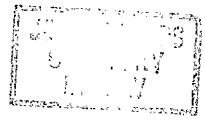
Deep Inelastic Scattering at HERA

A. W. E. Dann

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Department of Physics in the Faculty of Science.

February 1993

Contents



Abstract

The energy spectrum of the scattered electron in inclusive neutral current deep inelastic scattering is presented. The data analysed was taken from the first run period of the H1 detector at the HERA *ep* collider in which 26.7 GeV electrons were collided with 820 GeV protons, corresponding to an integrated luminosity of 1.3 nb^{-1}. A number of methods of identifying the scattered electron within a DIS event and of removing the photoproduction background are discussed in detail. The electron energy spectrum is compared with Monte Carlo predictions using different parametrizations of the proton structure functions. The results favour those parametrizations prediciting a steep rise in the structure functions as the Bjorken scaling variable $x \rightarrow 0$.

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other institute of learning.

Biographical Note

The author was educated at Carres Grammar School in Sleaford, Lincolnshire between 1979 and 1986. In 1989 he obtained a 1st class honours degree in Physics from the University of Manchester. The author joined the High Energy Physics group within the same university, and was awarded a Diploma in Advanced Studies in Science in 1990. The work in this thesis was conducted under the supervision of Dr R. J. Ellison both at Manchester and at DESY, Hamburg.

To my families, and my friends.

This work was supported financially by the Science and Engineering Research Council (S.E.R.C.) between October 1989 and October 1992.

> After a year's research, one realises that it could have been done in a week. *Sir William Henry Bragg, 1862–1942*

Chapter 1

Introduction

The analysis described in this work was performed on data taken with the H1 detector at the HERA experiment at DESY (Deutsches Elektronen SYnchrotron) in Hamburg. HERA (Hadron Electron Ring Accelerator) is an electron–proton collider with design energies of 35 GeV for the electrons and 820 GeV for the protons. The HERA machine is the first electron– proton collider, and was completed in late 1990 after six years design and construction. After extensive testing, collisions were first obtained between 12 GeV electrons and 480 GeV protons in October 1991. The H1 detector was largely complete a few months later, and the first period of data taking occurred in July 1992, lasting for one month. During this run, electrons at 26.7 GeV and protons at 820 GeV were collided, with ten out of a possible 220 bunches of each particle filling the HERA ring during most runs. Throughout the remaining chapters, it is this period which is referred to whenever the phrase 'current data' (etc.) is employed. The integrated luminosity for the data taken was ≈ 1.3 $\mathrm{nb}^{-1} \pm 7\%$, with the machine achieving an average luminosity $\overline{\mathcal{L}} \approx 7.5 \times 10^{27}$ cm^{-2} s^{-1}. A cutaway diagram of H1 is provided in figure 1.2.

A simplified view of the HERA complex is shown in figure 1.1. The two storage rings are located one above the other in a single tunnel of circumference 6.3 Km. The electron ring accepts 14 GeV electrons from the older PETRA II device and accelerates them by means of a series of r.f. cavities. The proton ring is less conventional; with such a high beam energy and relatively small radius, a field of $\sim$ 4.5 T is required, which is achievable only with liquid-helium cooled superconducting

bending/focussing magnets. Protons are injected into the HERA ring at 40 GeV from PETRA II, and accelerated to design energies by r.f. systems located along the straight segments of the HERA ring. The two beams are brought together with zero crossing angle in three locations with a bunch crossing interval of 96 ns. HERA was built to allow the beam energies to be varied between $10 < E_{e^-} < 35$ GeV, $300 < E_{p^+} < 1000$ GeV, although the intention is to perfect stable running at the nominal energies before attempting different energy values. Currently (1992) there exist experiments at two of these sites, H1 and ZEUS. The HERA design luminosity, with the full energy beams and 220 bunches of both electrons and protons in the filled rings is 1.5×10^{31} cm^{-2} s^{-1}. Each experiment is therefore expecting an annual accumulated luminosity of 100 pb^{-1}. A fuller account of HERA and the engineering challenges its construction entailed is provided by Wiik [1].

Dainton [2] has illustrated how HERA may be viewed as a logical extension of all the great physics experiments in the 20^{th} century, with electron scattering providing some of the most important questions and solutions to modern science, from J.J. Thompson onwards. Certainly, even restricting the field of interest to the standard model and development of Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD), the importance of electron scattering is great. The motivations for the building of HERA are summarised in [3], and in [4] a discussion of the physics aims of the experiments is provided. These aims are, in a few words, to provide tests of the QCD theory and to search for physics that goes beyond the standard model. This is not the place to cover that material again, but if nothing else is said of the QCD theory which HERA aims to explore, the uniqueness of the collisions between leptons and quarks for the first time at a centre of mass energy ($\sqrt{s} \approx \sqrt{4E_{e^-}E_{p^+}}$) of ≈ 296 GeV must be mentioned.

The aim of this work was to describe the analysis performed on the early data to collect a sample of Deep Inelastic Scattering (DIS) events. DIS processes are those where the electron scatters from a quark or gluon inside the proton with an exchange of energy and momentum mediated by a virtual particle. DIS physics lies at the heart of HERA, with this simple model of a lepton 'probing' the structure of the proton being the origin of the view of HERA as a large electron microscope. The process is characterized best by two variables–Q^2, the 4-momentum transfer and x, which here will be viewed as the fraction of the proton's momentum carried by the struck quark or gluon. These terms will be derived and explained in detail

in chapter 2. They are introduced here to emphasize the novelty of HERA. Values of up to 10^5 GeV^2 for Q^2 and down to 10^{-5} in x will be obtainable at HERA (not simultaneously), the high Q^2 limit corresponding to a spatial resolution of $\sim 10^{-18}$ m. These should be compared to existing limits of $x \geq 10^{-2}, Q^2 \leq 10^2$ GeV^2 from fixed target experiments such as those at EMC and NMC [5], [6]. HERA clearly extends into hitherto unexplored ranges. The exchanged virtual particle may be either neutral (i.e. a photon or Z^0) or charged ($W^{\pm}$) as allowed by the standard model, and so DIS physics is divided into Charged Current (C.C.) or Neutral Current (N.C.) processes. In N.C. events the electron is scattered into the detector, in C.C. events the final state lepton is a neutrino. This analysis is entirely concerned with N.C. DIS and, moreover, only with the case where a photon is exchanged. The reason for this is that the Q^2 value must be $\mathcal{O}(M^2_{Z^0}, M^2_{W^{\pm}})$ for the other vector bosons to be exchanged, i.e. the C.C. cross- section is negligible at $Q^2 < 10^3$ GeV^2 compared to that for N.C. DIS. It was always known that any early data would contain very few $W^{\pm}, Z^0$ exchange DIS events, and analysis must concentrate on the low Q^2 ($Q^2 \leq 100$ GeV^2) photon exchange case, where reasonable statistics could be rapidly accumulated . As will be demonstrated later, this has a kinematic consequence in that the scattered electron (whose detection is vital) goes into the backwards part of the H1 detector. This study then by no means covers the full range of DIS physics, merely the low energy range of the N.C. case, utilising just one angular region of the apparatus.

Having described the method by which DIS events were extracted from the data, the study then focuses on an attempt to discern which of a large set of parametrizations for the proton structure function fits the data best. Knowledge about the x and Q^2 evolution of certain terms in the DIS cross-section (the structure functions) has significant consequences for the theory of QCD. Existing parametrizations make assumptions about the extent to which gluons dominate the proton's dynamics, and the manner in which gluon–gluon interactions affect the scattering process, without the benefit of any data below $x \approx 10^{-2}$. It is below this x value that current knowledge ceases, with a large range of suggested forms for the structure functions. By presenting some of the first experimental results in a new, lower x region, a conclusion is reached which at least indicates that certain of the current models are unlikely to be correct.

Chapter 2 contains a summary of the current model of DIS physics, intro-

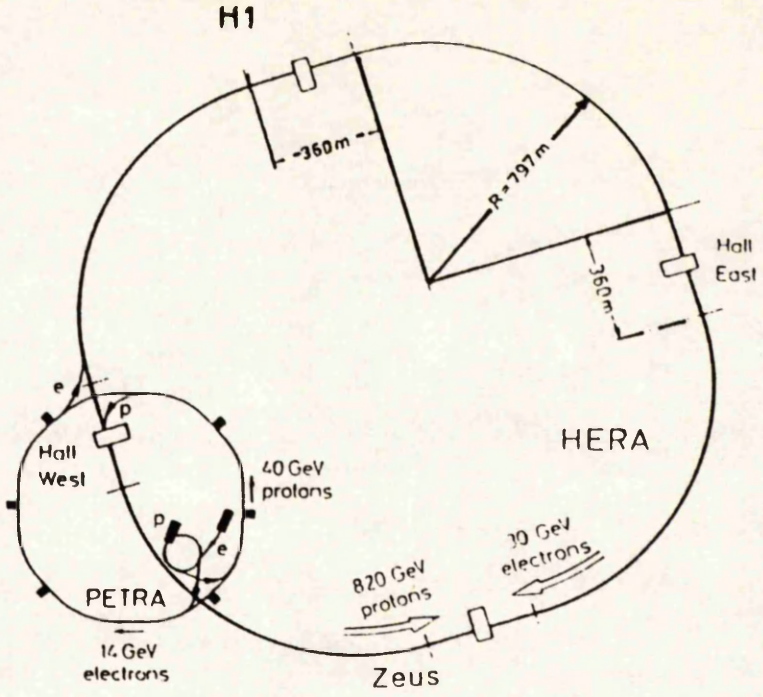


Figure 1.1: The HERA ring, showing the location of the two main experiments

ducing and explaining the terminology in which it is always couched, and discusses qualitatively the origin and form of the structure functions and their parametrizations. Chapter 3 describes the expected background processes giving rise to contamination of the DIS signal. The H1 detector is described in chapter 4, and the software tools used for the analysis are covered in the following section. Early work performed with Monte Carlo (i.e. simulated) results, prior to the advent of real data, on the problems of identifying the scattered DIS electron are presented in chapter 6. During the months immediately prior to the first runs of HERA and H1 work was undertaken within the ELAN working group to provide a strategy to filter data, the aim being to provide every member with a compacted set of events on which to perform their own analyses. This is described in the first section of chapter 7. This continues with an account of the work done to simulate the dominant source of background (photoproduction), and the techniques developed to remove this background. The final section shows how a comparison of the shape of the scattered electron's energy spectrum with Monte Carlo predictions shows some sensitivity to structure function parametrizations, and this is discussed in the concluding chapter 8.

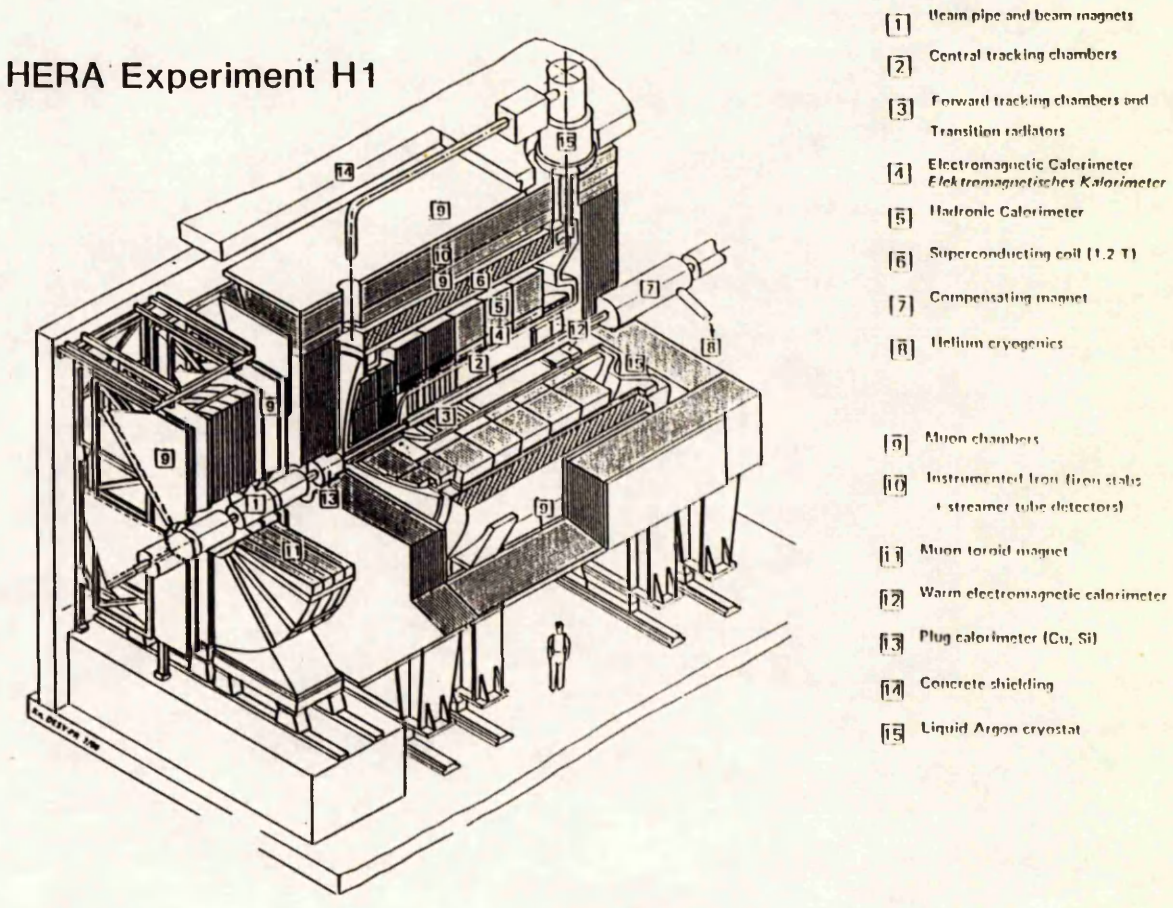


Figure 1.2: Cutaway view of the H1 experiment

Chapter 2

Theoretical Aspects of DIS

The standard model of high energy physics succesfully explains many aspects of the fundamental particles and forces which constitute the known universe. Quantum chromodynamics (QCD), that part of the standard model which describes the physics of strong interactions, gives rise to the *quark-parton model* of deep ineleastic scattering. In this chapter a qualitative review of the quark-parton model is given, and the origin and physical interpretation of the structure functions and gluon densities is described. Simple explanations of Bjorken scaling and the logarithmic Q^2 scaling violations are outlined after a detailed review of the kinematic equations used for DIS. The chapter continues with a description of some of the parametrization schemes adopted for the structure functions, and shows how a study of the energy spectrum of the DIS final-state electron can serve as a valuable indicator of the correct form.

There exists a wealth of technical literature in which the full, mathematical treatment of DIS is covered, either as a general theme within particle physics or specifically applying to electron-proton scattering at HERA. There is an extensive program of structure-function research planned for the H1 experiment, which can only commence in full after the luminosity has increased (from its value for the July 1992 run of 1.3 nb^{-1}) to around 100 pb^{-1} [7]. It is hoped that these sections will give some insight into the physical significance of the structure functions, and demonstrate why even the simple result obtainable with early data alone is of interest to physicists.

2.1 The Quark-Parton Model–Basic Principles

Two important aspects of the QCD description of strong interactions are embodied in the apparently contradictory observations that *a)* quarks are never observed to exist in isolation and *b)* that certain processes (such as DIS) have cross-sections which are successfully predicted by assuming that the quarks are non-interacting. These two phenomena are referred to as *confinement* and *asymptotic freedom* respectively. The model of inter-quark strong forces used in QCD caters for both, with a force which is weak at small separation and stronger at large distances. Asymptotic freedom lies at the heart of Feynman's Quark-Parton-Model (QPM) [8], in which the constituent elements of a hadron - the partons - are assumed to be independent, non-interacting particles. The partons are identified with the familiar quarks and gluons. To develop this model, assume that a virtual photon emitted from an electron scatters off one of these free partons, transferring some large momentum q. For the moment, only those partons which are quarks need be considered, since the gluons cannot couple to the exchanged photon. Each of the partons is in some state with a momentum $x_i P$, where P is the hadron's momentum, and $0 \leq x_i \leq 1$. The processes of time-dilation and Lorentz contraction suffered by the hadron ensures that the electron 'sees' the partons as being frozen in this state. The likelihood of there being a second quark in the vicinity which could take part in the hard scattering processes can be calculated from simple geometry. From the uncertainty relation $\Delta P \Delta x \sim 1$, the 'range' of the photon is $\sim 1/q$ and so the probability of finding a second quark is proportional to a factor $1/(q^2 \pi R^2)$, with $R \approx 1$ fm, a proton radius.

The time-scale for the scatter is similarly $\sim 1/q$, and this is assumed to be much shorter than the processes which occur after the basic interaction. The electron scatters through some angle, and the quarks and gluons undergo hadronization, converting into the streams of final-state hadrons which are eventually detected. In the QPM, these secondary reactions do not affect the cross-section at all, and happen over much longer time-scales.

Having given this brief picture of a deep-inelastic lepton-nucleon scatter, it is worth pausing to develop the well-established framework of kinematical variables and relations in which DIS physics is discussed.

2.2 DIS Kinematics

Throughout this work, a convention is maintained for labelling the interacting particles. According to this, the initial proton is said to have 4-momentum $P = (E, 0, 0, E)$, and the incoming electron (or more generally, lepton) has $E_l = (E_l, 0, 0, -E_l)$. The lepton scatters with energy E_e through an angle θ_e in the lab frame, with 4-momentum P_e. (N.B. The angle is here defined w.r.t. the incident proton beam direction. This is the convention for HERA, although many texts prefer to use the angle $\theta' = \pi - \theta$.) The proton is viewed in this quark-parton model as consisting of three asymptotically free quarks, one of which couples to the lepton via the exchanged vector boson, in this case a photon. The final state *struck* quark is denoted by its momentum P_j, and the whole hadronic final state (including the unstruck spectator quarks) is said to have momentum P'. The j subscript denotes *jet*, since the struck quark hadronizes into a stream of final state particles. This is summarized graphically in figure 2.1.

The total invariant mass squared is

$$s \equiv (P_l + P)^2 \quad \simeq \quad 2P \cdot P_l \quad = \quad 4E_l E$$

where the rest masses of the electron and proton have been ignored here as throughout the argument (entirely valid for HERA energies where $S \approx 9 \times 10^4$ GeV^2). The *momentum transfer variable* Q^2 is defined in terms of the 4-momentum transfer vector $q = (P_l - P_e)$

$$Q^2 \equiv -q^2 = -(P_l - P_e)^2$$

and 4-vector arithmetic [9] shows that

$$Q^2 \simeq 4 \cdot E_l E_e \cos^2 \frac{\theta_e}{2} \tag{2.1}$$

so that Q^2 is positive ($(P_l - P_e)^2 = -2P_l \cdot P_e$), and $0 \leq Q^2 \leq s$, i.e. at the nominal HERA beam energies Q^2s of 10^5 GeV^2 are attainable.

Restricting attention to the hadronic components, consider the quantity x', *the fraction of the incident proton 4-momentum carried by the massless struck quark.* The struck and unstruck quarks then have initial momenta $x'P$ and $(1-x')P$ as illustrated, and it follows that $q = (x'P - P_j)$. The *total* invariant hadronic mass

squared can then be written

$$W^2 = (P_j + (1-x')P)^2 \quad = \quad M^2 + 2P_j \cdot P(1-x')$$

with ${P_j}^2 = M^2$ the invariant mass (squared) of the final state *struck* quark. Using the definition of q in terms of P and P_j, it is easy to re-write this so that

$$x' = \frac{Q^2 + M^2}{W^2 + Q^2} \tag{2.2}$$

A more commonplace term to use is the scaling variable Bjorken x, always in the range $0 \leq x \leq 1$, defined as

$$x = \frac{Q^2}{2P \cdot q} \quad = \quad \frac{Q^2}{Q^2 + W^2} \tag{2.3}$$

An alternative, equivalent expression for x is

$$x = \frac{Q^2}{2m_p \nu} \tag{2.4}$$

With m_p the proton rest mass, and $\nu = P.q/m_p = E_e - E_l$. Comparing 2.3 and 2.2, it is clear that

$$x' = x\left(1 + \frac{M^2}{Q^2}\right) \tag{2.5}$$

In fact, it is usually x and not x' which is described as the fraction of the proton momentum carried by the struck quark, but this is strictly true only for the case $Q^2 \gg M^2$. For a fixed Q^2, the smallest x obtainable is given by $x_{min} = Q^2/s$, and so values of $x \sim 10^{-5}$ are opened up by HERA for the first ever time.

It is more useful to develop expressions as functions of directly measurable quantities, in particular the lepton energy and angle E_e and θ_e. Equation 2.3 can be re-written to produce

$$x = \frac{E_l E_e \cos^2 \frac{\theta_e}{2}}{E(E_l - E_e \sin^2 \frac{\theta_e}{2})} \tag{2.6}$$

and a second scaling variable y is introduced,

$$y = \frac{P \cdot q}{P \cdot P_l} \tag{2.7}$$

defined so that $Q^2 = sxy$, with $0 \leq y \leq 1$. More algebra [9] shows that

$$y = 1 - \frac{E_e - \hat{P}_{z_e}}{2E_l} \quad = \quad 1 - \frac{E_e}{E_l} \sin^2 \frac{\theta_e}{2} \tag{2.8}$$

with $\hat{P}_{z_e}$ the final state leptonic 3-momentum z component, i.e. the longitudinal momentum. In the rest-frame of the proton $P = M$ and equation 2.7 is alternatively expressed

$$y = \frac{E_l - E_e}{E_l} \tag{2.9}$$

which leads to the normal interpretation of y as the *fractional loss of energy of the incident lepton.* The three most useful variables (x, y, Q^2) have now been introduced; with just these terms and the energy and angles of the lepton and struck quark there are numerous permutations of equations which can be derived, giving any quantity as a function of any other two independent parameters. It can be convenient to write x and Q^2 in terms of the final state lepton's transverse (3) momentum $\hat{P}_{t_e}$,

$$Q^2 = \frac{\hat{P}_{t_e}^2}{1-y} \tag{2.10}$$

$$x = \frac{\hat{P}_{t_e}^2}{sy(1-y)} \tag{2.11}$$

of which use is made in chapter 6. (N.B. for the remainder of this work, which concentrates on measurable quantities, un-hatted terms such as P, P_x etc. will be used to represent 3-vectors, unless otherwise stated). Finally, for completeness, the preceeding equations can be re-arranged to give E_e and E_j in terms of (x, y, Q^2). Thus,

$$E_e = Exy + E_l(1-y) \tag{2.12}$$

$$E_j = E_l y + Ex'(1-y) \quad = \quad E_l y + Ex(1-y) + \frac{M^2}{4yE_l} \tag{2.13}$$

The plethora of formulae and interrelated terms is best understood graphically. Figure 2.2 shows the kinematic plane with the Q^2 and x axes on logarithmic scales. Iso-lines of constant y are shown. The shaded area under the leading diagonal $y = 1$ is the physical region where both x and y lie in the allowed range $0 \leq (x, y) \leq 1$. The darker shaded area covers that part of the plane which fixed target experiments such as EMC and NMC [5],[6] have already studied. Figures 2.3 (a), (b) and 2.4 (a), (b) show the Q^2, x plane with contours of constant $E_e, E_j, \theta_e, \theta_j$ marked. The angular lines have a quite simple pattern compared to the more complex form of the energy lines. The energy contours appear to 'change direction' about the line $x = E_l/E \approx 0.032$ where E_e or E_j exceed E_l.

2.3 Cross-sections and Structure Functions

The full derivation of the differential DIS cross-section is detailed in many standard texts, e.g. [10],[11],[12], and so will not be performed here. It should be stressed that all the cross-sections discussed are *inclusive* ones, i.e. no attention is paid to the make-up of the hadronic jet, and the cross-sections are those summed over all possible hadronic final states. It is worth remarking that the basis for all derivations is the treatment of large Q^2, inelastic lepton-nucleon scattering as *elastic* scattering of a lepton by free, pointlike partons. The cross-section for $ep \rightarrow eX$ may be written in terms of a contraction of a leptonic tensor with a hadronic tensor, i.e.

$$d\sigma \sim L_{\mu\nu} W^{\mu\nu}.$$

The form of the hadronic tensor may be deduced from invariance and symmetry principles and by analogy with elastic ep scattering; in this case, where the proton is treated as a pointlike particle, the cross section involves electric and magnetic form factors. Now, the structure of the proton is introduced by replacing the form factors with two arbitrary *structure functions* W_1 and W_2, corresponding to the transverse and longitudinal polarisation states of the exchanged photon. These are taken to be functions of two variables, in this case ν and Q^2. (There is a deal of variation in the relevant literature on the exact definitions of W_1 and W_2.) In fact there is a third structure function, W_3, for weak-current $(W^{\pm}, Z^0)$ exchange, but this is zero by parity invariance for the case of a mediating photon. The differential cross-section may now be written in the form

$$\frac{d^2\sigma}{dq^2 d\nu} = \frac{4\pi\alpha^2 E_e}{Q^2 E_l m_p} \left[2W_1(\nu, q^2) \sin^2(\frac{\theta_e}{2}) + W_2(\nu, q^2) \cos^2(\frac{\theta_e}{2}) \right] \qquad (2.14)$$

where $\alpha = e^2/4\pi$ is the fine-structure constant. The next important concept to summarise is Bjorken's *scaling hypothesis*. This states that in the limit $Q^2 \rightarrow \infty, \nu \rightarrow \infty$ the structure functions become functions not of ν, Q^2, but of their ratio $x = Q^2/2m_p\nu$ only, i.e. they scale with x. In this case, W_1, W_2 are replaced with $F_1(x)$, $F_2(x)$

$$W_1(\nu, q^2) \rightarrow F_1(x)$$
$$\frac{\nu W_2(\nu, q^2)}{m_p} \rightarrow F_2(x).$$

The important point is that in the limit of increasing Q^2 and ν, F_1 and F_2 are always finite. It was in fact the experimental confirmation at SLAC (see, e.g., [13]). of Bjorken scaling that led to the development of the QPM and the interpretation of DIS results in terms of elastic scattering from pointlike constituents of the nucleon. Another important result is the relation between F_1 and F_2 given by

$$2xF_1 = F_2$$

the Callan-Gross relation, a direct consequence of the spin-$\frac{1}{2}$ nature of the partons. There is also good experimental evidence for this from SLAC results. It is instructive to return to the earlier description of the QPM for insight into the physical significance of scaling. The virtual photon can be visualized as a 'probe', whose resolution increases with Q^2 and ν, and it is assumed that the partons contribute incoherently to the cross-section. If the partons exist in a number of types i (for quarks, these are simply the flavours) then the total partonic contribution is evaluated by integrating over all values of momentum fraction and summing over all types. Aitchison and Hey [11] show how these considerations lead directly to an expression

$$\nu W_2(\nu, q^2) = \sum_i {e_i}^2 x f_i(x) \equiv F_2(x)$$

where e_i is the charge in units of e, and $f_i(x)$ the probability for a parton of type i and momentum fraction x. (Strictly, this form is specific to the case where $Q^2 \ll {M_Z}^2$, and the full form replaces ${e_i}^2$ with a term involving the N.C. vector and axial couplings of the fermions). This introduces the idea of the structure functions as parton probability distributions. Again, assuming that the quarks alone can couple to the photon, and ignoring any contribution from heavy quarks, F_2 can be split into six unknown functions

$$F_2 = x\left(\frac{4}{9}[u(x) + \overline{u}(x)] + \frac{1}{9}[d(x) + \overline{d}(x) + s(x) + \overline{s}(x)]\right)$$

where $u(x)$ is the probability distribution for u quarks, i.e. the number density of up quarks with a fraction x of the proton's longitudinal momentum.

So far, the other partonic constituent of the proton - the gluons- has been ignored. Using a notation taken from [12], integrating over the momentum fraction,

$$\int_0^1 x(u + \overline{u} + d + \overline{d} + s + \overline{s})dx = 1 - \epsilon_g$$

with ϵ_g the momentum fraction of the gluons. This has been determined from electron-proton and electron-neutron scattering experiments to be $\sim 50\%$ [10]. The gluon density function $G(x, Q^2)$ can be defined so that, for momentum conservation,

$$\int_0^1 dx F_2(x, q^2) + \int_0^1 xG(x, q^2) = 1$$

Any detailed study of DIS must take QCD effects into consideration, and the effect on the scattering dynamics that e.g., gluon radiation from quarks and quark-antiquark pair production (i.e. $\gamma g \rightarrow q\bar{q}, \gamma q \rightarrow qg$) can have. These contributions are $\mathcal{O}(\alpha\alpha_s)$ (where α_s is the strong coupling constant) compared to the leading contribution $\mathcal{O}(\alpha)$. Including the gluon-bremsstrahlung and $g \rightarrow q\bar{q}$ diagrams into the QCD calculations for F_2 reveals that the scaling previously noted no-longer holds, and a logarithmic Q^2 dependence is introduced, i.e the scaling has been violated. An attempt can be made to justify this qualitatively; returning to the previous idea of the virtual photon probe, then previously the photon at some value ${Q_0}^2$ was capable of 'seeing' just the point-like partons. At a higher value of Q^2, the quark is observed to be surrounded by other partons arising, for instance, from a radiated gluon which has in turn formed a $q\bar{q}$ pair. As Q^2 increases, then the number of these resolved partons among which the proton's momentum is divided also increases. Since higher momentum quarks will radiate energy in the form of gluons, there is a reduced chance of the photon being absorbed by a high x quark, and the structure function is decreased. Conversely, the structure function is raised in the small x region, where there are more quarks with a lower momentum (called *softer* quarks) which have undergone gluon emission. Another result of taking the gluons into account is that the Callan-Gross relation no-longer applies; the quantity $F_2 - 2xF_1$ is now non-zero, and is often called the *longitudinal* structure function, F_L. It is in fact more conventional within H1 to express the differential cross-section in terms of F_2 and F_L, i.e.

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4 x}\left[(1 - y + \frac{y^2}{2})F_2(x, Q^2) - \frac{y^2}{2}F_L(x, Q^2)\right] \tag{2.15}$$

once again ignoring the F_3 term.

The Q^2 dependence of the structure functions is described by the Altarelli-Parisi (A–P) evolution equation [14], e.g. for F_2,

$$\frac{dF_2(x, q^2)}{d\log Q^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y}F_2(y, Q^2)P_{qq}\left(\frac{x}{y}\right)$$

The splitting function $P_{qq}(k) = 4(1+k^2)/3(1-k)$ is the probability of a quark reducing its momentum by a fraction k by gluon emission. In words, this equation states that a quark with a momentum fraction x can arise from one with a larger fraction y which has undergone gluon radiation. A similar equation exists for $dxG(x,Q^2)/d\log Q^2$. The splitting functions reflect the underlying processes such as $g \rightarrow q\bar{q}$, $q \rightarrow gq$, $g \rightarrow gg$ etc., i.e. the way in which high momentum quarks and gluons become quarks and gluons of lower momentum via splitting and radiating.

The existence, and violation of scaling are demonstrated graphically in figure 2.5. From the BCDMS and NMC (preliminary) results (see section 2.4 below), it shows F_2 plotted against Q^2 for a range of x values. It will be seen that F_2 increases with Q^2 at small x, and decreases with increasing Q^2 at large x. At $x \approx 0.3$, there is approximate independence of Q^2 (i.e. scaling holds). The graphs also demonstrate another point made below, that the NMC data lies above the KMRS fit to F_2 at low x.

2.4 Parametrizing the Structure Functions

At its simplest, the procedure for developing models of the structure functions to fit with existing data consists of i) select a low value of Q^2, ${Q_0}^2$ in the range of existing data (~ 4 GeV^2), ii) Parametrize the parton distributions by some (physically justifiable) form e.g. $xG(x,Q^2) \sim (1-x)^\alpha$, iii) use the A–P equations (solving numerically) to evolve the form to higher Q^2 and iv) perform a global fit to data in order to establish the value of the parameters. There are numerous complications however; the A–P equation is valid only in a range of Q^2, x where a 'Leading Log' approximation can be made, which holds only for $x \sim 1$ and $Q^2 > {Q_0}^2$, the reference values. Mathematically, the formation of the A–P equation requires the summation of terms $\mathcal{O}(\alpha_s \ln Q^2)$, $\mathcal{O}(\alpha_s \ln 1/x)$; with a value of $Q_0 \approx 4$ GeV^2, the summation can be performed, within the approximation that only those terms $\mathcal{O}(\alpha_s \ln Q^2)$ are included [4]. When the small x region ($x < 10^{-2}$) is explored, this is no longer valid. A second gluon evolution equation has been developed which is valid at small x and $Q^2 \sim {Q_0}^2$, the Lipatov equation [15]. The details of the Leading Log Approximation and the formation of the Lipatov equation lie far beyond this study (see [4] and [16]) but it is important to realise that it corresponds to a different kinematical domain.

The Lipatov equation includes sums over terms $\mathcal{O}(\alpha_s \ln 1/x)$. It is normally derived by considering a 'ladder diagram', where a series of gluon 'rungs' between two interacting quarks are summed up. It is this equation which extends the kinematic range to small x i.e. the HERA range, and allows the fitting procedure described above to be performed for the gluon distribution. Solving the Lipatov equation (numerically) reveals that in the limit $x \to 0$ the gluon density $xG(x,Q^2) \to C \cdot x^{-\lambda}$, with λ having a maximum value of $\approx 1/2$.

It is again worth pausing to contemplate a qualitative view of the mathematical operations described in the above paragraphs. The A–P and Lipatov equations describe a proton in which the virtual photon scatters from a quark that is the final result of a cascade of quark-gluon interactions. The equation for the gluon density reveals that at smaller x and Q^2 the gluons tend to dominate, and eventually the density rises until the proton becomes fully packed. This rise cannot continue indefinitely, and some damping mechanism is required.

Two mechanisms have been proposed to limit the rise of the gluon density. One is the suggestion [17] that the higher-order terms of the Lipatov equation would, if taken into account, produce the necessary dampening effect. The second, commoner, idea is that of saturation. The GLR equation [18] is a third evolution equation for quark and gluon distributions. It incorporates a linear term such that at large Q^2 it becomes the standard A–P equation, and at small x it behaves as the Lipatov equation. To the GLR equation a non-linear, negative term is added. The physical justification proposed is that there is some process of parton recombination which sets in at small x, which reaches an equilibrium with (or exceeds) the partonic decays. This recombination is often referred to as shadowing, and is usually characterized by a parameter R, interpreted as the size of the region in which saturation starts. This may be very localised ('hot spots', see [19]) or uniformly distributed throughout the proton, in which case R is taken as the proton radius. It is believed [20] that 2 GeV^{-1} < R < 5 GeV^{-1}. In the 'ladder picture', shadowing is equivalent to different ladders of gluons fusing together.

