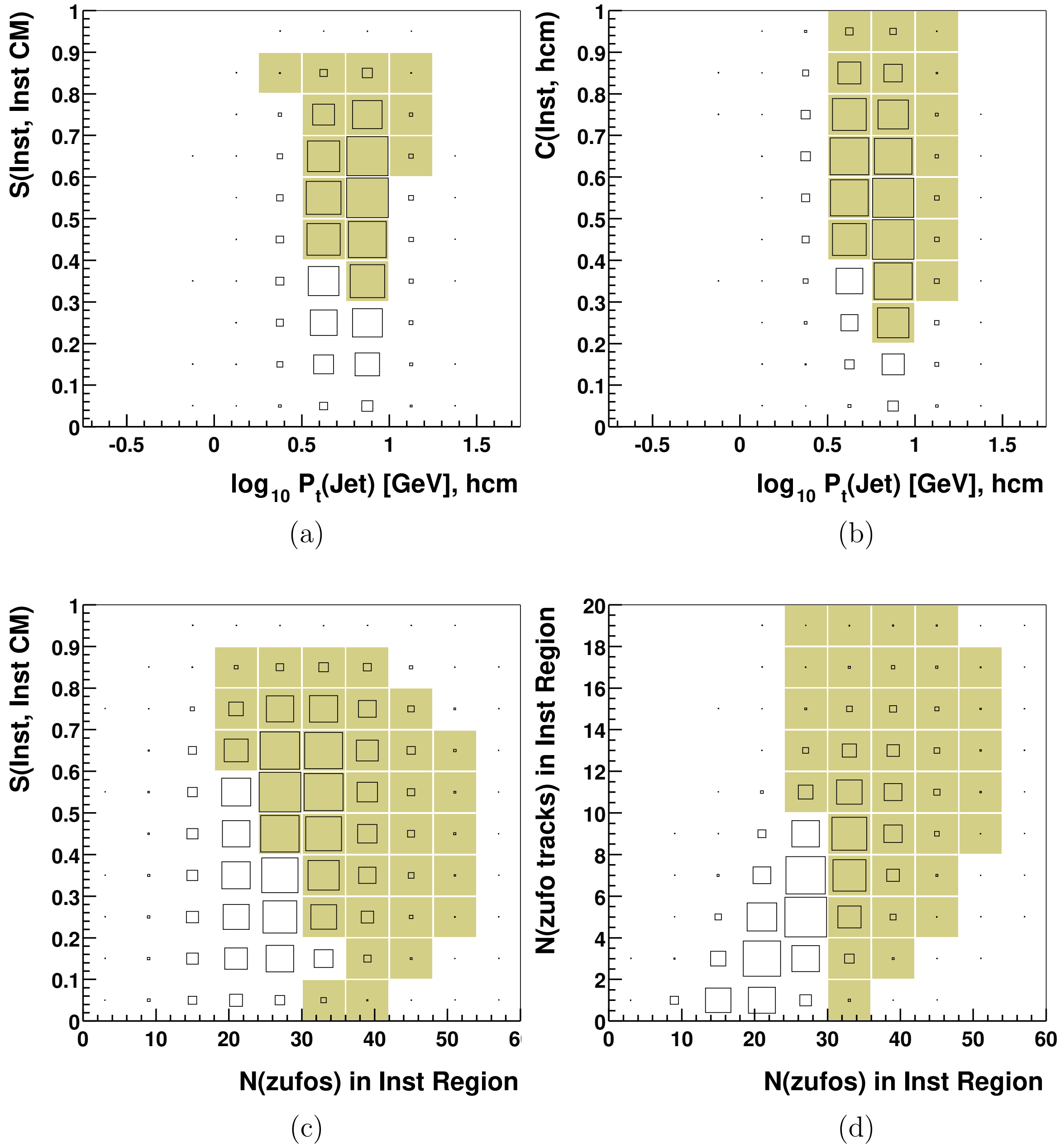
cut “A”	cut “B”	cut “C”
$S > 0.8 - \log_{10} p_T^{\mathrm{Jet}}$	$S > 1.0 - \log_{10} p_T^{\mathrm{Jet}}$	$S > 1.2 - \log_{10} p_T^{\mathrm{Jet}}$
$C > 0.55 - 0.5 \cdot \log_{10} p_T^{\mathrm{Jet}}$	$C > 0.75 - 0.5 \cdot \log_{10} p_T^{\mathrm{Jet}}$	$C > 0.95 - 0.5 \cdot \log_{10} p_T^{\mathrm{Jet}}$
$S > 2.5 - 0.1 \cdot N_{\mathrm{efo}}$	$S > 3.0 - 0.1 \cdot N_{\mathrm{efo}}$	$S > 3.5 - 0.1 \cdot N_{\mathrm{efo}}$
$N_{\mathrm{eft}} > 13. - 0.4 \cdot N_{\mathrm{efo}}$	$N_{\mathrm{eft}} > 16. - 0.4 \cdot N_{\mathrm{efo}}$	$N_{\mathrm{eft}} > 19. - 0.4 \cdot N_{\mathrm{efo}}$

Table 5.2: *Definition of two-dimensional straight line cuts.*



(a) (b) (c) (d)

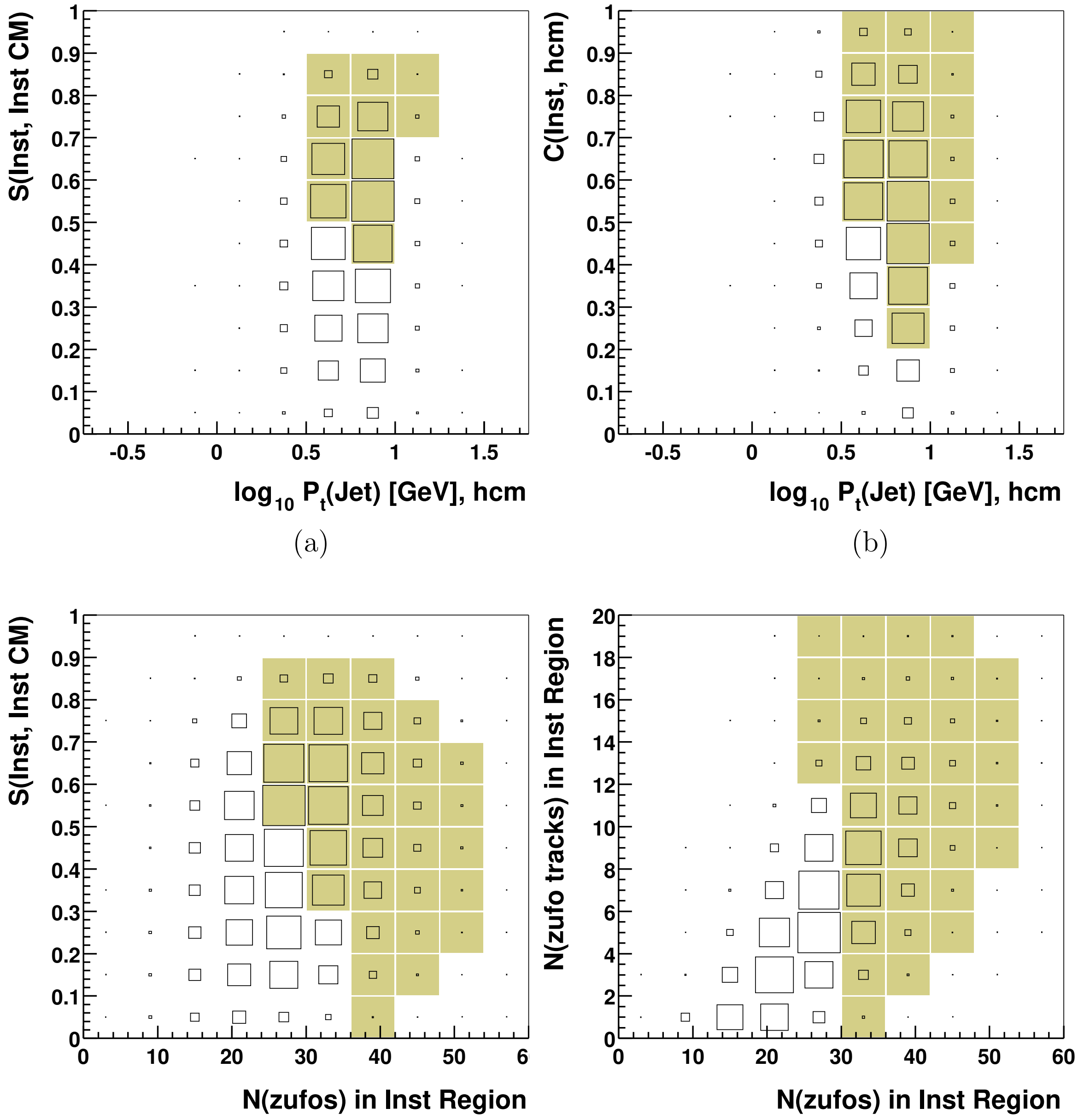
Figure 5.7: *Distributions used in the two-dimensional cut approach to enhance the instanton fraction. Plotted are the QCDINS distributions and the definition of cuts "A" of bin based two-dimensional cuts. The cuts select the shaded bins in each distribution. These correspond to the minimum separation power per bin* $p_{i,\min}$ *indicated for each plot, see also text and table 5.1. (a)* S *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 1.4$, *(b)* C *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 1.2$, *(c)* S *vs.* N_{efo}, $p_{i,\min} = 1.2$, *(d)* N_{eft} *vs.* N_{efo}, $p_{i,\min} = 1.4$.



(a) (b)

(c) (d)

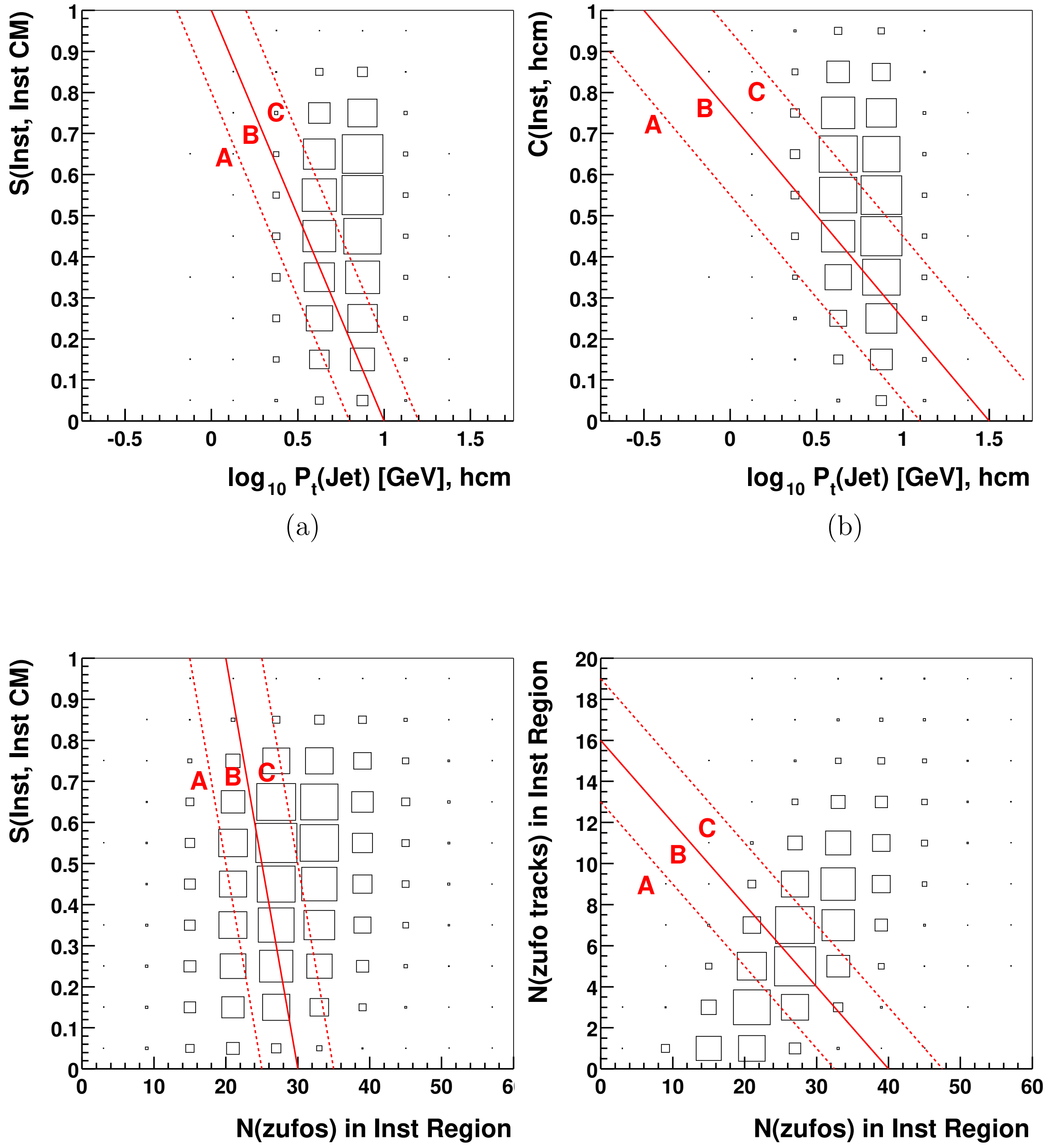
Figure 5.8: *Distributions used in the two-dimensional cut approach to enhance the instanton fraction. Plotted are the QCDINS distributions and the definition of cuts "B" of bin based two-dimensional cuts. The cuts select the shaded bins in each distribution. These correspond to the minimum separation power per bin* $p_{i,\min}$ *indicated for each plot, see also text and table 5.1. (a)* S *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 1.8$, *(b)* C *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 1.4$, *(c)* S *vs.* N_{efo}, $p_{i,\min} = 1.6$, *(d)* N_{eft} *vs.* N_{efo}, $p_{i,\min} = 2.0$.



(a) (b)

(c) (d)

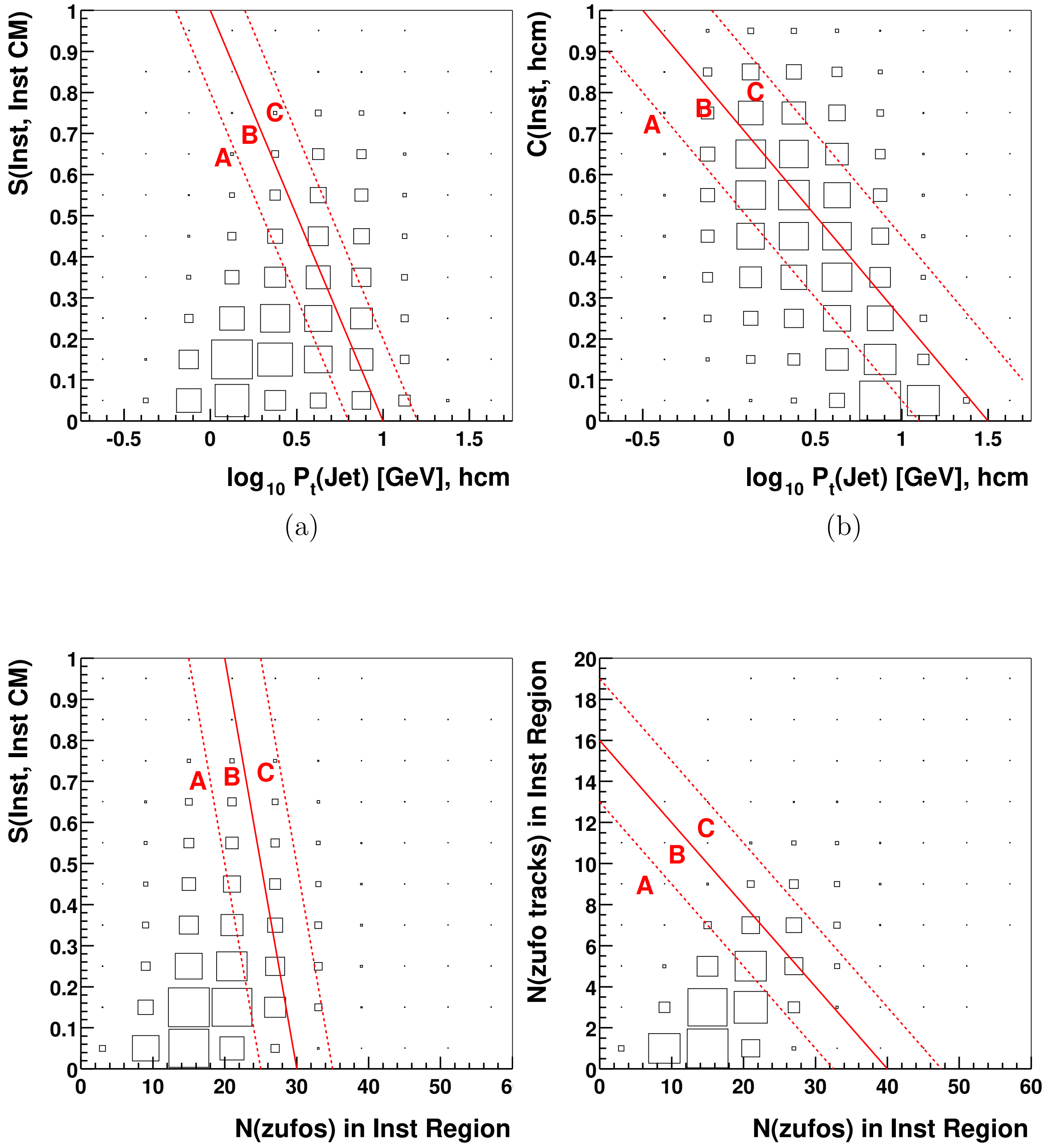
Figure 5.9: *Distributions used in the two-dimensional cut approach to enhance the instanton fraction. Plotted are the QCDINS distributions and the definition of cuts “C” of bin based two-dimensional cuts. The cuts select the shaded bins in each distribution. These correspond to the minimum separation power per bin* $p_{i,\min}$ *indicated for each plot, see also text and table 5.1. (a)* S *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 2.4$*, (b)* C *vs.* $\log_{10} p_T^{\mathrm{Jet}}$, $p_{i,\min} = 1.8$*, (c)* S *vs.* N_{efo}, $p_{i,\min} = 2.4$*, (d)* N_{eft} *vs.* N_{efo}, $p_{i,\min} = 2.4$.