It is the GLR evolution equation which applies to the relevant HERA domain of $x \to 0$ and $Q^2 > {Q_0}^2$. No complete, analytical solution to the GLR equation exists, but a number of schemes have been developed to predict the value of both the F_2 structure function and the gluon density at small x using approximate numer-

ical solutions to the GLR or Lipatov equations. (N.B., the longitudinal structure function F_L is heavily dependent on $xG(x,Q^2)$, and so these techniques apply to it also.) For instance, two of the more recent (and most often used) schemes are that of the KMRS group and the MRS group. Using data taken from the EMC and BCDMS [5], [21], (μ—protons and μ—deuterium) collaborations KMRS constructed two parametrizations, B0 and B- (1990). The '0' and '-' labels correspond in turn to the gluon distributions at $x = 0$ being either finite ($xG(x) \sim x^0$) or singular ($xG(x,q^2) \sim x^{-1/2}$). The sea-quark distributions are parametrized in similar forms, with an x^0 or $x^{-1/2}$ dependence, reflecting that they are expected to have the same $x \to 0$ behaviour as the gluons. (In the language of KMRS, 'the sea is driven by the gluons via the process $g \to q\overline{q}$'.) The singular forms then have shadowing terms added to prevent the gluon densities rising too fast, in the same way that the evolution equation $d(xG(x,Q^2))/d\ln Q^2$ was modified by a shadowing component. The shadowing term involves a 'saturation gluon density', the limiting value of $xG(x,Q^2)$ calculated from GLR. It must be stated that there is no *a priori* reason or physical justification to suppose the $x^{-1/2}$ form; numerical solutions to the Lipatov equation simply predict that the distributions behave as $x^{-\lambda}$ at small x, and the x^0 and $x^{-1/2}$ parametrizations both fit existing data at $x > 10^{-2}$. B- and B0 agree well with each other down to $x = 0.07$ but then diverge, with the B- set being higher at very small x. Reference [22] contains all the details of the full analysis. Later data produced by the NMC and CCFR [6], [23] experiments (scattering μ—protons/deuterons and ν—iron respectively) led to an adjustment of these methods by the MRS team [24] in 1992. The new NMC data was found to be consistently above the KMRS fits (by $\sim 30\%$ at $x = 0.0125$) when they were extrapolated to low x, and three new fits (S0, D0 and D-) were devised. The 'D' fits are different principally in that the density distributions of up and down quarks are allowed to differ, whereas $\overline{u}(x,Q^2)$ was previously fixed $= \overline{d}(x,Q^2)$ in the 'B' fits. Other schemes resulted in the older (1990) B1 and B2 structure functions of Morfin-Tung [25] and EHLQ [26].

The different families of structure functions are summarized graphically in figures 2.6 (a), (b) and 2.7 (a)-(c). These show plots of F_2 and xG versus x at fixed Q^2 for various parametrizations. For all the graphs, the following points are to be noted; the parametrizations agree at $x > 0.1$, disagree increasingly as x drops below this value, and all rise with decreasing x. The first plot shows 7 such curves with the MT-B2 and KMRS-B0 at the extremes. The second (from [24]) shows the later

results of the MRS group's D- and D0 results, fitted to the NMC and BCDMS data at high x. The top line corresponds to a parametrization with no shadowing term, and the other two D- plots represent different shadowing scenarios. $R = 5$ GeV^{-1} is the value for a uniform distribution of gluons, $R = 2$ GeV^{-1} assumes them to be localized. Figure 2.7 (a) shows the motivation for the MRS D0/D- set, with the extrapolated KMRS B0/B- results lying clearly below the new NMC data. The final two graphs are of the MRS gluon density parametrization at two Q^2 values.

Two important results are borne out in these plots; firstly, it will be observed from the D- and D0 plots that similar overall results for the steepness of F_2 can be achieved either by having a steep input distribution evolved with shadowing terms, or by having an un-shadowed flatter input form; the shadowed D- (with $R = 2$ GeV^{-1}) and un-shadowed D0 are not too dissimilar in shape. Secondly, from the third pair, it is clear that changing the fixed Q^2 alters the extent to which different parametrizations differ at a fixed x. This demonstrates the importance of making comparisons at a fixed Q^2, or at least over a range of Q^2 narrower than ~ 2 orders of magnitude.

In all cases, the two most critical input quantities are the starting distributions (i.e. the form of xG and F_2 at some initial, low ${Q_0}^2$, x_0) and the existence and size of shadowing corrections to the GLR equation, that is terms introduced to account for parton recombination processes.

2.5 Structure Functions and the H1 Experiment

References [4], [7] and [27] contain detailed accounts of the comprehensive structure function projects planned for H1. Their analyses on the extraction of F_2, F_L and the gluon distribution from the data, and the proposed tests on QCD that may be conducted (e.g. providing precise values for Λ_{QCD} or α_s) will be of central importance within the full physics programme.

All precise studies of structure functions require the extraction of F_2 and F_L from the differential cross-section, and the analyses proceed by, for instance, plotting F_2 as a function of x for a variety of fixed Q^2 values and plotting $dF_2/d\log Q^2$ against x. It may also be necessary to run HERA at a range of lower beam energies

in order to extract F_L and F_2 and to make contact with the results of fixed-target experiments.

These are however of little current relevance until increased luminosities are achieved, and so the question must be asked; what tests can be performed with the low level of data (and low level of precision, e.g. calibration of calorimeters) amassed by the cut-off date for this analysis?

At the most basic level, it will have been made clear by the previous discussion that a determination of the low x behaviour of F_2 will be of great interest. The very wide range of F_2 values at low x shown in figure 2.6 (a) indicates the path which the early-day analysis should follow. Such is the discrepancy which exists between the steepest structure function parametrizations (e.g. MT-B2) and the shallowest (e.g. KMRS B0) that it should not be necessary to extract F_2 from the cross-section in order to distinguish between the extremes; the shape of the spectrum of some directly measurable quantity should serve as a reflection of the underlying structure function. The most suitable quantity is the energy of the outgoing scattered electron. The likely form of the spectrum should be considered with reference to figures 2.3 and 2.4. In the range $Q^2 < 10^2$ GeV^2 where most of the initial data will occur (equivalently, in the angular range $\theta_e > 150°$) a large portion of the (x, Q^2) plane is occupied by events with $E_e \approx E_l$. Consequently, the energy spectrum will peak around this value, (the so-called kinematic peak), with tails extending either side. The cross-section is highest at low Q^2 and low x (witness the form of the F_2 structure function), and so the tail on the high-energy side of the peak will be small compared to that on the low-energy side. Figure 2.3 shows that the low E_e events correspond to low x values. So, it is to be expected that the shape and/or height (relative to the peak value) of the low energy tail will illuminate the low x structure function behaviour. The tail should be compared with a variety of Monte Carlo predictions, each generated with different input parametrizations. Flatter tails (i.e. a smaller low x cross-section) are expected to result from a choice of MT-B1, larger ones from MT-B2 and so forth.

It seems certain that no great accuracy can result from such a procedure, but that it inevitable for any early result; the aim is not to be able to make precise statements about the evolution equations and parton interactions, but to differentiate between broad sets of parametrization schemes. Provided that radiative

corrections (see chapter 3) are included into the Monte Carlos used, there is no need to separate the Born from the radiative cross-sections in this method.

The non-trivial task of obtaining a clean sample of DIS events at low energy forms the core of this work, and it is hoped that the motivation for this has been clearly illustrated. Of course, even if it were to prove impossible to gain any insight into the structure functions via this simple method, the techniques of electron selection and background removal are widely applicable and remain of paramount importance to future work.

2.6 Measuring x, y and Q^2

As explained above, it is the electron energy spectrum which is desired for this analysis, and the extraction of the kinematic variables is not of primary importance. Nevertheless, it is worth a short diversion into this topic since it is essential to any further work which is conducted, and leads also to an understanding of the limits which are placed on the accessible regions of the kinematic plane.

Equations 2.1, 2.6 and 2.8 show that by measuring the scattered electron's energy and angle, x, y and Q^2 can all be deduced. The calorimeter and tracking devices have finite resolutions however, and these lead to resolutions in the variables as follows;

$$\frac{\delta Q^2}{Q^2} = \sqrt{\left(\frac{\delta E_e}{E_e}\right)^2 + \left(\delta\theta_e \tan(\frac{\theta_e}{2})\right)^2}$$

$$\frac{\delta x}{x} = \sqrt{\left(\frac{1}{y}\frac{\delta E_e}{E_e}\right)^2 + \left(\tan(\frac{\theta_e}{2}) + (\frac{1}{y} - 1)\cdot \cot(\frac{\theta_e}{2})\cdot \delta\theta_e\right)^2}$$

$$\frac{\delta y}{y} = \sqrt{\left((1 - \frac{1}{y})\cdot \frac{\delta E_e}{E_e}\right)^2 + \left((\frac{1}{y} - 1)\cdot \cot(\frac{\theta_e}{2})\cdot \delta\theta_e\right)^2}$$

The presence of the $1/y$ term in the x and y resolutions makes these divergent at low y, and so introduces a restriction that y must be greater than ~ 0.1 for them to be reasonable (say, below $\sim 20\%$). The Q^2 measurement is free from such terms and is limited only by the detector resolutions. In an attempt to extend the y range,

the kinematic formulae in terms of the current jet may be considered.

$$Q^2 = \frac{1}{1-y} \cdot E_j \sin^2 \theta_j \tag{2.16}$$

$$y = \frac{E_j}{E_l} \sin^2 \frac{\theta_j}{2} \tag{2.17}$$

A similar exercise to the one above shows that only the $\delta y/y$ term is without any y divergencies; the Q^2 and x resolutions both have $1/(1-y)$ terms, leading to poor results at high y. Thus, one simple scheme is to derive a Q^2 measurement from the electron, a y measurement from the hadron, and so obtain x from $x = Q^2/sy$. There are additional limitations to such an approach, the most restricting of which is the beam pipe. This acts an angular cut of approx 5° and a study of the θ_e, θ_j contours in figure 2.4 show that the overall effect is twofold; a low Q^2 cut of $\sim$ 10 GeV^2 for $x \geq 10^{-4}$ because the electron is lost in the pipe (and at smaller x, the low electron energy means that the $\delta E_e/E_e$ term becomes sizeable), and a high x cut of $\sim 10^{-2}$ beyond which the hadrons are lost in the forward direction.

Measurements of the energy and angle of the current jet are complicated by its hadronization; although the above discussion implied that the struck quark fragmented into a well-defined pencil-like jet with angle θ_j, in practice the resulting spray of particles can form a wide cone. Also, the M term in equation 2.13, non-negligible at small y, means that fluctuations in the jet mass lead to fluctuations in jet energy and angle. Hoeger [28] states that M^2 may exceed 30 GeV^2 and showed that jet algorithms (i.e. an algorithm for identifying which particles belong to the current jet) could never lead to a better result than could be obtained by using the approach of Jaquet and Blondel [29]; this latter method is a technique for calculating the kinematic variables which uses the hadronic flow *regardless* of any jet structure. Rather, summations are made of the energy and momentum of the hadrons, and the following equations derived;

$$y_{jb} = \frac{\sum_i (E_i - P_{z_i})}{2 \cdot E_l} \tag{2.18}$$

$$Q^2{}_{jb} = \frac{\sum_i P_{t_i}{}^2}{1 - y_{jb}} \tag{2.19}$$

$$x_{jb} = \frac{Q^2{}_{jb}}{s \cdot y_{jb}} \tag{2.20}$$

In each case, the summation is over all observed hadrons. Reconstruction errors are the result of finite resolutions and the beam pipe angular cut-off. The Jaquet-Blondel technique, like the current-jet approach, is unable to overcome the fact that the Q^2 determination is singularly poor for all y values. The $(2-y)/(1-y)$ term in the Q^2 resolution rises at large y, and at small y particles lost down the beam hole deplete the P_t summation.

A number of attempts have been made to go beyond the Q^2 from electron and y from hadrons (Jaquet-Blondel) approach (now called the 'mixed method'). One fairly obvious alternative is to combine the lepton and hadron information more directly, and it has been proposed to do this by using the electron and current-jet angles [30]. Having expressed x, y and Q^2 in terms of θ_j, θ_e the problem of the ill-defined θ_j is overcome by using the Jaquet-Blondel formulae to write $\cos\theta_j$ as a function of ${Q^2}_{jb}$ and y_{jb}. This double-angle method was re-worked by Hoeger [28], who defined two pseudo-angles

$$\alpha_l \equiv \frac{E_e - P_{z_e}}{P_{t_e}} = \sqrt{\frac{(1-y)E_l}{xyE}}$$

for the electron and

$$\alpha_h \equiv \frac{\sum_i (E_i - P_{z_i})}{\sum_i P_{t_i}} = \sqrt{\frac{yE_l}{x(1-y)E}}$$

with summations over the hadrons. The two 'angles' are convenient parameters; the M^2 term present in 2.13 (also present in the expression for $\sum_i P_{z_i}$) no longer appears and α_l is independent of the electromagnetic energy resolution. A 'full fit' routine is also described in [28], essentially a χ-squared minimization using both the hadron- and electron- derived kinematic variables as inputs.

The five schemes outlined above–electron only, Jaquet-Blondel, mixed, double angle and 2α—have been studied in detail and summaries of this work are present in [28], [30], and [31]. The first of these contains a simplified summary of the best method to use in a given region, reproduced here;
For $y \geq 0.1$, best results from electron alone
For $y < 0.1, x > 0.01$, best results from 2α
For $y < 0.1, x < 0.01$, best results from fitting routine.

The above references confirm that the worst Q^2 measurement came from using the hadrons alone, and that in all combined methods the Q^2 determination

relied on the Q^2/θ_e constraint. Setting limits to the allowed level of migrations which occur between (x, Q^2) bins due to reconstruction errors, Monte Carlo studies in [28] showed that the 'mixed' method covered the largest part of (x, Q^2) space, but was not superior to all the other methods in any local region. Bernardi [31] in particular has provided an extensive analysis of the likely effect of the different reconstruction schemes on the differential cross-section measurement. No attempt will be made here to discuss this topic, but figure 2.8 is reproduced from this work, and shows the predicted domains where the systematic errors on $d\sigma/dxdQ^2$ are below 10% for 4 schemes.

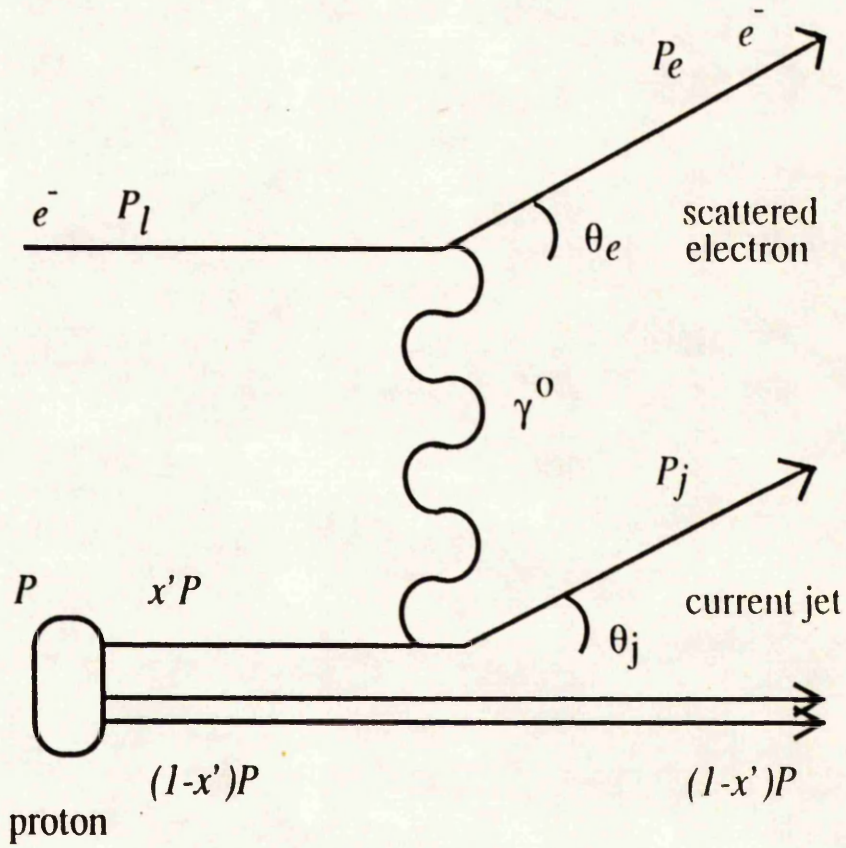


Figure 2.1: DIS scattering process

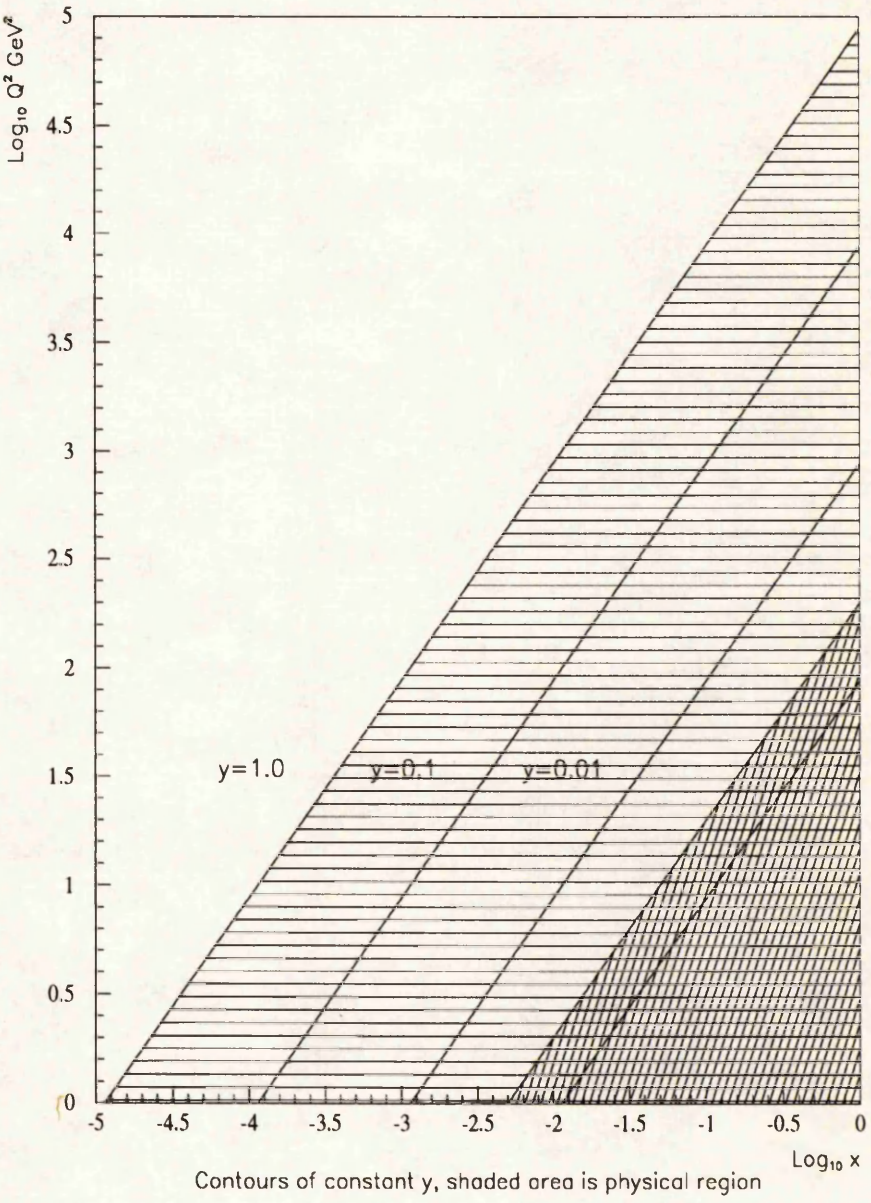
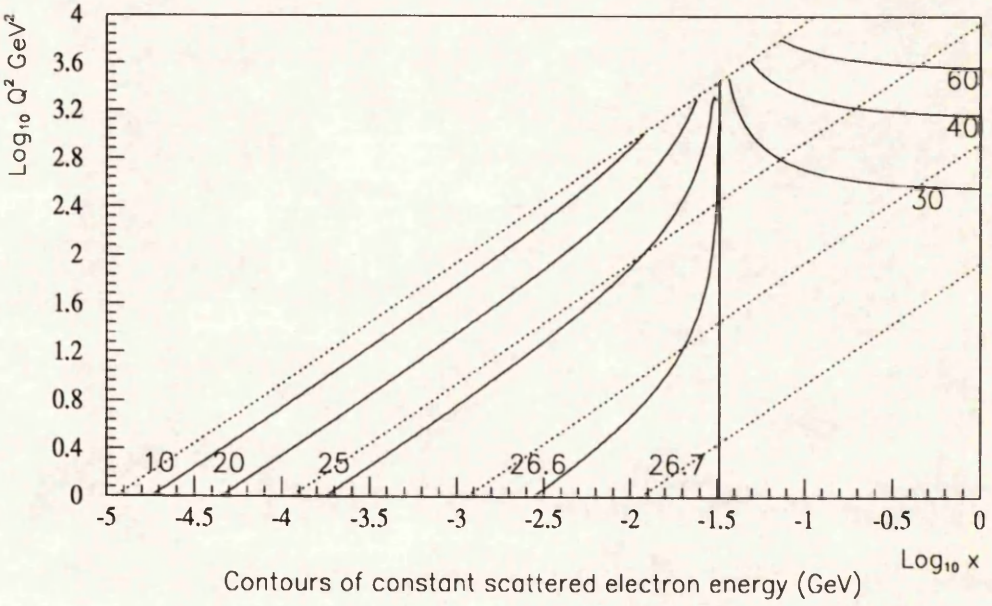
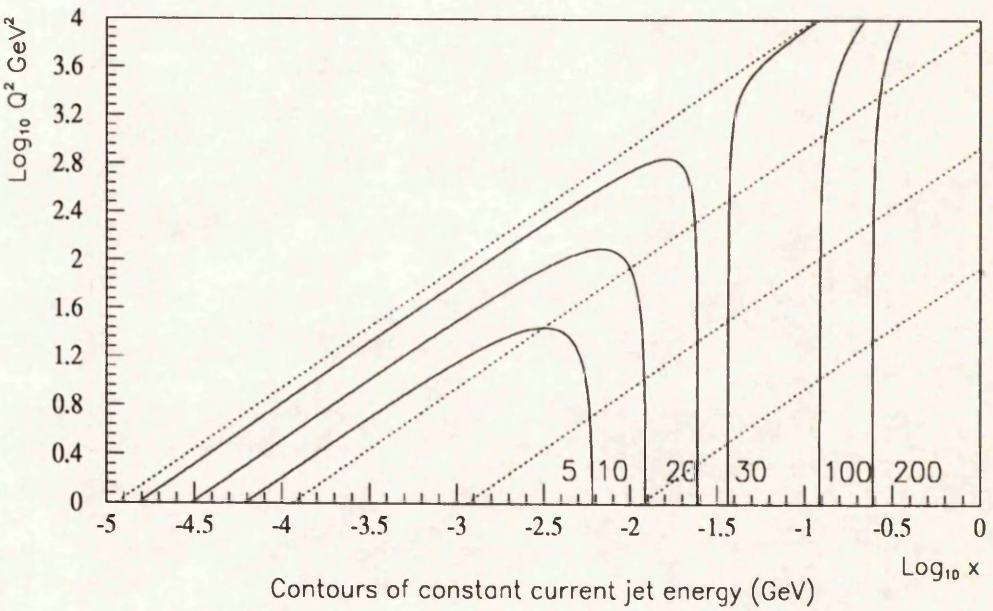


Figure 2.2: The (x, Q^2) plane, log axes. Double shaded area is region explored by fixed target experiments.

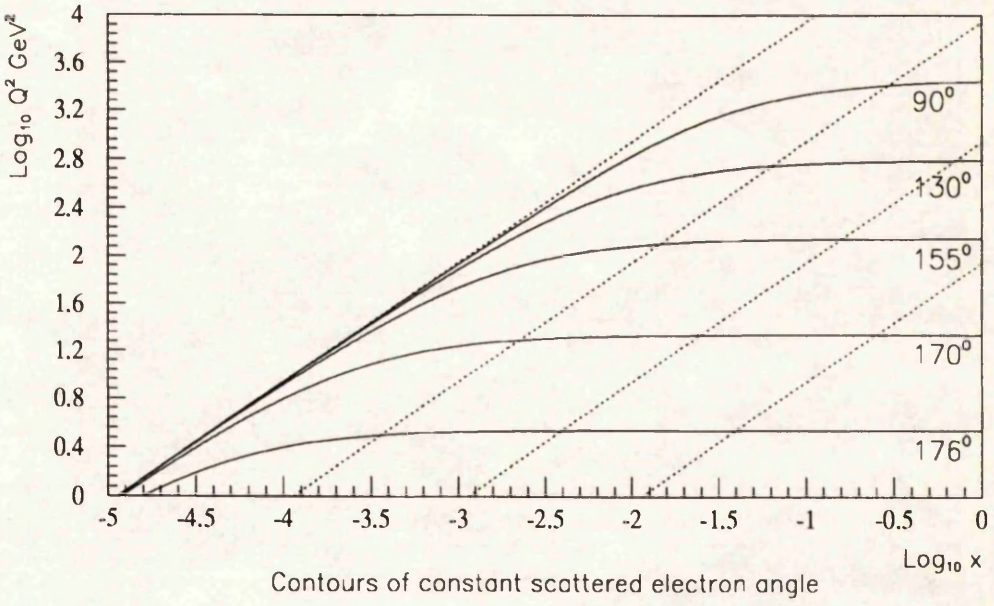


Contours of constant scattered electron energy (GeV)

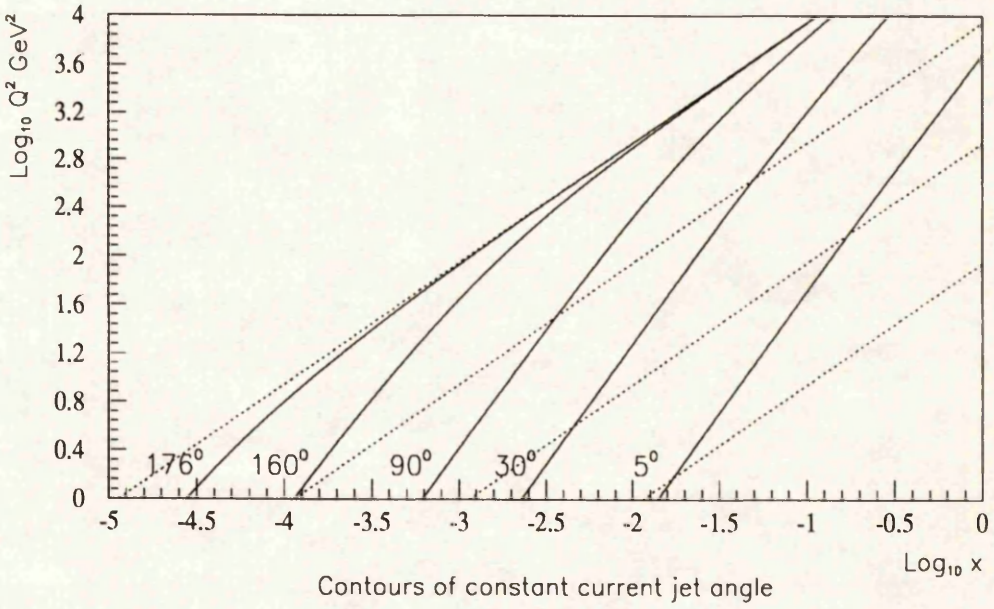


Contours of constant current jet energy (GeV)

Figure 2.3: Iso-lines on the (x, Q^2) plane



Contours of constant scattered electron angle



Contours of constant current jet angle

Figure 2.4: Iso-lines on the (x, Q^2) plane

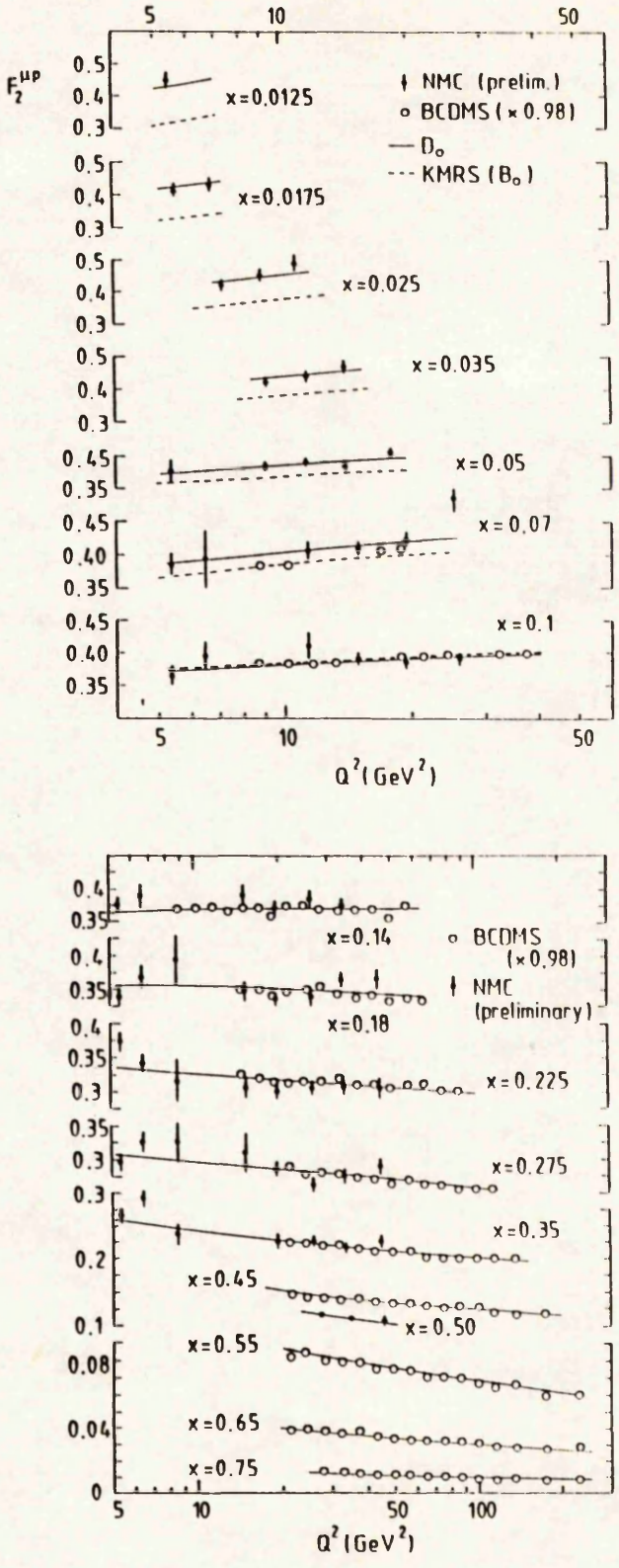
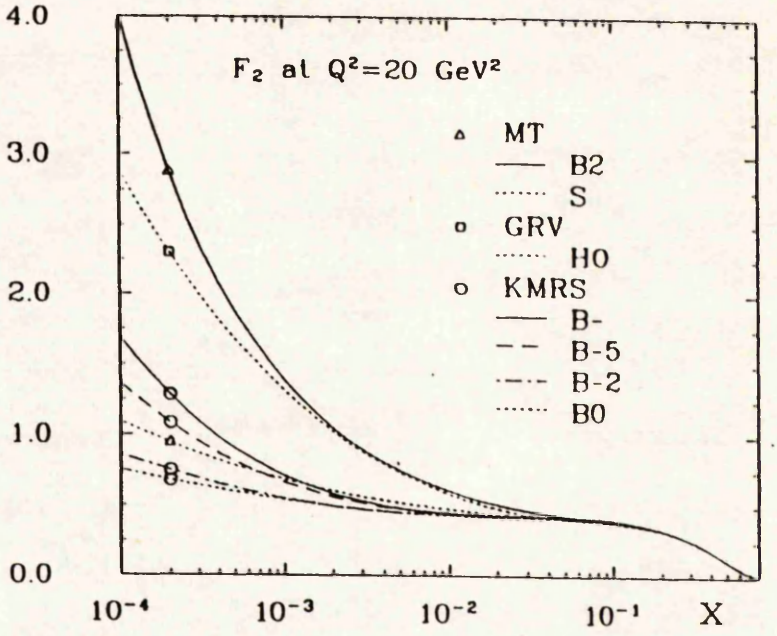
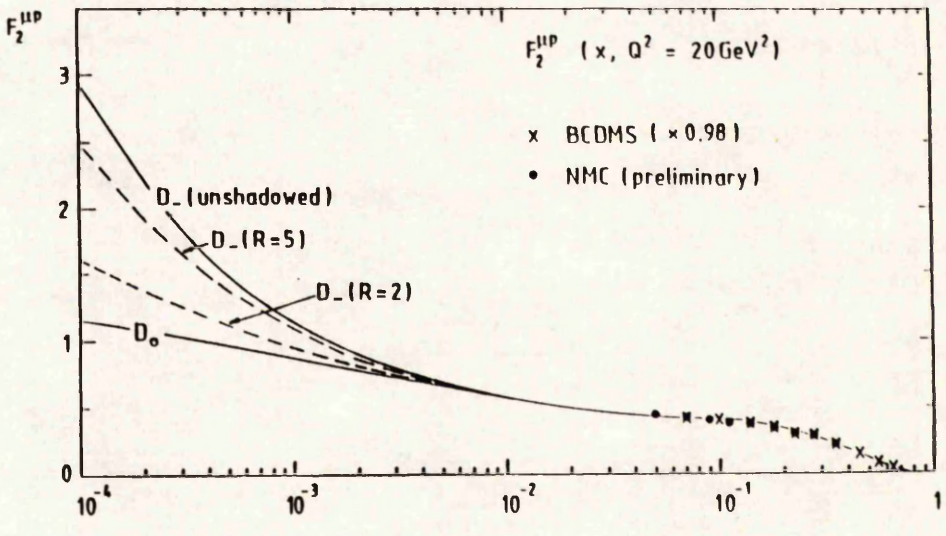


Figure 2.5: F_2 as a function of Q^2 and x.

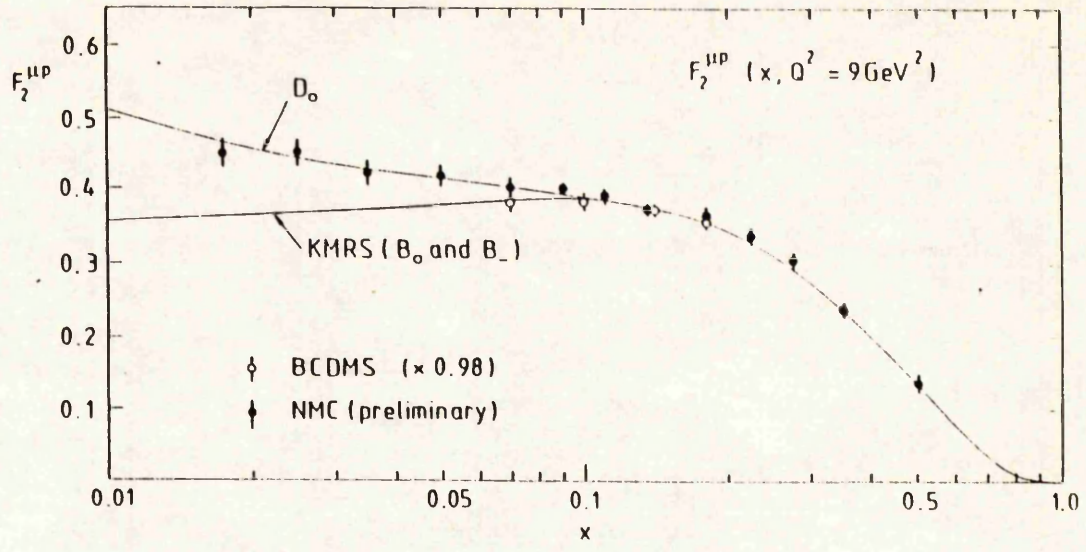


(a) Several F_2 parametrizations.

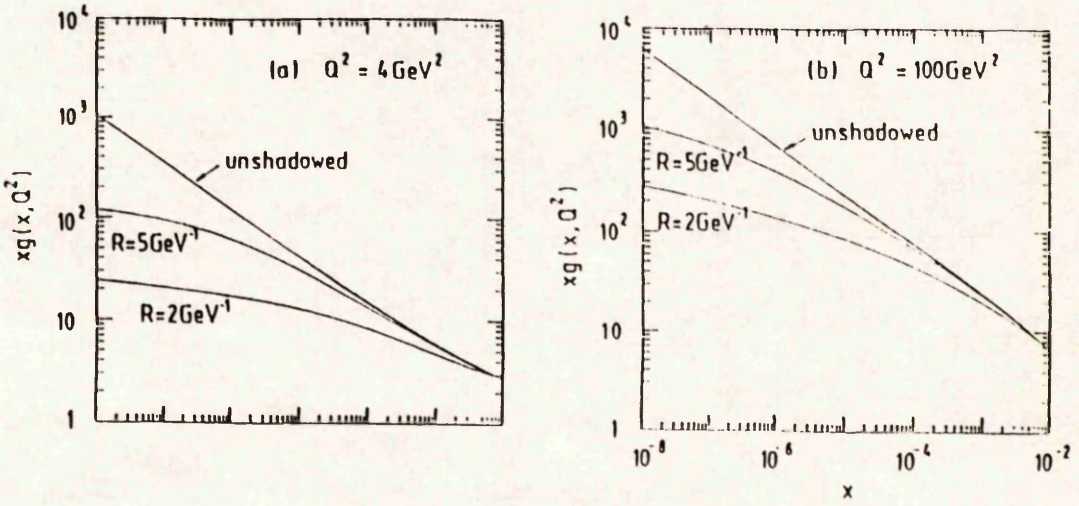


(b) the MRS D0/D- sets.

Figure 2.6: Structure function parametrizations



(a) MRS D0/D- cf KMRS B0/B- .



(b), (c) the MRS D- gluon distributions at two orders of Q^2.

Figure 2.7: Structure function parametrizations

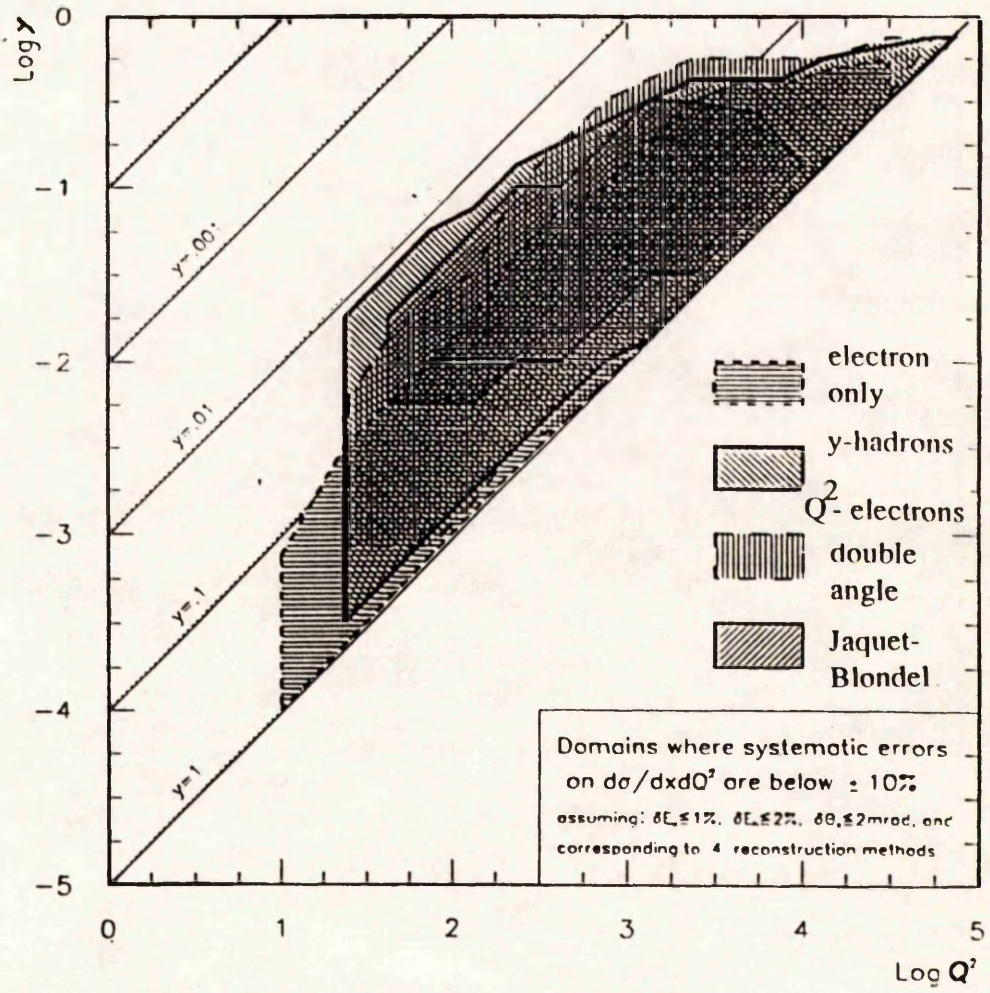


Figure 2.8: Reproduced from Bernardi et al

Chapter 3

Background Sources to DIS

The main sources of background events to DIS are briefly described in this chapter, and an attempt made to predict their likely magnitude.

3.1 Photoproduction

The regime of very low Q^2 electron-proton scattering, where the exchanged photon is very nearly on mass-shell is called *photoproduction*. It is a topic of much interest in itself, although interest here is confined merely to explaining why it constitutes a likely source of contamination.

The basic physical process is the interaction between the proton and a low Q^2 photon emitted by the scattering electron. The electron is deflected through only a very small angle, and if $\theta_e > 176°$ then it will escape from the detector down the backwards beam pipe. Although the final-state electron has a low transverse momentum, there can still be high other P_t particles produced. The boundary between photoproduction and DIS is somewhat arbitrary, the former being merely the low Q^2 part of the latter. In practice, it is conventional to label any event where the electron is lost in the beam pipe, corresponding to a Q^2 cut of ~ 3 GeV^2 as photoproduction, but this does not mean that the underlying physics has undergone any sudden change. Photoproduction events are separated into a number of classes,

the most interesting of which are

- Vector Meson Dominance (VMD)
- *Direct* point-like photon interactions
- *Resolved* photon interactions

These terms will now be explained.

3.1.1 Vector Meson Dominance

In this model the photon is viewed as a vector meson e.g. ρ, ω, ϕ, (provided Q^2 is $\mathcal{O}(\text{mass})^2$ of the meson in question). The scatter is then that of a meson and proton, and is generally soft (very low Q^2) and hence not calculable in perturbative QCD. VMD is the dominant contribution to photoproduction at $Q^2 \approx 1$ GeV^2, and by $Q^2 \approx 2$ GeV^2 its cross-section has fallen to be about equal to that of the direct and resolved photon contributions. A quasi-elastic diffraction of the meson through the proton may take place or, at the other extreme, hard scattering between the meson and a parton in the proton may occur and give rise to high P_t final-state particles. (In fact, the term 'hard' is here used somewhat loosely; if the scatter were truly hard, it would be between a quark of the vector meson and a quark of the proton, i.e. the same as resolved photon case described below. VMD is a term for the non-diffractive, softer tail of photoproduction, and all scatters are really low Q^2, i.e. soft by comparison with DIS. High P_t products can be found in the final-state as a result of the hadronization occuring when the vector meson breaks up.) It is qualitatively different from the resolved photon process described below in that the meson is considered as a single entity throughout the scatter, and is not treated as a separate quark pair. VMD is illustrated schematically in figure 3.1 (a).

3.1.2 Direct Photon Sub-Processes

Here, the photon is treated as point-like, and assumed to couple directly to a quark in the proton. The two lowest-order processes are those where the photon

couples directly to a quark, called QCD Compton scattering (figure 3.1 (b)), and where it couples to a gluon via a quark propagator, called photon-gluon fusion (figure 3.1 (c)). Cross-sections for these interactions are calculable within QCD.

3.1.3 Resolved Photon Sub-Processes

Due to the fact that the exchanged photon can undergo $\gamma \rightarrow q\bar{q}$, the idea of a photon structure function arises to describe its 'hadronic content'. This is sometimes referred to as the *anomalous* structure function, and these higher-order events described as anomalous photoproduction. The scatter then proceeds via the interaction of a partonic element (i.e. a quark or a gluon) of the resolved photon with one from the hadron. Possible resolved processes are shown in figures 3.2 (a) - (f). These events will all produce three jets (ignoring the unstruck quarks in the proton) compared to two for the direct photon case which will be observable if of sufficient P_t.

The range of available structure function parametrizations for the proton has been described at length in chapter 2, but for many years there was only the Drees-Grassie parametrization for the photon [32]. This has recently been supplemented by the LAC1 parton distribution [33] which was devised using more data than was used for Drees-Grassie. The principal difference between the two is the softer gluon distribution of LAC1 (i.e. greater at low x), raising the cross-section wherever the gluon from the resolved photon is involved in the interaction.

As for the direct photon case, the cross-sections for the illustrated resolved processes are all QCD calculable to first order and hence can be modelled by Monte Carlo generators, including the PYTHIA ([34] program used for this analysis. For both the direct and resolved components, a cut on the transverse momentum of the final state struck quark has to be introduced in such Monte Carlos because the QCD cross-section calculation is badly divergent at low P_t. The choice of the P_t cut is discussed in section 7.2.1, and is typically in the range 1-5 GeV/c.

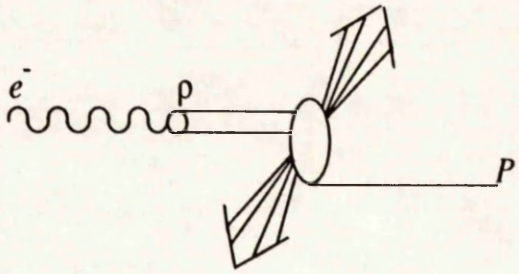
3.1.4 Significance for DIS Studies

The threat that photoproduction events pose to *charged* current DIS (where there is no electron in the final-state) has long been recognized, e.g. see [35]. The suggestion was that if high P_t fragments remained in the final-state, then they could be mis-construed as the jet arising from a C.C. DIS process. The risk of contaminating neutral current events was not initially considered, since the detection of an electron was all that was necessary to resolve the problem. However, once the task of electron identification in H1 was studied at length, it became clear that significant mis-identification of hadrons/photons as electrons was possible (see chapter 6). Thus, any type of event with a detectable (i.e. not lost in the beam pipe) hadronic jet was a potential source of background. Photoproduction was clearly a matter of concern, because of its large cross-section.

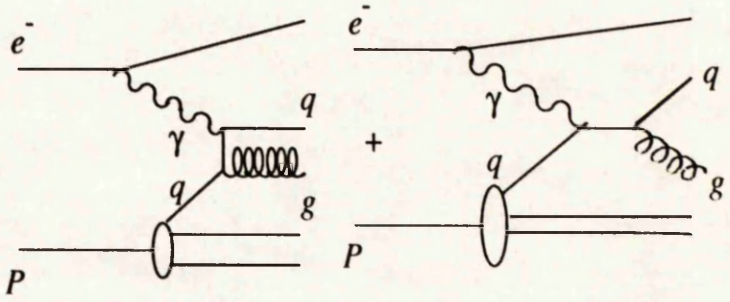
Calculation of the total photoproduction cross-section is a complex issue, and depends on factors such as the photon structure function, the value of P_t minimum, and the treatment of the VMD contribution. A number of estimates have been made of the relative contributions of the hard, soft and diffractive components, based on experimental results conducted at lower energies extrapolated to the HERA range [36]. These lead to a variety of predictions for the total cross-section, presented in [37] and [38], and whilst the various predictions vary by a factor 3, they are all of the order 100 μb. A comparison with N.C. DIS cross-sections of order 100 nb demonstrates the nature of the problem; if only a few percent of the photoproduction events are mis-identified as DIS, they will still completely swamp the sample.

3.2 Other Electron Sources

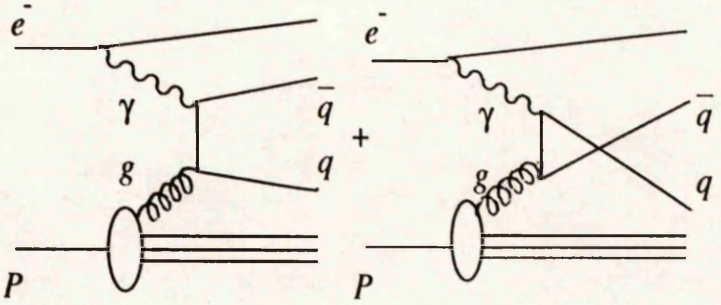
Photoproduction events are assumed to be a background source because of their high-rate production of *hadron* jets which may contain particles such as πs and/or γs, capable of being mistaken for electrons. Another way in which they may cause problems is by producing genuine electrons. This is postulated to occur via the semi-leptonic decay of a charm quark pair, which is itself formed from photon-gluon



(a) Vector Meson Dominance

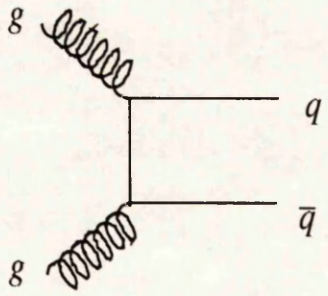


(b) QCD Compton scattering

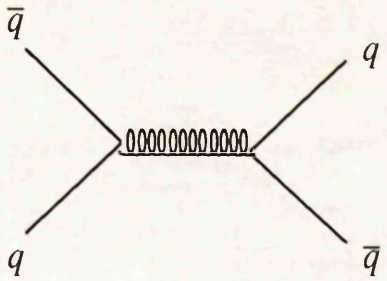


(c) Photon gluon fusion

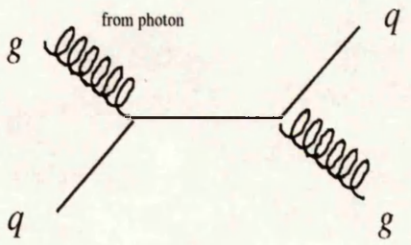
Figure 3.1: Photoproduction processes



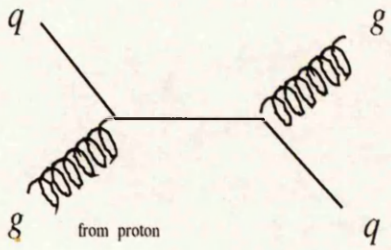
(a) gluon-gluon fusion



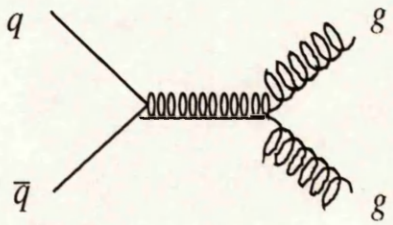
(b) quark-antiquark annihilation in $q\bar{q}$



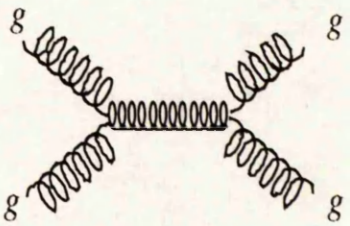
(c) quark-gluon scattering



(d) quark-gluon scattering



(e) quark-antiquark annihilation in gg



(f) gluon-gluon scattering

Figure 3.2: Resolved photoproduction processes

fusion, i.e.

$$\gamma g \rightarrow c\bar{c} \rightarrow e^+e^-X$$

The electron could then be detected and incorrectly thought to be the scattered electron, which in fact was lost in the beam pipe. The kinematics reconstructed from this electron would have no relation with the genuine x, y, Q^2 of the event, and would contaminate the true DIS sample. The detection of these dileptonic decays of heavy quarks has itself been studied with a view to extracting the proton's gluon density (see [39]), and so estimates of the cross-section have been already performed. Van Woudenberg [40] has calculated the dileptonic charmed cross-section to be ~ 20 nb with the requirement that the electron has $P_t \geq 0.8$ GeV/c and $\theta > 5°$. Studies in more detail using the AROMA Monte Carlo for heavy flavour production [41] and LEPTO for DIS events [42] (Monte Carlo descriptions are provided in chapter 5) were performed, and the results quoted in [27], [43] showed that for $y \leq 0.7$ the contamination was negligible. As will be demonstrated in chapter 6 $y \geq 0.7$ is a y region which is currently inaccessible, and in section 7.3.1 a y cut of ~ 0.7 is imposed due to a minimum electron energy requirement. On this basis the open (i.e., the $c, \bar{c}$ quarks do not form a bound state, such as a J/ψ meson) charm background is not considered further.

3.3 Non *ep* Backgrounds

In addition to 'genuine' physics events resulting from electron-proton collisions, there are backgrounds due to the particle beams. There are three sources of this variety, namely

- Synchrotron radiation
- Interactions of beam electrons or protons with residual gas in the beam pipe
- Collisions of beam protons with the beam pipe wall or beam structures such as synchrotron masks.

In order to bring the electron and proton bunches together at the interaction point, the electron beam is deflected through 10 mrad over a distance of 13.6

m. This causes an inevitable flux of synchrotron radiation, which it is the task of masks and collimators to reduce. Early predictions of the size of the synchrotron radiation [44] were 10^8 photons ($E > 20$ KeV) s^{-1} within ± 2.5 m of the interaction point, equivalent to 10 photons per beam crossing. This is not expected to pose any significant problem for drift chambers. The effects of long-term exposure of scintillators to synchroton radiation are not negligible however, and may limit severely a sub-detector's lifetime.

Electron or proton collisions with molecules of gas remaining in the beam pipe produce high multiplicity showers of singly charged hadrons and nuclear fragments. Proton interactions up to 100 m upstream of the crossing point can affect the detector, and the proton-nucleon rate is ~ 9000 $s^{-1}m^{-1}$ at a beam vacuum of 10^{-9} Torr, c.f. ~ 1000 $s^{-1}m^{-1}$ (over a sensitive length of ~ 5 m) for eN collisions. Electron-gas strikes can create electrons with $E_e \geq 10$ GeV at low angles to the beam, a possible N.C. DIS background.

Off-momentum protons striking the beam wall have a larger rate of 3×10^5 $s^{-1}m^{-1}$. Many are removed by scrapers which can be moved into or withdrawn from the proton beam pipe a long way from the detector. The masks installed to reduce synchrotron radiation are, unfortunately, sources of this background, as are the quadruploes in which the pipe diameter is reduced and certain intruding elements of the vacuum system.

Beam *Halo* is a term used to describe the cloud of particles which results from beam wall interactions. The halo concentrates around the beam pipe and fills the whole of the beam tunnel, consisting mainly of muons from the decay of pions produced by the original proton-wall impact. Beam halo muons, travelling at the same speed as the proton bunch and parallel to the beam may constitute a major source of false triggers within the detector.

As an indication of the size of the beam wall/gas background, table 3.1 from [45] shows expected background rates in the BEMC calorimeter as a function of θ, with energy depositions per stack typically of a few GeV.

The shielding walls around H1 protect against a large fraction of beam wall and beam gas noise, but some particles (particularly muons) can penetrate all shielding. The job of removing events which consist of either just beam gas/wall

Polar angle (°)	Rate (kHz)
175-170	77
170-165	59
165-160	21
160-155	9
155-150	4

Table 3.1: Estimated BEMC background rates

derived particles or some overlap of these with a genuine *ep* collision falls to triggers and software selections. The use of the ToF scintillator and level 5 cuts imposed after event reconstruction (see section 7.1) are designed to eliminate the remaining non *ep* background.

3.4 Radiative Corrections

Radiative corrections are distinct from the other backgrounds in that they are not due to some different type of event, but to the higher-order electroweak effects of DIS itself. The major corrections to the basic Born cross-section originate from the emission of photons from either the incoming or scattered electron, illustrated in figure 3.3 (a), (b). These processes are referred to as initial- and final- state radiation, for obvious reasons. Numerous other $\mathcal{O}(\alpha)$ diagrams can be drawn which contribute to the complete set of corrections, but these are found to be small by comparison with the lepton line bremsstrahlung. Kwiatowski [46] states that (along with just the additional contribution from the 'photon self-energy', i.e. a fermion loop inserted into the virtual photon exchanged between the electron and proton), the differential cross-section can be described to an accuracy of better than 1% provided that $Q^2 < 2 \times 10^3$ GeV^2. The leptonic corrections dominate because of two reasons; firstly, a term in $\ln(Q^2/{m_e}^2)$, where m_e is the electron mass, is found to occur in the detailed calculation of the correction (see, e.g. [47], [48]) and secondly; it is easy to understand that if a photon of momentum k is emitted from the (incoming) lepton, then the resulting momentum transfer with the proton is reduced to $Q'^2 = -(l - l' - k)^2$, with $l, (l')$ the initial (final) electron 4- momenta.

This may be much smaller than $Q^2 = -(l - l')^2$, the non-radiative situation, and because of the presence of the $1/Q^2$ factor in the cross -section formulae such as 2.14, the radiative cross-section is enhanced. It is the Born cross-section σ^{Born} which is needed for the information it carries on the proton structure function, and any thorough analysis using DIS will need to unfold this from the measured, radiative form σ^{exp}. The term 'radiative correction' can be expressed as δ_{rad} where

$$\frac{d^2\sigma^{exp}}{dxdy} = (1 + \delta_{rad})\frac{d^2\sigma^{born}}{dxdy}.$$

3.4.1 Simple Radiative Considerations

The kinematic regions for which the radiative corrections are most significant are large y and small x. This can be appreciated by considering the initial-state radiation of a photon by the incoming electron. The ep collision which follows is then the same as the non–radiative one but with an effectively lower incident electron beam energy, i.e. $E_{l'} < E_l$. The Q^2 is degraded as discussed above, and so measurement of the electron energy and angle would imply that the electron was scattered from a lower momentum quark than it in fact was, i.e. there would be a migration from high x to low x. Equivalently, the electron would be viewed as having lost a larger fraction of its momentum in the scatter than was true, and events would be reconstructed at a higher y than was correct. Failure to identify an event as radiative would result in a shift towards the high y, low x corner of the kinematic plane. Implicit here was the notion that the event would be reconstructed from the electron-based measurements. For previous (fixed-target) experiments this has usually been the case, but H1 allows the hadron jet to be observed, as seen in section 2.6. This is important, as it provides a new and independent handle for the treatment of radiative events. Consider the final-state radiation case, shown in 3.3 (b). Here, the momentum transfer is different at the lepton and hadron vertices. At the former it is written ${Q^2}_e = (l - l')$ and at the latter ${Q^2}_h = (p' - p)$, which are not equal since momentum conservation requires $l+p = l'+k+p'$. Using the appropriate value for Q^2 and q, two unequal x values, x_e and x_h can be derived according to 2.3, and similarly for y. This difference between electronic and hadronic kinematics is exhibited by initial-state radiative events also. The expression for y (equation 2.8)

should be compared with its hadronic form

$$y = \frac{E_j}{E_l} \sin^2 \frac{\theta_j}{2}$$

As explained above, for these radiative events where a photon of energy E_γ has been emitted, $E_l \neq 26.7$ GeV (the nominal beam energy) but some lower value $E_{l'} = 26.7 - E_\gamma$. So, if y_e and y_h are calculated as if they were non-radiative, i.e. with the 'wrong' beam energy, then $y_e \neq y_h$, and $x_e \neq x_h$.

In their description of the HERACLES generator, discussed below, Kwiatowski et al. have produced some useful spectra which are worth reproducing here to gain an insight into the physical nature of radiative events. The photon-exchange contribution is split into three factors in their model, corresponding to initial and final state radiation and a third, Compton part. (This separation into three component cases is a result of taking lowest orders only in the QED calculations; in reality, there are not three distinct processes, and the exact calculation makes no distinction between initial and final state radiation.) This last part becomes dominant as $Q^2 \to 0$ and corresponds to the case where a quasi-real photon is emitted from the *quark* line, followed by a Compton-like scatter $e\gamma \to e\gamma$, portrayed in 3.3 (c). It is this part which causes a $1/(1-y)$ term to appear in the differential cross-section, leading to large corrections at high y. Figure 3.4 (a) is the spectrum of E_γ, with a low energy peak of soft photons mainly from final-state events. Compton events are responsible for the peak at ~ 35 GeV. The middle peak at ~ 20 GeV is formed from the overlap of two maxima due the hard, initial state photons; one occurs when $E_\gamma = E_l$, and one when $\theta_\gamma = 0$. Figure 3.4 (b) shows the polar angle θ_γ spectrum of the radiated photon. This was produced over a fairly narrow range at high x ($0.075 \leq x \leq 0.125$), high y ($0.8 \leq y \leq 0.9$) with $E_\gamma \geq 0.5$ GeV, i.e. hard photons. The peak at $\theta_\gamma = 0$ comes from the initial state events, where the photon is emitted parallel to the incoming electron and passes down the beam pipe. The width of this peak is found to be $\sim \sqrt{\frac{m_e}{E_l}}$, with E_l the incoming (pre-radiation) electron energy of (here) 30.0 GeV. In the given kinematic range the electron scatters at angles θ_e between 142° and 160°, and so those photons emitted from the final-state electron are found with this same angular spread, explaining the high angle peak. For a unique scattered electron energy E_e the final-state peak has a width $\sim \sqrt{\frac{m_e}{E_e}}$. The broader mid-range peak in the backwards direction comes from the Compton events. The final graph is of the difference in azimuthal angle ϕ between the emitted photon

and scattered electron. The peak at zero is from the final-state radiative events, collinear with the scattered electron, the broad spread is from initial state events where the photon has a small θ value and essentially random ϕ distribution, and Compton events account for the 180° peak where the electron and photon balance each other in P_t.

The size, and importance, of radiative corrections is made plain from figure 3.5, which portrays the leptonic corrections for a variety of x values as a function of y at fixed $s = 10^5$ GeV. The quantity δ is defined as

$$\delta = \frac{d^2\sigma/dxdy^{\mathcal{O}(\alpha)}}{d^2\sigma/dxdy^{Born}} - 1$$

and the Duke-Owens [49] parton distributions have been used. The trend for large corrections at high y and small x is clear.

3.4.2 Removing Radiative Events

The tagging and removal of radiative events is a vital step in any full DIS analysis. The details of this are omitted here, with a fuller review to be found in [50].

Final-state radiative events constitute less of a problem because, as illustrated by the angular spectra, most of the radiated photons are close to the scattered electron in both θ and ϕ. The granularity of the calorimeters ensures that the majority of these final-state radiation events pose no problem, since the photon and electron energies are then automatically combined in the calorimeter cell, and it is as if no radiation from the electron took place.

If the (initial- or final-state) radiated photon is at a reasonable angle, within the acceptance range of the calorimeters, then its identification and energy/angle measurements can lead directly to the tagging and reconstruction of the DIS event. Problems arise when the photon is undetectable, either because (for initial-state) it is lost down the beam pipe or because it is lost (for example) in amongst the hadron jet debris.

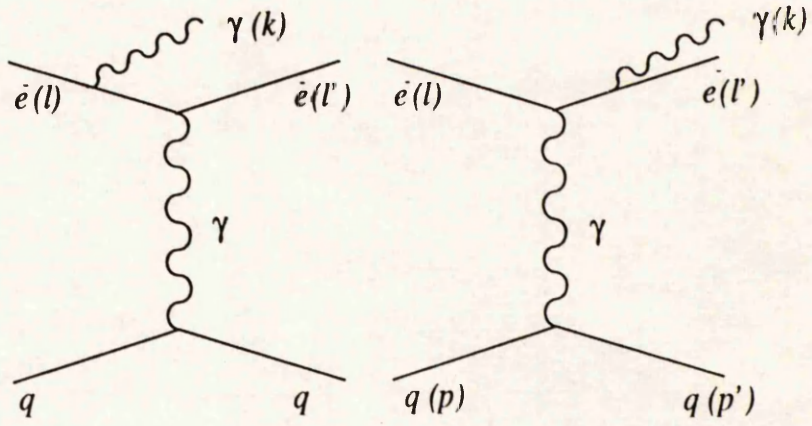
The luminosity detector (see 4.13) is designed to tag photons and electrons from the elastic $ep \rightarrow ep + \gamma$ process. Mounted at the rear of the H1 detector it

has an angular acceptance of $\pi - 0.005 \leq \theta_\gamma \leq \pi$, and is thus capable of detecting a large fraction of initial-state photons (c.f. 3.4 (b)). This tagging may be confused by random coincidences of a non-radiative DIS event with an elastic scatter. This is likely to be true only at higher luminosities than have been currently achieved (over the summer, a luminosity of 7.5×10^{27} cm^{-2} s^{-1} was achieved compared to a nominal value of 1.5×10^{31} cm^{-2} s^{-1}.). The luminosity monitor has an electron detector which is sensitive to electron energies between 12 and 25 GeV, and so only some of the elastic scattering events could be separated from radiative, inelastic events of interest. The presence of a photon in the luminosity monitor's photon tagger with *no* electron in the electron tagger does not necessarily identify an event as radiative DIS; it could be an elastic scatter with the electron being outside the tagger's energy acceptance window.

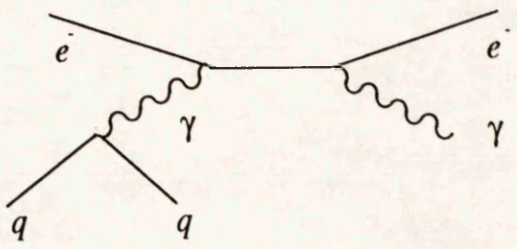
The second approach, if the photon is undetectable, is to use the difference between the hadron and electron x or y values. By requiring that they agree to within some range, the radiative events could again be isolated. Moreover, by substituting for E_e between hadronic and electronic formulae, the actual energy of the initial electron can be recovered; the event is then obviously radiative if this is very different from the electron beam energy.

3.4.3 Significance for this Analysis

This summary of the origin, characteristics and identification of radiative events has aimed to provide a general view of the topic within the wider context of DIS physics. With respect to this first-day analysis, the details of how to extract the Born cross-section from the measured one are not relevant. As explained in section 2.5, the aim here is to obtain an indication on the proton structure function by examining the low E_e tail of the electron energy spectrum, and comparing it to different Monte-Carlo predictions. It has now been shown that radiative corrections are large at small x, exactly the region which in which structure function sensitivity manifests itself. Therefore, it is possible to proceed only if a Monte Carlo event generator is employed which accurately predicts radiative effects. Only then is it possible to assume, within limits, that any observed difference between data and simulation is due to the adopted F_2 parametrization and not because of inaccuracies

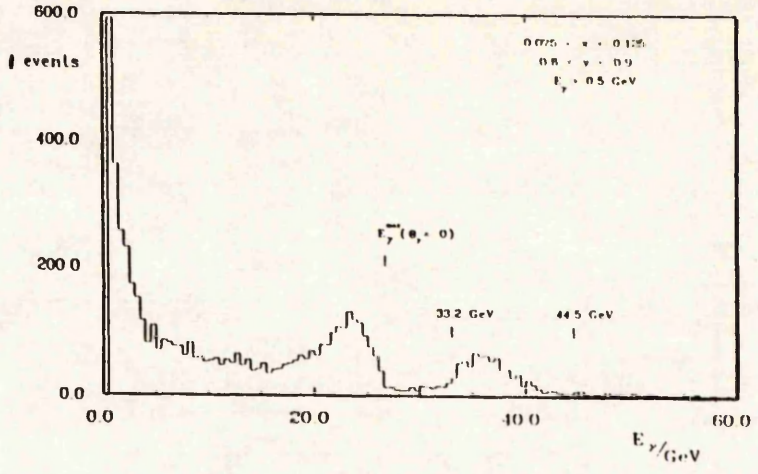


(a) Initial state radiation (b) Final state radiation

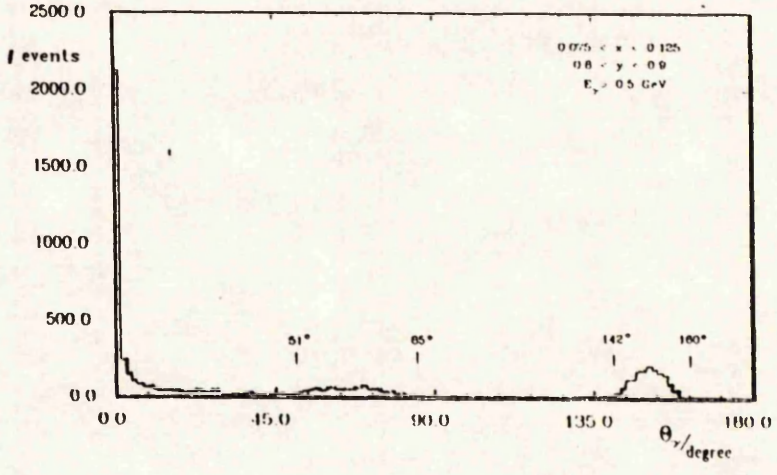


(c) Compton scattering

Figure 3.3: Some first-order radiative processes.