(a) (b)

(c) (d)

Figure 5.10: *Combinations of variables used in the two-dimensional cut approach to enhance the instanton fraction. The straight line cuts are shown for the QCDINS MC distributions, see also text and table 5.2. (a) S vs.* $\log_{10} p_T^{\mathrm{Jet}}$*, (b) C vs.* $\log_{10} p_T^{\mathrm{Jet}}$*, (c) S vs.* N_{efo}*, (d)* N_{eft} *vs.* N_{efo}*.*



(a) (b)

(c) (d)

Figure 5.11: *Combinations of variables used in the two-dimensional cut approach to enhance the instanton fraction. The straight line cuts are shown for the Djangoh with Rapgap MC distributions, see also text and table 5.2. (a)* S *vs.* $\log_{10} p_T^{\mathrm{Jet}}$*, (b)* C *vs.* $\log_{10} p_T^{\mathrm{Jet}}$*, (c)* S *vs.* N_{efo}*, (d)* N_{eft} *vs.* N_{efo}.

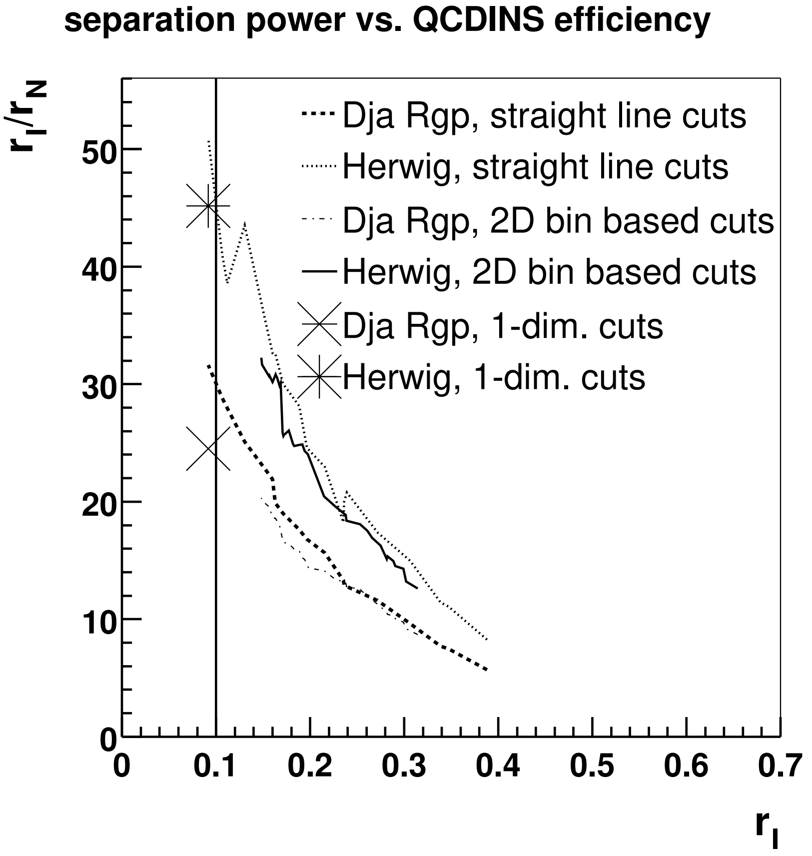


Figure 5.12: *Comparison of performance of straight line cuts and bin based cuts on two-dimensional distributions in terms of the separation power* $P_s = r_I/r_N$*, plotted as a function of the QCDINS efficiency* r_I*, see text. The separation power of the combination of one-dimensional cuts is also shown. "DjaRgp" stands for Djangoh with Rapgap MC.*

The second approach to define cuts uses straight lines to separate the part of the two-dimensional distributions with high QCDINS content from the one with high nDIS background. Again, three cuts "A", "B", "C" have been defined for each distribution, cf. table 5.2. They are drawn in Fig. 5.10 and 5.11, showing the QCDINS and Djangoh/Rapgap input distributions, respectively. In the same way as for the bin based cuts, P_s and r_I have been found for each combination. In Fig. 5.12, P_s vs. r_I is shown for the straight line cuts and compared to the separation power found for the bin based cuts. One can see that the straight line cuts cover a wider range of QCDINS efficiencies r_I. In the range of r_I values covered by both methods, both perform equally well. The plot also shows the result of the combination of one-dimensional cuts. One can clearly see that the combination of two-dimensional cuts rejects more nDIS events at the same QCDINS efficiency of $\approx 10\%$. For determining limits, three combinations of the straight line cuts as stated in table 5.3 have been used.

	fit approach	cut approach	
S vs. $\log_{10} p_T^{\mathrm{Jet}}$	A	C	C
C vs. $\log_{10} p_T^{\mathrm{Jet}}$	B	C	C
S vs. N_{efo}	C	C	C
N_{eft} vs. N_{efo}	C	B	C
label	"ABCC"	"CCCB"	"CCCC"
r_I	17.6%	10.7%	9.2%
P_s (DjaRgp)	15	29	32
P_s (Hwg)	23	41	51

Table 5.3: *Sets of two-dimensional straight line cuts used for determining limits by the fit and cut approaches and their instanton efficiency, r_I, and separation power, P_s, see text, for Djangoh with Rapgap ("DjaRgp") and for Herwig ("Hwg") MC.*

5.2.3 The Fisher Algorithm

Instanton-induced events can also be separated from nDIS by the Fisher algorithm [Fish36, BlLo98], which takes into account all correlations between a set of characteristic observables. In this approach, the combination of cuts on separate distributions is replaced by one cut on a combined variable, the Fisher discriminant t. The algorithm has been shown to yield an optimal separation of signal from background in the case of Gaussian input distributions and equal correlation matrices between observables for signal and background [MaOd93].

For n input variables x_i, $i = 1, 2, \ldots, n$, the mean values $\overline{x_i}^s$ for the signal and $\overline{x_i}^b$ for the background are determined. The correlation matrices between the variables are calculated according to

$$V_{ik,s} = \frac{1}{N} \sum_{\text{events}} (x_i^s - \overline{x_i}^s)(x_k^s - \overline{x_k}^s)$$

for the signal sample, and equivalently for the background ($V_{ik,b}$). These are averaged, $\overline{V}_{ik} = \frac{1}{2}(V_{ik,s} + V_{ik,b})$. The resulting matrix is inverted. The weights w_i are given by the expression

$$w_i = \sum_k (\overline{V}^{-1})_{ik} (\overline{x_k}^b - \overline{x_k}^s) \quad .$$

These are used in the definition of the Fisher discriminant,

$$t = \sum_i w_i x_i \quad .$$

One can think of the variable t as defining an optimised axis, onto which the QCDINS

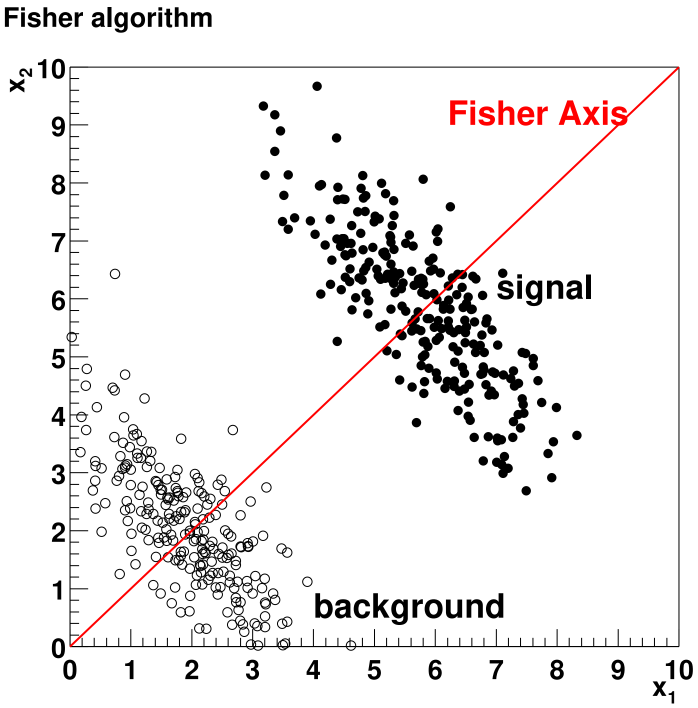
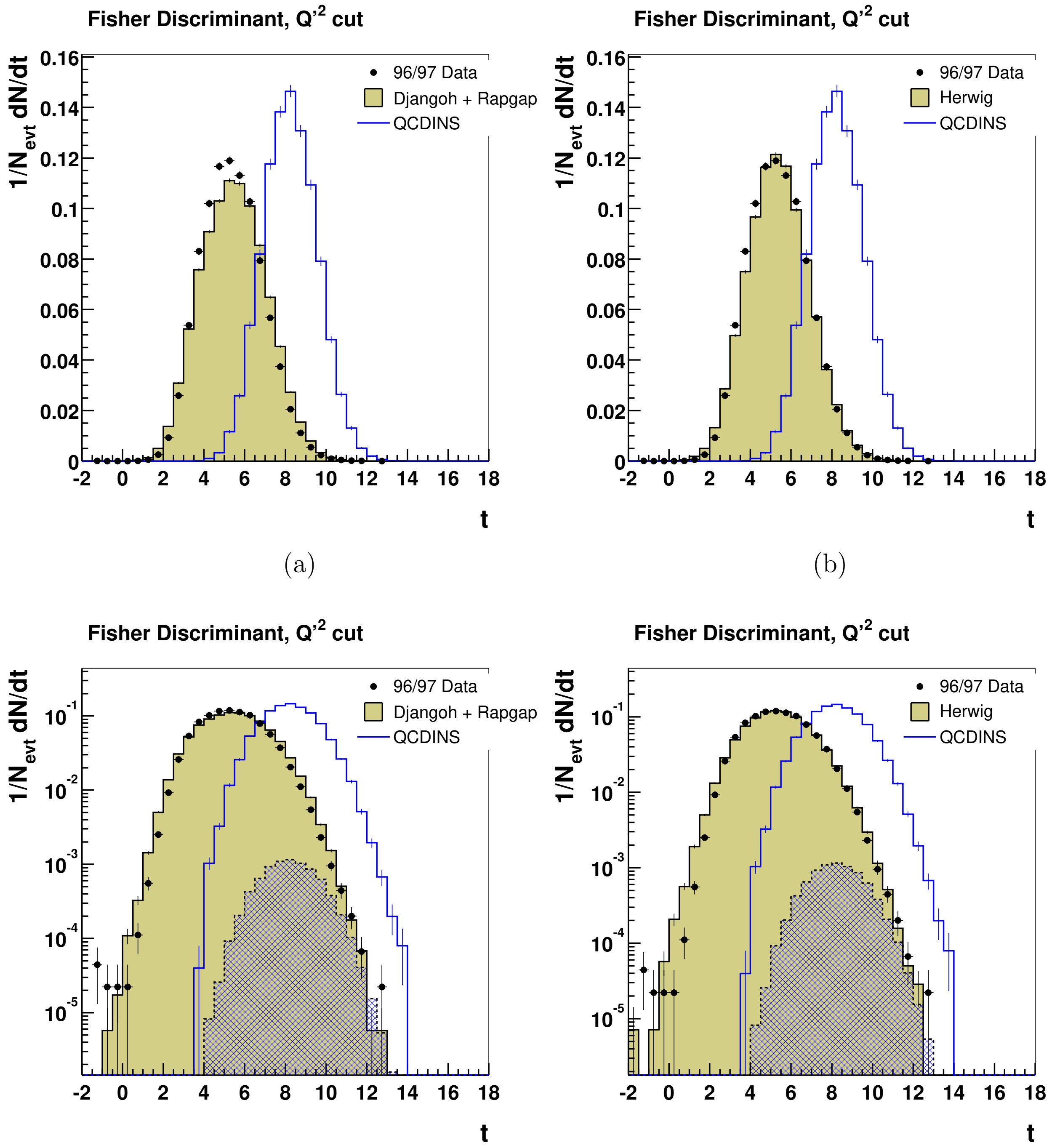


Figure 5.13: *Illustration of the Fisher algorithm for two input variables x_1 and x_2.*

and nDIS distributions are projected, see Fig. 5.13. Cutting on t correspond to cuts perpendicular to this Fisher axis.

Two different Fisher discriminants are used for the different approaches to set limits. For the fit approach, the discriminant is calculated from the same variables S, C, $\log_{10} p_T^{\mathrm{Jet}}$, N_{efo} and N_{eft}, which are used for the two-dimensional cuts. This variable is denoted by t'.
Fig. 5.15 (c) shows that the explicit use of all correlations yields an improvement in the instanton enhancement at a given QCDINS efficiency. For the fit approach, it is required that $t' > 8.0$ and additionally $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$.

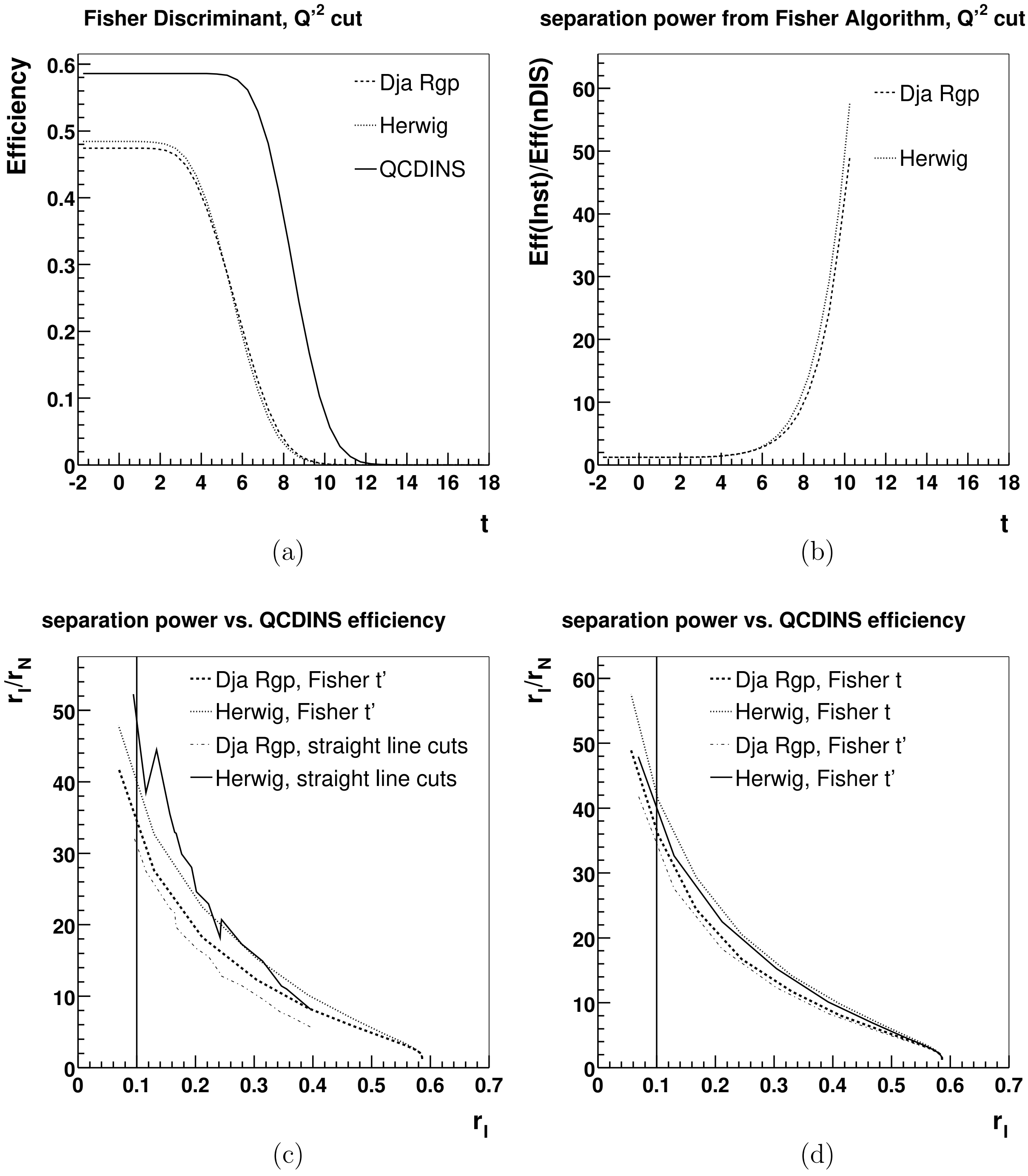
For the cut approach, another Fisher discriminant, denoted by t, is used, which is calculated by including the variable ϵ' in addition to S, C, $\log_{10} p_T^{\mathrm{Jet}}$, N_{efo} and N_{eft}. Requiring to keep about 10% of the QCDINS sample yields a cut value of $t = 9.5$, with the cut $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$ applied in addition. A distribution of this discriminant with $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$ applied is shown in Fig. 5.14. Fig. 5.15 shows the QCDINS efficiency r_I and the separation power P_s as functions of the cut value t. In Fig. 5.15 (d), values of the separation power of the two Fisher discriminants t' and t are compared. The inclusion of ϵ' gives a slightly better instanton enhancement.



(a) (b)

(c) (d)

Figure 5.14: *Fisher discriminant* t *calculated from the variables* S, C, $\log_{10} p_T^{\mathrm{Jet}}$, N_{efo}, N_{eft} *and* ϵ' *with an additional one-dimensional cut* $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$*. Shown are* t *distributions for (a) Djangoh/Rapgap as nDIS MC, (b) Herwig as nDIS MC, (c) Djangoh/Rapgap as nDIS MC (log plot) (d) Herwig as nDIS MC (log plot). The QCDINS distribution, normalised to the predicted fraction of* 7.9% *within the* $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$*-sample, is shown as shaded histogram on the logarithmic plots.*



(a) (b)

(c) (d)

Figure 5.15: *(a) QCDINS efficiency r_I as function of the cut on the Fisher discriminant t, (b) separation power P_s vs. cut value t, (c) comparison of performance of straight line cuts on two-dimensional distributions and a one-dimensional cut on t' calculated from the same set of variables, S, C, $\log_{10} p_T^{\mathrm{Jet}}$, N_{efo} and N_{eft}, (d) comparison of performance of the Fisher discriminant t' used for the fit approach to set limits and the Fisher discriminant t used for the cut approach.*

Chapter 6

Results

By restricting the inclusive sample as described in the previous chapter, it is possible to considerably increase the contribution of instanton-induced events to the remaining sample as compared to the inclusive one, thus increasing the sensitivity towards a possible instanton signal.

In order to determine the contribution of instanton-induced events to the enhanced samples, two approaches are made: the fit approach aims at extracting information on the instanton contribution from the shape of characteristic observables. The sphericity and the circularity of the instanton part of the hadronic final state are used for that purpose. Upper limits on the fraction of instanton-induced events in the enhanced samples are set.

However, it turns out that this approach is very sensitive to the exact description of the background. Therefore, another approach is taken, which yields a limit on the number of instanton-induced events, that is completely independent from the nDIS Monte Carlos. The idea is that the number of instanton-induced events cannot exceed the number of data events in a given region of phase space (plus a certain statistical fluctuation of that number). Extremely hard cuts, including cuts on sphericity and circularity are applied. A most conservative limit on the number of instanton-induced events in the remaining sample is set, assuming the nDIS background to be zero. This approach is referred to as cut approach in what follows.

6.1 QCDINS prediction for fraction of instanton-induced events

The predicted fraction $f_{\mathrm{I,th}}$ of instanton-induced events within a given data sample is defined as the ratio of the cross sections for instantons predicted by the QCDINS MC, σ_I, and the cross section of the data, σ_D:

$$f_{\mathrm{I,th}} = \frac{\sigma_I}{\sigma_D} \quad .$$

The QCDINS predicted cross section is calculated from the number of events, N_c, within the cuts that constrain the sample, the number of events generated, N_g, and the generated cross section, σ_g, as

$$\sigma_I = \frac{N_c}{N_g}\sigma_g \quad .$$

The data cross section is obtained from the number N_D of data events in the sample and the luminosity $\mathcal{L}$ according to $\sigma_D = N_D/\mathcal{L}$.

In the inclusive NC DIS sample, the predicted fraction of instanton-induced events is 0.65%. The predictions for the different samples used for the fit approach are given in table 6.1.

6.2 The Fit Approach

For an instanton enhanced sample, the shapes of distributions of characteristic variables, like sphericity and circularity, contain information on the instanton contribution within that sample. In order to extract that information, the maximum likelihood approach is applied: Distributions are fitted by a combination of nDIS and QCDINS Monte Carlo with the instanton contribution f_I as a free parameter of the fit.
For different values f_I, the sum of normal DIS Monte Carlo and QCDINS is compared with the data. The resulting χ^2 is given by:

$$\chi^2(f_\mathrm{I}) = \sum_{i=1}^{n_\mathrm{bins}} \frac{\{n^*_{iD} - [f_\mathrm{I} \cdot n^*_{iI} + (1-f_\mathrm{I}) \cdot n^*_{iN}]\}^2}{{\sigma^*_{iD}}^2 + f_\mathrm{I}^2 {\sigma^*_{iI}}^2 + (1-f_\mathrm{I})^2 {\sigma^*_{iN}}^2} \quad .$$

In this expression, n^*_{iD}, n^*_{iN} and n^*_{iI} are the number of data, normal DIS and QCDINS events within bin i, σ^*_{iD}, σ^*_{iN} and σ^*_{iI} the respective statistical errors and the asterisk indicates, that each of the distributions has been normalised to 1 within the cuts considered. The statistical error σ^*_i for each of the samples is given by the expression $\sigma^*_i = \sqrt{n_i}/n$, where n_i is the number of events in bin i and n the total number of events in the sample considered.
For the maximum likelihood approach, the function $\chi^2(f_I)$ is regarded as negative log likelihood function. The 2σ limit is obtained according to the maximum likelihood method. That method assumes that the Monte Carlo describes the data correctly and allows the 2σ limit, $f_\mathrm{I,lim}$, for the fit parameter, i.e. the instanton fraction, to be determined from:

$$\chi^2\left(f_\mathrm{I,lim}\right) = \chi^2_\mathrm{min} + 4$$

or

$$\chi^2\left(f_\mathrm{I,lim}\right) = \chi^2(f_I = 0) + 4$$

in case the minimum χ^2 is found to be at a negative, hence unphysical, f_I-value.

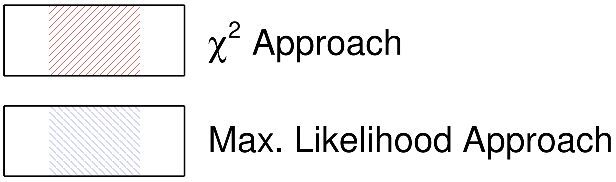


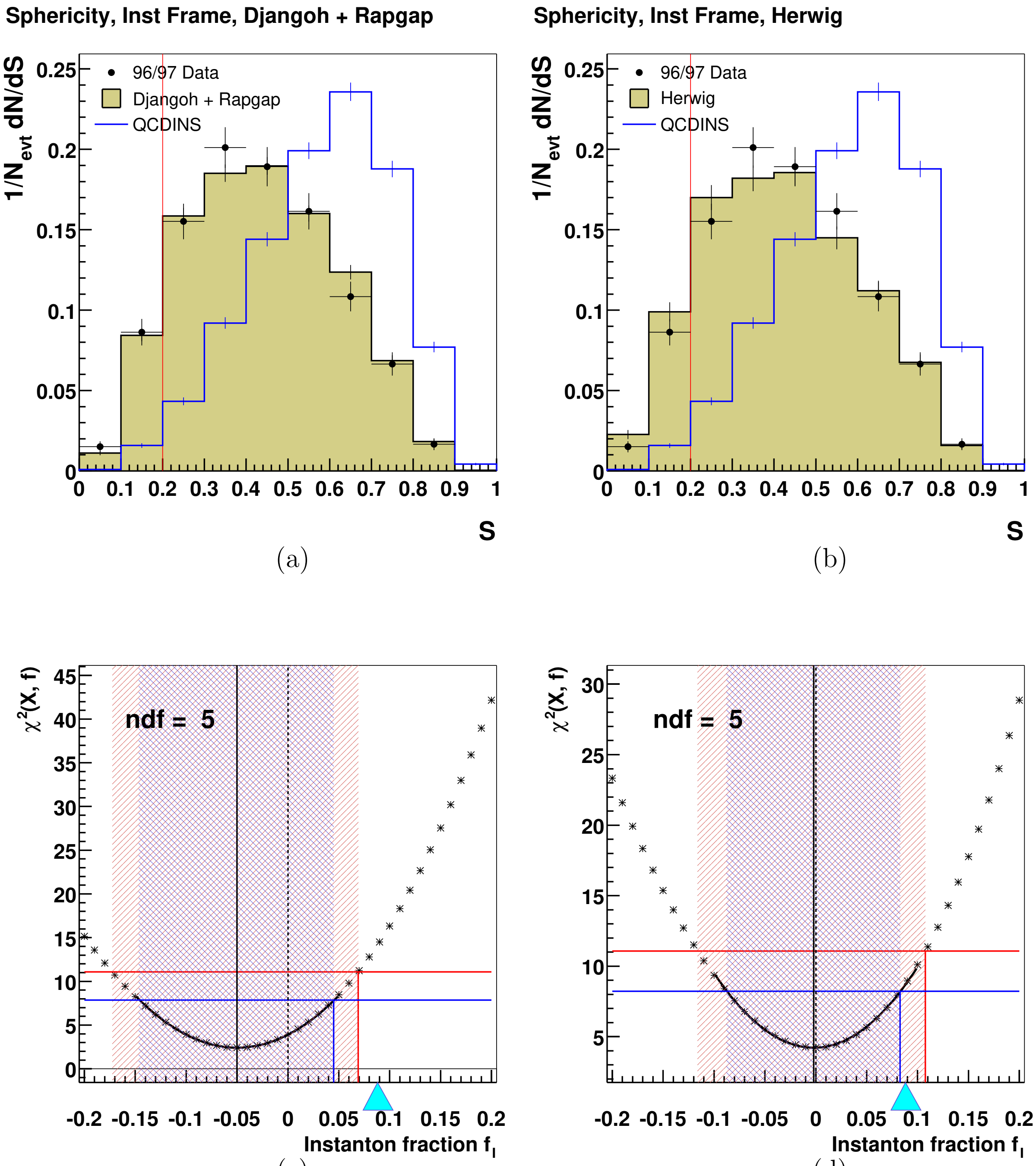
Figure 6.1: *Legend for the fits shown in Figures 6.2 to 6.4.*

Distributions fitted are the sphericity S of the instanton region within the three different instanton enhanced samples described in sections 5.2.1, 5.2.2 and 5.2.3, and the instanton circularity C , as well as C vs. S for the sample obtained from the combination of one-dimensional cuts, cf. section 5.2.1. As the first two bins of both the S and C distributions for the sample from one-dimensional cuts were not well described when the analysis was begun, a cut $C > 0.2$ is applied in the fits of the samples obtained from one-dimensional and from two-dimensional cuts. A cut $S > 0.2$ was applied in all the fits presented in this section. Fig. 6.2 to 6.4 (a) and (b) show the sphericity distribution in the different instanton-enhanced samples. The cut $S > 0.2$ has only been applied in the fit, but not when plotting the distributions.

By adjusting the diffractive contribution, see section 3.3, and the width of the P_t distribution of primary hadrons, cf. section 6.4.1, it has been ensured that the sphericity distribution is well described for the inclusive sample. Yet, because of correlations with other quantities, it might not be described within the instanton-enhanced sample. In order to check the hypothesis, that the Monte Carlo describes the data correctly, a χ^2 test has been performed for different fractions of instanton-induced events and it has been required that in the range of instanton fractions $f_I > 0$ the minimum χ^2 lies within the boundaries obtained from this χ^2 test. This check will be referred to as χ^2 approach in what follows.

The χ^2 functions for the different S distributions are shown in Fig. 6.2 to 6.4 (c) and (d). Two hatched regions are plotted, see the legend in Fig. 6.1. One of them indicates the 2σ limit, which one obtains from the χ^2 test of the hypothesis that the MC prediction describes the data. The other one shows the allowed region corresponding to a 2σ variation of the fit parameter, i.e. the instanton fraction, obtained from the maximum likelihood method. The predicted instanton fraction is indicated by the filled triangle on the f_I axis.