(a) Energy of radiative photon.



(b) Polar angle of radiative photon; (c) (below) Difference in azimuthal angle between photon and electron.

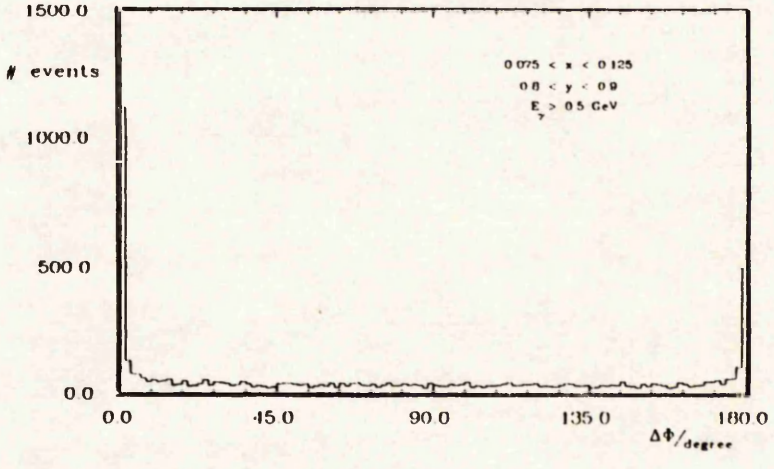


Figure 3.4:

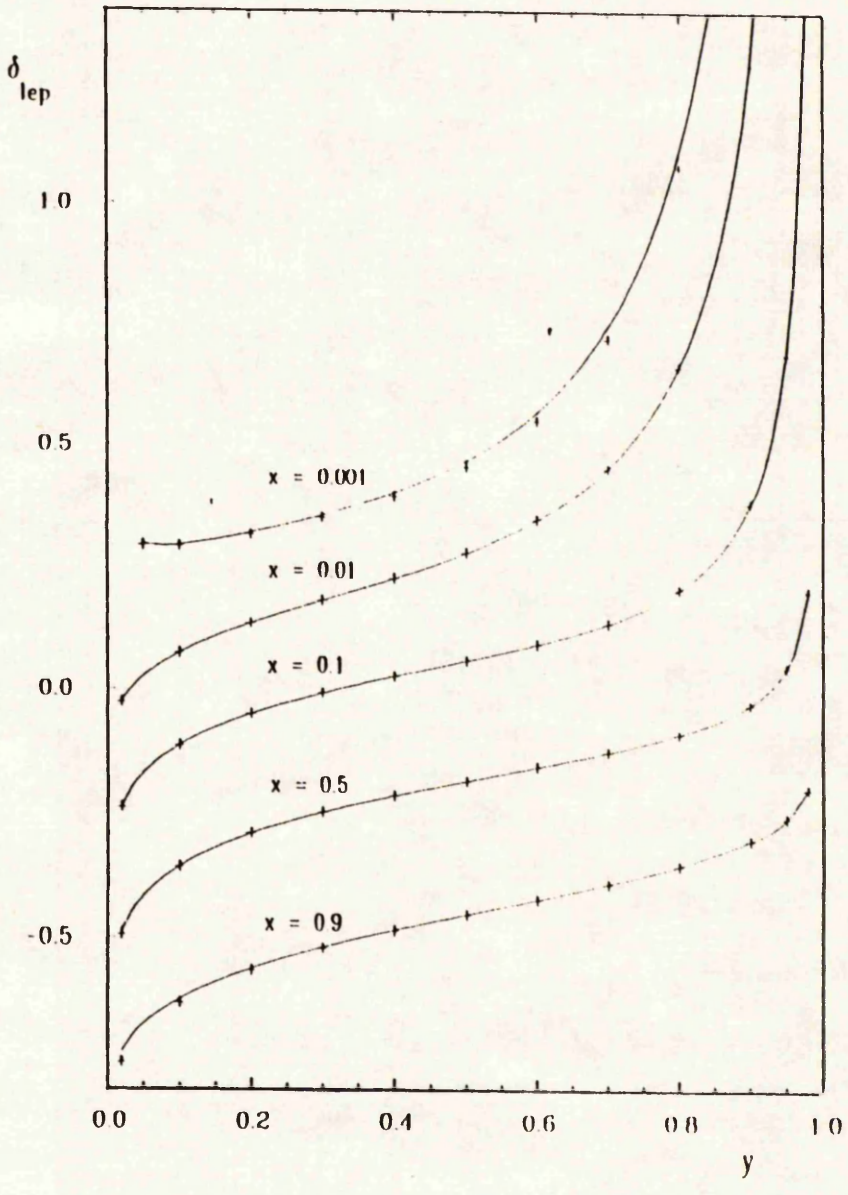


Figure 3.5: Predicted radiative corrections

in the modelled radiative corrections.

The generator used here was DJANGO [51]. This is in fact a general interface which can be used for a number of Monte Carlos, but it is synonymous within H1 with the true generator at its heart, HERACLES [52].

HERACLES includes the leptonic corrections, quark-line radiation, and the complete one-loop virtual corrections, i.e. a comprehensive set of corrections; it is limited to the range $y(1-x^2) \geq 0.004$, and the quoted accuracy at $Q^2 \leq 2000$ GeV^2 is 1%. In reference [46] this statement is enlarged upon, and it is reported that HERACLES is reliable in the approximate range $10^{-3} \leq x \leq 0.9$, $0.02 \leq y \leq 0.95$, producing output which agrees to within 3% of results obtained by Spiesberger [48]. The low photon energy cut-off is chosen at $\mathcal{O}(10)$ MeV, i.e. below a detectable value.

The predicted accuracy of HERACLES is made more believable upon the realisation that, for the same reasons as outlined in section 3.2, an upper y limit of $y \sim 0.7$ is deployed in this work. Returning to figure 3.5, this shows that the radiative corrections are ≤ 0.5 for most values of x and y, becoming large only in the region $y > 0.7$, $x \sim 10^{-3}$. This is unlikely to change by more than a factor of 2 for the MT or KMRS structure function sets. HERACLES is therefore accurate at the single percent level in its calculations of a correction that is almost always less than 50%. When the low level of statistics used in this analysis is introduced in subsequent chapters, it becomes obvious that any errors introduced by the radiative corrections are unimportant by comparison.

It seems certain then that although radiative corrections are in general never negligible, occasionally huge, and always important for DIS physics, the use of HERACLES/DJANGO and the avoidance of high y events will enable a valid comparison between data and Monte Carlos to be performed.

Chapter 4

The H1 Detector

4.1 Overview

In this section a brief description of the main elements of the H1 detector is provided. As for most modern high energy physics detectors, the H1 experimental apparatus was built to be a general-purpose device, with a mixture of calorimeters and tracking devices. The main design aims were to provide the following features.

- High hermiticity, with complete coverage over the solid angle. Only the carbon-fibre beam pipe (with inner and outer radii of 9.2 and 9.4 cm) spoils this.
- Calorimeters to study energy flows. There is an obvious need to measure jet energies in any (e.g. C.C. or N.C. DIS) reaction where they are produced. For neutral particles, this must be provided by calorimeters, since momentum measurements from trackers will be unavailable. Calorimeters' energy resolutions improve with increasing energy for both charged and neutral particles, whereas that of trackers decreases, as higher momentum tracks bend less in a given magnetic field. In addition, neutrinos can only be inferred from missing energy, so the total final-state energy must be found.
- Trackers to detect particles, help resolve jet topologies and study charged track kinematics.

- Muon identification and tracking must be provided for. Muons are unlikely to be contained in the calorimeters, and are an important indicator of heavy-flavour physics and postulated new processes.
- A magnetic field to enable the trackers to make momentum and charge sign measurements from the radius and direction of curvature of tracks.
- An overall design that takes into account the highly asymmetric lab-frame kinematics present at HERA. With the imbalance in incident electron and proton energies, many particles will be boosted into the very forward direction. A greater depth of calorimeters and devices capable of tracking dense jets of particles are needed at low angles.

The H1 detector is shown schematically in figure 4.1, with its main elements (sub-detectors) labelled. The co-ordinate system used for the whole of this work is also illustrated. The z axis is taken as positive in the proton beam direction (always called the *forward* direction), and negative in the electron (*backwards*) direction, the nominal interaction point being the origin. The polar angle θ is defined with $\theta = 0$ along the positive z axis. This chapter has been sub-divided into individual sections for each of the main sub-detectors, with special consideration paid to those components in the backward region which play a dominant role for this DIS analysis.

4.2 Calorimetry

The passage of an electromagnetic particle through some material will cause that particle to lose energy via the familiar process of bremsstrahlung. The high-energy photons thus liberated undergo pair-production, and the electrons/positrons in turn will suffer bremsstrahlung. The spray of electrons and photons is referred to as an electromagnetic shower, and the showering propagates until the particles' energy drops below $\sim$ 10 MeV. Below this energy, ionization and excitation processes become increasingly dominant. The shower development and energy loss are best characterized by the *radiation length* parameter, X_{rad}, the distance it takes for an electron's energy to drop to $1/e$ of its incident value. This is conventionally

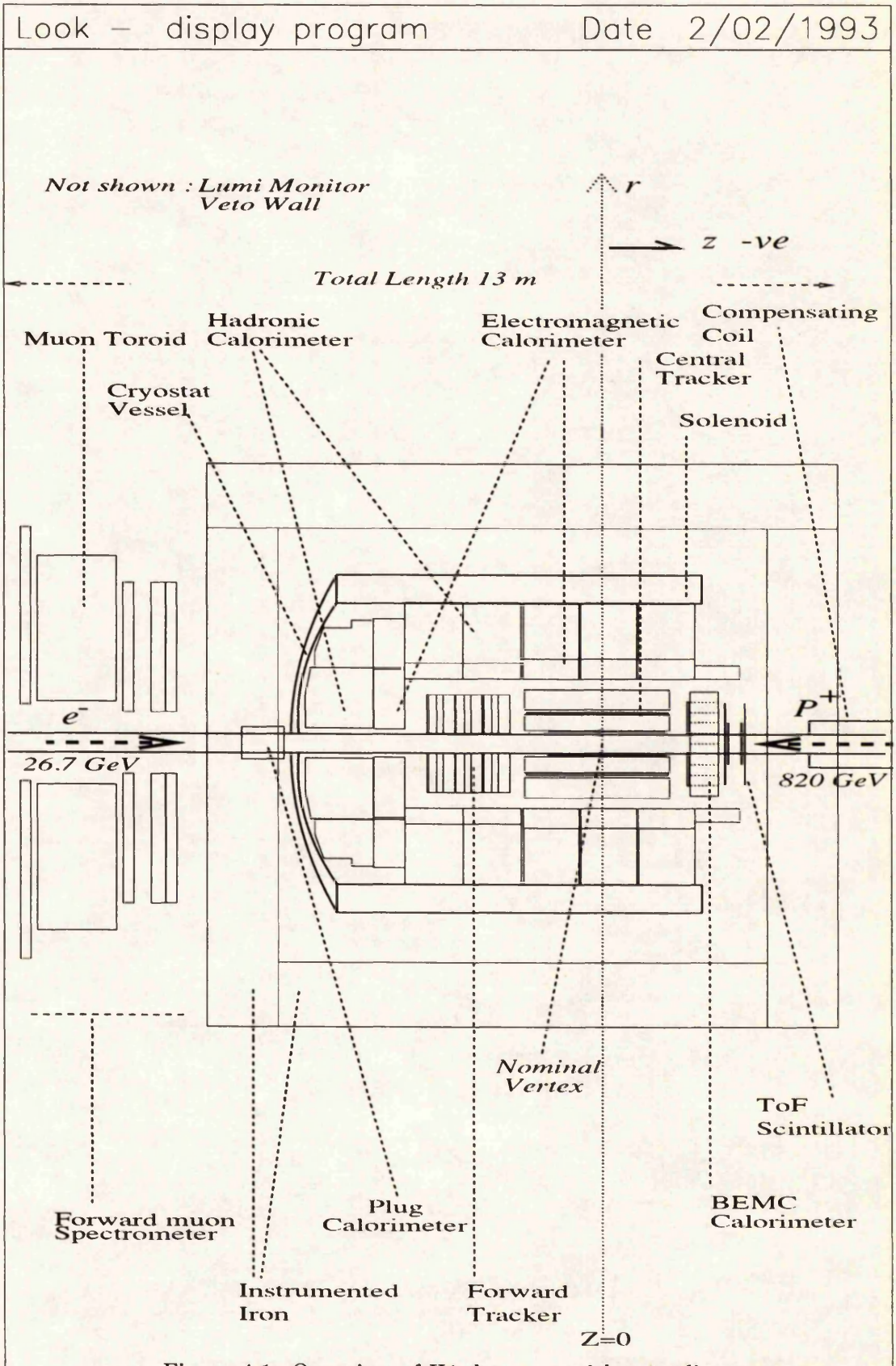


Figure 4.1: Overview of H1 detector, with co-ordinate system.

expressed in units of g cm^{-2}. X_{rad} is approximately proportional to A/Z^2 where A, Z are the atomic mass and atomic numbers of the material being traversed. For carbon, X_{rad} = 63.05 g cm^{-2}, for lead X_{rad} = 6.37 g cm^{-2}. Converting X_{rad} to an absorption length in cm shows that a dense material is desirable in order to make a physically compact detector. There exists a well-established model of electromagnetic interactions, described for example in [53], and experiments show that a shower can be effectively contained within a length of $\approx 20X_{rad}$ = 11.2 cm for lead. Crucially, the total ionization produced, (or the total charged track length, since the two are proportional), provides a measure of the incident particle's energy.

There is a great deal of difference between the formation and development of electromagnetic and hadronic showers. There are several energy loss mechanisms which take place this time. An incident proton, say, will collide with the absorber's nuclei and liberate secondary, fast protons, pions, kaons. The π^0's decay electromagnetically to photons (and thence to electrons). The secondary protons and other charged particles will in turn scatter off nuclei, and also cause ionization as they strip electrons from atoms. In fact, in decreasing order of importance, the dominant losses are a) ionization of secondary protons and charged pions b) electromagnetic cascade from π^0's and c) nuclear binding energy. Together these account for $\sim$ 70% of the energy deposited. The remainder is accounted for by the kaons and neutrons liberated from nuclei, ionization caused by the initiating proton etc. Hadronic showers are characterised by the nuclear interaction length λ (g cm^{-2}), which is $\mathcal{O}(10)$ times the magnitude of X_{rad} for a high Z material such as iron or lead. Approximately 5 interaction lengths are needed to contain hadronic showers up to $\sim$ 40 GeV according to [53]. Dividing by the absorber's density to convert to real distances shows that hadronic showers develop over longer distances than electromagnetic ones, and so hadronic calorimeters must be deeper than their electromagnetic counterparts. Whereas X_{rad} drops with rising Z, λ rises (though more slowly). A high Z absorber is still desirable though, since the more rapid rise in density ensures that substances like iron and lead can contain showers in smaller volumes. For example, for iron λ = 131.9 g cm^{-2} $\equiv$ 16.7 cm. The conversion of measured ionization to an energy value is more complicated in the hadronic case. About 20% of the energy deposited by a proton in a liquid argon/iron calorimeter is converted to nuclear binding energy and the production of neutrinos. This is largely undetectable, and so some scale factor has to be applied to the measured energy (or rather, to the hadronic

fraction of the measured energy) to account for this loss. (One way of making a *compensating* calorimeter is to use a uranium absorber; fast neutrons present in the shower induce fission in the ^{238}U, and the resulting fission energy in the form of photons is detectable). The separation of the pion-induced electromagnetic part of the hadron showers is achieved, in H1, by having a fine longitudinal segmentation of the calorimeter. The fact that X_{rad} is much smaller than λ means that electromagnetic interactions show up as large depositions over small distances. Once identified, a different scaling is applied to these showers to convert the measured ionization into an energy, which is then summed with the true hadronic part. This is performed by on-line software described in [54].

4.3 The Liquid Argon (Lar) Calorimeter

The H1 detector uses a large liquid argon calorimeter supplemented by an instrumented iron tail catcher, with separate plug (very forward) and BEMC (backward) calorimeters which work on different principles. Showers of particles initiated by dense metal plates traverse a liquid argon filled gap, and cause ionization. The charges thus generated drift in an electric field (650 V mm^{-1}) applied across the gap, and the read-out signal is formed by electrons drifting towards anode pads. The inner layer of the calorimeter is the electromagnetic part, with lead absorber plates, the outer layers constitute the hadronic part with stainless steel absorbers.

The liquid argon/metal absorber approach was chosen because of the good long-term stability and ease of calibration, and homogenous response offered [44]. High granularity and longitudinal segmentation are also possible. This allows shower profiles to be studied to enable the different electromagnetic/hadronic weightings to be applied. The segmentation of a calorimeter chosen represents a compromise between the needs to minimize the dead material and limit the number of read-out electronic channels, and the need for high granularity for e/π separation and energy flow measurements. The Lar calorimeter covers the polar angle range $4° < \theta < 152°$. This range is sub-divided into forward ($4° < \theta < 20°$) and central ($20° < \theta < 152°$) regions. All the calorimeter modules and liquid argon are contained within a large steel cryostat vessel connected to a storage dewer and refrigeration system.

The basic calorimeter unit is the *stack*. The exact design of a stack depends upon its location (forward, central) due to the structural limits imposed by the need to support large metal plates and geometrical considerations. The calorimeter is divided along the z (beam) axis into 8 'wheels' each of which is then divided in ϕ into 8 stacks. Figure 4.2 (b) shows ϕ segmentation of two of those wheels. The stacks are located so that the cracks between them point to the beam line in the case of the electromagnetic part, whereas the hadronic cracks are offset at an angle of $\approx 20°$. In this way, no hadron can avoid passing through some hadronic stacks, and any electron which passes completely along a crack will be detected in the first layer of the hadronic calorimeter. A typical electromagnetic stack is a pile of cells, each cell being composed of a read-out board/lead plate sandwich (the lead being 2.4 mm thick) with a central Lar-filled gap 2.35 mm thick (i.e. the sequence is lead-Lar-lead). The hadronic cell sequence is 19 mm thick stainless steel plate, read-out board, 2.3 mm of Lar, read-out board, 2.3 mm of Lar, read-out board, steel plate. In the central part of the detector an electromagnetic stack is formed from 43 cells, a hadronic stack from 48.

The absorber plates are arranged so that the incidence of particles to them is as close as possible to 90°, and never less than 45°, because the performance is optimum at normal incidence. Thus, the plates are oriented vertically in the forward direction and horizontally in the central part.

The total thickness of the Lar calorimeter is 20-30 X_{rad} and 6-8 λ [55]. Isolines of X_{rad} and Λ are drawn on figure 4.2 (a). The calorimeter was designed with an increased depth in the forward direction to cope with the strong forward-boosting of events at HERA.

The read-out pads are independent of the calorimeter stacks themselves, i.e. many anode read-outs are distributed over one stack. According to [44] a minimum of three longitudinal segments is required for effective e/π discrimination in the electromagnetic calorimeter. In the hadronic part, one segment per λ depth is sufficient to allow the software compensation method to work. The size of the pads varies in θ, larger pads are used at higher angles where particles are better separated. The basic electromagnetic pad dimension is $\approx 2\times$ the Molière radius (4.3 cm), the size characteristic of an electromagnetic shower's lateral spread. The basic hadronic pad size is double this. The total number of read-out channels is

$\sim 60,000$. Sets of pads are associated together by the read-out electronics to form 'towers', projecting towards the axis in ϕ, 'pseudo-projecting' in θ. The longitudinal and transverse segmentation of some of these towers is shown in figure 4.2 (c). This illustrates their pseudo-projecting, i.e. the way in which they point towards the ill-defined interaction point.

Tests conducted at CERN, in which electron (3.7 - 80 GeV) and pion (3.7-170 GeV) beams were fired into single stacks showed that the resolutions could be parametrized by the form $\sigma_E{}^2/E^2 = (A^2/E) + B^2/E^2 + C^2$. For the electromagnetic calorimeter, the results yielded were

$$A = (11.16 \pm 0.05)\%\sqrt{\text{GeV}} \quad B = 152 \pm 4 \text{ MeV} \quad C = (0.64 \pm 0.07)\%$$

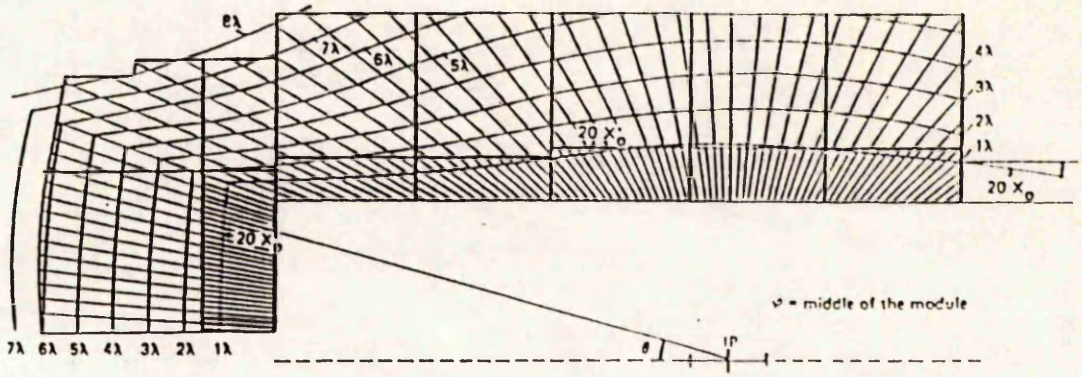
and for the hadronic calorimeter,

$$A = (46.1 \pm 0.7)\%\sqrt{\text{GeV}} \quad B = 730 \pm 30 \text{ MeV} \quad C = (2.6 \pm 0.2)\%$$

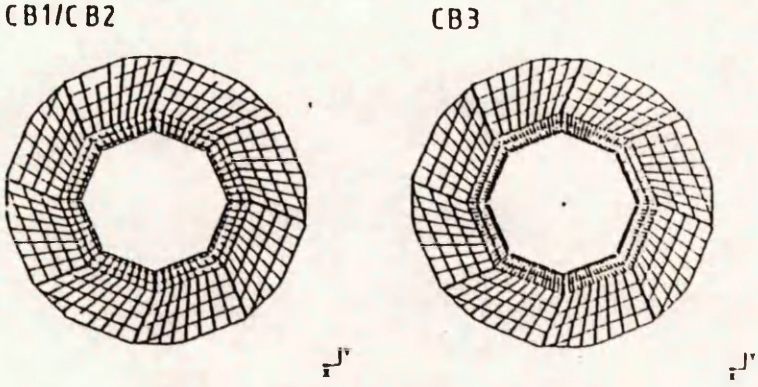
(these results from [55]). The e/π rejection ratio was found to be $\leq 10^{-3}$. Particles which penetrate through the entire Lar calorimeter have their energies measured by the tail catcher calorimeter situated in the instrumented iron, described below.

4.4 The Plug Calorimeter

The plug calorimeter covers the angular region in the very forward direction $0.7° \leq 4°$, covering that part of the solid angle range obscured from the Lar by the beam-pipe. A silicon-instrumented sandwich design was chosen for compactness, with eight 75 mm copper absorber plates and silicon-pad read-out boards, with a total interaction depth of 4.3λ. The calorimeter's function is to provide information on the transverse momentum of forward travelling particles, mainly the proton beam fragments, which would otherwise be missing. This requires a good angular resolution and moderate energy resolution. The expected performance of the plug is an energy resolution of $\sigma_E/E = 110\%/\sqrt{E}$ and an angular resolution of 4 mrad [44]. Susceptibility of the plug to radiation damage (from particle impact and synchrotron radiation) is an important and as yet not fully understood factor.



(a) Lar calorimeter, iso-lines of λ and X_0.



(b) Lar calorimeter wheels and stacks; (c) (below) Lar calorimeter projecting tower structures.

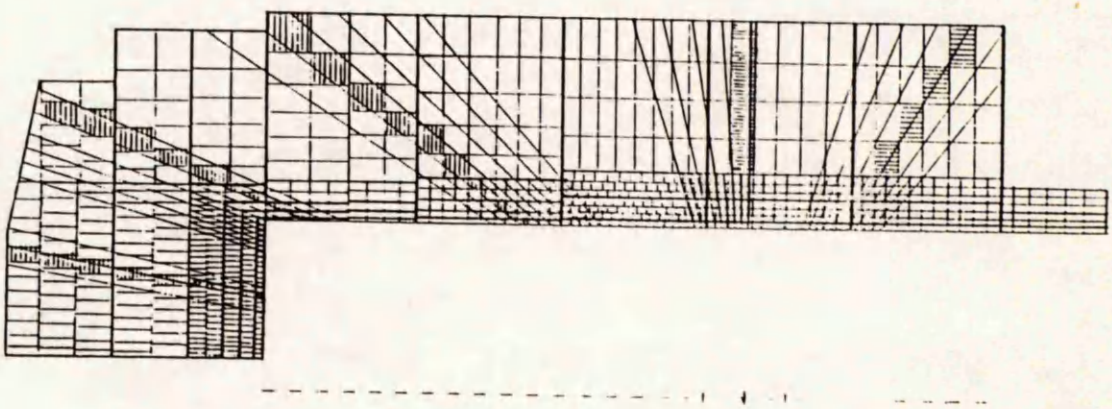


Figure 4.2:

4.5 The Backward Electromagnetic Calorimeter

The backward electromagnetic calorimeter (BEMC) completes the solid-angle range, covering the angles $151.4° - 176.5°$, i.e. up to the beam pipe limit. It consists of 88 lead-scintillator stacks with their long axes parallel to the beam line, mounted in an aluminium cylinder. A calorimeter of this type was used rather than extending the Lar apparatus, despite the poorer stability and homogeneity expected, for reasons of economy, ease of installation and compactness. There existed the need to insert and withdraw the central trackers, which would not have been an easy task if a large Lar calorimeter/ cryostat had to be moved first.

The calorimeter's primary aim is to measure the energy of electrons scattered in the backwards direction by low Q^2 neutral current DIS events. A secondary aim is to provide an energy measurement of hadrons produced by photoproduction and low x hadronic final states. Good energy resolution for electrons and good e/π separation are therefore the major requirements.

The BEMC is sited 144 cm behind the nominal interaction point, with an outer diameter of 162 cm and a total thickness of 43.9 cm.

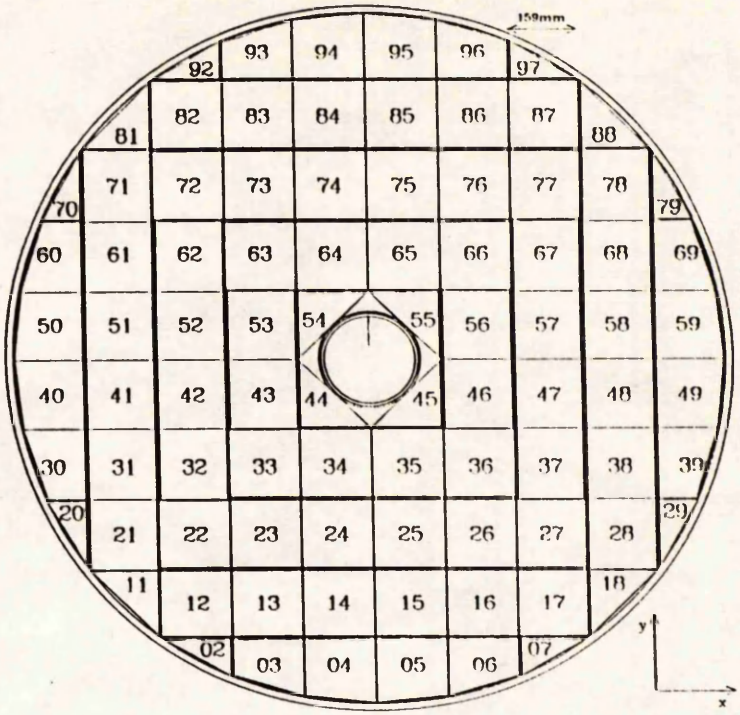
The individual stacks are made of 50 sampling layers of lead-scintillator sandwich, with a total sensitive depth of 22.5 $X_{rad} \equiv 0.97\lambda$ for normally incident particles. This low value means that the BEMC is effectively transparent to hadrons, which will pass through to be tagged in the iron tail catcher (q.v.). Kuhn [56] reports that 5 GeV electrons have a shower depth of 5-6 X_{rad}, 30 GeV electron showers are contained in 6-7 X_{rad}. Each sampling layer is formed from 2.5 mm lead layer and a 4 mm scintillator layer. A face-on view of the BEMC is shown in figure 4.3 (a) which shows how the 88 stacks are arranged to approximate a circle. Apparent from this is the number of different stack shapes which exist, 8 in total, summarised in figure 4.3 (b). The commonest is the simple quadratic stack, 15.9×15.9 cm^2. The scintillating light is collected by wavelength–shifter (WLS) bars 3 mm thick arranged around each stack. These pipe the light to photodiodes mounted at the rear of the BEMC along with the preamplifiers and associated electronics. The WLS bars are of two types. *Long* bars cover the full stack length, *short* bars cover just the last 15 sampling layers. Their location is shown in figure 4.3 (b), taken from [45] WLS types 1-4 being long, types 5 and 6 are short. It will be observed that the short bars

are twice the width of the long ones, and mounted only on the quadratic and big trapezoid stacks. Each long WLS is connected to one photodiode, each short WLS to two. The total number of channels is 472 = 344 (long) + 128 (short).

The motivation for the large number of WLS bars was to improve the spatial resolution of the detector and provide information on shower development to assist in particle identification. In particular, it was intended that a comparison of the signals from the long and short bars would help separate hadronic signals from electromagnetic ones. Unfortunately, this separating power was found in practice to be extremely limited.

4.5.1 Calibrating the BEMC

The fact that lead-scintillator calorimeters were known to offer a less homogeneous response than Lar devices was appreciated from the inception of the design. It was planned to compensate for this by performing an in-situ calibration using the kinematic peak electrons from DIS. As discussed in the theory chapter, the DIS spectrum has an almost mono-energetic peak at the incident beam energy which is highly suitable for this task. With an integrated luminosity of 100 pb^{-1} it is reported [57] that a stack-to-stack intercalibration can be performed which will cross-calibrate the stacks to within 1%. Using a test beam facility at DESY, each stack was calibrated with 1-5 GeV electrons, and measurements made of the stability of the calibration over time. Electron (5-30 GeV) and muon (120 GeV) beam tests made at the CERN SPS were also performed. Based on these measurements, the electromagnetic resolution due to sampling fluctuations was determined to be $\sigma_E/E = 10 \pm 4\%/\sqrt{E}$ [57]. Other terms in the resolution are discussed in detail in section 7.4.1. As will be shown in the analysis chapters, the calibration of the BEMC with the present, low integrated luminosity remains uncertain, with an effective peak resolution of just $\sigma_E/E = 24\%/\sqrt{E}$. Differently shaped stacks are known to have different performances. Tests performed on the photodiodes by illuminating the last scintillator plane of a number of stacks showed that variation in performance was almost entirely a function of stack shape. The charge-energy conversion errors are also known to vary with shape, being a maximum for some of the trapezoid shapes (12%, cf 1% for quadratic stacks). In summary, the exact calibration is subject to



(a) BEMC viewed from interaction point with numbered stacks; (b) (below) stack shapes and WLS positions.

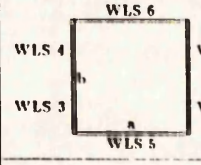
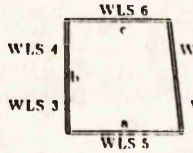
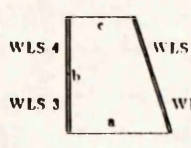
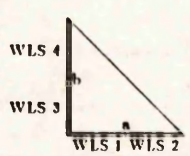
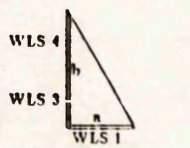
shape of stack/type	cross section	dimensions (mm)
quadratic-stack Type A 1-61		a= 149.8 b= 156.8 (149.8) stack dimensions: A = 158.9 B = 158.9
big trapezoid-stack Type B1, B2 101-105, 201-205		a= 149.6 (149.3) b= 156.8 (149.8) c = 133.7 (133.8) stack dimensions: A = 158.9 B = 158.9 C = 142.4
small trapezoid-stack Type C1, C2 301-305, 401-405		a=133.2 b=156.8 c=83.5 stack dimensions: A=142.3 B=158.9 C=92.1
big triangle-stack Type E 501-512		a=118.4 b=148.4 stack dimensions: A=158.1 B=158.1
small triangle-stack Type D1, D2 601-605, 701-705		a=84.2 b=144.6 stack dimensions: A=91.9 B=157.9

Figure 4.3:

many differing and poorly understood factors, and is currently known only at the $3-5\%$ level [55].

4.5.2 The BSET Trigger

Expected beam-wall and beam-gas rates are $\approx$ 40 kHz and $\approx$ 70 kHz [58] compared to DIS rates $\mathcal{O}(1)$ Hz. The first-level (L1) trigger designed to accept DIS events in the BEMC is the 'BEMC Single Electron Trigger,' BSET. The basic philosophy of BSET is to find isolated high energy deposits in the BEMC stacks. A set of stacks which have a deposited energy above some threshold is called a cluster. Being a first-level trigger, BSET is a hardware device. It takes as its inputs the summed signals of all the long WLS channels of each stack. (Bearing in mind what was said about the difference in electromagnetic / hadronic shower development in 4.2, it is clear that the short stacks are designed mainly to tag the more extended hadronic showers, and so it is reasonable that a trigger looking for electrons should use the long WLS signals.)

The BSET operation and electronics is detailed in [58], and is summarised here. For any stack a low and high energy threshold can be set, and up to four conditions (C1-C4) can be satisfied by an event.

- C1: number of low energy clusters. Clusters are formed from those stacks with more than the low threshold energy (2 - 3 σ above noise, typically 0.8 - 1.2 GeV), i.e. $E > Lt$. They are found by looking at the (x, y) projections of the $E > Lt$ stacks and comparing these with allowed patterns held in memory look-up-tables. Clusters formed with the innermost triangle stacks are treated separately, and may be de-selected from the total if desired.

- C2: number of high energy clusters. This is as C1 but looking at stacks whose energy exceeds the high threshold i.e. $E > Ht$, set in the range 3-15 GeV. C1 (C2) is set to 'KEEP' if any there are between 1 and 8 low (high) threshold clusters.

- C3: Energy sums. The energy of each high threshold cluster is found, if any exist according to C2. The cluster energy is the sum of the energies of any

stacks with $E > Lt$ surrounding a *seed* stack with $E > Ht$, and the seed stack's energy. If there is more than 1 high threshold cluster, the sum of the cluster energies is formed. Three thresholds exist for the setting of three C3 bits to 'KEEP'.

- C4: total BEMC energy. The sum of all stacks with $E > Lt$ is formed. There is one threshold total for C4 to be set to keep.

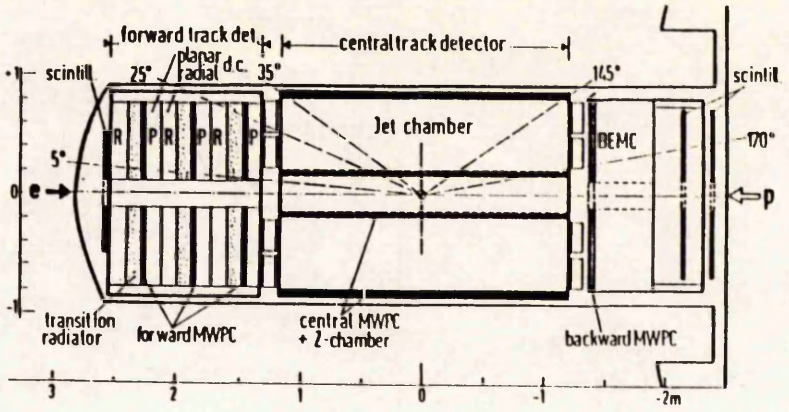
The C1-C4 output bits can then be combined in the H1 central trigger logic. All the events which were eventually reconstructed and classed by the ECLASS routine as DIS candidates (see section 7.1) had the BSET trigger set, as an essential first requirement for the presence of an electron located in the BEMC.

4.6 Tracking

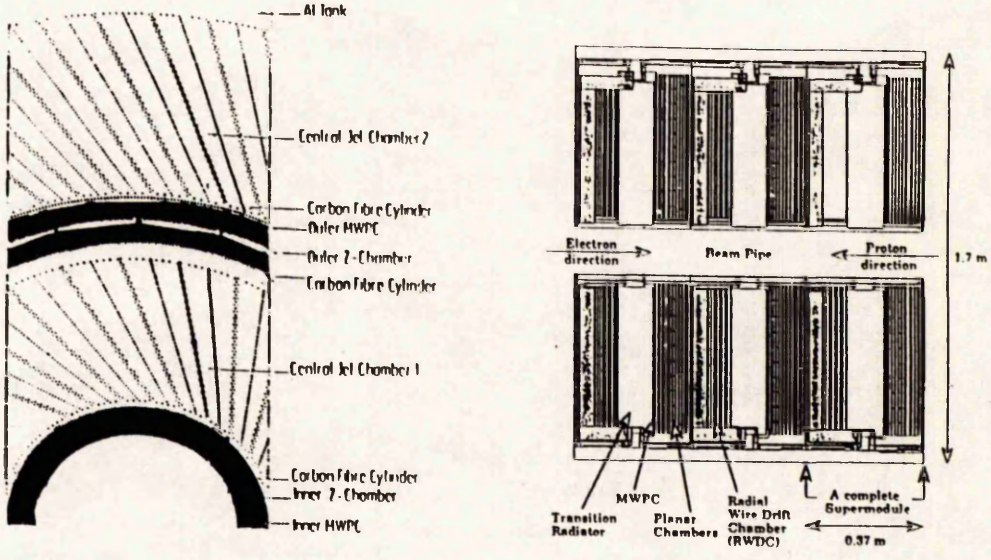
The aims of the H1 tracking devices are the normal ones of providing momentum and angle measurements for multiple tracks, assisting in particle identification via e/π discrimination, and the provision of fast triggers. A side view of the tracking system is shown in figure 4.4 (a), mounted in the centre of the U-shaped cryostat. It can be divided into the central ($25° < \theta < 160°$), forward ($5° < \theta < 25°$) and backward trackers. Like the calorimeter, it reflects the asymmetry of HERA events by having an extensive array of forward devices, with just a single multi-wire proportional chamber (MWPC) at the rear. All the individual chambers have independent gas systems. The physics of drift and multi-wire chambers can be found in many standard texts (e.g. [53]), and since tracking is much less vital for this early DIS study than calorimetry, it will not be repeated in this chapter.

4.7 The Central Trackers

The Central Tracking Detector (CTD) consists of six cylindrical, co-axial volumes located around the beam pipe. Proceeding outwards from the interaction point, a particle traverses the inner MWPC (CIP), inner z-chamber (CIZ), inner



(a) Central and forward tracking devices.



(b) Cross-section through CTD; (c) Side view of Forward Trackers; (d) (below) Forward muon spectometer.

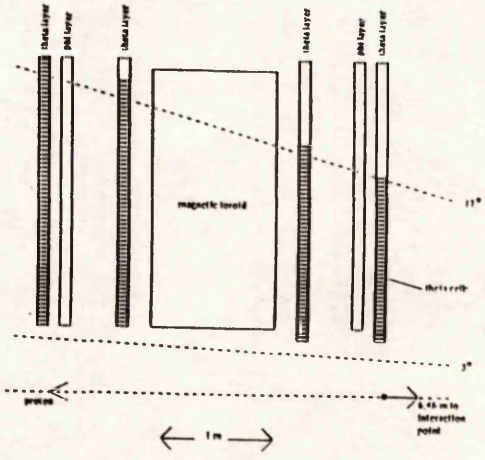


Figure 4.4:

jet chamber (CJC1), outer z-chamber (COZ), outer MWPC (COP) and finally the outer jet chamber (CJC2). The jet-chambers are designed to provide an $r - \phi$ measurement, the z-chambers determine the z co-ordinate and hence θ. The CTD was designed to reconstruct jets with high particle densities.

4.7.1 Central Jet Chambers

Figure 4.4 (b) shows a cross-section through the CTD. The jet chambers are built of cells running along the chambers' full length, parallel to the axis. Each cell contains a plane of parallel sense wires, separated by potential wires. The sense wires are staggered by ± 150 μm to help resolve left/right ambiguities. A cell is separated from its neighbours by planes of cathode wires, and field shaping wires run along the inner and outer radii of the cell. The cells are tilted by angles of $\approx 30°$; this is so that the combined effect of the electric field in the cell and the magnetic field provided by the H1 solenoid forces secondary (ionization) electrons from stiff (high momentum) tracks to drift at $\approx 90°$ to the track. This optimizes the resolution. The tilt also ensures that a stiff track crosses the sense wire plane at least once in both jet chambers, and will pass through several cells. In this way, tracks are reconstructed by combining segments from different cells; wrong reconstructions will not point to the vertex. The sense wires are read out at both ends with the $r - \phi$ measurement being derived from the drift times, but charge division also provides a z value $\sigma_z = 25$ mm to supplement the z-chambers' results. Collecting the charge also makes possible a dE/dx measurement (10% resolution) since the ionization produced by a particle is proportional to its energy loss. This helps distinguish, e.g. electrons from π^-s. (This is of no real use in the case of tracks entering the BEMC, because those high angle tracks cross only a small section of the CTD).

The performance achieved, with a mainly argon gas mixture, is $\sigma_{p_\perp}/{p_\perp}^2 \approx 10^{-2}$ GeV^{-1}, with the spatial resolution $\sigma_{r-\phi} \approx 160$ μm [55].

4.7.2 CIZ/COZ, CIP/COP

The inner and outer Z-chambers surround CJC1. They are divided into 24 (15 for CIZ) sections in z, each ring then divided into 24 (17 for CIZ) sections in ϕ.

Four sense wires are strung in an $r-\phi$ plane perpendicular to the beam direction in the COZ case, with the CIZ having the wire plane at a 45° tilt. The drift of electrons is then in the z direction within each of the elemental drift chambers. Drift time measurements fix z for the COZ with an accuracy of $\sigma_z \approx 300$ μm, and charge division along each wire yields ϕ, with $\sim 2\%$ accuracy. The two-track resolution is $\approx 2-3$ mm. Corresponding figures for the CIZ are $\sigma_z \approx 320$ μm, $\sigma_{2-track} = 3.2$ mm. The inner/outer proportional chambers are each divided into a number of chambers, 60 rings in z, 8 chambers per ring in ϕ. Sense (anode) wires run parallel to the beam, but it is the cathode pads which are read out; the separating ions and electrons induce a fast signal of these. They are used for z-vertex and t_0 (bunch crossing time) determination. All possible paths, connecting (fired) inner- and outer-layer pads are connected, forming 'roads' pointing to a z-vertex. The possible vertices are histogrammed, the true vertex appearing as a peak. The fast read-out enables the MWPCs to establish the beam crossing for any event and thus supply a t_0 for the drift chambers.

4.8 The Forward Tracker

The forward detectors optimise tracking in the difficult low angle region, into which many particles are sent. The high track density makes pattern recognition particularly hard, so a special set of trackers was designed and installed rather than just extending the CTD in z. The tracker is made of three identical *supermodules*, each consisting of (moving from the interaction point) a planar module, a MWPC, a transition radiator and a radial module. This is illustrated in figure 4.4 (c).

The MWPCs serve just as the central ones, for vertex and timing determination. They are built with 2 layers of sense wires 4 mm apart, the wires offset between the two planes. The resulting drift time is just 50 ns, which enables a track to be associated uniquely to its beam crossing. Three cathode planes separate the two layers of anode wires, and have ring-shaped pads etched into them. In the same way that *towers* are formed from groups of calorimeter pads, *rays* are defined by the forward MWPC pads, which match the towers in acceptance.

Transition radiators help in e/π discrimination. Within 96 mm are spaced

400 polypropylene films, the radiator separated from the adjacent radial unit by a thin window. When charged particles cross the boundary between different materials, with different dielectric/magnetic qualities, the associated electric field changes, and the charge re-distribution in the material causes the emission of X-rays. Detection of these by the radial chambers then distinguishes between a charged electron and a neutral pion.

4.8.1 The Radial Chambers

In these modules, each of which is split into 48 ϕ wedges, sense wires arranged 12 deep in z radiate outwards from a central hub. Charge division measurements supply a crude r, $\sigma_r \approx 2.9$ cm and the drift time then gives the second spatial co-ordinate, $\sigma \approx 210$ μm. The large number of sense wires within a wedge means that a track's angle through a cell is well determined, and the change in ϕ between the three radial chambers allows the curvature (and hence momentum) to be found. The chambers operate in the proportional region, i.e. deposited charge $\propto$ output signal produced for an ionizing particle, so dE/dX can be calculated for passing tracks.

4.8.2 The Planar Chambers

Placed in front of the radials, the planar chambers provide a more accurate r value to supplement the radial's measurement, allowing the polar angle θ to be calculated. Good momentum measurements require that a track is detected at many points along its path, and so the planars complement the radials in this function. A planar chamber is made of three drift chambers, each with 32 parallel drift cells 4 sense wires deep in z. The three planes are oriented at 60° to one another. Thus, θ and r can be found from drift-time measurements alone (no charge-division), and the wires are read out at one end only. The spatial resolution is $\sigma \approx 210$ μm.

The combined forward trackers have a momentum resolution of $\sigma_p/p^2 \sim 0.003$ GeV^{-1}.

4.9 The BPC

H1's tracking is completed (ignoring for the moment the muon spectrometer) by the Backwards Proportional Chamber (BPC). This MWPC consists of 4 wire planes, each rotated by 45° w.r.t. its neighbours. A plane is formed from 312 wires spaced 2.5 mm apart. Every other anode wire is read out, resulting in a spatial precision of $\approx$ 1.5 mm. The BPC is designed to detect and provide a space-point for backward charged tracks (principally DIS electrons) going into the BEMC which pass through only a very small fraction of the CTD. Currently, a BPC hit-point (x, y) can be assigned to the vertex to provide (θ, ϕ) values. In the long run it is intended to combine BEMC, BPC and central tracking information to improve the angular measurement of high angle tracks.

4.10 The Instrumented Iron

The iron structure surrounding the trackers, calorimeters and cryostat serves as a magnetic flux return yoke for the superconducting solenoid and as an absorber for the tail catcher calorimeter and muon detectors. Muons will penetrate through the calorimeter and need to be tagged and tracked before passing completely out of the detector. The iron also serves a structural rôle, supporting the Lar calorimeter and solenoid. It is surrounded by a 50 cm thick shell of concrete shielding. The iron is 'instrumented' with layers of streamer tubes inserted into gaps. For instance, in the barrel part, the iron is divided into 10 iron slabs of thickness 7.5 cm with 25 mm gaps in between (apart from the fourth slot which is 50 mm wide). Similar figures apply to the forward and backward endcaps. Streamer tubes are used for both the tail catcher and muon detectors. In principle, these are similar to drift chambers, in that a charged particle ionizes gas in a cell, and the electrons/ions accelerate towards/away from an anode wire. The gas, voltage and anode diameter are selected so that the electron avalanches develop into conducting plasma cascades, called streamers, and the size of pulse induced on the anode or cathode is such that no electronic amplification is required before readout. Streamer tubes are used because they offer a way of covering a very large area of detector at much lower cost than drift chambers.

In H1, sets ('profiles') of seven or eight 6 m long tubes with square cross-sections are operated at potentials of 4500 V. The gas used means that the tubes' performance is very dependent upon atmospheric pressure and temperature, and the exact voltage is altered according to the daily conditions. Outside the tubes are mounted 4 mm wide strips of aluminium on one side (some perpendicular to the wires, some parallel), and aluminium pads on the other. The pads vary in size with θ from 5×5 cm^2 in the forward direction to 30×30 cm^2 at 90° and are grouped by the electronics to form towers matching those of the Lar calorimeter. Pads are read out in an analogue manner, digital read out is used for the strips.

The muon detector part of the iron consists of three staggered double-layers of streamers situated in front of the iron, behind the iron and at a depth of 30 cm within the iron. The wire and strip read-outs provide space-point measurements with resolutions of ~ 2 mm [44] allowing muons to be tracked. Muons require ≈ 2.5 GeV energy to be able to penetrate the iron in the forward direction.

The tail catcher calorimeter is used to give an approximate measure of the energy of hadronic showers that are not contained within the Lar (or BEMC in the rear). The pad signals from the first five and last 6 layers through which a particle passes are summed, the result being proportional to the total number of streamers produced and thence the energy lost in the iron. The 7.5 cm iron slabs correspond to about 1 λ, and the resulting energy resolution is $\sigma_E/E = 100\%/\sqrt{E}$ [59]

4.11 The Forward Muon Detector

The central trackers and instrumented iron are not sufficient for measuring the momenta of forward ($\theta < 17°$) muons with high accuracy. The Forward Muon Detector (FMD) is composed of six large, octagonal, double planes of drift chambers, three either side of a toroidal magnet (which produces a field of ≈ 1.5 T). The first plane is located at 6.35 m from the nominal vertex. Four of the layers measure the polar angle θ, two measure the azimuthal angle ϕ, as yet uninstrumented. Each layer has ≈ 250 cells varying in length from 40 to 240 cm, arranged in 8 octants. The exact number of cells per layer changes, increasing with distance from the vertex. Drift times are recorded which gives the position of a particle across the width of

a cell with an accuracy of $\sigma \approx 250$ μm, whilst charge division gives the particle's ordinate along a cell to within ~ 3 cm. A schematic view of the FTD is in figure 4.4 (d), showing the ordering of the layers. Each drift cell has a single anode wire running down the center.

Muons passing through the FTD are bent in the magnetic field of the toroid, allowing the momenta to be reconstructed by determining the radius of curvature of the tracks.

The momentum resolution varies between $\sigma_p/P = 25\%$ for 25 GeV/c muons, $\sigma_p/p = 32\%$ for 150 GeV/c tracks [55].

4.12 Superconducting Solenoid

H1's magnetic field is produced by 4 superconducting coils of 6 m diameter placed outside the Lar calorimeter. The field is 1.2 T in the central region with variations of $\leq 3\%$ over the central tracking volume. The field strength was chosen on the basis of cost and likely tracker performance. Placing the solenoid outside the calorimeter, as opposed to surrounding just the trackers, has several advantages. The field is made homogeneous, dead material in front of the Lar calorimeter is minimized, and the return yoke is able to act as a tail catcher as described. The solenoid's main elements are a cylindrical aluminium former inside which is wound a superconductor (niobium/titanium composite) cooled by liquid helium circulation. The cryostatic system is separate from that of the Lar calorimeter. Placed just outside the backwards iron endcap, around the beam pipe, is a small *compensating* magnet which develops a field of 6 T. Beam particles travelling exactly along the beam axis would experience no force when passing through the magnetic field, but there are always some with non-zero P_t; these would be deflected on passing through the detector, the deflection building up on each circuit, leading to beam losses. The compensating magnet produces a field equal in magnitude and opposite in direction to the main one, so that over a full ring circuit $\int B.dl \approx 0$.

4.13 The Luminosity Monitor

Outside the bulk of the detector, behind the BEMC, is the luminosity monitor, composed of an electron tagger and a photon tagger. The process used for luminosity measurement is the bremsstrahlung reaction $ep \rightarrow ep\gamma$. This has a large and precisely calculable cross-section, totalling $\approx$ 71 mb at HERA [60]. The detector is designed to measure an electron and photon in coincidence. The photon 'arm' is a hodoscope of 25 Cerenkov total absorption crystals arranged in a 5×5 pattern, each $20 \times 20 \times 200$ mm^3 in volume. The tagger is placed at -107 m, along the proton beam pipe, and has a polar angle acceptance of $\approx$ 0.44 mrad. A scintillator is placed just in front of the crystals to veto the entry of charged particles, and a sheet of carbon absorber (thickness $\approx 3X_{rad}$) protects against the low energy photon background. A similar arrangement of identical crystals, this time in a 7×7 grid forms the electron tagger. Sited at -37 m, along the electron beam pipe, this hodoscope detects electrons deflected through angles < 5 mrad. The electron beam bending magnets act like a spectrometer, causing electrons with energies other than the 26.7 GeV beam value to be deflected out of the pipe. The taggers have an energy resolution of $\sigma_E/E \approx 10\%/\sqrt{E}$ and a spatial resolution of $\approx$ 0.5 mm. A good event consists of an electron and photon in coincidence, no charge in the veto photon counter, $12 \leq E_{e^-} \leq 25$ GeV and $5 \leq E_\gamma \leq 18$ GeV, the optimum energies for detection. The acceptance of the electron detector is $\approx 31\%$, i.e. it detects a cross section of $\sigma \approx 23 \pm 1$ mb [60]. At the design luminosity of 1.5×10^{31} cm^{-2} s^{-1} this corresponds to a rate of $\sim$ 300 kHz. The accuracy of the luminosity measurement is currently known to a precision of $\approx 7\%$ [57]. The principal backgrounds to the monitor are proton beam halo, electron bremsstrahlung off residual beam gas and electron synchrotron radiation.

4.14 The ToF Scintillator and Veto Wall

Beam particle induced backgrounds have been discussed in chapter, 3 and it was made clear there that this background was high rate, threatening to dominate over interesting physics. Upon receiving an L2 keep (see trigger section below), the trigger initiates read-in of data from sub-detectors, halting the read-in of new events

- the dead time; if beam background particles cause triggers to fire, there is a danger that the detector will be locked continuously in extensive dead time periods. The energies deposited by these events may be quite compatible with genuine physics, and so cannot be used to distinguish between them. One main difference, exploited by the Time of Flight scintillator (ToF), between beam background particles and genuine *ep* reaction products is the *time* at which they reach the detector. The backward region of H1 is the most sensitive to background, because it is from that direction that beam halo particles will arrive. Clearly, genuine interaction particles travelling from the vertex will arrive at a given point behind the BEMC *later* than any undesired ones accompanying the incoming proton bunch. The ToF's function is to veto events if particles pass through it in a time window which indicates that they are too early, i.e. are travelling in the wrong direction.

The ToF detector is formed from two scintillator planes with its center at -2.1 m, i.e. behind the BEMC but inside the iron. The two planes, ToF 0 (most rearward) and ToF 1 are formed from 8 and 16 scintillator pieces respectively, approximating the BEMC granularity. Lacking the room to extend light guides to reach a magnetic field-free region outside the detector, photomultiplier tubes are mounted on the rear of the planes. High field PM's have to be used, with dynodes arranged so that some photoelectrons will travel from one dynode to the next before being deflected into the PM walls. The reason for having two planes is to prevent false triggers from (electronic) noise in the PM tubes, cosmic events etc. A coincidence between ToF walls is required. If a hit in both walls occurs, it does so within one of three time windows, and a trigger bit is assigned to each hit pad according to the window in which it happened. These are

1. Interaction: if a hit occurs within a 6 ns window of the expected interaction time given by the HERA clock.
2. Background : if a hit occurs within a 6 ns window of the time of arrival at the ToF of the proton bunch (and any associated secondaries/halo), as defined by the HERA clock.
3. Global : a hit occurs at any time within a 96 ns wide window.

No hit should occur outside the third window. The signals produced by the ToF logic come too soon for the central trigger logic (3 bunch crossings after particles arrive

at the ToF, c.f. 20 for other subdetectors), and so have to be pipelined and delayed [61]. For each pad, the three trigger bits are grouped together (moving from least to most significant placing) in the order 'physics bit', 'background bit','global bit' to form a ToF word. Thus, ToF word = 101 indicates that a global and interaction bit were set, but no backround. These trigger bits can then be used by the central trigger logic to veto whole events. For the early data collected in June 1992 the requirement was kept very loose; the only veto condition was that there was only ≥ 1 background bit set. The ToF was used again in the final steps of data reduction after the formation of DSTs, as described in section 7.1.5, by looking at the individual scintillator signals in more detail.

4.14.1 The Veto Wall

This consists of two planes of scintillators at -6.5 m, roughly 2.7 m across. The wall covers a larger area of the detector's rear than the ToF scintillator and is designed to trigger on beam halo muons which have emerged from the beam tunnel, thus providing a global veto.

4.15 The H1 Trigger

The new difficulties faced by any H1 trigger system are those of rate of data taking, and volume of data produced; both are higher here than at any previous experiment. With a bunch crossing every 96 ns, and a raw digitised information volume of 3 Mbytes [62] from $\sim 250,000$ channels, the challenge presented to the trigger and data acquisition systems are considerable. The aims of the trigger are to accept as many interesting physics events as possible, whilst rejecting the dominating background, thus requiring it to make a decision about the value of an event once every bunch crossing. The initial load and rate of data must eventually be reduced by stages of event rejection and data compression to a final rate of 5 events s^{-1} with an average event size of 120 KBytes. The H1 trigger has three important features;

- Independence of *Subsystems*; all subdetectors (e.g. calorimeter, radial chamber, ToF scintillator..) have standalone trigger and readout mechanisms, and

the overall trigger system is capable of running with several subsystems turned off.

- Standardization of hardware and software, with well-defined data structures. This simplifies the task of combining information from sub-detectors and controlling all data flows.
- The trigger is multi-level, with 4 layers operating at different levels of sophistication and with different quantities of data.

The 4 levels of the H1 trigger (L1 to L4) are as follows.

4.15.1 Level 1

This uses a hardware frontend data pipeline; that is, the analog signals from each subdetector are digitised and converted into trigger elements, (simple binary conditions such as 'energy > some threshold' , 'hit in pad') which are synchronised to the HERA clock and grouped into blocks of 8. These elements are then fed into the pipeline, a chain of shift registers, and are moved along the chain by one register every HERA clock cycle. The different groups of elements produced by one subdetector are synchronised as the controlling Central Trigger Logic (CTL) deliberately imposes various signal delays. The chain is 25 beam crossings deep, so that after ≈ 2 μs the elements reach the end of the line; if they satisfy the set L1 trigger condition, then an L1 keep signal is produced, the pipeline is frozen and dead time begins. If not, the event 'falls off' the line, and the process continues. In fact, every subdetector produces its own subtrigger, the global L1 being the OR of 64 of these.

4.15.2 Level 2

L2 refines the L1 decision and produces its keep or reject signal 20 μs after the bunch crossing. It too is a hardware trigger, with the individual subdetectors producing their L2 subtriggers which are combined in the L2 CTL. An L2 keep extends dead time, and initiates the read-out of detailed subdetector data. A reject

decision ends the dead time and frees the L1 pipeline. In fact, the global L2 trigger does not currently exist, although individual sub-detectors are capable of producing their own L2 outputs.

4.15.3 Level 3

The read-out of the data from the frontends of the subdetectors takes 800 μs, a large fraction of the total dead time. Level 3 is a software trigger whose task is to make yet more refined cuts with more detailed information in parallel to this task. In this way it can, if necessary, reject the current event before read-out has completed. This saves subsequent levels from having to reconstruct uninteresting events. An L3 reject aborts the read-out and frees L1.

An account of the methods used to digitise the calorimeter, drift chamber, MWPC and streamer tube data is presented in [62]. The digitised information, encoded in standard format data banks, is fed into one of many multievent buffers (MEB), each event uniquely tagged by an 'L2 Keep number'. The following stages of the trigger now proceed asynchronously, not governed by the HERA clock. A central event builder reconstructs events by collating all banks from the MEBs with the same event number, forms tracks, clusters and vertices and places the assembled event into a full event buffer (FEB). Events may be reconstructed in a different order to that in which they genuinely occurred depending on the volume of data and hence processing time needed. The average rate for event building is 50 Hz.

4.15.4 Level 4

The FEB events are passed to the L4 filter farm, a set of processor boards running software to analyse the full event in detail. Pattern recognition for tracks and clusters may be performed, illegal vertices are isolated etc. The 3-D vertex reconstruction is a particularly powerful beam wall/gas rejection method. The surviving events are passed out as a 5 Hz rate from data logging. A dedicated μ-processor sends events to the central DESY IBM along a fibre optic link.

4.16 Data Acquisition

Data acquisition (DAQ) is the task of controlling and monitoring the streams of data which pass between the subdetectors, trigger stages, and main computers. It is also provides the interface between the detector elements and the human operators, allowing them to monitor and set (for instance) operating voltages, gas pressures etc. The whole DAQ system is summarised in [63]. The principles and vocabulary of the DAQ are those of electronics and software engineering, and so no further considerations will be paid to it here. It should nevertheless be noted that the DAQ is just as an important a part of the H1 experiment as any calorimeter or tracker.

Chapter 5

Software

A short description of the main programs provided by DESY and used in this study is provided in this chapter. In every case, more details are to be found in the appropriate, referenced manual.

An overview of the entire set of programs is provided by figure 5.1, taken from [64]. It shows how there are four stages to the analysis chain - the physics event, the response of the detector, the reconstruction of the event from the detector's outputs and the physics analysis itself. Every step in the 'real' chain, with a real event, the real H1 detector etc., has its simulation equivalent; An *event generator* models some particular physics process, e.g. DIS or photoproduction. The output is passed to a *detector response* program which simulates the performance of the detector, and whose output is a set of calorimeter, tracker (etc.) response banks. The real and simulated chains then converge, with the *reconstruction* code, which runs on either the genuine detector's output or the simulated data, and which builds the event— forming tracks, clusters and vertices. Graphical views of the simulated or reconstructed event within the detector are also provided by the H1ED package [65]. All programs' input/output and data handling, as well as that of the complete DAQ system and subdetector outputs is standardized and processed in exactly the same way.

5.1 Monte Carlos

The event generators all use Monte Carlo techniques, i.e. random numbers are generated for the values of physical parameters according to known distributions. Physics processes are simulated with quantities such as the angle or energy etc. of individual particles or events chosen randomly (within some limits), but in a way such that the overall distribution of that quantity for many events follows experimentally observed distributions. The lifetimes of particles, cross-sections for reactions etc. are built in or calculated as necessary. For well-known reactions (i.e. where the cross-section is exactly calculable) the results of a Monte Carlo generator should accurately describe real data. This Monte Carlo approach is also employed by the detector simulators, with (for instance) resolutions and noise levels generated randomly but in accordance with the form of real measurements.

5.2 The event Generators

The programs used in this study included LEPTO [42], DJANGO/ HERACLES [51],[52], and PYTHIA [34].

5.2.1 LEPTO

LEPTO 5.2 models lepton-nucleon DIS. The basic hard parton level scatter is modelled using electroweak theory with some QCD corrections (gluon radiation, photon-gluon fusion, higher twists) applied. Neutral and charged current interactions are simulated. The user may specify the desired range in Q^2, x, y and the cross section is calculated for the region bounded by those limits. The parton distribution of the proton is specified in a library. Events are generated and weighted by the cross-section for that point in the kinematic plane. An output bank is produced for each event, with every particle occupying one row, its 4-momentum and particle type, history and status (stable, unstable, beam particle, etc.) recorded in columns. The manner in which the struck quark fragments into the observed hadrons after the scatter is modelled by JETSET 6.3 [70]. This uses a phenomenological approach

to perform the hadronization, because although in theory calculable within QCD, the computational power necessary for a complete calculation of all the interactions occurring between many partons is impossible to realize. The soft fragmentation is based on strings; the gluon-gluon interaction between two partons is visualized as a colour flux tube stretched between them. The force increases as the string placed between two diverging partons stretches, and breaks when $\approx 2-5$ fm long, when it is energetically more favourable to form a new $q, \bar{q}$ pair and two shorter strings. The fragmentation continues for those strings with sufficient invariant mass. Many other aspects of jet phenomenology are incorporated (e.g. momenta distribution, suppression of heavy flavours etc) into the full model. An approach called parton showers [71] is used to include higher than first order QCD effects (emission of partons from the struck quark before or after the photon vertex).

5.2.2 DJANGO/HERACLES

HERACLES is the program used to simulate ep collisions with initial and final state radiation of photons from the lepton line included. Like LEPTO, the user specifies a range in x, y, Q^2, and the program performs a numerical integration over this region to calculate the cross-section. The program has a range of possible proton structure function sets available in the form of libraries. Not only are leptonic QED corrections incorporated, but quarkonic ones too, and further refinements such as corrections for γ, Z, W self-energies, γ, Z interference. DJANGO provides an interface between the HERACLES model of $ep \rightarrow e'X\gamma$, JETSET for the hadronization of the quarks and the various input/output routines.

5.2.3 PYTHIA

This program is a general purpose generator for any lepton/hadron - lepton/hadron collision. Provided with the appropriate structure functions it can model a large range of hard scattering processes. The user selects the choice of structure functions (e.g. for the proton, as normal, and for the photon – see section 3.1.3) and a number of hard scattering mechanisms, such as N.C. DIS or QCD Compton scattering etc. At HERA it is used mainly for hard photoproduction physics, with

the DIS option turned off, and a variety of direct and resolved photon processes selected. PYTHIA simulates in effect all but the soft VMD contribution (see section 3.1, 7.2.1 for a fuller discussion). Parton showers are also used for the initial and final state showers, i.e. for $e \rightarrow e\gamma$, $q \rightarrow qg$, $q \rightarrow q\gamma$, $g \rightarrow gg$ and $g \rightarrow q\bar{q}$. Each $a \rightarrow bc$ process is characterized by a splitting kernel $P(z)$. The branching rate is determined by $\int P.dz$, with the daughter products having fractions z, $(1-z)$ of the original energy. Each daughter may then decay and so forth. Simplified enormously, each emission is governed by a $Q^2 = m^2$, m being the mother parton's invariant mass. Branching continues until a lower cut off $Q^2 \sim 1$ GeV^2 is reached, at which point the phenomenological quark fragmentation of JETSET comes in.

5.3 The Detector Response Programs

5.3.1 H1SIM

The full detector simulator is H1SIM [66]. It takes the 4-vectors of each event produced by an event generator and simulates the effect that they have passing through H1. The basis for H1SIM is the CERN program GEANT (GEometry And Tracking) [67]. Each particle is stepped through the detector, with the effect of the materials and magnetic fields it passes included. A description of the detector (with its geometry, material properties, locations of gaps etc.) is stored in a database. A set of 'hit banks' results, with the co-ordinates and magnitudes of energy/charge deposits etc. recorded. The program then decides how the detector would have responded to such hits. In the database are stored the locations, resolutions, method of readout etc. for all subdetectors, and 'digi banks' are created, containing the simulated digitisations of each device in exactly the same format as the real hardware's output. A simulation of the trigger is also performed. Either full shower production algorithms (GEANT/GEISHA) or parametrized ones (ALGORIX /GFLASH) may be used; the former are more accurate, and require typically $\sim$ 200 s per event on the DESY IBM 3090 mainframe. The latter (not employed for the Monte Carlo datasets used in this analysis) are far quicker, $\sim$ 50 s per event.

5.3.2 H1PSI

Before large-scale production of standard Monte Carlo outputs by the H1 collaboration was completed, and before the reconstruction code H1REC was written, early analyses were performed with the simplified, fast H1PSI (Parametrized SImulation) program [68]. H1SIM uses detailed descriptions of the detector's construction and performance, H1PSI uses parametrizations for faster processing, and simplified geometrical models of each subdetector. For instance, central tracks all had a helix fitted to them, a number of hit points was then allocated to the helix, and the track discarded if this was below a threshold. Parametrizations of the tracker resolutions were then used to smear the hit points and produce a 'reconstructed' track. Similarly, the BEMC calorimeter was modelled not with rectangular stacks but as a series of wedges in ϕ subdivided radially.