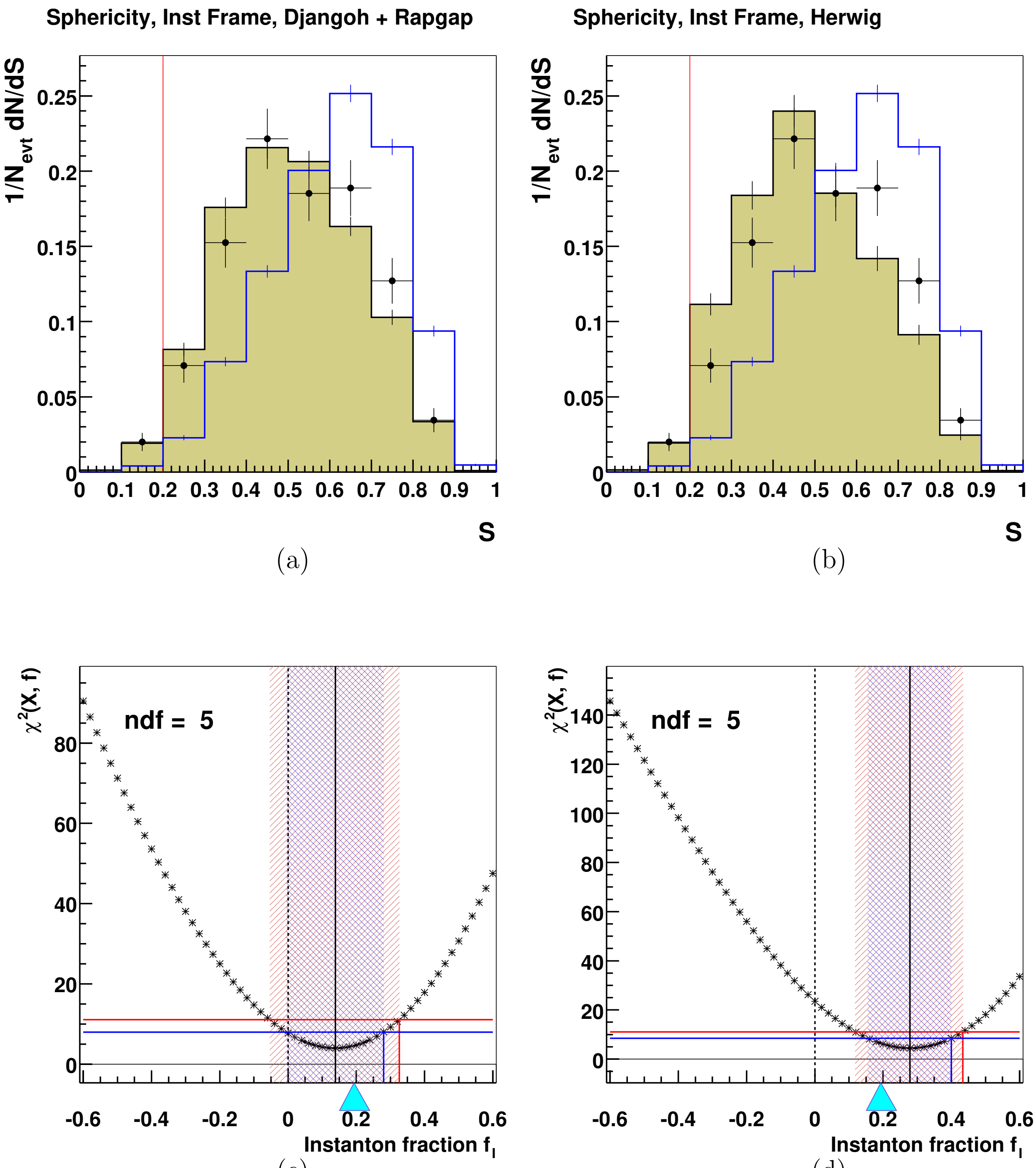
An overview over the results from the sphericity fits, not taking systematic effects into account, is given in table 6.1, sorted with respect to the predicted instanton fraction, $f_{\mathrm{I,th}}$, in each of the samples considered. All distributions yield good fits according to the χ^2 test. While an instanton fraction of zero is favoured by the sample obtained from one-dimensional cuts, the other two samples, with a higher $f_{\mathrm{I,th}} \approx 20\%$ and $\approx 30\%$,



(a) (b)

(c) (d)

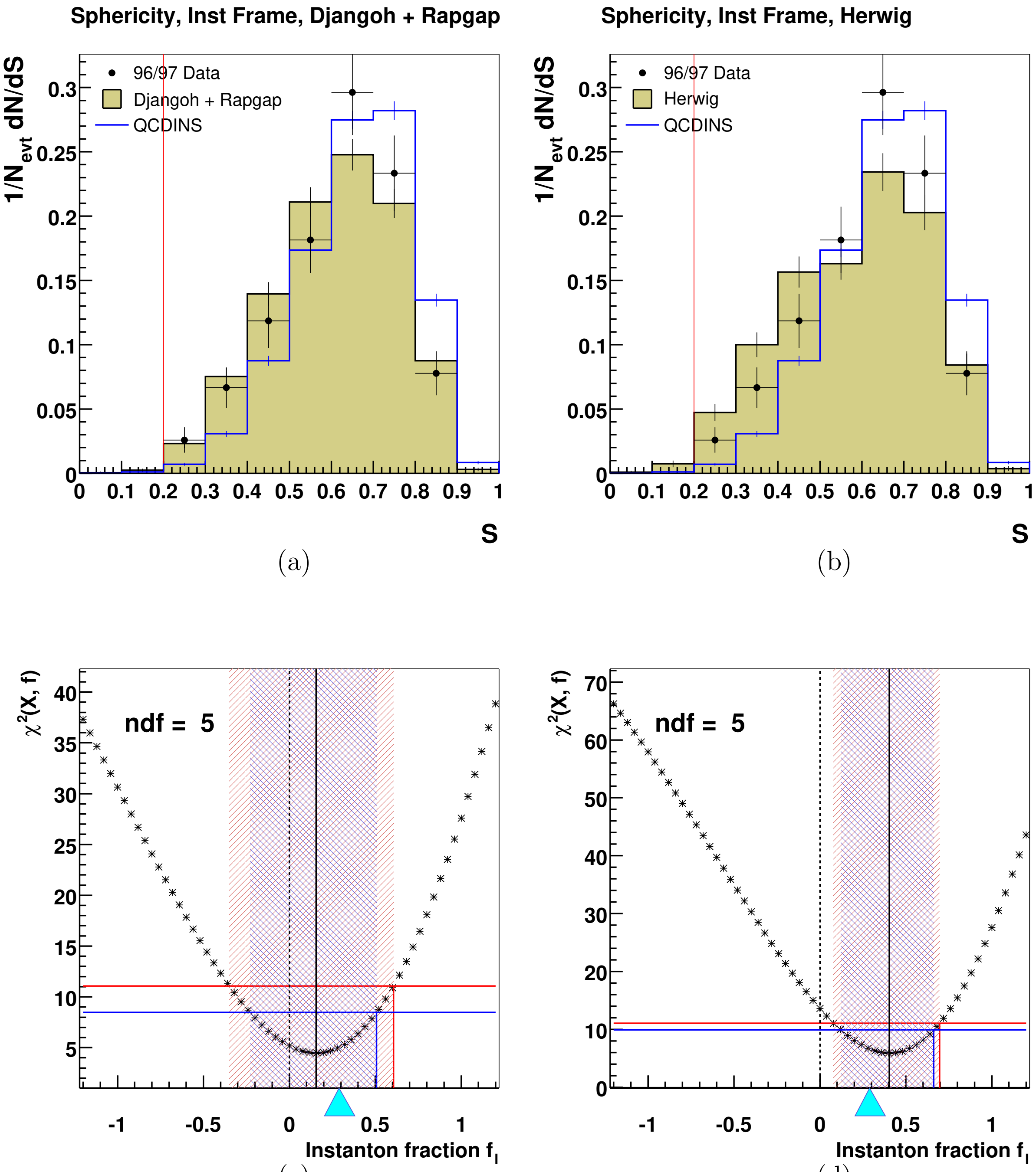
Figure 6.2: *Sphericity S of the instanton part of the hadronic final state in its CM system for the instanton-enhanced sample obtained from a combination of one-dimensional cuts given in section 5.2.1 for (a) Djangoh with Rapgap and (b) Herwig as nDIS MC (The cut on S itself is not applied in plots (a), (b), but the first two bins are not used in the fit). The lower plots show the χ^2 value obtained from a fit of sphericity as function of the instanton fraction in the sample for (c) Djangoh with Rapgap and (d) Herwig.*



(a) (b)

(c) (d)

Figure 6.3: *Sphericity S of the instanton part of the hadronic final state in its CM system for the instanton-enhanced sample obtained from the two-dimensional straight line cut combination "ABCC", cf. section 5.2.2, a cut $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$ and cuts $S, C > 0.2$. (The cut on S itself is not applied in plots (a), (b), but the first two bins are not used in the fit). As nDIS MC (a) Djangoh with Rapgap and (b) Herwig is used. The lower plots show the χ^2 value obtained from a fit of sphericity as function of the instanton fraction in the sample for (c) Djangoh with Rapgap and (d) Herwig.*



(a) (b)

(c) (d)

Figure 6.4: *Sphericity S of the instanton part of the hadronic final state in its CM system for the instanton-enhanced sample, obtained from a cut on the 5 variable Fisher discriminant t', cf. section 5.2.3, a cut $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$ and a cut $S > 0.2$. (The cut on S itself is not applied in plots (a), (b), but the first two bins are not used in the fit). As nDIS MC (a) Djangoh with Rapgap and (b) Herwig is used. The lower plots show the χ^2 value obtained from a fit of sphericity as function of the instanton fraction in the sample for (c) Djangoh with Rapgap and (d) Herwig.*

				2σ upper limits on f_I			
	$f_{\mathrm{I,th}}$	fit minimum		χ^2 approach		max. likelihood	
		Hwg	DjaRgp	Hwg	DjaRgp	Hwg	DjaRgp
1D	0.088	−0.002	−0.050	0.108	0.069	0.082	0.045
2D	0.19	0.279	0.139	0.434	0.326	0.400	0.280
F.	0.28	0.404	0.156	0.697	0.507	0.662	0.735

Table 6.1: *Results of fits of the sphericity S of the instanton part of the event within different sets of instanton enhancing cuts. Fit minima and limits on the instanton fraction f_I are given and compared to the predicted fraction $f_{\mathrm{I,th}}$. "1D" corresponds to the combination of one-dimensional cuts, "2D" to the combination "ABCC" of straight line cuts on two-dimensional distributions and "F." to a cut on the Fisher discriminant, $t' > 8.0$. For "2D" and "F.", a one-dimensional cut $Q'^2_{\mathrm{DA}} < 250\,\mathrm{GeV}^2$ has been applied in addition. "Hwg" stands for "Herwig", "DjaRgp" for Djangoh with Rapgap and "max. likelihood" for the maximum likelihood approach.*

respectively, favour a contribution from instanton-induced events over a pure nDIS description. For the samples obtained from one-dimensional cuts, the 2σ limit from the maximum likelihood method is below the QCDINS prediction for both nDIS Monte Carlos. However, this result is found without taking any systematic uncertainties into account. For the fits of S, C and C vs. S within the combination of one-dimensional cuts, systematic effects have been studied, see the following section. The results can be found in tables 4 to 7 in the Appendix. For a discussion of the fit result including systematic effects, cf. section 6.5.

6.3 Systematic Studies

Various sources of systematic uncertainty have been considered and checks on the method performed. In each case, the fits within one-dimensional instanton-enhancing cuts have been repeated under changed conditions for both the nDIS Monte Carlos. The effect of each of the systematic variations of cuts and checks of the method on the fit of sphericity S, circularity C and C vs. S is given in tables 4 to 6 in the Appendix. The χ^2 values at the fit minimum or at $f_I = 0$ (in case χ^2 is minimal at a negative, hence unphysical, instanton fraction f_I) are compared to the numbers of degrees of freedom in table 7, see Appendix. A graphical overview over the predicted fraction and the 2σ upper limits from the maximum likelihood approach is shown in Fig. 6.5. The resulting upper limit is given by the largest upper limit obtained in these studies. Errors introduced when cutting on a variable, which is not perfectly described, were estimated by varying the cut value by a small amount.

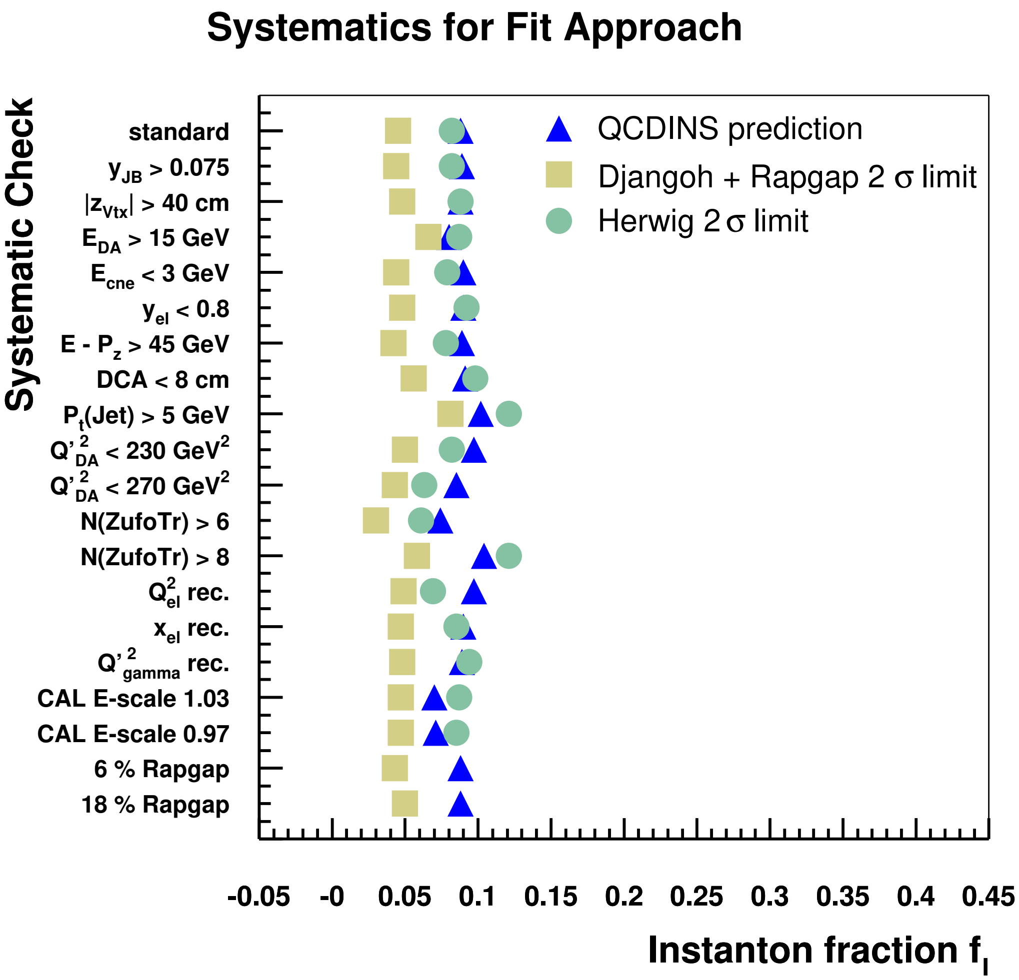


Figure 6.5: *Graphical overview over the systematic studies. Shown are the maximum likelihood* 2σ *limits on the instanton fraction within the instanton enhanced sample obtained from a combination of one-dimensional cuts, cf. section 5.2.1, found from the fit of the sphericity of the instanton region.*

At offline cut level, the following cuts were taken into account:

- The cut on y_{JB} was increased from 0.05 to 0.075.
- The vertex cut was relaxed from 50 cm to 40 cm.
- The electron energy cut was increased from $E_{\mathrm{DA}} > 10\,\mathrm{GeV}$ to $E_{\mathrm{DA}} > 15\,\mathrm{GeV}$.
- The cut on E_{cne} was lowered from 5 GeV to 3 GeV.
- The cut on y_{el} was lowered from $y_{\mathrm{el}} > 0.9$ to $y_{\mathrm{el}} > 0.8$.
- The $E - P_z$ balance cut was made more restrictive by increasing the lower cut value from 35 GeV to 45 GeV.

- The distance of closest approach of the electron , DCA, was required to be larger than 8 cm instead of 10 cm.

For the two-dimensional fit of circularity vs. sphericity, the variation of the electron energy and of the DCA had the largest effect. For E_{DA}, the maximum likelihood 2σ limit decreased by $\approx 20\%$, while increasing by $\approx 15\%$ for the modified DCA cut. One should note, however, that for the E_{DA} variation also the predicted instanton fraction is smaller by $\approx 10\%$.
Among the offline cuts, the DCA-cut variation also gave rise to the largest change for the one-dimensional sphericity fit, increasing the 2σ limit by 25%. For the circularity fit, a decrease of $\approx 15\%$ in the 2σ limit was found for the increased E_{DA} cut.

Not only the offline cuts, but also the instanton enhancing cuts have been varied, giving rise to larger changes in the limit:

- The cut on the current jet transverse momentum has been increased from 4 GeV to 5 GeV. This results in the largest change in the limit encountered, namely to a 50% increase in the 2σ maximum likelihood limits from the sphericity fit and a 30% increase in the limit obtained from circularity. For the combined, two-dimensional fit, the limit for Djangoh with Rapgap is increased by about 75%, yielding the highest upper limit found for that Monte Carlo from that distribution. For Herwig the upper limit for the two-dimensional distribution only rises by 15% when the P_t cut is increased.

- The upper cut on Q'^2_{DA} has been varied by $\pm 20\,\mathrm{GeV}^2$ around its central value of $250\,\mathrm{GeV}^2$. The maximum likelihood limit obtained from sphericity and from the two-dimensional fit increases when the Q'^2 cut is decreased and vice versa. For the circularity fit, the limit becomes slightly smaller in both cases, maybe indicating a statistical fluctuation at the central value. While the upward shift is $\approx 10\%$ for the sphericity fit and for circularity vs. sphericity, the decrease is of the order $\approx 25\%$ for the sphericity fit.