5.3.3 H1REC

H1REC is the complete event building program which can process either genuine or simulated events. The manner in which it turns tracker hit points and calorimeter channel data back into 'real' artifacts of use to the end-user is very complex, and varies greatly between each detector component. No overall manual exists, and the user is referred to a set of incomplete reports compiled in [69].

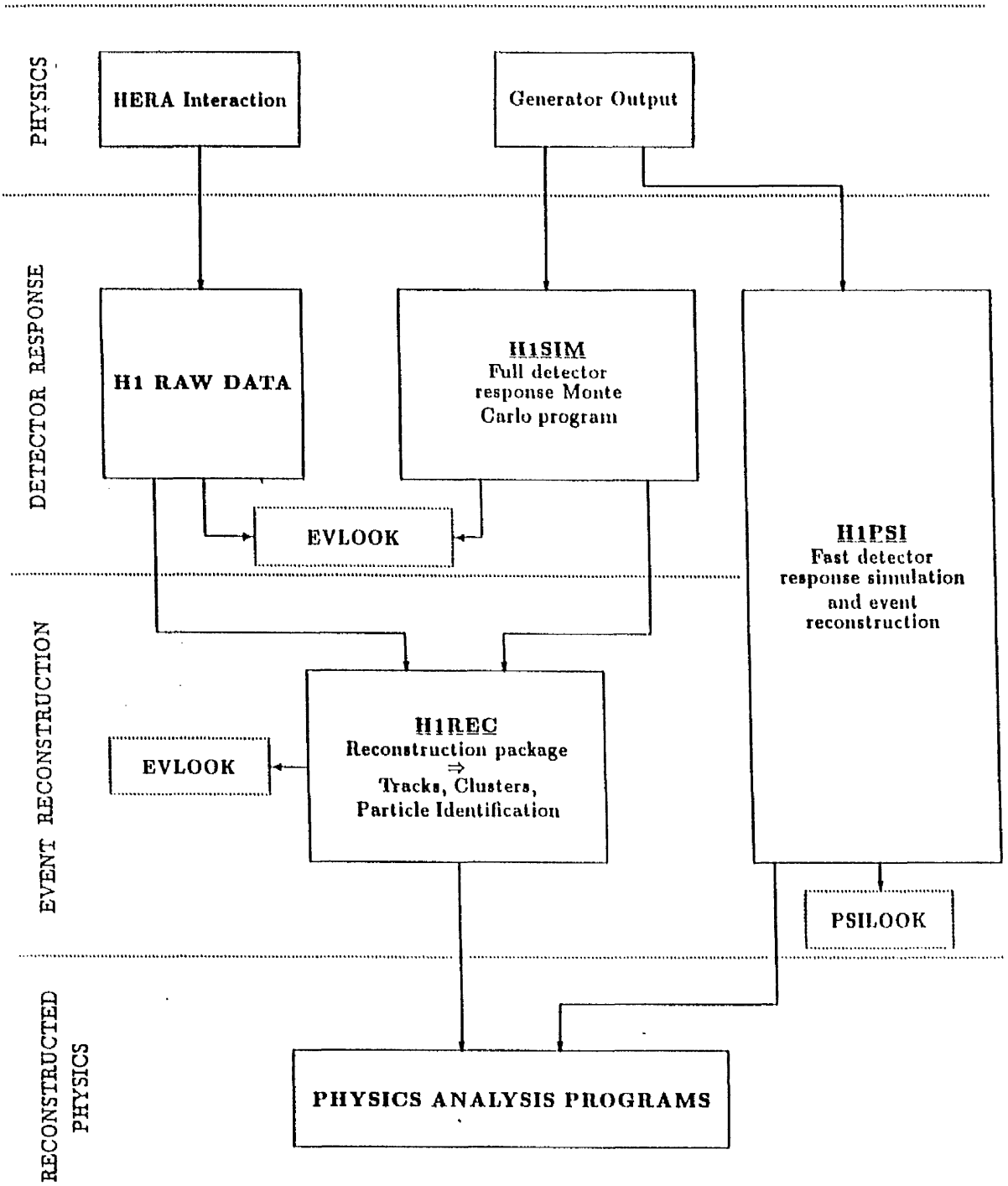


Figure 5.1: H1 software chain

Chapter 6

Identifying the Scattered Electron

Essential to the study of N.C. DIS physics is the identification of the scattered electron and the accurate reconstruction of its energy E_e and polar angle θ_e, both for the determination of the electron energy spectrum and for obtaining the kinematic variables Q^2, x and y. The preceeding introduction to the kinematic plane will have made clear that there is no domain in which the hadrons alone can be used to accurately reconstruct the variables (and even if such a method was used, the summations over the hadrons would still require the isolation and exclusion of the electron). The Q^2 determination is always reliant upon the electron angle. The discovery of an electron in the final state also helps to classify an event as a N.C. DIS process, as opposed to a Charged Current or background event.

This chapter describes a number of *kinematic* means with which to select the electron, that is using a kinematic feature (energy, momentum etc.) rather than the response of some detector component. The motivation for this approach was firstly that it allowed studies to be performed independently of the detector response and event reconstruction software tools H1SIM and H1REC (described in section 5), inevitably unreliable for much of their development time, and secondly that simple physics considerations revealed that such methods should be successful for a large fraction of the events.

The techniques covered here are only applicable to the case where $\theta_e > 90°$, i.e. in the backward hemisphere of the detector. The issue of electron identification is

clearly highly correlated with that of background rejection; to fail to find a final state electron is to possibly reject a DIS event as background, and so any adopted lepton i.d. scheme may simultaneously be a photoproduction rejection scheme (accepting the later arguments of this chapter that photoproduction events constitute a major background source). In fact this is not strictly true. It is possible to imagine a method whereby an event is found by some means (topological?) to be DIS, but where the electron still cannot be found, but this more advanced possibility is not considered here and no such method has yet been successfully employed. For the sake of clarity, the topics of background rejection and electron identification have been separated here, but it is important to realise that they are essentially two aspects of the same problem, and were developed together.

6.1 Kinematic Algorithms

6.1.1 Maximum P_t

The description of the N.C. DIS scattering process in section 2 revealed how the final state consists of the scattered electron and a stream of hadrons (the current jet) resulting from the struck quark. The two unstruck (spectator) quarks can be ignored here, since they travel in the forward direction with high momentum and at low angle ($P_t \sim 1$ GeV/c), and are thus effectively lost in the forward beam pipe. Figures 6.1 (a)-(c) show Monte Carlo predictions for the numbers of backwards travelling neutral hadrons, charged hadrons and photons, figures 6.2 (a)-(d) displays the total backwards hadronic energy, total backwards energy (including the scattered electron) and mean hadronic P_t, P_z. Particles with $\theta \geq 176°$ were omitted, since these are supposed to have passed down the beam pipe and be undetectable.

If the initial assumption about the low transverse momentum of the spectator quarks is correct, then for momentum conservation the electron and current jet will be back-to-back in the $r - \phi$ plane. The P_t of the current jet is divided between many particles, and so the electron is usually the particle with the largest P_t.

6.1.2 Minimum P_z

For that part of the kinematic plane where $\theta_j < 90°$ (see figure 2.4) i.e. where the current jet points forwards, the electron will be the sole particle travelling backwards, and identification is trivial. At lower values of x / higher values of y the current jet will point into the backward region also, but a study of figure 2.3 shows that the jet energies are then lower than the electron energies, with $E_j \leq 5$ GeV. It is therefore to be expected that the particle with the highest energy in the backward region, or alternatively the one with the largest, negative ('most negative') P_z will be the electron.

6.1.3 Maximum $E - P_z$

A third algorithm is deduced from considerations of the quantity $E - P_z$, a term which occurs in the expressions for the y variable and again later in the background reduction discussion. Since

$$E - P_z \approx E \cdot (1 - \cos\theta)$$

this term will be small for the spectator quarks, emitted at small angle ($\leq 5°$). Recalling the formulae for y, 2.8 and 2.18 then for $y < 0.5$

$$\sum_{hadrons} E_i - P_{z_i} < E_e$$

and $E_e - P_{z_e} > E_e$ since P_z is negative, and the electron is the particle with the greatest value of $E - P_z$. This is still possible for $y > 0.5$ because the *individual* values of hadronic $E_i - P_{z_i}$ may still be less than $E_e - P_{z_e}$

The problematic regions can already be anticipated. At $Q^2 < 10^3$ GeV^2, as y increases beyond 0.5 the current jet angle and energies θ_j, E_j also increase, whilst the electron energy E_e and angle θ_e fall. At $Q^2 < 10^2$ GeV^2 both θ_j and θ_e are $\geq 150^0$. This overlap in angular range coupled with the fact that the electron may have less energy than a hadronic fragment is bound to constitute a challenge for any scheme.

	x	y	Q^2
min	4×10^{-4}	0.0	4 GeV^2
max	0.999	0.99	2500 GeV^2
cross section			165.2 nb

Table 6.1: LEPTO 5.2 generated kinematics

6.2 Testing the Algorithms

Having arrived at three simple identification mechanisms, a comparison of their performances is necessary. These studies were performed at the generator level, that is by using Monte Carlo event generators to produce the expected output 4-vectors of the DIS final state and applying a very simple detector model to them. Extensive work was also done using the H1PSI fast detector response Monte Carlo (see section 5). This fast, parametrized model of the H1 detector was used to develop many of the approaches (including some of the ultimately successful ones) incorporated into the final analysis. H1PSI was however at best a crude testbed, particularly in its approach to clustering particles and its simulation of the BEMC calorimeter, and little more confidence could be placed in the results obtained from it concerning electron identification than in those obtained from a generator level approach.

The two aspects of the real detector included here were the beam pipe and finite energy resolutions. The effect of the beam pipe, with a radius of $\approx$ 10 cm was incorporated by neglecting events with $\theta > 176°$. The energy resolutions were simulated by applying a Gaussian smear of $\sigma_E/E = 10\%/\sqrt{E}$ for electromagnetic particles with $\theta_e > 152°$, $12\%/\sqrt{E}$ for those with $\theta_e < 152°$, and for hadrons in the same angular ranges, $\sigma_E/E = 100\%/\sqrt{E}$ and $\sigma_E/E = 50\%/\sqrt{E}$ respectively. (The values being based on expected BEMC, liquid argon and instrumented iron performances (see chapter 4)).

An event file of 15,000 LEPTO 5.2 (incorporating the JETSET 6.3 jet fragmentation routine) [42],[70] events was generated, using the Morfin-Tung B1 structure function set [25], with the kinematic limits listed in table 6.1.

6.3 Results of the Analysis

Figures 6.3, 6.4 and 6.5 show acceptances for the LEPTO file using the simple detector model on $\log Q^2/\log x$, $\log y/\log x$ and $y/\log x$ planes. Figures 6.6 (a) - (d) show projections of these planes onto $\log Q^2$ and y planes.

The term 'acceptance' is here defined very simply as the percentage of correctly identified electrons in a given bin. For the two-dimensional graphs, the size of a shaded square is proportional to the acceptance. Thus, a completely empty square represents a bin in which the true scattered electron was never found. The *number* of events in a given bin changes across the plane according to the cross section, highest at low Q^2. In every case *generated* kinematics are used, not those reconstructed from the identified electron.

The acceptance could in fact be defined in one of two ways, because there are two distinct classes of falsely identified electron.

- The case where the simulated electron was lost in the beam pipe, and therefore *any* particle identified as the electron is the wrong one.
- The case where the electron was present in the detector ($\theta_e < 176°$) but where some other particle was selected.

Acceptance could therefore be defined alternatively with the denominator being the total number of events generated in a bin *regardless of whether it was lost in the beam pipe*. Events of the first class could be labelled as photoproduction; the boundary between the two processes is an arbitrary low Q^2 value, as was said in section 3.1. For this study, all electrons were generated at $\theta_e \leq 176°$, inside the detector, and so the distinction is irrelevant.

The significance of the two classes of misidentified events is the nature of the migrations across the (x, Q^2) plane which results. The issue of migrations between bins on the kinematic planes is a complex one, and no attempt is made here to discuss it in detail (and indeed, detailed detector simulations are necessary for any discussion, see [30], [31]). Migrations may occur for a large number a reasons - misidentification of the electron, finite detector resolutions resulting in smeared

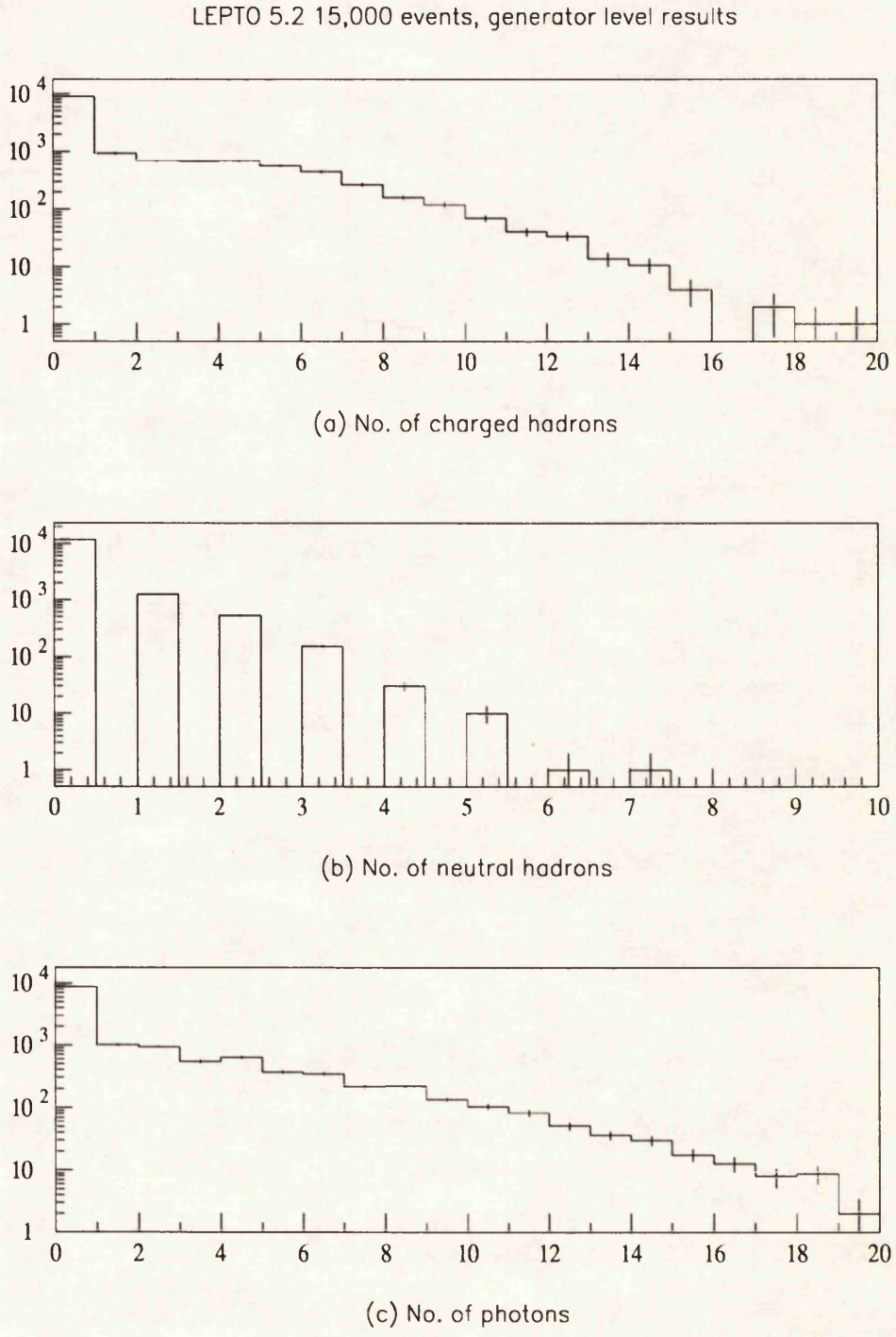


Figure 6.1: Multiplicities of hadrons & photons with $\theta \leq 176°$

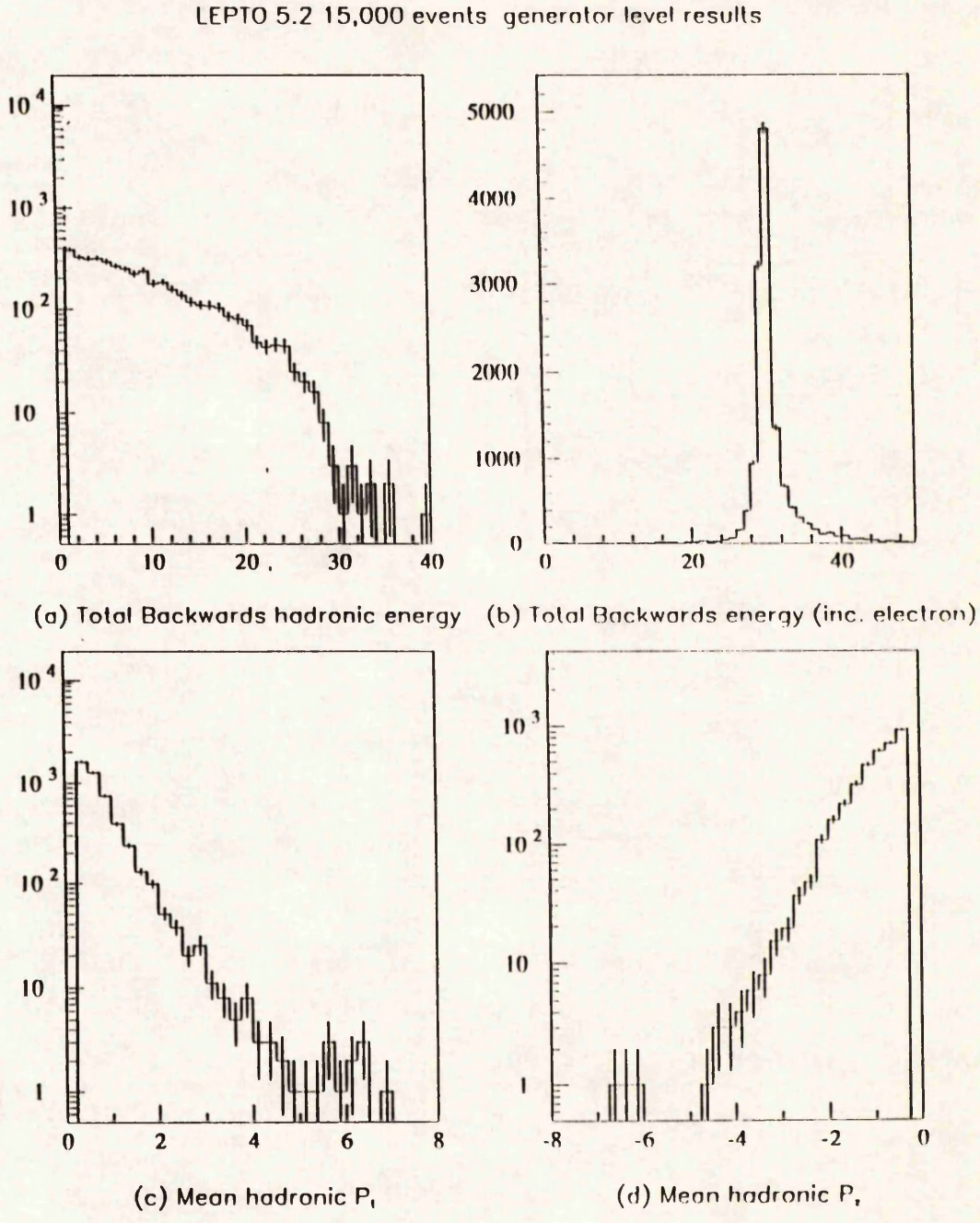


Figure 6.2: Total and mean energies of hadrons with $\theta \leq 176°$

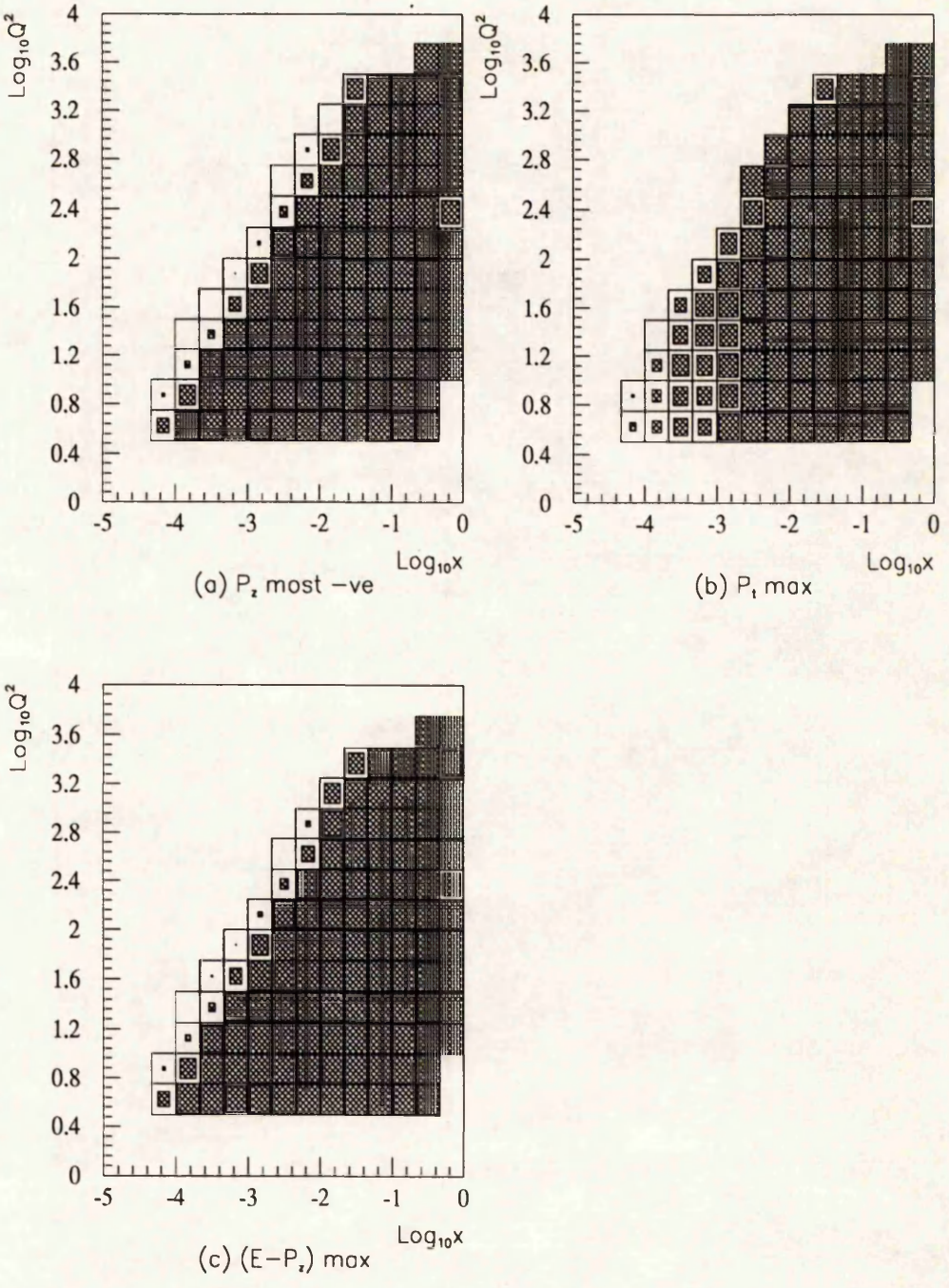


Figure 6.3: (Q^2, x) acceptance plots for LEPTO MT-B1 Monte Carlo

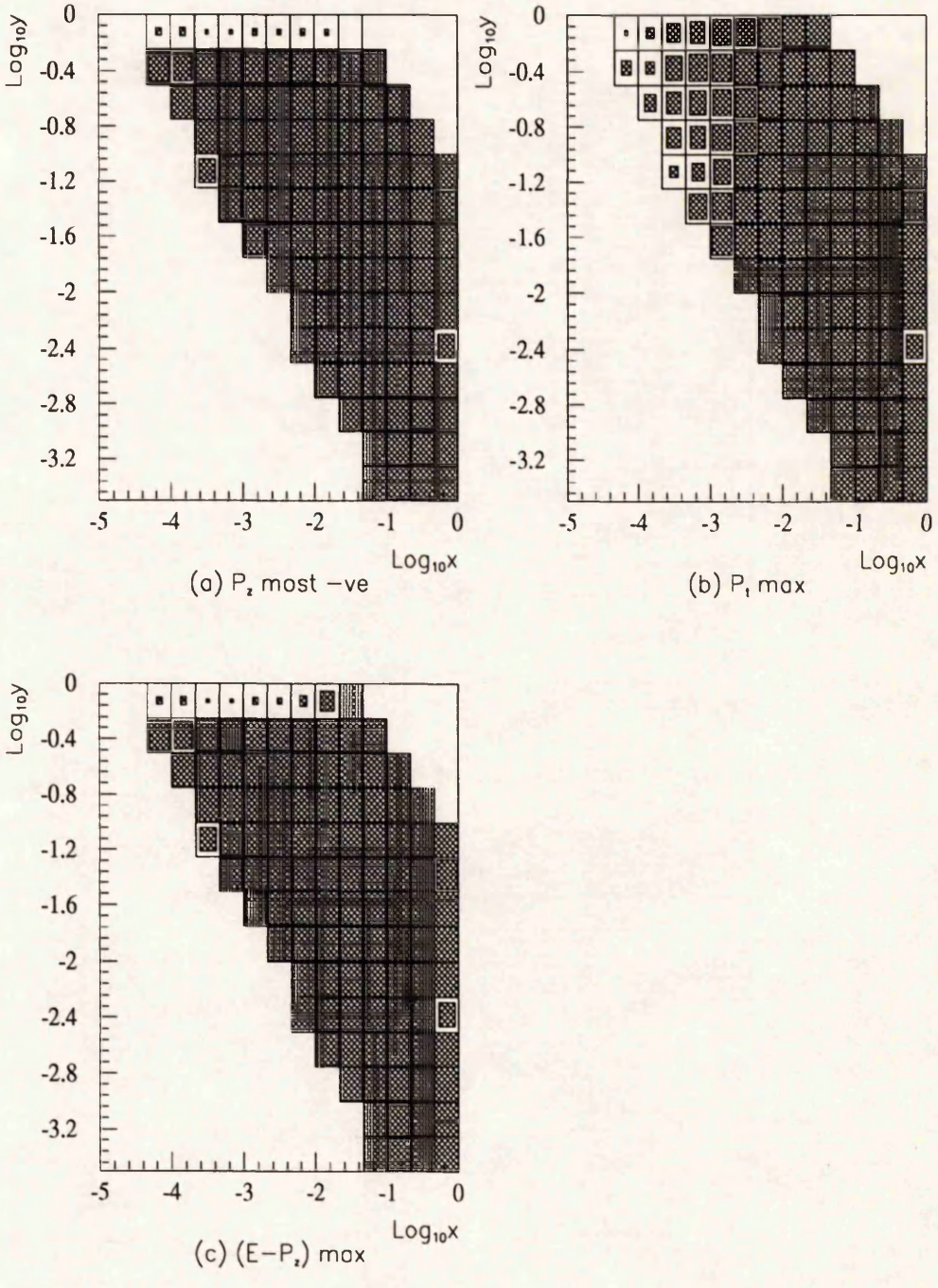


Figure 6.4: (y, x) acceptance plots for LEPTO MT-B1 Monte Carlo, log scale

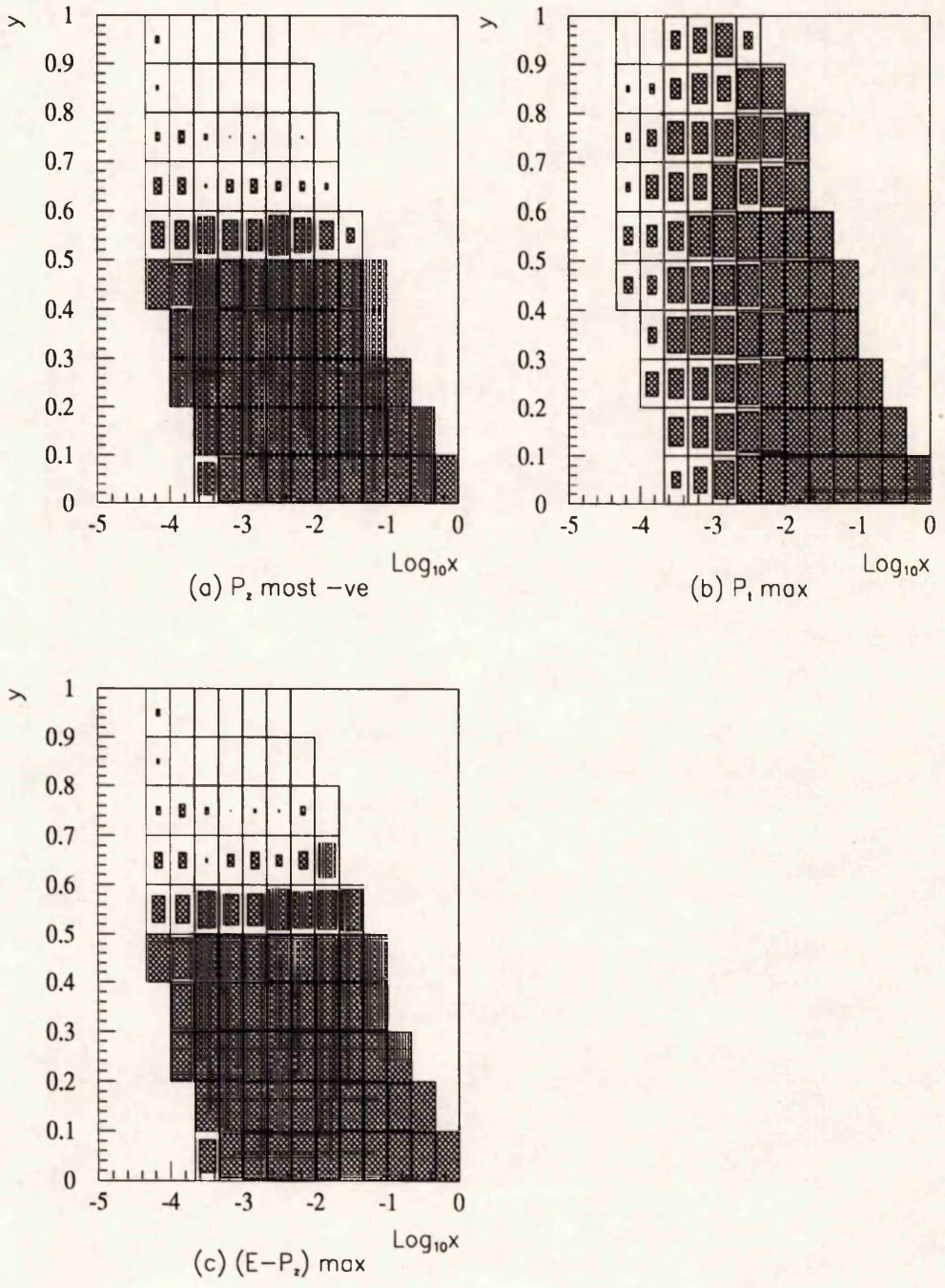


Figure 6.5: (y, x) acceptance plots for LEPTO MT-B1 Monte Carlo, linear scale

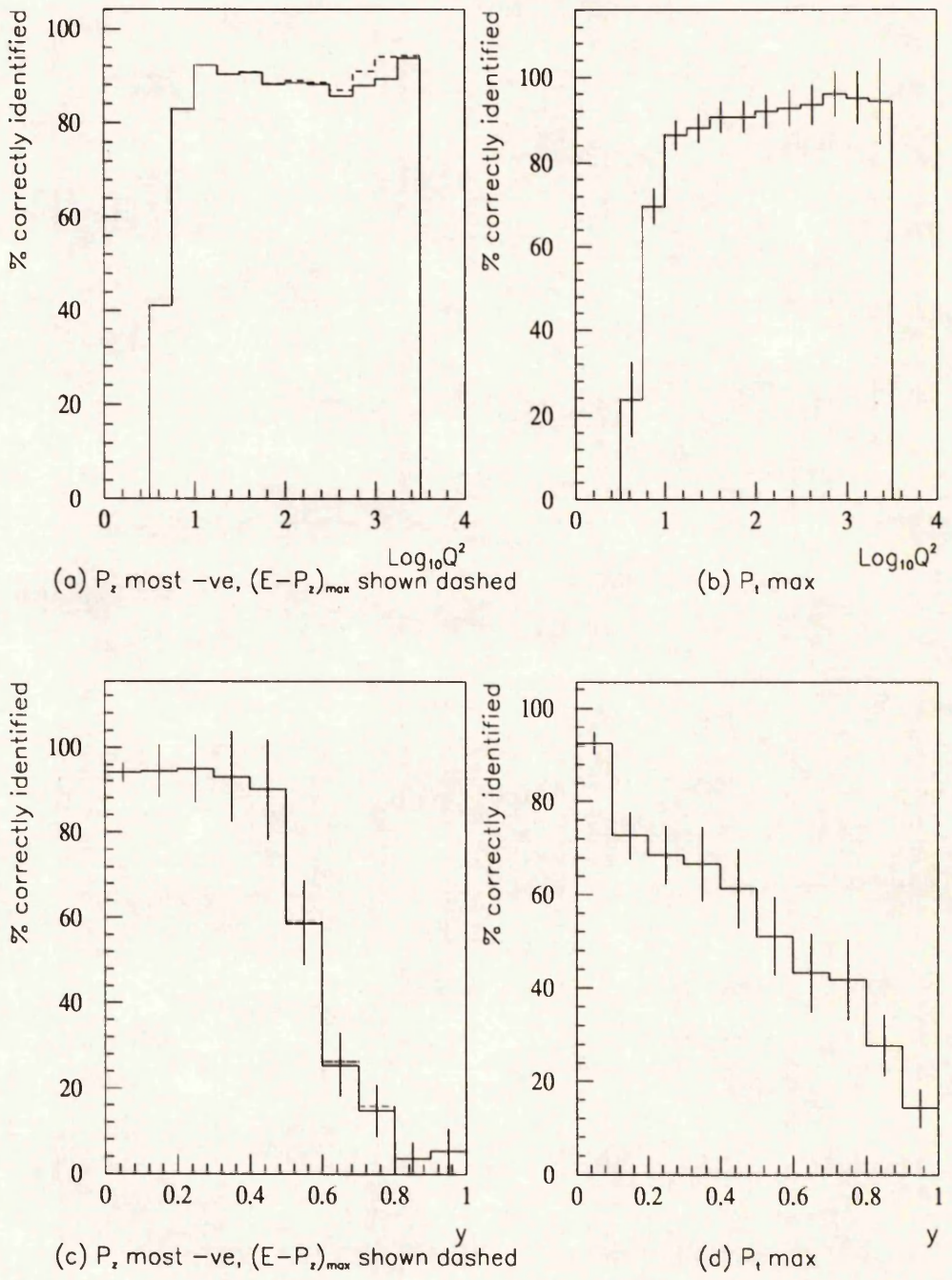


Figure 6.6: LEPTO MT-B1 e^- identification rates as a function of Q^2, y

x, y, Q^2 values, energy leakage from the detector ([72]); and this in turns leads to the equally involved subject of choosing the best set of physical measurements (energy, angle,.. from electrons or hadrons or some combination) from which to construct the kinematic variables, touched upon in section 2.6.

Events from the second class will clearly reconstruct to some wrong location on the (x, Q^2) plane, depleting some other point on that area above the $\theta_e = 176°$ contour. Those from the first class will however migrate *into* this area from below that contour, and so will inevitably reconstruct to a higher Q^2 than with which they were generated. All this is mentioned to emphasize the threat of the high rate, low Q^2 background; even if an algorithm could be found which always correctly selected the electron in a Monte Carlo event sample where $\theta_e > 176°$, there would be a contamination arising from those events (whether called 'low Q^2 DIS' or 'photoproduction') where the true electron was undetectable.

The first observation to be made is that the three algorithms do not produce very different results; no one graph is considerably better or worse than the corresponding graph for the other methods. The P_z minimum and $E - P_z$ maximum results are closer to each other than to the P_t maximum results. The most striking trends to be noted are

- For much of the (x, q^2) plane, acceptances are high, typically $\geq 90\%$.
- All techniques show a drop in acceptance at high y (visible on the $\log Q^2 / \log x$ plane as the band of lower acceptances along the leading diagonal).
- P_t max is poorer at $x \leq 10^{-2.5}$ (see 6.3, 6.4).
- P_t max is superior at high y (see 6.4, 6.5). Figures 6.6 (c), (d) show that the acceptance for this technique drops suddenly at $y > 0.1$ and thereafter continues to fall across the y range; for $E - P_z$ max and P_z min the y performance is $> 90\%$ until $y = 0.5$ and then drops catastrophically to just $\sim 20\%$ at $y = 0.8$.
- The P_t max acceptance drops at low Q^2 to $\approx 70\%$, whereas the other methods vary little across the Q^2 range (see 6.6 (a), (b)). Note that the error bars on the superimposed graphs have been omitted for clarity, but are identical to those on 6.6(b), i.e. the graphs are constant in Q^2 at high Q^2 to within errors.

The drop in performance at high y and low Q^2 for the P_t max algorithm is most easily understood with reference to equation 2.10, with P_t as either electron or summed hadron value. At low Q^2, P_t is low for fixed y, and the earlier assumption, that the spectator quark system is negligible now no longer holds, since it will have a comparable P_t. The electron alone no longer balances the current jet, and need not be the particle with greatest P_t. Similarly, for $y \sim 1$, P_t is small because of the $(1-y)$ factor, and there is a greater risk of some hadron (etc.) having the largest transverse momentum. A review of the angle contour plots (2.4) reveals that the low x, low Q^2 corner of the figure 6.3 (b) corresponds to the region where the quark jet and electron are both at high angle; the electron's P_t is thus small, and there are rival particles around the same θ value, so it is perhaps not surprising that the acceptance drops here. This also explains the steep drop visible in figure 6.6 (d) at $y > 0.1$, since at lower y values the current jet is confined to be at $\theta_j < 160°$, increasing the isolation of the electron.

The reason for the drop in performance at $y > 0.5$ of $E - P_z$ max has already been covered in section 6.1, with $E - P_z$ decreasing with rising y, and $\sum_{hadrons} E_i - \sum_{hadrons} P_{z_i}$ increasing. The acceptance of P_z most negative at high y is similarly low, but unlike that of $E - P_z$ max drops as Q^2 rises. This is clear from figure 6.4, where the performance in the top y band drops from left to right for P_z min and rises for $E - P_z$ maximum. (Lines of constant Q^2 run parallel to the diagonal edge, increasing in magnitude towards the top-right corner.) Equation 2.8 again explains why this is so; for a fixed y, $E - P_z$ is constant and, since E_e increases with Q^2, $|P_z|$ must drop, and a hadronic particle with a higher backwards P_z is misidentified as the electron. Once again, the explanation for the poor results at high y is the presence of high angle hadrons from the struck quark providing a wealth of rival electron candidates with high, negative P_z. This difference at high y, high Q^2 between the two algorithms is not visible on the integrated y acceptance plot (figure 6.6 (c)); this is because the large numbers of *low* x events in each y bin (where the cross-section is highest) dominate, and mask the effect completely.

6.4 Topological Cuts

Phillips [73] extended the topic of electron identification to include topological cuts; by maximising a quantity T_1,

$$T_1 = \sum_{j \neq i} \left(\frac{(E - P_{z_i}) - (E - P_{z_j})}{2 \cdot E_e} \right)^2 \left(\frac{P_{x_i} - P_{x_j}}{P_{x_{max}}} \right)^2 \left(\frac{P_{y_i} - P_{y_j}}{P_{y_{max}}} \right)^2$$

an attempt was made to combine the elements of the $E - P_z$ max and P_t max routines. Another strategy tested used the minimisation of I, defined as

$$I = \sum_{i \neq j} (E_i E_j - P_i P_j \cos \theta_{ij})$$

The electron was defined to be that particle whose exclusion from the sum resulted in the minimum value of I. This formula was derived from the observation that any pair of particles within a jet contributes an amount

$$W^2 = (P_1 + P_2)^2 \approx 2 \cdot (E_1 E_2 - \hat{P}_1 \hat{P}_2 \cos \theta_{12})$$

to the total invariant mass of the jet, and this mass would be raised if a non-jet particle (in this case the true electron) was included.

Although not detailed here, Phillips reported that these topological algorithms - considerably more complicated - did not out-perform the $(E - P_z)$ maximum method for $y < 0.8$, but offered the possibility of being useful for higher y events.

6.5 Source of Misidentified Particles

The crude detector model employed so far could clearly be improved upon. A breakdown of the those particles misidentified as electrons reveals that almost any final state hadron or photon could be selected on the basis of its kinematics. The H1 detector possesses several detector elements capable of providing information useful to the separation of hadrons and electrons; the BEMC calorimeter, with its sensitivity to electromagnetic particles and transparency to hadrons (sensitive length of BEMC scintillator stack $\sim$ 22.5 radiation lengths, $\sim$ 0.97 hadronic interaction lengths [45]); the iron (hadronic) tail catcher, and the central tracking detectors and Backwards Proportional Chamber (BPC).

Work was performed using the H1PSI fast detector simulator and with early versions of the full H1SIM/H1REC analysis chain ([66],[69] and section 5). The important conclusion arrived at was that no electron identification strategy based on kinematic, topological or detailed detector response techniques which was superior to the simple kinematic means outlined here could be found; and that the most important task in view of the need to produce physics results from early data was not to work upon the isolated task of electron selection, but to combine this with background reduction.

One other important result arising from the inclusion of realistic BEMC granularities, resolutions and hadronic/electromagnetic responses was the identity of particles misidentified as electrons. It was found (by looking at the LEPTO generator banks) that the majority of false electrons in the $152° - 176°$ range were either pions or the overlap of a charged pion and photon. The processes by which these then mimic DIS electrons are

$$\pi^0 \rightarrow \gamma\gamma \quad \text{then} \quad \gamma \rightarrow e^+e^-,$$

the π^0 decaying almost instantly, the pair production occurring either in the beam wall or central jet chamber end-flange. Another possibility is the interaction of a charged pion with matter (e.g. the BPC or BEMC)

$$\pi^{\pm} \rightarrow \pi^0 X$$

and the subsequent decay of the π^0. The third case, of a $\pi^{\pm}/\gamma$ overlap is where a charged pion and a photon (possibly from the decay of a π^0) both lie within a narrow solid angle range. This case is particularly dangerous because the charged particle leads to a detected track or BPC hit and the photon deposits the required electromagnetic energy in the BEMC. Single photons or groups of photons travelling within a narrow solid angle cone were an additional common source, with some of the photons converting to an e^+e^- pair leading to a BPC hit, and some mixture of electrons and photons registering as a BEMC cluster. Overlaps of up to three photons and a single charged hadron were found to occur in these simulations.

6.6 Preliminary Background Studies

Having established that an electron selection scheme based on kinematic means alone would have a failure rate of between 40% and 80% at $0.7 < y < 0.5$ and with the knowledge that pions and photons in the final state could be chosen in preference to the true electron, it became apparent that photoproduction processes would indeed provide a major source of contamination. The relative cross-sections of DIS and photoproduction were introduced in section 3.1.4. Figures 6.7 (a)-(c), 6.8 (a)-(c) show the multiplicity of backwards travelling hadrons and photons, the total backwards hadronic energy and total backwards energy (with $\theta < 176^0$ for all particles) predicted by the PYTHIA 5.6 Monte Carlo [34] (cf figures 6.1, 6.2). The importance of the observation that a π^0 or $\pi^{\pm}/\gamma$ were the largest source of false electrons now becomes apparent, as photoproduction events are a rich, high-rate source of such energetic particles in the relevant angular range. The much greater number of hadrons and photons in found in the BEMC compared with DIS events is to be noted, and the different distributions of backwards energies. A typical PYTHIA event has ~ 5 hadrons of mean energy ~ 1.5 GeV and a similar number of photons within the BEMC.

The observation that pions and photons could be chosen in preference to the true electron event within a pure DIS sample, where the electron was always within the detector has another significance. Since the same types of particles misidentified as electrons in DIS events are likely to lead to false electrons being found in photoproduction events, any cuts designed to remove wrongly reconstructed DIS events will presumably also function as anti-photoproduction cuts, and vice-versa.

This idea is strengthened by a study of event display pictures produced using the H1ED programme [65] of Monte Carlo DIS and photoproduction events (generated with DJANGO and PYTHIA respectively). Figure 6.12 shows four $r-\phi$ projections of the H1 detector showing the inner (electromagnetic) liquid argon calorimeter, the central and forward trackers, the BEMC calorimeter and BPC. Each display depicts a different event, the first and third being from DJANGO, the second and fourth from PYTHIA. (The kinematics for the DIS events are (x, y, Q^2) $= (8 \times 10^{-5},\ 0.48,\ 3.6\ \text{GeV}^2)$, $(2.6 \times 10^{-3},\ 0.53,\ 12.7\ \text{GeV}^2)$ respectively.) The electron candidate in each case is depicted by the largest of the dark rectangular

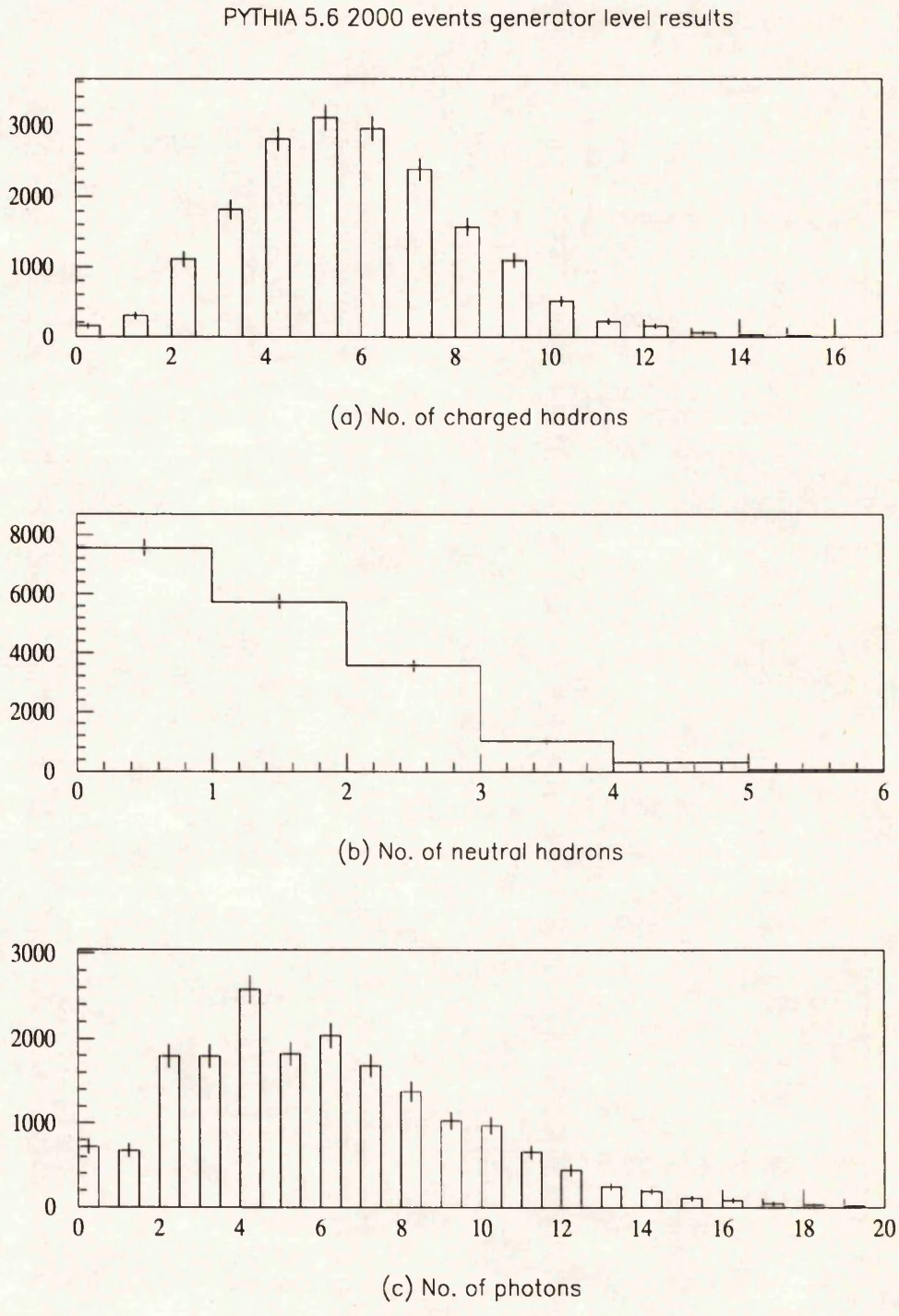


Figure 6.7: Multiplicities of hadrons & photons with $\theta \leq 176°$

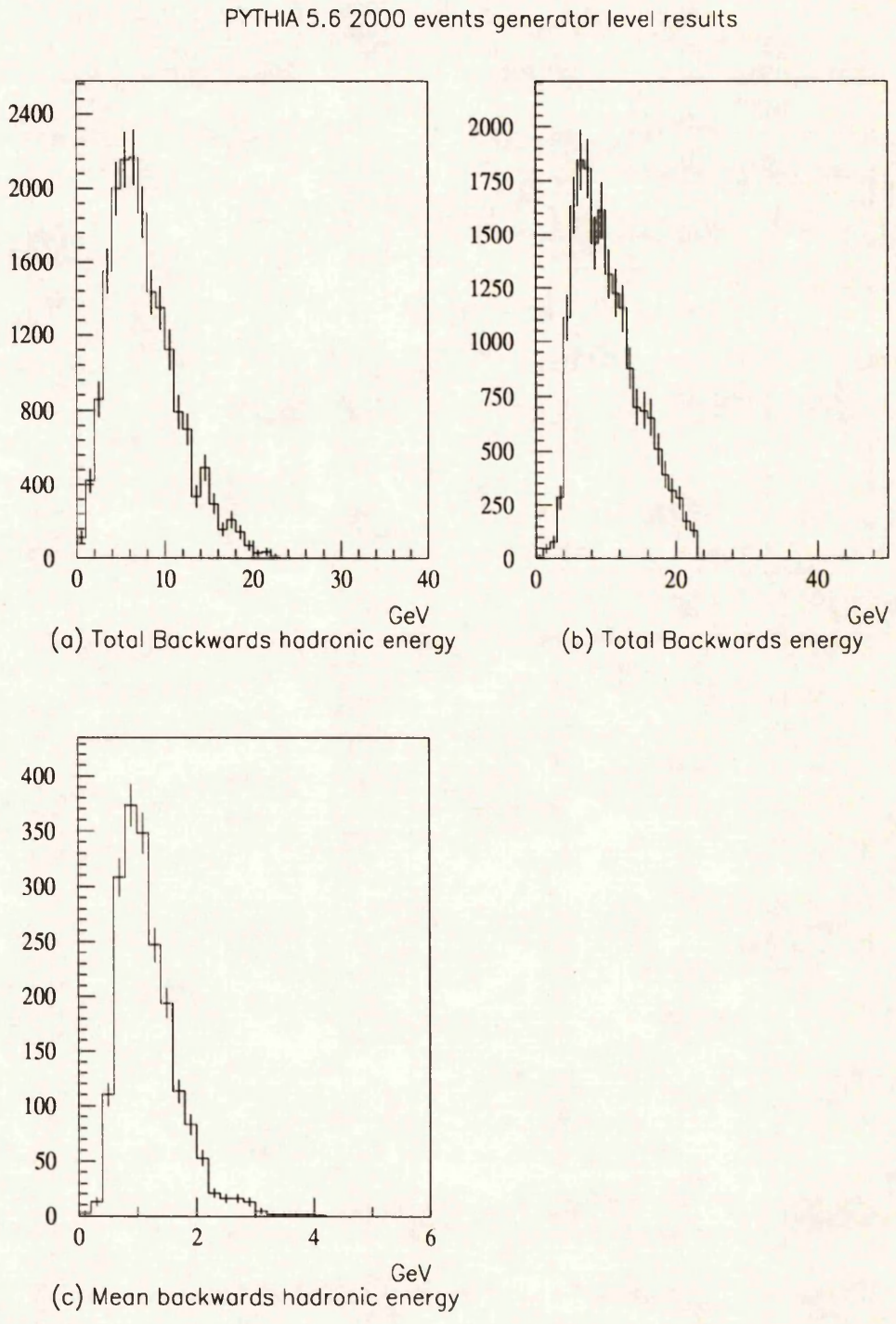


Figure 6.8: Total backwards hadronic energy, total backwards energy, $\theta \leq 176^0$

marks in the BEMC calorimeter, which is drawn as a vertical 'ladder' on the right hand side of the event. In the case of both PYTHIA events, the true electron passed down the beam pipe at very small angle, and the false electron is formed from 2 photons (tracks not illustrated). It is clear that there are no *obvious* features based on event topology, track multiplicity, angular isolation of the electron etc. which distinguishes between the PYTHIA and DJANGO events.

The scale of the background from photoproduction was found to be worryingly large, and initially uncertain. The difficulties of assessing its size were the results of two uncertainties - poor detector simulations and a lack of understanding of the correct photoproduction cross section.

Figures 6.9 (a), (b) show early predictions of the scattered electron energy spectrum for DIS (DJANGO with KMRS B– structure function) and photoproduction (PYTHIA 5.6) Monte Carlo outputs. Figures 6.10 (a), (b) are the corresponding y spectra, the y variable having been obtained from the scattered electron according to 2.8 again. (Here, deliberately no attempt has been made to scale the spectra according to the relative cross section). With x and Q^2 similarly obtained from the electron only, figures 6.11 (a), (b), display the location on the (x, Q^2) plane of photoproduction events incorrectly identified as DIS, and the distribution of DJANGO DIS events. H1SIM/H1REC were used to model the detector and reconstruct the events.

The essential points revealed by such exercises were that the background was largest at low electron energy, equivalently at high y, and that great care would be needed to remove the non-DIS events without severely depleting the lower end of the energy spectrum, so important for distinguishing between structure function parametrizations. Photoproduction events reconstructed to high y, low x across the Q^2 range.

The acceptance rates predicted by the different schemes are now worth reconsidering. Given the initial likelihood of some form of high y cut, figure 6.6 shows that the $E - P_z$ method is likely to be most useful, given its superior acceptance at $y < 0.7$ over P_t max and marginally improved acceptance at high Q^2 range compared P_z most negative.

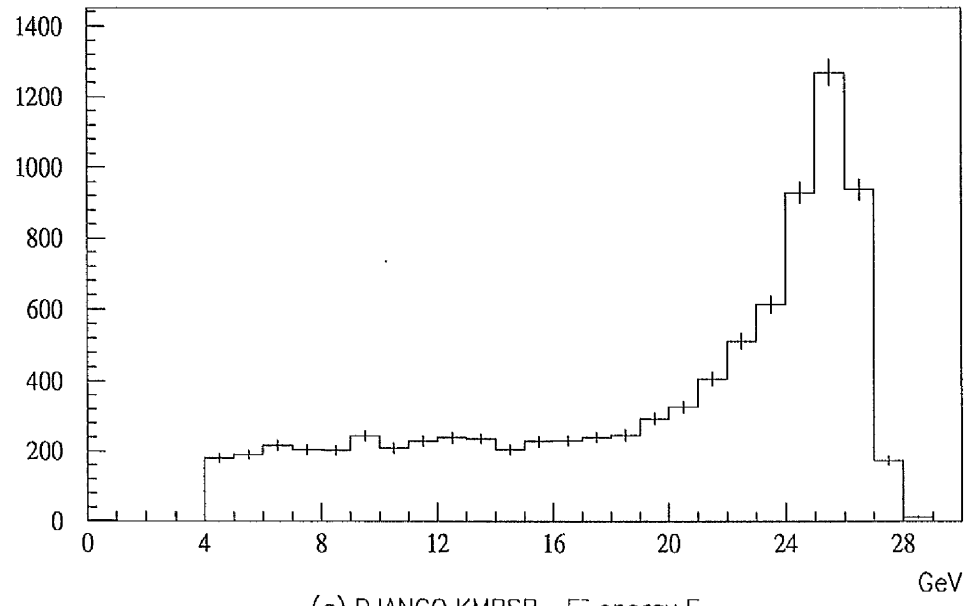
For the remainder of the analysis conducted in this study, the initial step in

the processing of any event was to identify an electron according to this algorithm.

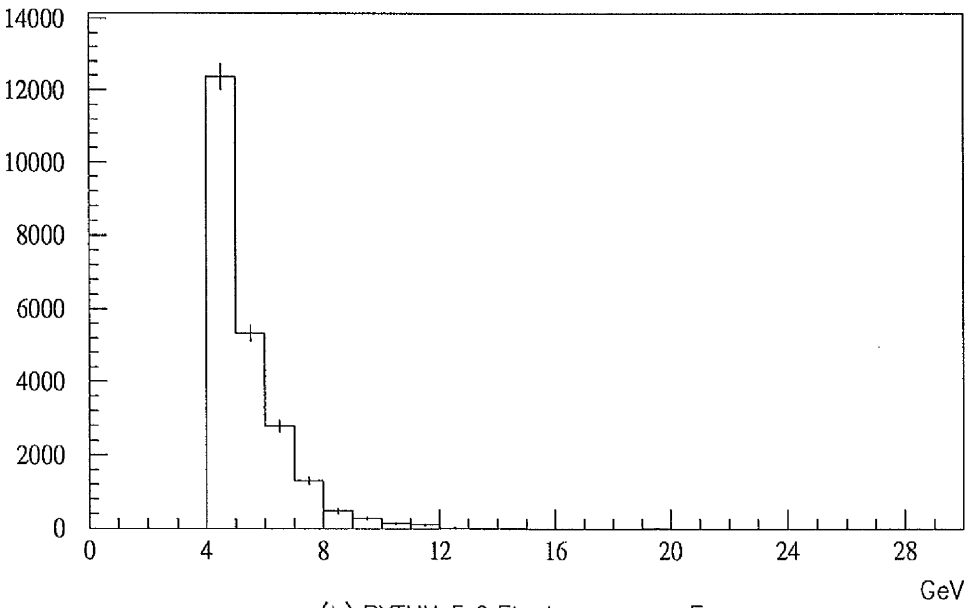
6.7 Conclusions

Monte Carlo studies of identifying the scattered electron in N.C. DIS in the angular range $\theta_e > 152^0$ corresponding to the low x, Q^2 region revealed a number of crucial points despite the simple detector models employed.

- Kinematic algorithms, looking at each of the final state 4 vectors in turn, can provide an effective means of identifying the scattered electron.
- Such means offer acceptance rates $\geq 90\%$ provided that $Q^2 \geq 10$ GeV^2 and $y \leq 0.7$.
- The most effective algorithm for $y < 0.7$ is to accept the particle with the greatest value of $E - P_z$ as the electron.
- No topological method or more complex combination of the simple techniques tested here proved more successful, though they may be of use in detecting the electron at higher y ranges.
- The largest source of fake electrons are pions or pion/photon overlaps.
- Photoproduction processes with a predicted cross section $\sim$ 1000 times σ_{DIS} constitute a large background at low E_e , high y. The pions and photons which can be confused with electrons in a pure DIS sample are present in large numbers in photoproduction final-states.

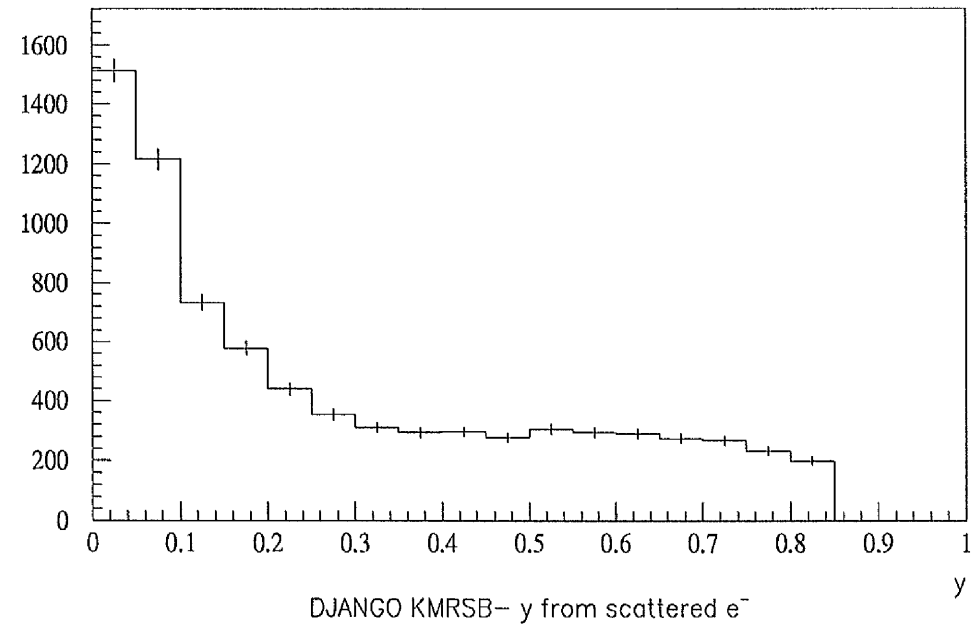


(a) DJANGO KMRSB– E^- energy E_{BCLR}.



(b) PYTHIA 5.6 Electron energy E_{BCLR}

Figure 6.9: Early energy spectra



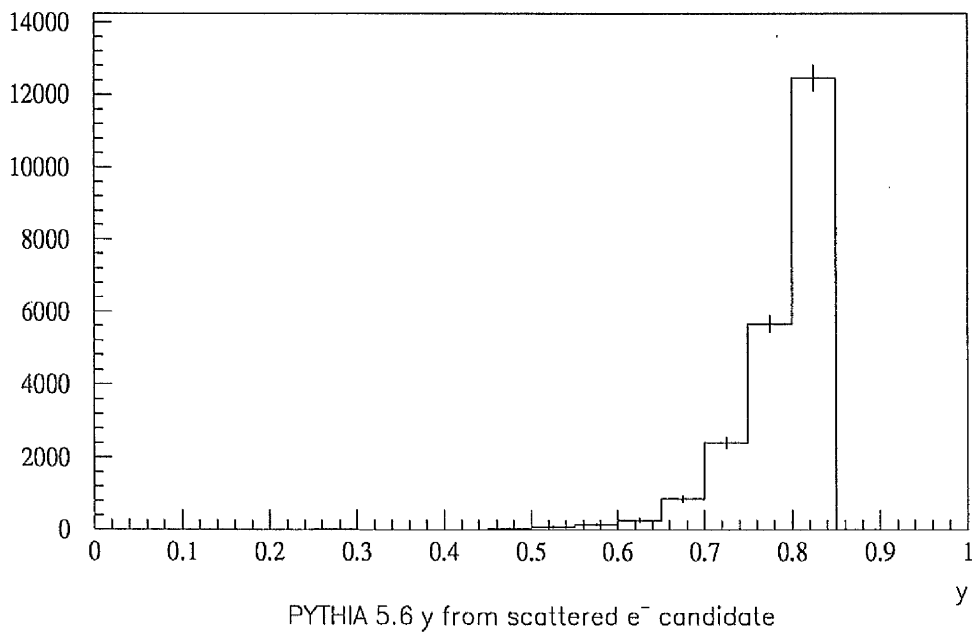


Figure 6.10: Early y spectra

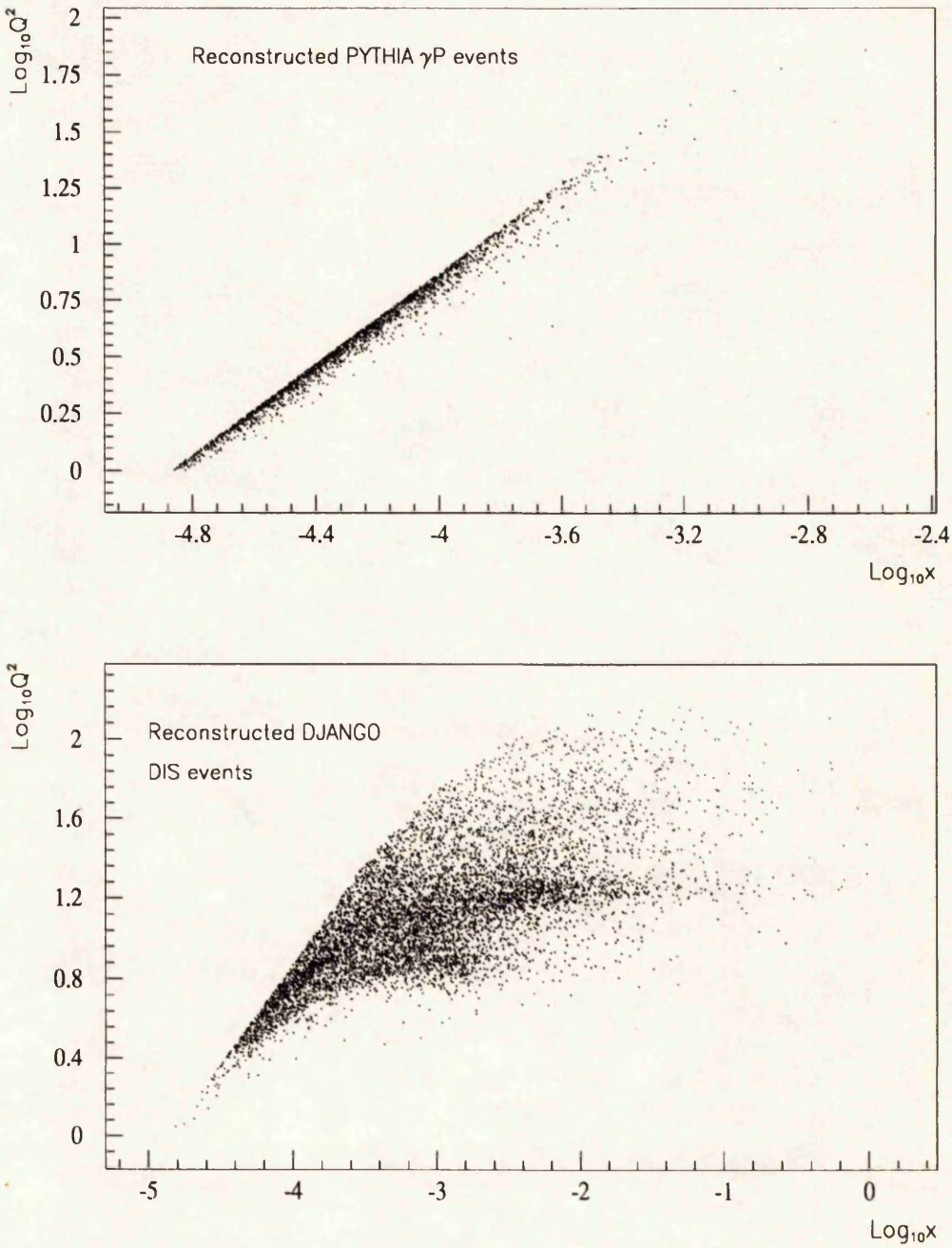


Figure 6.11: Early (x, Q^2) planes

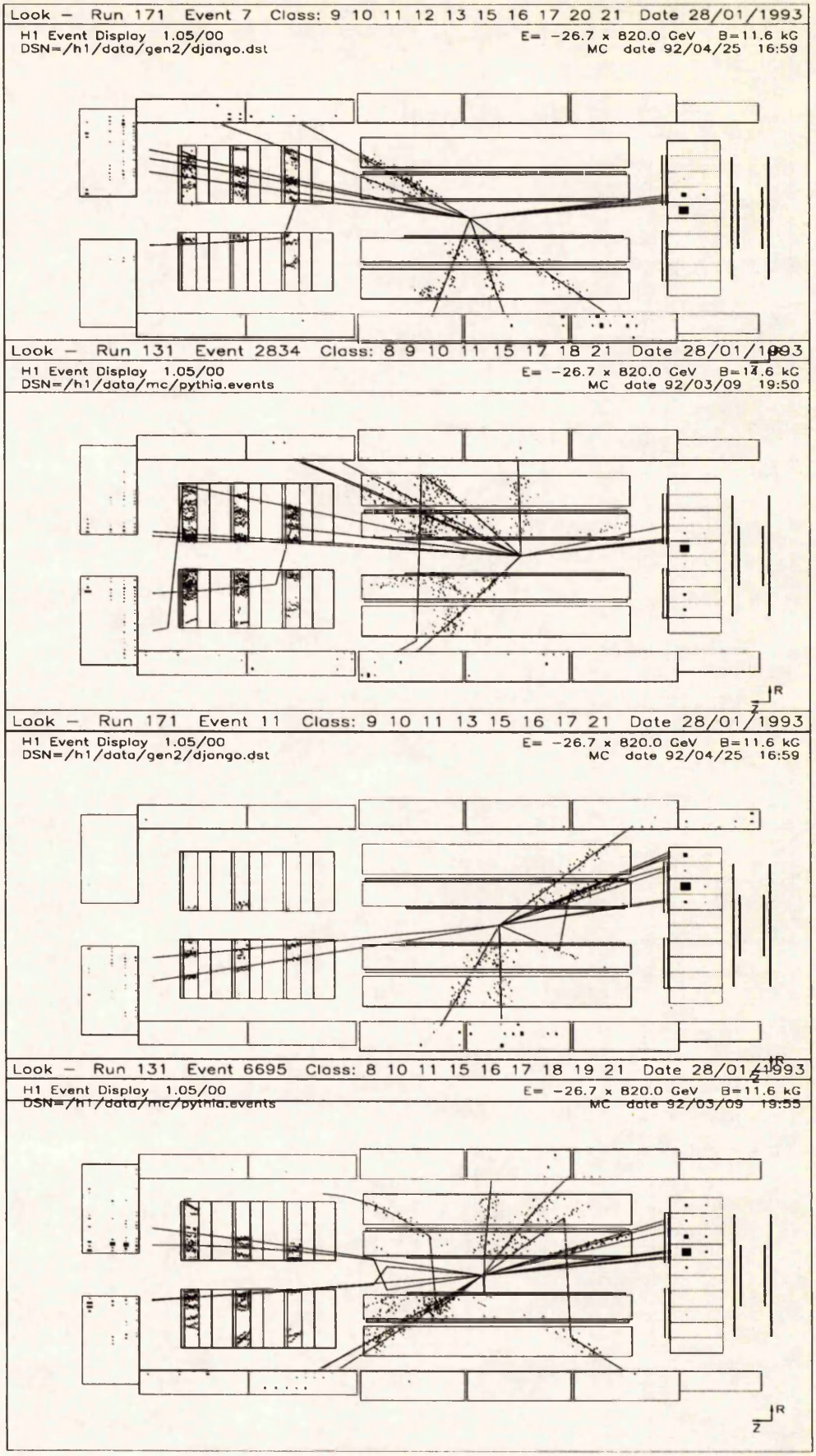


Figure 6.12: H1ED outputs

Chapter 7

Data Analysis

7.1 Production of Real Data

7.1.1 The L5 event Classification

So far, all the discussion has centred on Monte Carlo simulations of physics events. The H1 detector commenced data taking during the late spring of 1992, enabling new physics analysis to proceed. The Trigger/Data Acquisition system employed by H1 is summarised in section 4.15, and it was stated there that the output from the level 4 filter farm (L4) trigger stage operates at a rate of ≈ 5 Hz, resulting in the filling of a 200 MByte IBM cartridge once every seven minutes. These raw data cartridges are clearly too unwieldy for data analysis by individual collaboration members. An additional processing chain has therefore been developed, taking the raw data as its initial input, and producing a series of output data sets according the wishes of physicists. This is the level 5 (L5) event classification scheme.