- The cut on the number of zufo tracks, N_{eft}, has been varied around its central value 7 by ±1. Despite this being the smallest variation possible for an integer type variable, the increase of this cut value gives rise to the largest systematic effect encountered for the fit of sphericity and the two-dimensional distributions, yielding the maximum values for the limit, of 12.1% for S and 7.3% for C vs. S. For the circularity fit, the variation has the opposite effect than for the other distributions, yielding a larger limit at a smaller cut value.

In addition to the systematic variation of cut values, the sensitivity of the resulting limit to some other aspects of the analysis has been studied:

- The variables Q^2 and x, that define the kinematic region, in which the search is performed, were reconstructed using the electron method instead of the double angle method. While the change in the reconstruction of x has hardly any influence on the fits, the usage of Q^2_{el} instead of Q^2_{DA} shifts the instanton fraction, at which χ^2 is minimal, to much lower values for the two-dimensional fit: The minimum is shifted from -2.4% to -9.3% for Herwig and from -4.5% to -10.2% for Djangoh. Although this does not increase the limit and therefore has no influence on the result, it casts some doubt on the reliability of the fit approach and of the position of the fit minimum.

- Similarly, ${Q'}^2$ was reconstructed using the "photon-method" instead of the double angle method. It results in an increase of the Herwig limit by $\approx 15\%$ for sphericity and for C vs. S.

- The uncertainty in the energy scale of the calorimeter was accounted for by varying the energy of all those zufos which use calorimeter energy information by $\pm 3\%$. The variation was only applied in the data, leaving the MC samples unchanged. Interestingly, the resulting limit from the sphericity fit increases for both an increase and a decrease of the energy scale, for both nDIS Monte Carlos. In contrast, for circularity and the two-dimensional distribution, one obtains smaller limits for both an increase and a decrease of the energy scale.

- For the Djangoh with Rapgap sample, the fraction of diffractive Monte Carlo has been varied by $\pm 50\%$, yielding an increase in the 2σ maximum likelihood limit of $\approx 10\%$ for the sphericity fit and of $\approx 15\%$ for the two-dimensional one, in case of an increased Rapgap contribution.

6.4 Tuning of the Djangoh Monte Carlo

6.4.1 Variation of the width of the inital hadron P_t distribution

A limit on the instanton contribution has been set from a fit of the sphericity distribution within instanton-enhancing cuts. However, this limit is valid only, if the Monte Carlo describes the data correctly. A minimum requirement is therefore, that for the inclusive sample a pure normal DIS MC prediction describes the data. In such a sample, the predicted fraction of $\approx 0.5\%$ is below the measuring accuracy and hence negligible. As one can see from Fig. 5.4, the sphericity distribution of the inclusive sample is well described for the nDIS MC sample used. However, to obtain such a good agreement, it was necessary to adjust one of the input parameters of the Djangoh Monte Carlo. The results from the preceding study are presented in this section.

The Djangoh sample formerly used had been generated with the ZEUS standard settings of input parameters, which had been obtained by one of the LEP experiments[1].

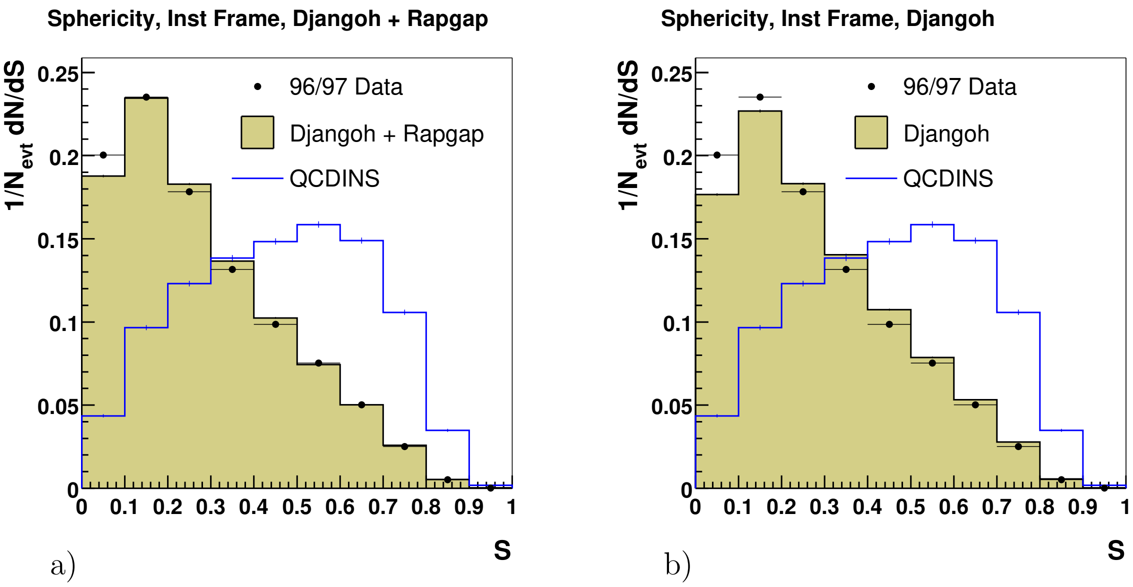


Figure 6.6: *Sphericity distribution of the instanton region of the hadronic final state for (a): untuned Djangoh(Ariadne) MC with Rapgap contribution, (b): untuned Djangoh(Ariadne) MC without Rapgap contribution.*

The corresponding sphericity distribution is shown in Fig. 6.6 (a). Discrepancies between data and MC are found in many bins. Especially, the maximum region is poorly described, but also the slope of the falling edge, where the instanton distribution peaks, is not correctly reproduced. In order to solve that problem, different input parameters of the Djangoh MC were changed.

As it would have taken too much time to run the detector simulation for each of these settings, the hadron level sphericity distribution was reweighted to detector level and compared to the data. Besides, instead of cutting on the zufo multiplicity, a cut on the multiplicity of zufo tracks, $N_{\text{eft}} > 2$, was applied for the data, and a cut on the multiplicity of charged particles for the Djangoh sample. A multiplicity cut is necessary, because sphericity is only defined for systems of at least three four-momenta. The cut variable was changed for this study, because there is no hadron level quantity directly corresponding to the zufo multiplicity. Another difference from the plots shown in former sections is that no Rapgap contribution was added. Without Rapgap, the discrepancies between data and Monte Carlo are even larger, see Fig. 6.6 (b). This is acceptable, as the aim of this study is only to determine the sign and to estimate the size of changes in the distribution due to changes in the input parameters.

---

[1] Technically, this set is labelled “EMC/DELPHI” in the ariadne routine artune.fpp steering the Monte Carlo generation.

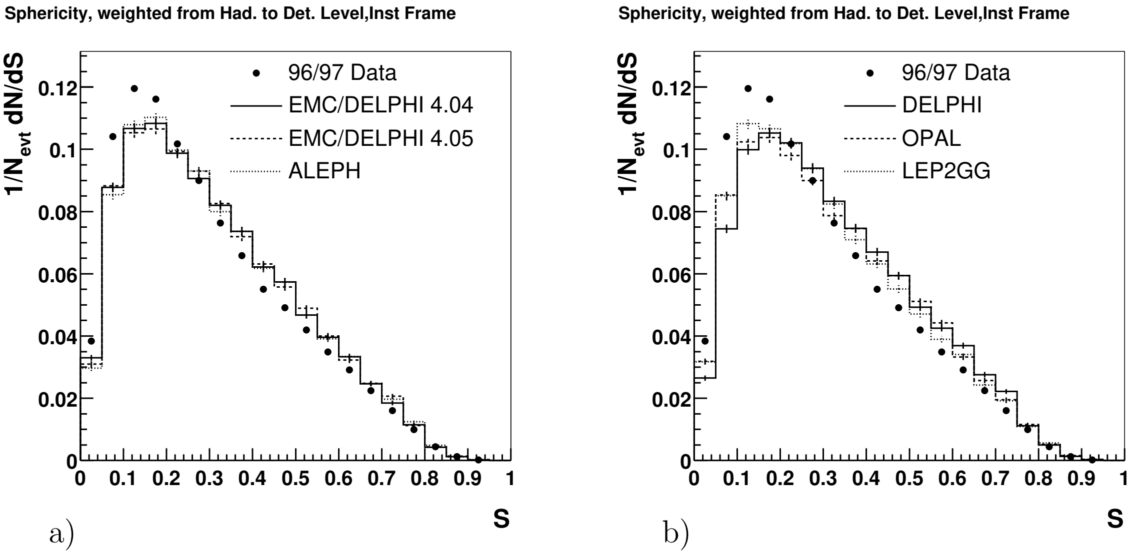


Figure 6.7: *Sphericity distribution of instanton region of the hadronic final state for different LEP tunings of Djangoh(Ariadne) MC. Hadron level MC is reweighted to detector level.*

The following changes were studied: Firstly, the default set of Djangoh parameters was replaced by other LEP tunings, cf. Fig. 6.7. No significant improvement was found.

Secondly, two parameters labelled PARA(10) and PARA(25) were varied. The parameter PARA(10) corresponds to the power in the soft suppression for the proton remnant. The parameter PARA(25) governs the probability of emissions outside the soft suppression cut-off [Sjö94].
Fig. 6.8 shows the result obtained for different values of PARA(25), namely (1.2), 1.5 and 1.8 for two different values of PARA(10). In Fig. 6.8 (a), PARA(10) = 1.0, in (b) PARA(10) = 1.8. This is to be compared to the default values of PARA(10) = 1.0 and PARA(25) = 2.0. None of the combinations investigated gives sufficient improvement.

In a third attempt to improve the description, the parameter PARA(15) describing the power in the soft suppression for the struck quark was varied to 0.5 from its default value of 1.0 . This also did not improve the description. All the parameters PARA(10), PARA(25) and PARA(15) have in common that they steer the MC generation at parton level.

Finally, a hadronisation parameter, PARJ(21), representing the width σ of the Gaussian p_x and p_y transverse momentum distributions of primary hadrons, was varied. Fig. 6.8 (d) shows the resulting sphericity distributions for the parameter values 0.10, 0.28 and 0.42.
The ZEUS default value of this parameter is 0.405, while the default value given in the hadronisation MC manual is 0.36 [Sjö94]. One can see from the plot that the

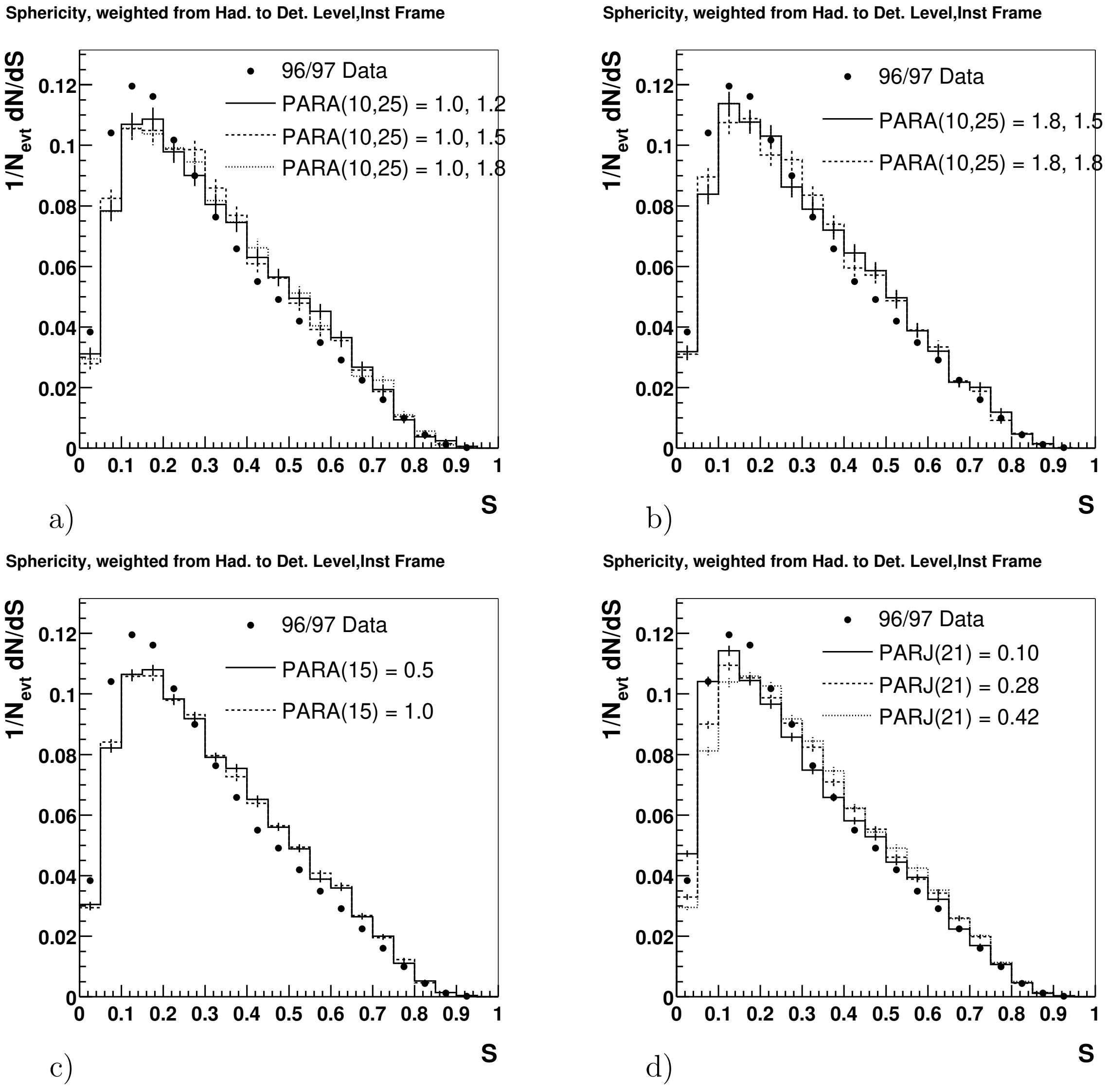


Figure 6.8: *Sphericity for different values of Djangoh tuning parameters, see text: (a): PARA(10) = 1.0 const., PARA(25) varied, (b): PARA(10) = 1.8 const., PARA(25) varied, (c): PARA(15) varied, (d): PARJ(21) varied. Hadron level MC is reweighted to detector level.*

sphericity distribution is sensitive to changes of this parameter and that lowering its value improves the description of the data. Using the value PARJ(21) = 0.28, a new Djangoh MC sample was generated and passed through the detector simulation. This is the sample used for the search, yielding the sphericity distribution shown in Fig. 5.4.

6.4.2 Variation of the fraction of BGF events in the Djangoh MC

In instanton-induced events, the hard subprocess involves a gluon emitted by the proton, cf. Fig. 1.9. In order to check the sensitivity of the sphericity fit to the fraction of boson-gluon fusion events in the Djangoh MC, χ^2 values have been calculated on a grid of values of the fraction f_I of instanton-induced events and of the fraction f_{BGF} of BGF events in the Djangoh sample within instanton-enhancing cuts. The study has been performed at an earlier stage of the analysis, than what has been presented so far. Therefore, a pure Djangoh sample without Rapgap contribution was used. The width of the transverse momentum distributions of primary hadrons, PARJ(21), was still set to 0.405 instead of 0.28, see previous section. Also, older versions of the data reconstruction code and of the "orange" code to write the ntuple were used. The Q'^2 offline level cut was at 130 GeV^2 instead of 140 GeV^2. Fig. 6.9 (a) shows the resulting χ^2 distribution as a function of f_I and the Djangoh BGF fraction normalised to its default value $f_{\mathrm{def}} \approx 0.58$ within the instanton-enhanced sample obtained by applying the one-dimensional enhancing cuts, $f_{\mathrm{BGF}}/f_{\mathrm{def}}$. In Fig. 6.9 (b), the according contour lines are drawn. The straight lines correspond to the default BGF fraction and to the predicted instanton contribution. The cross indicates the position of the absolute χ^2 minimum. One can see that the χ^2 minimum is found at the default BGF fraction value. The limit on the instanton fraction is stable even for very high BGF fractions.

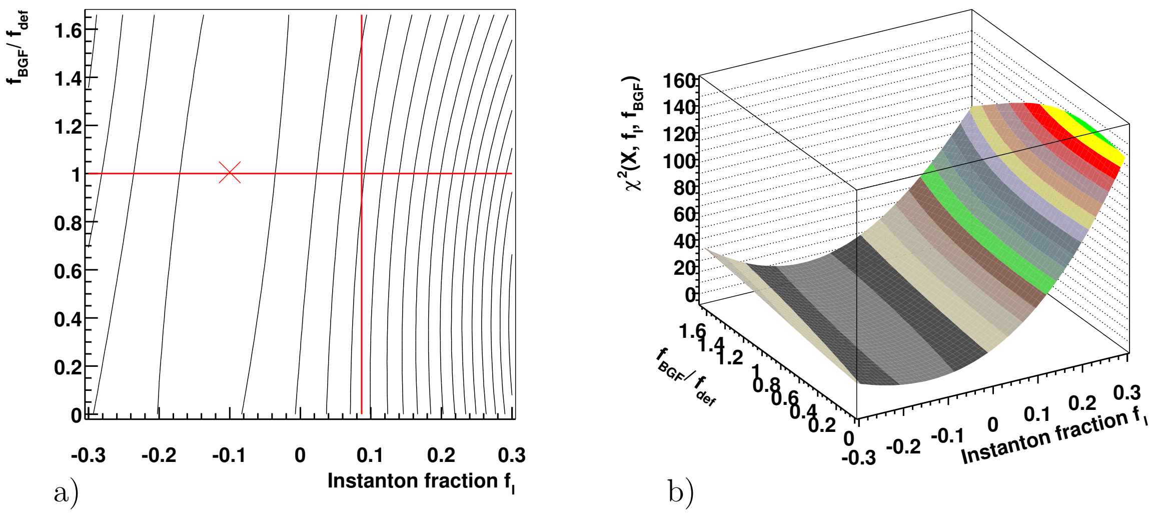


Figure 6.9: *Contour lines and χ^2 distribution obtained from sphericity fit with instanton fraction f_I and the normalised fraction of BGF events in the Djangoh MC, $f_{\mathrm{BGF}}/f_{\mathrm{def}}$, as free parameters. One-dimensional instanton enhancing cuts as described in section 5.2.1 have been applied to obtain the fitted distribution.*

One can therefore conclude that the sphericity fit shows hardly any sensitivity to the BGF fraction.

6.5 Discussion

In the search for instanton-induced events, two problems are encountered. The first is that the predicted instanton fraction contributes only $\approx 0.5\%$ to the inclusive sample. If that instanton-contribution were located in a region of phase space with no or little background, that would not matter, and the available luminosity would be sufficient to obtain a conclusive result. This line of argument requires the background to be known precisely. That this is not the case, is the second, more severe problem. The results presented in chapter 5 show that the two normal DIS Monte Carlos differ strongly in the instanton-enhanced region of phase space. This can be seen from Fig. 5.12 and 5.15 (d), in which the separation power of the different enhancement methods is plotted for both nDIS Monte Carlos. Neither of the Monte Carlos describes the data correctly, as can be seen from the distribution of the Fisher discriminant t, depicted in Fig. 5.14. The Djangoh with Rapgap t distribution is wider than that of the data. As a consequence, the nDIS MC distribution overshoots the data in the region, where the QCDINS MC has its maximum. For Herwig, the overall description of the Fisher discriminant seems to be better, but still differs from the data distribution, where no instanton contribution is expected. The correctness of the HERWIG description at the QCDINS peak can therefore also be doubted. These caveats have to be kept in mind when considering the fit approach results.

The circularity distribution of the inclusive sample is not well described, see Fig. 5.4 (b). The circularity fit is not as sensitive to the instanton contribution as the sphericity fit, while the combined fit of C vs. S fails the χ^2 test.
The sphericity distribution is well described within the inclusive sample. Also for the three instanton-enhanced distributions considered in section 6.2, the description by a combination of nDIS and QCDINS is good according to the χ^2 test.
The three sets of instanton-enhancing cuts differ in the predicted fraction, $f_{\mathrm{I,th}}$, of instanton-induced events, see table 6.1. One would expect a distribution with a larger predicted fraction to be more sensitive to a possible signal. However it turns out that the combination of one-dimensional cuts, which has the lowest predicted instanton-fraction, yields the lowest upper limit, compared to the prediction - it seems to be the most sensitive one. Including systematic effects, a 2σ upper limit of 12% (8.1%) is found for Herwig (Djangoh with Rapgap) as nDIS MC, compared to a predicted fraction of 8.8%. For all three sets of instanton-enhancing cuts, the fit minimum from the Djangoh with Rapgap MC is below that obtained from Herwig with the difference being comparable (up to factors 0.5) to the size of the predicted QCDINS contribution. That the distribution obtained from one-dimensional cuts seems to be most sensitive might be due to the fact that for weaker cuts the difference in shape of the S distribution between nDIS and QCDINS is larger than it is for the stronger cuts, as can

	r_I [%]	DATA	QCDINS	DJA, RGP	HERWIG
1D	9.1	254 ± 16	55.1 ± 0.9	344.5 ± 9.6	186.8 ± 8.0
2D sl: "CCCB"	10.7	223 ± 15	64.2 ± 1.0	336.3 ± 9.4	235.6 ± 8.9
2D sl: "CCCC"	9.1	163 ± 13	55.1 ± 0.9	264.3 ± 8.4	165.5 ± 7.5
Fisher: $t > 9.5$	10.3	180 ± 13	61.9 ± 0.9	264.5 ± 8.4	230.1 ± 8.8
Fisher: $t > 10.0$	5.6	76 ± 8.7	33.8 ± 0.7	105.9 ± 5.3	90.0 ± 5.5
Fisher: $t > 10.5$	2.8	33 ± 5.7	16.8 ± 0.5	35.2 ± 3.1	35.5 ± 3.5

Table 6.2: *Numbers of events within instanton enhancing cuts chosen such that a fraction $r_I \approx 10\%$ of the QCDINS sample within offline cuts is kept. For the Fisher discriminant, harder cuts are also tried. The nDIS MC is normalised such that the number of events in the inclusive sample agrees with the data, cf. section 4.2. Statistical errors are given. "DJA, RGP" stands for the combination of Djangoh and Rapgap nDIS Monte Carlos. "1D" corresponds to the combination of one-dimensional cuts, "2D sl" to the combinations of straight line cuts stated, and "Fisher" to the cut on the Fisher discriminant including the ϵ' variable.*

be seen from Fig. 6.2 to 6.4. This is not the only reason, however. The fits for the different sets of cuts also differ in the position of the fit minimum: while the minimum is obtained at negative, hence unphysical values of the instanton fraction for the one-dimensional cuts, favouring an instanton contribution of zero, the minimum is found at positive values of around the size of the prediction for the stronger cuts. The fits at larger predicted fractions seem to suggest that there is indeed a contribution from instanton-induced events in this region of phase space.
However, the differences between the two nDIS MC models are of about the same size as the predicted effect. All fits are also consistent with an instanton fraction of zero for at least one of the nDIS MC models. It is therefore neither possible to prove that instanton induced-events contribute nor to exclude such a contribution.

6.6 The Cut Approach

As has been explained in the previous section, there are indications that the normal DIS Monte Carlos are not reliable in the extreme region of phase space, in which the instanton-induced events are located. In order to be independent of the nDIS prediction, an extremely conservative limit is set by assuming that the nDIS background is zero. The idea is that within a given set of cuts, there cannot be more instanton-induced events than there are events observed in the data. As these limits are determined only by the cuts, that method will be referred to as cut approach.

While only moderate cuts could be used for the fit approach, in the cut approach

the only requirement is that the number of events predicted by QCDINS should be sufficiently above zero. To be able to compare with H1, cuts have been chosen such that about 10% of the inclusive QCDINS sample are kept. The results obtained from the different enhancement methods described in chapter 5, with the normalisation given in section 4.2, are listed in table 6.2. At an instanton efficiency of 10%, the difference between the two nDIS Monte Carlos is about three times the size of the QCDINS prediction for the combination of one-dimensional cuts and about the size of the prediction for the combinations of two-dimensional cuts. For cuts on the Fisher discriminant t, the discrepancy is smaller, but still about half the instanton prediction at 10% QCDINS efficiency.
For a very extreme cut on the Fisher discriminant, $t > 10.5$, the nDIS Monte Carlos are in agreement with each other and with the data. For all other sets of cuts, the Djangoh with Rapgap prediction is above the data by several standard deviations. HERWIG is below the data for the combination of one-dimensional cuts by approximately the predicted number of QCDINS events. For the combination of two-dimensional cuts, the number of HERWIG events is in agreement with the data, while HERWIG is above the data for the first two cuts on the Fisher discriminant listed in table 6.2.

When assuming the background to be zero, the cut on the Fisher discriminant at $\approx 10\%$ QCDINS efficiency yields an extremely conservative 2σ limit of $180+2\cdot 13 = 206$ events. This limit is a factor 3.3 above the QCDINS prediction of 62 events.

The factor of 3.3 is conservative, not only because the background has been assumed to be zero, but also because the QCDINS prediction underestimates the actual instanton-induced contribution that might be contained in the data: The QCDINS prediction of 62 events is obtained with cuts $x' > 0.35$ and $Q'^2 > 113\,\mathrm{GeV}^2$ at event generation. Since these variables cannot be reconstructed well, the data might include some instanton-induced events with real x', Q'^2 values below these cut values. Removing these events from the data sample would lower the factor of 3.3.

6.7 Comparison with H1

Results from a search for instanton-induced events have recently been published by the H1 collaboration [Ad02, Mi00]. In an instanton enhanced region of phase space, H1 find an excess in the data, which is of about the same size as the difference between the nDIS Monte Carlos: 549 events are selected in the data, while the nDIS Monte Carlos predict 363^{+22}_{-26} and 435^{+36}_{-22} events, respectively. Based on these results, H1 exclude an instanton-induced cross section of 109 pb assuming the "matrix element with parton shower" (MEPS) Monte Carlo model to provide the correct description of the nDIS background, and 47 pb, if the "color dipole model" (CDM) is used instead, compared to a QCDINS predicted cross section of 43 pb.
The shape of the difference between data and nDIS resembles the shape of the QCDINS distributions for four of the six characteristic observables investigated by H1, while this is not the case for the other two distributions.

In order to be independent from the nDIS Monte Carlo, H1 set a most conservative 95% confidence level upper limit of 221 pb assuming the nDIS background to be zero.

Both the H1 and the ZEUS analyses find large discrepancies between the nDIS Monte Carlo models, which limit the significance of the differences between the nDIS Monte Carlo and the data. The fact that the discrepancies between nDIS MC and the data generally have opposite sign in the H1 and the ZEUS analysis might suggest that they are due to a poor data description rather than an instanton contribution.

Summary

A search for QCD-instanton induced processes in deep-inelastic *ep* scattering has been presented. The analysis was performed using a neutral current sample of 1996/1997 data corresponding to an integrated luminosity of $38.6\,\mathrm{pb}^{-1}$. The kinematic range of the search was defined by the photon virtuality, $Q^2 > 120\,\mathrm{GeV}^2$ and the Bjorken scaling variable $x > 10^{-3}$.
Instanton-induced events were modelled by the Monte Carlo generator QCDINS 2.0, interfaced to Herwig for the description of fragmentation and hadronisation. However, instanton-induced events are predicted to contribute only $\approx 0.5\%$ to the inclusive sample. Background events from "normal DIS" were simulated using the colour dipole model as implemented in Ariadne, interfaced to Djangoh. A hadronisation parameter was tuned in order to obtain a good description of sphericity of the instanton region in the inclusive sample. Besides, a contribution of 12% of Rapgap MC was added to simulate diffractive events.

The search strategy was to increase the sensitivity towards instanton-induced events by restricting the inclusive sample to an instanton enhanced subsample. In order to do so, cuts were applied on quantities characterising the hadronic final state of the event. The following variables were found to be sensitive and have therefore been used:

- the transverse momentum of the current jet, p_T^{Jet};
- the virtuality Q'^2 of the quark participating in the hard subprocess;
- the multiplicity of energy flow objects in the instanton region, N_{efo};
- the multiplicity of tracks assigned to the instanton, N_{eft};
- the sphericity S of the instanton region of the hadronic final state;
- its circularity C;
- a variable depending on correlations between four-momenta of one event, ϵ'.

Four different methods to enhance the instanton fraction have been used, namely enhancement via

- a combination of one-dimensional cuts;
- a combination of two-dimensional cuts, defined by the binning of two-dimensional histograms;
- a combination of two-dimensional cuts defined by straight lines;

- a cut on a linear combination of characteristic input variables, with coefficients obtained from the Fisher algorithm.

A comparison of the separation power, P_s, of the different methods, plotted as a function of QCDINS efficiency, r_I, showed that the cut on the Fisher discriminant performed best.

For the enhanced samples, the sphericity of the instanton region of the hadronic final state was fitted by a combination of nDIS and QCDINS MC with the fraction of QCDINS as a free parameter. The fit was not sensitive enough to establish the predicted instanton contribution or to exclude it, because it strongly depended on the exact description of the background. Therefore upper limits on the contribution of instanton-induced events were set. Including systematic effects, the 2σ limit on the instanton fraction in the enhanced sample obtained from one-dimensional cuts is 12% (8.1%) for Herwig (Djangoh with Rapgap) as normal DIS Monte Carlo, compared to a predicted fraction of 8.8%.
From a comparison of the fits within different enhanced samples there were indications that the normal DIS contribution is not reliably described in the extreme region of phase space probed in the search. Therefore, most conservative upper limits on the number of instanton-induced events in an enhanced sample were set assuming the nDIS background to be zero. For a sample with a QCDINS efficiency of $\approx 10\%$, obtained from a cut on the Fisher discriminant, a 2σ limit of 206 events has been set, where 62 events are predicted by QCDINS.

Both the recently published H1 analysis [Ad02] and the ZEUS analysis presented in this thesis show, that the main caveat in searches for QCD-instanton induced events at HERA is the poor description of the data by the normal DIS Monte Carlo around the maximum of the QCDINS distribution in the multi-dimensional phase space. Future progress of searches for instantons therefore requires either the nDIS description to be improved or the sensitivity towards the quality of the nDIS Monte Carlo simulation to be reduced.

Appendix

variables	n(bins)	r_I	r_{DjaRgp}	P_s	$< P_s >$	$\Delta r_N / r_{\mathrm{Hwg}}$
η_{Inst}, M_{Inst}	8×8	—	—	—	2.34	—
	10×10	0.58	0.28	2.34		0.12
	12×12	—	—	—		—
$\log_{10} p_T^{\mathrm{Jet}}$, η_{Inst}	8×8	0.72	0.33	2.27	2.57	0.03
	10×10	0.52	0.20	2.64		0.03
	12×12	0.49	0.18	2.79		0.06
γ_{h}, M_{Inst}	8×8	0.48	0.18	2.80	2.80	0.05
	10×10	0.49	0.18	2.80		0.05
	12×12	0.49	0.18	2.79		0.05
ϵ', $\log_{10} p_T^{\mathrm{Jet}}$	8×8	—	—	—	2.81	—
	10×10	0.51	0.19	2.81		0.03
	12×12	—	—	—		—
$\log_{10} p_T^{\mathrm{Jet}}$, γ_{h}	8×8	0.61	0.22	2.93	3.01	0.07
	10×10	0.62	0.22	3.04		0.07
	12×12	0.62	0.22	3.05		0.07
ϵ', γ_{h}	8×8	0.46	0.16	3.10	3.10	0.04
	10×10	—	—	—		—
	12×12	—	—	—		—

Table 1: *Quantities characterising the distributions, which are considered as candidates for the definition of two-dimensional cuts and which pass the criteria of section 5.2.2. Efficiencies r_I of QCDINS and r_{DjaRgp} of the Djangoh with Rapgap normal DIS MC, the separation power P_s for distributions with 8×8, 10×10 and 12×12 bins and the mean separation power $< P_s >$, averaged over the three bin sizes are given, as well as relative differences $\Delta r_N / r_{\mathrm{Hwg}} = (r_{\mathrm{DjaRgp}} - r_{\mathrm{Hwg}})/r_{\mathrm{Hwg}}$ between the two normal DIS MCs, see section 5.2.2.*

variables	n(bins)	r_I	r_{DjaRgp}	P_s	$< P_s >$	$\Delta r_N/r_{\text{Hwg}}$
C, M_{Inst}	8×8	0.45	0.15	3.12	3.10	0.04
	10×10	—	—	—		—
	12×12	0.46	0.15	3.09		0.04
S, $\log_{10} p_T^{\text{Jet}}$ $*$	8×8	0.58	0.17	3.65	3.62	0.04
	10×10	0.59	0.17	3.66		0.05
	12×12	0.62	0.18	3.56		0.05
C, $\log_{10} p_T^{\text{Jet}}$ $*$	8×8	0.61	0.17	3.64	3.69	0.03
	10×10	0.62	0.17	3.68		0.03
	12×12	0.63	0.17	3.74		0.03
C, S	8×8	—	—	—	3.72	—
	10×10	—	—	—		—
	12×12	0.45	0.12	3.72		0.04
$\log_{10} p_T^{\text{Jet}}$, M_{Inst}	8×8	0.71	0.22	3.70	3.90	0.14
	10×10	0.63	0.17	4.15		0.14
	12×12	0.70	0.21	3.84		0.14
ϵ', N_{eft}	8×8	—	—	—	4.01	—
	10×10	—	—	—		—
	12×12	0.51	0.14	4.01		0.14
ϵ', M_{Inst}	8×8	0.47	0.11	4.47	4.07	0.09
	10×10	0.63	0.19	3.66		0.08
	12×12	0.57	0.15	4.07		0.09
N_{efo}, M_{Inst}	8×8	—	—	—	4.12	—
	10×10	—	—	—		—
	12×12	0.54	0.14	4.12		0.10
C, N_{efo}	8×8	—	—	—	4.12	—
	10×10	—	—	—		—
	12×12	0.54	0.14	4.12		0.08

Table 2: *Quantities characterising the distributions, which are considered as candidates for the definition of two-dimensional cuts and which pass the criteria of section 5.2.2. Efficiencies r_I of QCDINS and r_{DjaRgp} of the Djangoh with Rapgap normal DIS MC, the separation power P_s for distributions with 8×8, 10×10 and 12×12 bins and the mean separation power $< P_s >$, averaged over the three bin sizes are given, as well as relative differences $\Delta r_N/r_{\text{Hwg}} = (r_{\text{DjaRgp}} - r_{\text{Hwg}})/r_{\text{Hwg}}$ between the two normal DIS MCs, see section 5.2.2.*

variables	n(bins)	r_I	r_{DjaRgp}	P_s	$< P_s >$	$\Delta r_N / r_{\mathrm{Hwg}}$
S, γ_{h}	8×8	0.47	0.11	4.33	4.28	0.06
	10×10	0.49	0.12	4.22		0.05
	12×12	0.49	0.12	4.30		0.05
$\log_{10} p_T^{\mathrm{Jet}}$, N_{efo}	8×8	—	—	—	4.46	—
	10×10	—	—	—		—
	12×12	0.69	0.18	4.46		0.15
γ_{h}, N_{efo}	8×8	—	—	—	4.49	—
	10×10	—	—	—		—
	12×12	0.49	0.12	4.49		0.09
S, N_{eft}	8×8	0.45	0.11	4.70	4.57	0.15
	10×10	—	—	—		—
	12×12	0.46	0.12	4.44		0.14
N_{efo}, η_{Inst}	8×8	0.55	0.13	4.80	4.80	0.12
	10×10	0.54	0.12	5.08		0.13
	12×12	0.62	0.16	4.52		0.15
ϵ', N_{efo}	8×8	0.67	0.18	4.19	4.86	0.10
	10×10	0.55	0.11	5.59		0.11
	12×12	0.61	0.14	4.79		0.12
N_{eft}, N_{efo} $*$	8×8	0.46	0.10	4.95	4.86	0.12
	10×10	0.46	0.10	5.00		0.11
	12×12	0.48	0.11	4.64		0.10
S, M_{Inst}	8×8	0.64	0.15	4.77	5.20	0.14
	10×10	0.60	0.13	5.34		0.15
	12×12	0.60	0.12	5.49		0.15
S, N_{efo} $*$	8×8	0.51	0.10	5.80	5.54	0.14
	10×10	0.53	0.11	5.48		0.14
	12×12	0.54	0.11	5.33		0.13

Table 3: *Quantities characterising the distributions, which are considered as candidates for the definition of two-dimensional cuts and which pass the criteria of section 5.2.2. Efficiencies r_I of QCDINS and r_{DjaRgp} of the Djangoh with Rapgap normal DIS MC, the separation power P_s for distributions with 8×8, 10×10 and 12×12 bins and the mean separation power $< P_s >$, averaged over the three bin sizes are given, as well as relative differences $\Delta r_N / r_{\mathrm{Hwg}} = (r_{\mathrm{DjaRgp}} - r_{\mathrm{Hwg}})/r_{\mathrm{Hwg}}$ between the two normal DIS MCs, see section 5.2.2.*

				2σ upper limits on f_I			
	$f_{\mathrm{I,th}}$	fit minimum		χ^2 approach		max. likelihood	
		Hwg	DjaRgp	Hwg	DjaRgp	Hwg	DjaRgp
standard	0.088	-0.002	-0.050	0.108	0.069	0.082	0.045
$y_{\mathrm{JB}} > 0.075$	0.089	-0.003	-0.051	0.105	0.066	0.082	0.044
$\lvert z_{\mathrm{Vtx}}\rvert < 40\,\mathrm{cm}$	0.088	0.002	-0.045	0.111	0.075	0.088	0.048
$E_{\mathrm{e,DA}} > 15\,\mathrm{GeV}$	0.080	-0.020	-0.038	0.139	0.106	0.087	0.066
$E_{\mathrm{cne}} > 3\,\mathrm{GeV}$	0.090	-0.008	-0.057	0.098	0.056	0.079	0.044
$y_{\mathrm{el}} < 0.80$	0.090	0.006	-0.046	0.112	0.074	0.092	0.048
$45\,\mathrm{GeV} < E - P_z$	0.089	-0.007	-0.056	0.106	0.061	0.078	0.042
DCA $< 8\,\mathrm{cm}$	0.091	0.011	-0.034	0.117	0.082	0.098	0.056
$p_T^{\mathrm{Jet}} > 5\,\mathrm{GeV}$	0.102	0.017	-0.021	0.113	0.065	0.121	0.081
$Q'^2_{\mathrm{DA}} < 230\,\mathrm{GeV}^2$	0.097	-0.011	-0.054	0.130	0.072	0.082	0.050
$Q'^2_{\mathrm{DA}} < 270\,\mathrm{GeV}^2$	0.085	-0.023	-0.046	0.080	0.061	0.063	0.043
$N_{\mathrm{eft}} > 6$	0.074	-0.013	-0.066	0.095	0.049	0.061	0.030
$N_{\mathrm{eft}} > 8$	0.104	0.023	-0.051	0.152	0.089	0.121	0.058
use Q^2_{el} instead of Q^2_{DA}	0.097	-0.028	-0.054	0.083	0.053	0.069	0.049
use x_{el} instead of x_{DA}	0.090	-0.001	-0.047	0.111	0.073	0.085	0.047
use Q'^2_{γ} instead of Q'^2_{DA}	0.089	0.008	-0.048	0.119	0.072	0.094	0.048
CAL energy scale 1.03	0.070	0.003	-0.045	0.110	0.070	0.087	0.047
CAL energy scale 0.97	0.071	-0.001	-0.049	0.097	0.059	0.085	0.047
6% Rapgap	0.088	—	-0.055	—	0.063	—	0.043
18% Rapgap	0.088	—	-0.038	—	0.077	—	0.050
maximum	—	—	—	0.152	0.106	0.121	0.081

Table 4: *Systematic variation of cuts and checks of method for the fit of sphericity S of the instanton part of the event as described in section 6.3. The S distribution is obtained by applying the one-dimensional instanton-enhancing cuts stated in section 5.2.1. Fit minima and limits on the instanton fraction f_I are given and compared to the predicted fraction $f_{\mathrm{I,th}}$. "Hwg" stands for "Herwig", "DjaRgp" for Djangoh with Rapgap and "max. likelihood" for the maximum likelihood approach.*

				2σ upper limits on f_I			
	$f_{\mathrm{I,th}}$	fit minimum		χ^2 approach		max. likelihood	
		Hwg	DjaRgp	Hwg	DjaRgp	Hwg	DjaRgp
standard	0.088	-0.145	0.036	0.127	0.237	0.083	0.174
$y_{\mathrm{JB}} > 0.075$	0.089	-0.147	0.037	0.125	0.235	0.082	0.174
$\lvert z_{\mathrm{Vtx}} \rvert < 40\,\mathrm{cm}$	0.088	-0.139	0.034	0.130	0.241	0.086	0.173
$E_{\mathrm{e,DA}} > 15\,\mathrm{GeV}$	0.080	-0.200	-0.015	0.114	0.233	0.086	0.146
$E_{\mathrm{cne}} > 3\,\mathrm{GeV}$	0.090	-0.162	0.020	0.121	0.226	0.081	0.162
$y_{\mathrm{el}} < 0.80$	0.090	-0.158	0.017	0.112	0.219	0.079	0.156
$45\,\mathrm{GeV} < E - P_z$	0.089	-0.146	0.034	0.102	0.214	0.083	0.172
DCA $< 8\,\mathrm{cm}$	0.091	-0.165	0.022	0.108	0.225	0.079	0.161
$p_T^{\mathrm{Jet}} > 5\,\mathrm{GeV}$	0.102	-0.115	0.083	0.183	0.296	0.107	0.228
$Q'^2_{\mathrm{DA}} < 230\,\mathrm{GeV}^2$	0.097	-0.194	-0.007	0.119	0.205	0.081	0.148
$Q'^2_{\mathrm{DA}} < 270\,\mathrm{GeV}^2$	0.085	-0.134	0.034	0.126	0.243	0.079	0.164
$N_{\mathrm{eft}} > 6$	0.074	-0.066	0.081	0.117	0.235	0.094	0.197
$N_{\mathrm{eft}} > 8$	0.104	-0.250	-0.083	0.075	0.139	0.077	0.106
use Q^2_{el} instead of Q^2_{DA}	0.097	-0.251	-0.055	0.011	0.086	0.066	0.108
use x_{el} instead of x_{DA}	0.090	-0.164	0.022	0.104	0.222	0.079	0.161
use Q'^2_{γ} instead of Q'^2_{DA}	0.089	-0.170	0.012	0.119	0.222	0.077	0.154
CAL energy scale 1.03	0.070	-0.159	0.025	0.109	0.200	0.076	0.162
CAL energy scale 0.97	0.071	-0.184	0.002	0.106	0.213	0.072	0.142
6% Rapgap	0.088	—	0.021	—	0.222	—	0.162
18% Rapgap	0.088	—	0.062	—	0.262	—	0.195
maximum	—	—	—	0.183	0.296	0.107	0.228

Table 5: *Systematic variation of cuts and checks of method for the fit of circularity C of the instanton part of the event as described in section 6.3. The C distribution is obtained by applying the one-dimensional instanton-enhancing cuts stated in section 5.2.1. Fit minima and limits on the instanton fraction f_I are given and compared to the predicted fraction $f_{\mathrm{I,th}}$. “Hwg” stands for “Herwig”, “DjaRgp” for Djangoh with Rapgap and “max. likelihood” for the maximum likelihood approach.*

				2σ upper limits on f_I			
	$f_{\mathrm{I,th}}$	fit minimum		χ^2 approach		max. likelihood	
		Hwg	DjaRgp	Hwg	DjaRgp	Hwg	DjaRgp
standard	0.088	-0.024	-0.045	0.106	0.055	0.051	0.034
$y_{\mathrm{JB}} > 0.075$	0.089	-0.023	-0.045	0.100	0.050	0.052	0.033
$\lvert z_{\mathrm{Vtx}} \rvert < 40\,\mathrm{cm}$	0.088	-0.024	-0.048	0.101	0.059	0.052	0.033
$E_{\mathrm{e,DA}} > 15\,\mathrm{GeV}$	0.080	-0.043	-0.078	0.140	0.087	0.049	0.027
$E_{\mathrm{cne}} > 3\,\mathrm{GeV}$	0.090	-0.028	-0.051	0.090	0.039	0.050	0.033
$y_{\mathrm{el}} < 0.80$	0.090	-0.017	-0.052	0.103	0.014	0.057	0.031
$45\,\mathrm{GeV} < E - P_z$	0.089	-0.030	-0.056	0.106	0.027	0.046	0.029
$\mathrm{DCA} < 8\,\mathrm{cm}$	0.091	-0.016	-0.036	0.107	0.069	0.058	0.039
$p_T^{\mathrm{Jet}} > 5\,\mathrm{GeV}$	0.102	-0.032	-0.019	0.166	0.120	0.059	0.060
$Q'^2_{\mathrm{DA}} < 230\,\mathrm{GeV}^2$	0.097	-0.037	-0.047	0.065	—	0.050	0.038
$Q'^2_{\mathrm{DA}} < 270\,\mathrm{GeV}^2$	0.085	-0.029	-0.037	0.127	0.075	0.045	0.035
$N_{\mathrm{eft}} > 6$	0.074	-0.017	-0.033	0.141	0.082	0.045	0.032
$N_{\mathrm{eft}} > 8$	0.104	-0.013	-0.086	—	—	0.073	0.029
use Q^2_{el} instead of Q^2_{DA}	0.097	-0.093	-0.102	0.013	0.063	0.022	0.019
use x_{el} instead of x_{DA}	0.090	-0.024	-0.044	0.112	0.066	0.051	0.035
use Q'^2_{γ} instead of Q'^2_{DA}	0.089	-0.013	-0.043	0.135	0.076	0.061	0.036
CAL energy scale 1.03	0.070	-0.032	-0.051	0.112	0.067	0.044	0.029
CAL energy scale 0.97	0.071	-0.028	-0.045	0.124	0.100	0.049	0.034
6% Rapgap	0.088	—	-0.053	—	0.050	—	0.031
18% Rapgap	0.088	—	-0.032	—	0.070	—	0.039
maximum	—	—	—	0.166	0.120	0.073	0.060

Table 6: *Systematic variation of cuts and checks of method for the fit of circularity of the instanton part of the event C vs. its sphericity S as described in section 6.3. The C vs. S distribution is obtained by applying the one-dimensional instanton-enhancing cuts stated in section 5.2.1. Fit minima and limits on the instanton fraction f_I are given and compared to the predicted fraction $f_{\mathrm{I,th}}$. "Hwg" stands for "Herwig", "DjaRgp" for Djangoh with Rapgap and "max. likelihood" for the maximum likelihood approach.*

	χ^2/ndf for S		χ^2/ndf for C		χ^2/ndf for C vs. S	
	Hwg	DjaRgp	Hwg	DjaRgp	Hwg	DjaRgp
standard	4.2/5	3.9/5	5.6/6	3.9/6	50.6/47	62.2/52
$y_{\mathrm{JB}} > 0.075$	4.5/5	4.2/5	5.6/6	4.1/6	51.9/47	63.0/52
$\lvert z_{\mathrm{Vtx}}\rvert < 40\,\mathrm{cm}$	4.5/5	3.7/5	5.6/6	3.5/6	52.1/47	61.2/52
$E_{\mathrm{e,DA}} > 15\,\mathrm{GeV}$	1.7/5	2.7/5	6.8/6	2.8/6	40.1/44	51.3/51
$E_{\mathrm{cne}} > 3\,\mathrm{GeV}$	5.0/5	5.6/5	5.8/6	3.9/6	52.6/46	64.7/52
$y_{\mathrm{el}} < 0.80$	4.8/5	3.7/5	6.3/6	3.9/6	52.5/47	68.2/52
$45\,\mathrm{GeV} < E - P_z$	3.9/5	4.6/5	7.3/6	5.6/6	49.1/47	66.1/52
DCA $< 8\,\mathrm{cm}$	5.0/5	3.8/5	6.6/6	3.9/6	52.3/47	60.7/52
$p_T^{\mathrm{Jet}} > 5\,\mathrm{GeV}$	7.6/5	8.3/5	3.8/6	4.1/6	36.9/43	52.8/49
$Q'^2_{\mathrm{DA}} < 230\,\mathrm{GeV}^2$	1.6/5	4.5/5	6.0/6	4.8/6	55.8/45	74.4/51
$Q'^2_{\mathrm{DA}} < 270\,\mathrm{GeV}^2$	5.2/5	4.6/5	5.2/6	2.1/6	43.8/48	57.8/52
$N_{\mathrm{eft}} > 6$	2.1/5	3.8/5	7.0/6	5.4/6	36.6/49	54.2/52
$N_{\mathrm{eft}} > 8$	3.9/5	3.7/5	8.7/6	6.6/6	70.2/44	76.2/51
use Q^2_{el} instead of Q^2_{DA}	5.6/5	6.6/5	13.1/6	9.7/6	60.8/46	78.0/52
use x_{el} instead of x_{DA}	4.0/5	3.7/5	6.9/6	4.1/6	49.3/47	60.1/52
use Q'^2_{γ} instead of Q'^2_{DA}	4.3/5	4.0/5	5.7/6	3.6/6	464/47	58.6/52
CAL energy scale 1.03	4.4/5	4.0/5	6.2/6	5.9/6	46.5/47	57.6/52
CAL energy scale 0.97	5.8/5	5.6/5	6.0/6	3.3/6	45.7/46	51.8/52
6% Rapgap	—	4.5/5	—	4.2/6	—	62.4/52
18% Rapgap	—	3.6/6	—	3.3/6	—	60.4/52

Table 7: *The* χ^2 *value compared to the number of degrees of freedom,* ndf*, is a measure of the quality of a fit. The values in the table correspond to the fit minimum, whenever* f_I *at the minimum is above* 0*. For fits with* $f_I(\chi_{\min}) < 0$*, i.e. at unphysical values,* $\chi^2(f_I = 0)$ *is given instead. For the positions* f_I *of the fit minima, cf. tables 4 — 6. "Hwg" stands for "Herwig", "DjaRgp" for Djangoh with Rapgap.*

Bibliography

[Ab95] H. Abramowicz *et al.*, Nucl. Instr. and Meth. **A 365** (1995) 508;
R. Sinkus and T. Voss, Nucl. Instr. and Meth. **A 391** (1997) 360

[Ad02] H1 Collab., C. Adloff *et al.*, *Search for QCD Instanton Induced Processes in Deep-Inelastic Scattering at HERA*, Preprint DESY-02-062 (hep-ex/0205078), accepted by Eur. Phys. J. **C**

[Af02] CDF Collab., T. Affolder *et al.*, Phys. Rev. Lett. **88** (2002) 042001

[Am79] D. Amati and G. Veneziano, Phys. Lett. **B 83** (1979) 87;
A. Bassetto *et al.*, Phys. Lett. **B 83** (1979) 207;
G. Marchesini *et al.*, Nucl. Phys. **B 181** (1980) 335

[AlS] H1 Collab., T. Ahmed *et al.*, Phys. Lett. **B 346** (1995) 415;
H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 5** (1998) 625;
H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 6** (1999) 575;
H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 19** (2001) 289;
ZEUS Collab., M. Derrick *et al.*, Phys. Lett. **B 363** (1995) 201;
ZEUS Collab., J. Breitweg *et al.*, Phys. Lett. **B 507** (2001) 70;
ZEUS Collab., S. Chekanov *et al.*, submitted to Phys. Rev. **D**, Preprint DESY-02-105, (hep-ex/0208023);
ZEUS Collab., S. Chekanov *et al.*, submitted to Phys. Lett. **B**, Preprint DESY-02-112, (hep-ex/0208037)

[And92] J. Andruskow *et al.*, DESY 92-066 (1992)

[Art] X. Artru and G. Mennessier, Nucl. Phys. **B 70** (1974) 93;
M. G. Bowler, Z. Phys. **C 11** (1981) 169;
B. Andersson *et al.*, Z. Phys. **C 20** (1983) 317, Nucl. Phys. **B 264** (1986) 29;
B. Andersson *et al.*, Phys. Rep. **97** (1983) 31

[Bal93] I. Balitsky and V. Braun, Phys. Lett. **B 314** (1993) 237

[Bar02] A. Barakbaev, private communication (2002)

[Beh93] U. Behrens *et al.*, Nucl. Instr. and Meth. **A 332** (1993) 253

[Ben92] S. Bentvelsen *et al.*, *Proc. Workshop on Physics at HERA*, Vol. 1, W. Buchmüller and G. Ingleman (eds.), DESY (1992), p. 23;
K. C. Hoeger, *ibid*, p. 43

[Bel75] A. Belavin *et al.*, Phys. Lett. **B 59** (1975) 85

[Ber79] C. Bernard, Phys. Rev. **D 19** (1979) 3013

[Beth00] S. Bethke, J. Phys. **G 26** (2000) R27

[BlLo98] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*, Teubner (1998)

[Bot93] H. Boterenbrood *et al.*, Nucl. Instr. and Meth. **A 332** (1993) 263

[Br87] R. Brun *et al.*, *GEANT3*, Technical Report CERN-DD/EE 84-1, CERN (1987)

[Br98] G. Briskin, *Diffractive Dissociation in ep Deep Inelastic Scattering*, PhD Thesis, Tel Aviv University (1998)

[Ca93] S. Catani *et al.*, Nucl. Phys. **B 406** (1993) 187

[CAL91] A. Andresen *et al.*, Nucl. Instr. and Meth. **A 309** (1991) 101;
M. Derrick *et al.*, Nucl. Instr. and Meth. **A 309** (1991) 77;
A. Bernstein *et al.*, Nucl. Instr. and Meth. **A 336** (1993) 23

[Car99] T. Carli *et al.*, Proc. Monte Carlo Generators for HERA Physics DESY-PROC-1999-02, p.329 (hep-ph/9906441)

[Cha94] K. Charchuła *et al.*, Comput. Phys. Commun. **81** (1994) 381

[Chu94] M. C. Chu *et al.*, Phys. Rev. **D 49** (1994) 6039

[Col85] J. C. Collins *et al.*, Nucl. Phys. **B 261** (1985) 104

[CTD] N. Harnew *et al.*, Nucl. Instr. and Meth. **A 279** (1989) 290;
B. Foster *et al.*, Nucl. Phys. Proc. Suppl. **B 32** (1993) 181;
B. Foster *et al.*, Nucl. Instr. and Meth. **A 338** (1994) 254

[DGLAP] G. Altarelli and G. Parisi, Nucl. Phys. **B 126** (1977) 298;
Y. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641;
V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438

[Dus02] S. Dusini *et al.*, *ORANGE USER MANUAL* (2002), (http://www-zeus.desy.de/ZEUS_ONLY/analysis/orange/orange.ps)

[Ell93] S. D. Ellis and D. E. Soper, Phys. Rev. **D 48** (1993) 3160

[Ell96] R. K. Ellis *et al.*, *QCD and Collider Physics*, Cambridge University Press (1996)

[FiPa94] S. M. Fisher and P. Palazzi, *ADAMO Entity-Relationship Programming System, Version 3.3, Reference Manual*, CERN PTG, Programming Techniques Group, ECP Division (1994)

[Fish36] R. A. Fisher, Ann. Eug. 7 (1936) 179

[For00] H. Forkel, *A primer on Instantons In QCD*, contribution to the "12th Workshop on Hadronic Interactions", Rio de Janeiro (2000) (hep-ph/0009136)

[Ger98] J. Gerigk, *QCD-Instanton-induzierte Prozesse in tiefunelastischer* $e^{\pm}p$*-Streuung*, diploma thesis, MPI München (1998), MPI-PhE/98-20

[Gib95] M. Gibbs *et al.* Z. Phys. **C 66** (1995) 285

[Glae91] R. Glaeser *et al.*, *ZEUS software update and supplement*, ZEUS-Note 91-070 (1991)

[Gus] G. Gustafson and U. Petterson, Nucl. Phys. **B 306** (1988) 746;
G. Gustafson, Phys. Lett. **B 175** (1986) 453;
B. Andersson *et al.*, Z. Phys. **C 43** (1989) 625

[Haa92] T. Haas, *Generating Monte Carlo events with MOZART*,
ZEUS-note 92-021 (1992)

[Har97] G. F. Hartner *et al.*, *VCTRAK (3.07/04): Offline Output Information*,
ZEUS-Note 97-064 (1997);
G. F. Hartner, *VCTRAK Briefing: Program and Math*,
ZEUS-Note 98-058 (1998)

[Her90] L. Hervás, *The Pipelined Readout for the ZEUS Calorimeter*, Thesis, Universidad Autónoma de Madrid (1990)

[HERA81] *HERA: A Proposal for a Large Electron Proton Colliding Beam Facility at DESY*, DESY-HERA-81/10, Hamburg (1981)

[HERA98] *The HERA Luminosity Upgrade*, DESY-HERA 98-05, Hamburg (1998)

['t Hooft] G. 't Hooft, Phys. Rev. Lett. **37** (1976) 8;
Phys. Rev. **D 14** (1976) 3432;
Phys. Rev. **D 18** (1978) 2199 (Erratum);
Phys. Rep. **142** (1986) 357

[HWt99] R. Hall-Wilton *et al.*, *The CTD Tracking Resolution*,
ZEUS-note 99-024 (1999)

[Jacq79] F. Jacquet and A. Blondel, in *Proc. Study for an ep Facility for Europe*, U. Amaldi (ed.), DESY (1979), p. 391, also in preprint DESY 79/48 (1979)

[Jun95] H. Jung, Comp. Phys. Commun. **86** (1995) 147

[Kho91] V. V. Khoze and A. Ringwald, Phys. Lett. **B 259** (1991) 106

[Kl01] K. H. Klimek, *Cross Section Measurement of Vector Meson Quasi-Photoproduction at High Four-Momentum Transfer Using the ZEUS Detector at the HERA Collider*, PhD Thesis, Inst. of Nucl. Phys., Cracow (2001)

[Kuhl97] M. Kuhlen, *QCD and the Hadronic Final State in Deep Inelastic Scattering at HERA* (hep-ph/9712505),
Habilitationsschrift, Universität Hamburg (1997)

[Kwi92] A. Kwiatkowski *et al.*, Comput. Phys. Commun. **69** (1992) 155

[Lai97] H. L. Lai *et al.*, Phys. Rev. **D 55** (1997) 1280

[Lön92] L. Lönnblad, Comput. Phys. Commun. **71** (1992) 15

[Mar92] G. Marchesini and B. R. Webber, Nucl. Phys. **B 310** (1988) 461;
G. Marchesini *et al.*, Comput. Phys. Commun. **67** (1992) 465

[MaOd93] P. Mazzanti and R. Odorico, Z. Phys. **C 59** (1993) 273

[Mi00] S. Mikocki, talk presented at DIS 2000;
E. de Wolf, talk presented at ICHEP 2000, April 2000;
H1 Coll., contributed paper to ICHEP 2000

[Moch97] S. Moch *et al.*, Nucl. Phys. **B 507** (1997) 134

[PDG98] Particle Data Book, D. Caso *et al.*, Eur. Phys. J. **C 3** (1998) 1

[PDG00] Particle Data Book, D. E. Groom *et al.*, Eur. Phys. J. **C 15** (2000) 1

[Piot94] K. Piotrzkowski and M. Zachara, *Determination of the ZEUS Luminosity in '93*, ZEUS-note 94-167 (1994)

[Rin91] A. Ringwald *et al.*, Nucl. Phys. **B 365** (1991) 3

[Rin94] A. Ringwald and F. Schrempp, *Towards the Phenomenology of QCD-Instanton Induced Particle Production at HERA*, (hep-ph/9411217),
In: *Quarks '94*, Proc. 8th Int. Seminar, Vladimir,Russia, 1994, ed. by D. Grigoriev *et al.* (World Scientific, Singapore 1995), p. 170

[Rin98] A. Ringwald and F. Schrempp, Phys. Lett. **B 438** (1998) 217

[Rin99] A. Ringwald and F. Schrempp, Phys. Lett. **B 459** (1999) 249

[Rin00] A. Ringwald and F. Schrempp, long write-up (hep-ph/9911516),
Comput. Phys. Commun. **132** (2000) 267;
M. Gibbs, A. Ringwald and F. Schrempp, (hep-ph/9506392),
In: Proc. DIS 1995 (Paris), J.-F. Laporte and Y.Sirois, Eds., p. 341

[Rin00b] A. Ringwald and F. Schrempp, Proc. DIS 2000, Eds. J. A. Gracey and T. Greenshaw, Liverpool (2000) 318, (hep-ph/0006215)

[Rin01] A. Ringwald and F. Schrempp, Phys. Rev. Lett. **B 503** (2001) 331

[Rin02] A. Ringwald, *Vacuum structure and high-energy scattering*,
Preprint DESY-02-158, (hep-ph/0210209) and references therein

[Schä98] T. Schäfer and E. Shuryak, Rev. Mod. Phys. **70** (1998) 323
(hep-ph/9610451)

[Schr98] F. Schrempp, *Tracking QCD-Instantons in Deep-Inelastic Scattering*, review talk, Ringberg Conference on Trends in Theoretical Particle Physics, (1998)

[Schr99] F. Schrempp, *Auf der Spur von Instantonen - grundlegenden Tunnelprozessen in der Teilchenphysik*, general colloquium for non-experts, Würzburg (1999)

[Schr02] F. Schrempp, J. Phys. **G 28** (2002) 915 (hep-ph/0109032)

[Siev00] M. Sievers, *Search for QCD-Instanton-Induced Processes in Deep Inelastic Scattering at HERA*, DESY-THESIS-2000-027, Universität Hamburg (2000)

[StrFct] H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 13** (2000) 609;
H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 21** (2001) 33;
H1 Collab., C. Adloff *et al.*, Eur. Phys. J. **C 19** (2001) 269;
ZEUS Collab., S. Chekanov *et al.*, Eur. Phys. J. **C 21** (2001) 3, p. 443;
ZEUS Collab., M. Derrick *et al.*, Z. Phys. **C 72** (1996) 399;
ZEUS Collab., S. Chekanov *et al.*, Preprint DESY-02-113
(hep-ex/0208040), to be published in Eur. Phys. J.;

[Sjö94] T. Sjöstrand, Comput. Phys. Commun. **82** (1994) 74;
T. Sjöstrand, *Pythia 5.7 and Jetset 7.4*, CERN-TH.7112/93,
LU TP 95-20

[Tun01] N. Tuning, *Hadronic final state reconstruction with calorimeter, tracking and backsplash correction*, ZEUS-Note 01-021 (2001)

[Ver91] J. Verbaarschot, Nucl. Phys. **B 362** (1991) 33

[Wai93] L. Wai and R. Yoshida, *Running MOZART-ZGANA-ZEPHYR chain using FUNNEL*, ZEUS-Note 93-065 (1993)

[Web84] B. R. Webber, Nucl. Phys. **B 238** (1984) 492

[YoSm] C. Youngman, *The ZEUS Data Aquisition System*, DESY 92-105 A; W. H. Smith *et al.*, *The ZEUS Trigger System*, ZEUS 89-084

[Zha02] Z. Zhang for the H1 Coll., *Structure Function Results From H1*, contribution to ICHEP 02, Amsterdam (2002)

[ZEUS93] ZEUS-Collaboration, *The ZEUS Detector, Status Report 1993*, Ed. U. Holm (1993)
(http://www-zeus.desy.de/bluebook/bluebook.html)

Acknowledgements

Many have contributed to the work presented in this thesis, giving advice, help and encouragement, and I would like to take the opportunity to thank them.

First and in particular I am grateful to my thesis advisor and *Doktorvater* Prof. Dr. E. Lohrmann. The care he has taken in both setting the scope of the analysis as a whole and in considering all its details have always made me feel to be on safe ground.

Various aspects of the analysis have been discussed in the weekly meetings of the instanton subgroup of the ZEUS collaboration. I thank its members Prof. Dr. P. Bussey, Dr. S. Chekanov, Dr. C. Foudas, Dr. B. Levchenko, Dr. N. Pokrovskiy and Dr. M. Sutton for advice and helpful criticism. A. Barakbaev and Prof. Dr. E. Boos I would like to thank for developing an algorithm to calculate one of the sensitive variables used in the search.

I am thankful to Dr. T. Carli and to Dr. A. Geiser for commenting on the analysis and for discussing results and methods of the instanton search performed by the H1 collaboration.

It is a pleasure to thank Dr. F. Schrempp for explanations regarding the theoretical description of instanton effects and for commenting on parts of the manuscript.

Reports on the progress of the analysis have been given to the ZEUS “QCD/hadronic final states” physics group at intervals of a few months. Its members, notably Prof. Dr. J. Terrón, I thank for comments made and questions asked at these occasions. Dr. G. Hartner I thank for information regarding the ZEUS tracking.

I am grateful to present and former PhD students of Hamburg University for having introduced me into ZEUS software and for helping when computers would not do what they ought to.

Finally, I wish to thank all members of the ZEUS collaboration for their interest in the instanton search.