Figure 7.2 illustrates the L5 operation schematically. Raw data is processed by two fast Silicon Graphics computers, which undertake the full event reconstruction using the H1REC package. Two selection modules, operating within the framework of H1PHAN (The general purpose PHysics ANalysis software used to access all output banks from H1 software - generators, H1PSI, H1SIM, H1REC) then apply selection criteria to the reconstructed events, and 'tag' them with an L5 Event Classification Word. The ECLASS routine has

a softer set of criteria and produces output for the so-called POT tapes, the lowest level of data on which it is remotely possible to perform any detailed data analysis. ECLDST consists of all the ECLASS cuts plus some additional ones, and events which satisfy them are written to the Data Summary Tapes (DSTs), which are in fact usually stored on disk. Not all event classes have this additional set of cuts, and some (e.g. for those with low transverse energy) which are high rate but of low physics interest are simply downscaled. The event classification words are not unique, that is any one event may be tagged as belonging to several classes, and so will be recorded on several different DSTs.

It should be stressed that the ECLASS, ECLDST modules are *software* routines, and so easily changed when necessary.

The motivation for this approach is obvious, it allows a very rapid channelling of events into a number of disk (i.e. fast access) files, each one tuned for a different area of H1 physics. At the request of H1 working groups a new set of DSTs can be produced if the original ECLDST definitions were found to be too lax (leading to too large a fraction of uninteresting events) by re-processing the POT tapes belonging to that group. The POT tape definitions within ECLASS are sufficiently broad to ensure that the need to return to the raw data (extremely time consuming) is highly unlikely.

The DSTs thus produced, although representing a substantial saving in data volume over the raw data, often contain too many background events, and a further set of cuts is then applied to produce a micro DST (μDST) containing a sub-set of DST events. In such cases it is the μDSTs which are the starting point for an individual physics analysis. Their construction is necessarily something of a compromise between the differing approaches favoured by individuals and the need for a common, tightly defined set of data to work on.

7.1.2 L5 Selection Rules

The ELAN (ELectron ANalysis) group works with L5 classes 10 ('NCHAD', DIS candidates for hadron flow studies) and 11 ('NCLQSQ', N.C. low Q^2) [75].

The selection criteria used for these classes are listed below, and an explanation is provided afterwards.

CLASS 11

POT selection : ((≥ 1 RCLU energy cluster > 4 GeV) *AND* ($\theta_{RCLU} > 150°$))

DST selection : ((≥ 1 BCLR energy cluster > 4 GeV with a BPC hit within a radius of 15 cm of the cluster centre of gravity) *OR*
((≥ 1 BCLR cluster with 22 GeV < E_{BCLR} < 32 GeV) *AND*
(ToF cut : (no ToF activity) *OR*
(≥ 1 ToF word = X01 *OR* X11) *AND*
(< 2 ToF words = X10))

CLASS 10

POT selection : As for class 11 but with ≥ 1 track present

DST selection : (≥ 1 BCLR energy cluster > 4 GeV) *AND* (ToF cut as for class 11) *AND* ((≥ 1 good CJC (Central Jet Chamber) track) *OR* (≥ 1 good forward track)) [76]

RCLU (Reconstructed CLUster) clusters are energy clusters (i.e. groups of calorimeter cells or scintillator stacks) reconstructed by the H1REC code; BCLR (Bemc CLusters Reconstructed) are energy depositions in the BEMC only.

7.1.3 Energy Requirement

The energy demand of these selection rules is quite simple, an insistence upon some rearward deposition ≥ 4 GeV, met by all but very high y ($y > 0.85$) DIS events.

The value of 4 GeV was chosen on the basis of the BCL2 trigger, an L2 trigger requiring a BEMC cluster above some threshold energy in anti-coincidence with a Time of Flight scintillator veto. This trigger is, as will become clear, a sub-set of the L5 requirements. (The BSET trigger described in 4.5.2 was the L1 trigger used, with a high energy cluster threshold value of 3.5 GeV. This corresponded to an event acceptance rate of 60 – 180 Hz [60].)

Studies of the BCL2 efficiency were performed with Monte Carlos and on early data, the trigger acceptance being determined by using monitor triggers. The 4 GeV value was found to give an efficiency of $\approx 80\%$, accepting events at a rate of between 1 and 3 Hz. For $E_{BCLR} > 10$ GeV, the efficiency rose to $> 99\%$ [57]. The insistence upon a BEMC (as opposed to a liquid argon calorimeter) cluster also fixes the angular range over which the study is conducted to be $152° \leq \theta_{e^-} \leq 176°$. Figure 6.8 (c), a graph of the mean energy of backwards-travelling hadrons from Monte Carlo PYTHIA events, also justifies this value, $\sim 99\%$ of all events have a mean hadronic energy < 4 GeV. The effect of this initial energy cut is clearly illustrated in figures 7.3 (a)-(c), showing the region of the x, y spectra and (x, Q^2) planes made inaccessible as shaded. Only the very low x region $x < 10^{-4.4}$ and that with $y \geq 0.85$ is completely removed. Figure 7.1 was derived similarly from a DJANGO Monte Carlo DIS file (10,000 events, generated according to table 7.2), and a PYTHIA 5.6 photoproduction file, (30,000 events see 7.2.1 for generation details) with a graph composed from real data for comparison, and shows the histogrammed number of BEMC clusters with $E_{BCLR} > 4$ GeV for those events satisfying the class 11 cuts (i.e. all events have at least one such cluster), heavily peaked at 1 in all cases.

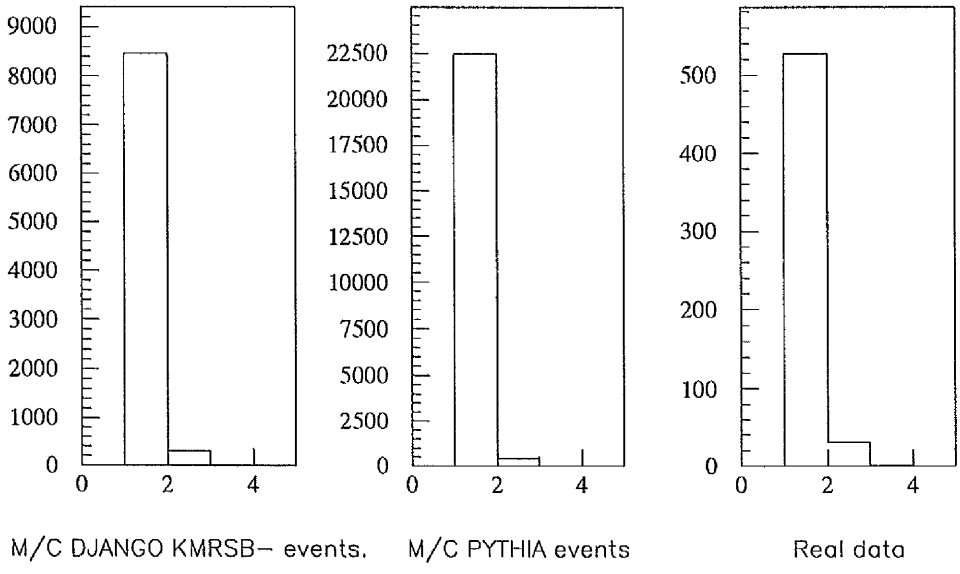


Figure 7.1: Number of BEMC clusters above 4 GeV, after $E_{BCLR} > 4$ GeV cut

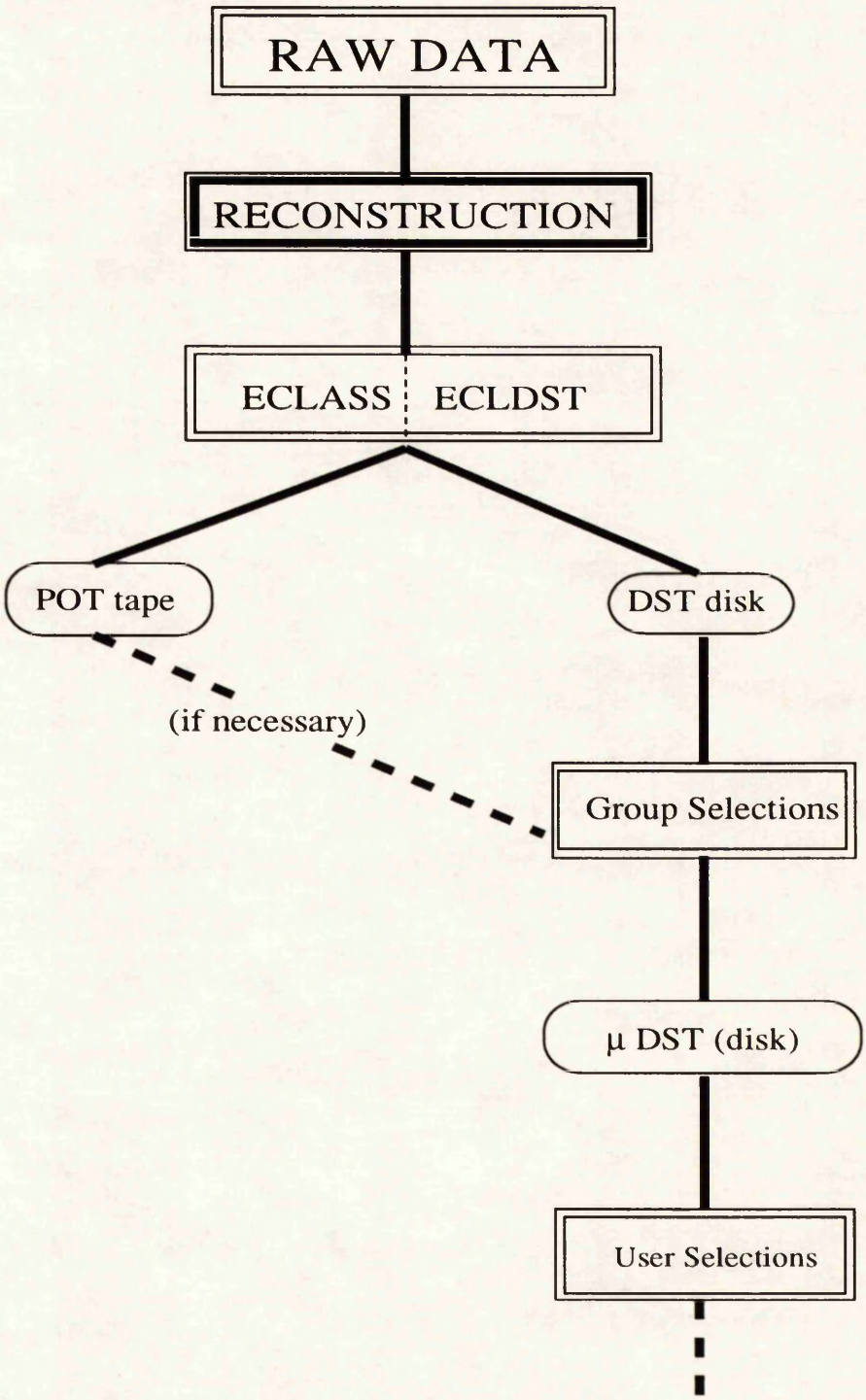


Figure 7.2: L5 Event Classification Scheme

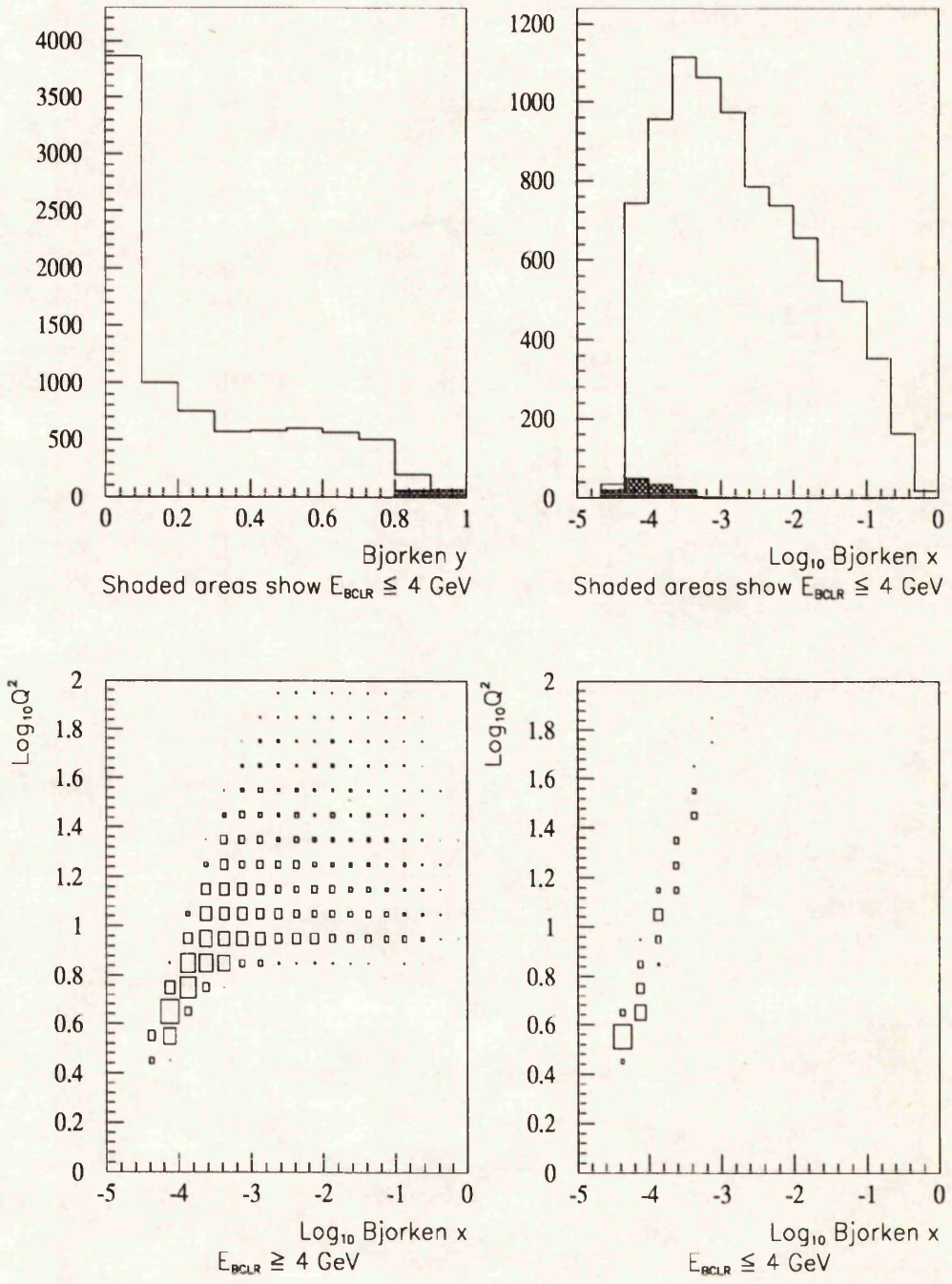


Figure 7.3: Effect on x, y of e^- energy cut at 4 GeV, DJANGO DIS

7.1.4 BPC Requirement

The use of the Backward Proportional Chamber (BPC) (described in section 4.9) is a powerful requirement which, when combined with the energy cut in the manner described acts as an anti-photoproduction selection. The BPC is the sole tracking device for angles $\theta > 170°$, and is itself limited to $\theta < 175°$.

Figures 7.4 (a)-(c) and (d)-(f) are histograms (created using the same Monte Carlo files as above) of the number of BPC hit points within a 15 cm radius of the electron candidate BEMC cluster and of the shortest radius between a BPC point and candidate cluster. In every case the radius is calculated by projecting the BEMC cluster hit onto the BPC plane using the knowledge of the detector geometry and (x, y, z) co-ordinates of the cluster centre of gravity. The sharp peak present on the DIS (DJANGO) histogram 7.4 (a), as opposed to the more continuous spread for the photoproduction events demonstrates the usefulness of this cut. For true DIS events, the passage of the charged electron will register as a hit on the BPC; the peak at 0 on the PYTHIA histogram is presumably caused by those events where a neutral pion (decaying $\pi^0 \rightarrow \gamma\gamma$) or photon was selected as the electron. The large proportion of events with ≥ 1 BPC hit within the critical radius must be explained by the higher multiplicity of charged hadrons present within a narrow cone of the $E - P_z$ maximum candidate, or the higher photon flux converting to e^+e^- pairs at the BPC. The need for $N_{R\leq 15} \geq 1$ removes $\approx 35\%$ of PYTHIA events. The value of 15 cm was selected on the basis of 7.4 (d)-(f), where the radius BPC-BEMC for DJANGO is ≤ 15 cm for all events. This is clearly over-safe, and later on is changed to 8 cm as suggested by the sharp drop in the DJANGO spectrum.

The requirement of a BPC hit is dropped for events with $E_{BCLR} > 22$ GeV since these are already few in number of sufficient interest to warrant inclusion into the μDST alone.

7.1.5 ToF Cuts

The ToF cuts (refering to the Time of Flight detector) are of a different nature to the previous ones since they do not distinguish between DIS

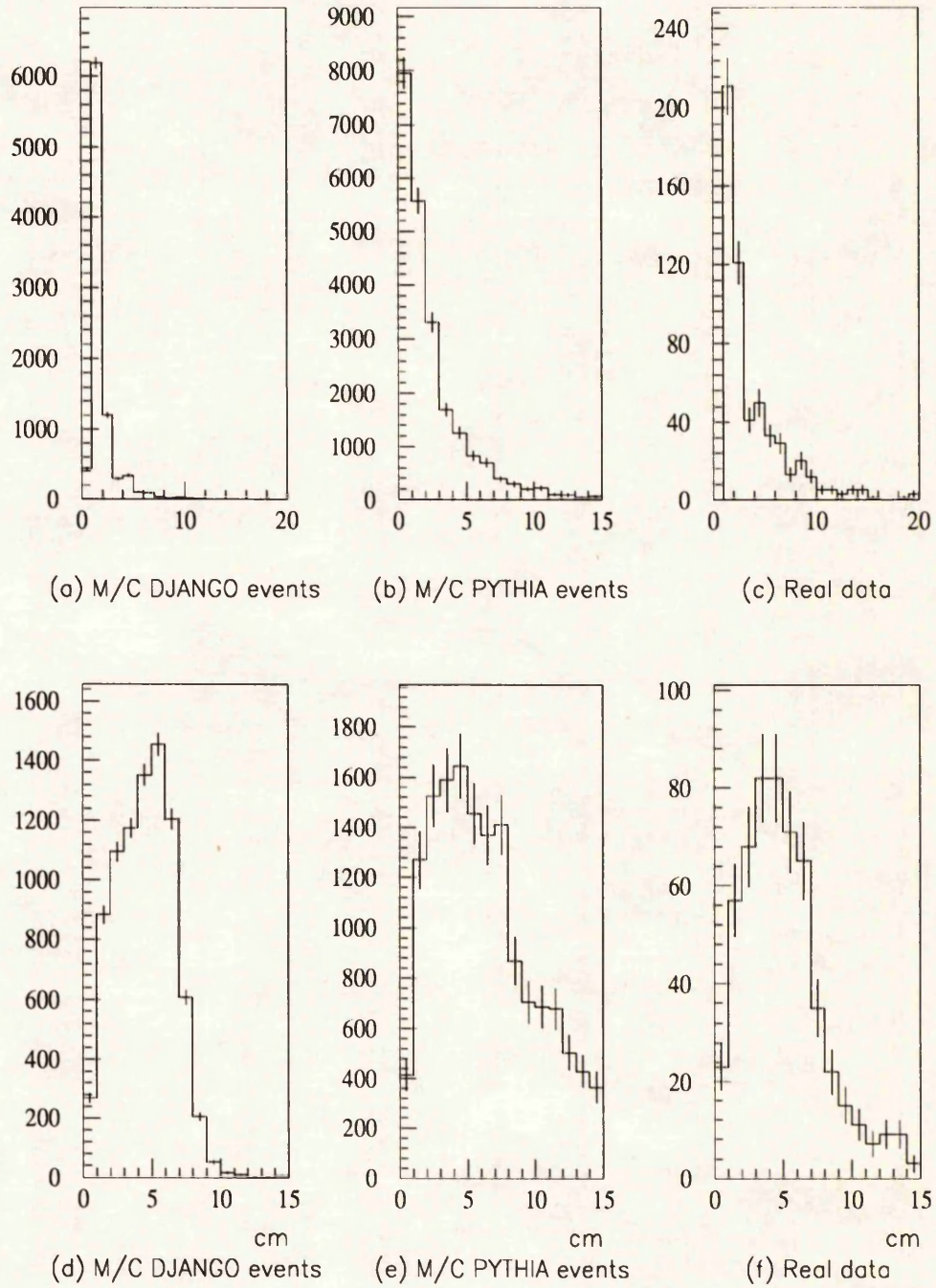


Figure 7.4: (a)-(c) No. of BPC hits within 15 cm of BEMC e^- cluster, (d)-(f) shortest distance BPC hit - e^- cluster

and photoproduction but between ep interactions occurring within the detector and events arising from e/p-beamwall, e/p-beamgas or even ep collisions downstream of it. These occur about 10 ns earlier than ep events inside the H1 volume, and the ToF scintillator logic assigns a 3 bit word to each hit ToF pad as explained in section 4.14, which should be referred to for a decomposition of the words.

The 'No activity' condition is trivial, with the scintillator producing no signal at all. The remaining conditions then state that either at least one 'physics' bit was set with at most one accompanying background flag, and that no event has more than one background bit set.

Any ToF cut is somewhat harder to justify quantitatively, as it is impossible in practice to simulate beamwall/beamgas background effects. For the July data runs used for this analysis there was no great confidence in the performance of the hardware itself. The recipe of ToF bits was chosen on grounds of safety (i.e., erring on the side of accepting more background than was necessary rather than rejecting too many good events) and after visually inspecting some hundreds of real events using the H1ED display package.

This analysis used a slightly different ToF selection. Regarding a physics bit as 'positive' and a background bit as 'negative' the events of early DSTs were inspected and placed on a ToF plane, with axes labelled positive (x) and negative (y). Events lying in one half of the plane, below the line $x = y + 2$ were determined to be poor (i.e. almost certainly not from a ep collision) and those above it good. Exceptions were events in the (0,1) and (0,2) bins, which were largely bad. The ToF rule then was
((No. of -ve bits - No. of +ve bits ≤ 2) *OR*
NOT (No +ve bits *AND* No. of -ve bits ≥ 1)) *OR*
No ToF activity
Although not adopted by the ELAN group as a μDST selection, this ToF rule was used for this analysis as an additional 'user' selection criterion.

7.1.6 Tracking Cuts

The track requirements are similar to the ToF conditions in that they principally bias against up/down stream interactions, although the definition

of a track is designed to assist in photoproduction rejection. Like the ToF selection, its final form was the result of visual scanning of early events. Such scanning revealed the existence of many events which satisfied the BEMC 4 GeV cluster and accompanying BPC hit conditions but which exhibited no track. A portion of these were almost certainly genuine physics events, especially if the BEMC cluster was at $\theta > 170°$ where the central tracking devices have zero acceptance. Clusters with $E_{BCLR} \approx 26.7$ GeV at high angle could be signatures of low Q^2, low y DIS events. Many however were not interesting physics and exhibited a linear shower of hit points nearly parallel to the beam axis from which the reconstruction code had been unable to form a track, or a shower of low (< 2 GeV) BEMC deposits (with at least one ≈ 4 GeV to pass DST cuts), or both. These indicated the passage of a particle or particles from outside of the detector which did not for some reason trigger the ToF scintillator. There existed two complications however. On occasions the reconstruction code would form a track(s) or track segments from what, by eye, appeared to be a near-random scatter of hit points in the central tracker. The simple insistence on a central track would not eliminate these events. Secondly, events were found with a high energy BEMC deposit and BPC hit but at such an angle which indicated that any track linking it to a reasonable vertex must have passed through the entirety of the tracker, and yet no track existed. These problems were possibly the result of details of the reconstruction code. Alternatively, the conversion of a π^0 in the CJC end wall could explain these events and a partial solution was adopted which was to impose at the ECLASS level tighter definitions of a good track. For the central region, a good track was said to exist if it had;

$P_t > 0.2$ GeV

Number of CJC hit points > 16

Distance of Closest Approach (DCA) to the (0,0,z) beam axis < 2 cm

Start radius < 30 cm (start radius is radial distance from z axis to the first point of a track)

The minimum number of hits requirement prevented short track segments from being classed as tracks. The DCA and start radius values were chosen to de-select tracks which did not originate from or near to the *ep* interaction point inside the beam pipe. Non *ep* events were found to give rise to low angle tracks passing through the the outer jet chamber (inner radius = 52.7

cm) which would no longer be accepted. These cuts also discriminate slightly against photoproduction events. Typical of these events are low momentum spiral tracks which do not originate from the axis (and are the the result of a decay), as shown in figure 7.5, from a PYTHIA Monte Carlo simulation. Those spirals which neither start at nor pass near enough to the beam axis fail the new track definition. If no other track exists, as may be the case, the event will accordingly fail. For forward tracks (i.e. tracks registered in the

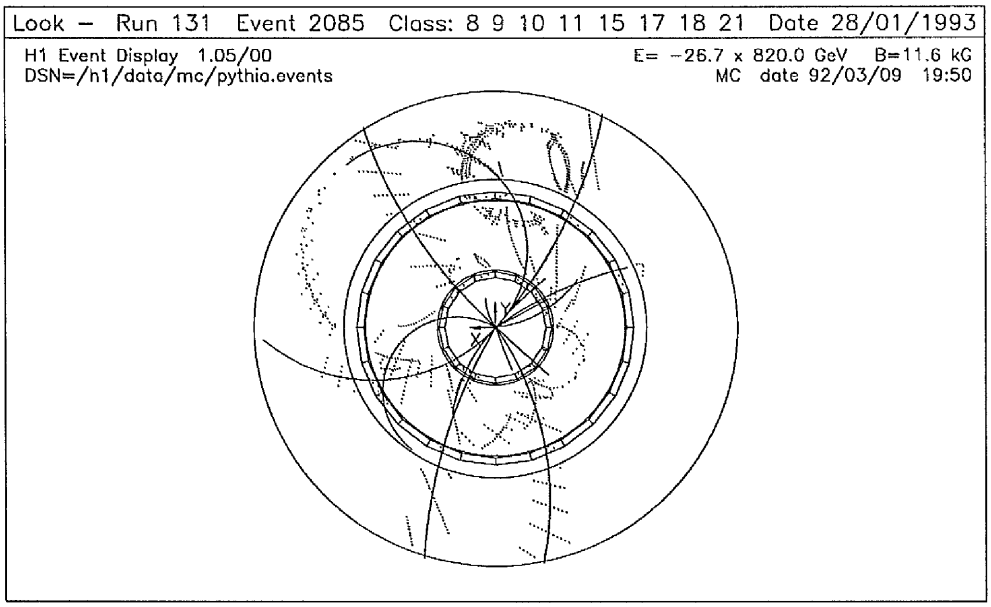


Figure 7.5: Spiral tracks in photoproduction event, view of Central Jet Chamber

Forward Tracking Detector), a cut on the minimum number of hit points set to 10 was imposed, and the final rule for class 10 events was that there should exist at least one good central or forward track.

There exist many more possibilities for more sophisticated track-related cuts. The linking together of central and forward tracks, calorimeter clusters, and BPC hit points and the assignment of vertices is a highly detailed task which is not yet complete. The simplicity of the 'one good track' requirement in the ELAN DST selection code ensured that a reasonably pure physics sample could be assembled which was not reliant upon poorly understood and unstable definitions of tracks, clusters and vertices within H1REC.

The μDST selection for July 1992 data was applied to class 10 and 11

events, and consisted of the class 11 DST definition, and an additional vertex component
(One vertex from either a central or forward track) *OR*
(22 GeV $< E_{BCLR} <$ 32 GeV)
which allows for the low y, Q^2 events with no track discussed earlier. The z vertex is created from central tracks as follows; all reconstructed tracks are projected back to $r = 0$, and the intersection of each with the z axis is histogrammed. A 5 cm wide window is then passed over the histogram, and the region of greatest point density located. The z values within this window are averaged to find the overall z vertex. For the July 1992 data the (x, y) vertex was fixed for a given run, after having determined some average position over a large number of events. This was done so that the stability of the beam from run to run could be observed. It is the z vertex which is important for determining the θ of any track, and this may vary over the length of the central trackers from $\approx +0.5$ m to ≈ -0.5 m, whereas the (x, y) position varies by only a few mm from the (0,0) position. For those events where no z vertex could be assigned, it was assumed to be at $z = 0$. The θ angles of tracks for an event with no proper z vertex then have a large error. The only particle for which this is significant is the electron candidate, either for finding the value of y or for evaluating $E_e - P_{z_e}$.

The BPC part of the μDST selection guarantees that only ECLASS 11 events were processed (it must be again remembered that many events belonged to both classes). The definition of a good track as described was however maintained, and the insistence upon at least one track was included.

From the approximately 2×10^6 events, 6×10^4 satisfied the DIS L1 trigger [57], and the resulting set of DSTs contained 7363 events. Application of the μDST selection reduced this to 649 events, and visual scanning further reduced this to 559 DIS candidates. The events de-selected by eye were those thought to be obvious non ep backgrounds, usually with a high degree of activity in the central tracking detector (CTD).

Albert De Roeck [76] reports that pilot bunch studies reveal the level of beam gas and beam wall events existing in the μDST sample to be <10%, and $< 5\%$ for $E_e \geq 10$ GeV, the level of accuracy obtainable by this method. (Pilot bunches are 'monitor' bunches of either electrons or protons which do

not collide with other bunches.)

The runs used to compile the original POT tapes were those deemed 'good' by data acquisition personnel, i.e. with all systems operational or 'medium' (some systems non-operational or poor reconstruction).

7.1.7 Some Real Events

Diagrams of three real events (from a different μDST data sample to the one used for this analysis) are shown in 7.6. These views, from the H1ED program, are of $r-\phi$ projections of the inner layer of the Lar calorimeter, trackers and BEMC. The electron candidate is the largest deposit in the BEMC calorimeter. The three have been placed in an order of 'increasing doubt' as to their likelihood of being DIS; the first event shows a single, strong track pointing to the BEMC cluster, though this track has been put in by the reconstruction code, and is not the result of any hits in the CTD. The second event has a large amount of tracker activity, characteristic of a photoproduction event's higher multiplicities. The two tracks pointing to the BEMC cluster are not necessarily a cause for concern; H1REC has simply attached two nearby BPC hit points to the event vertex with separate tracks. The third event has a very low (4 GeV) BEMC deposit, and a large amount of CTD hits which are not matched to a track (often, but not in this case, associated with a low momentum spiral). These events are typical of those selected by the ELAN μDST cuts. It is a fruitless task to try to accept or reject large numbers events such as these by visual inspection (although this was done to some extent for the early data), and the work described in the following chapters is an attempt to de-select events in an automatic and more rigorous manner.

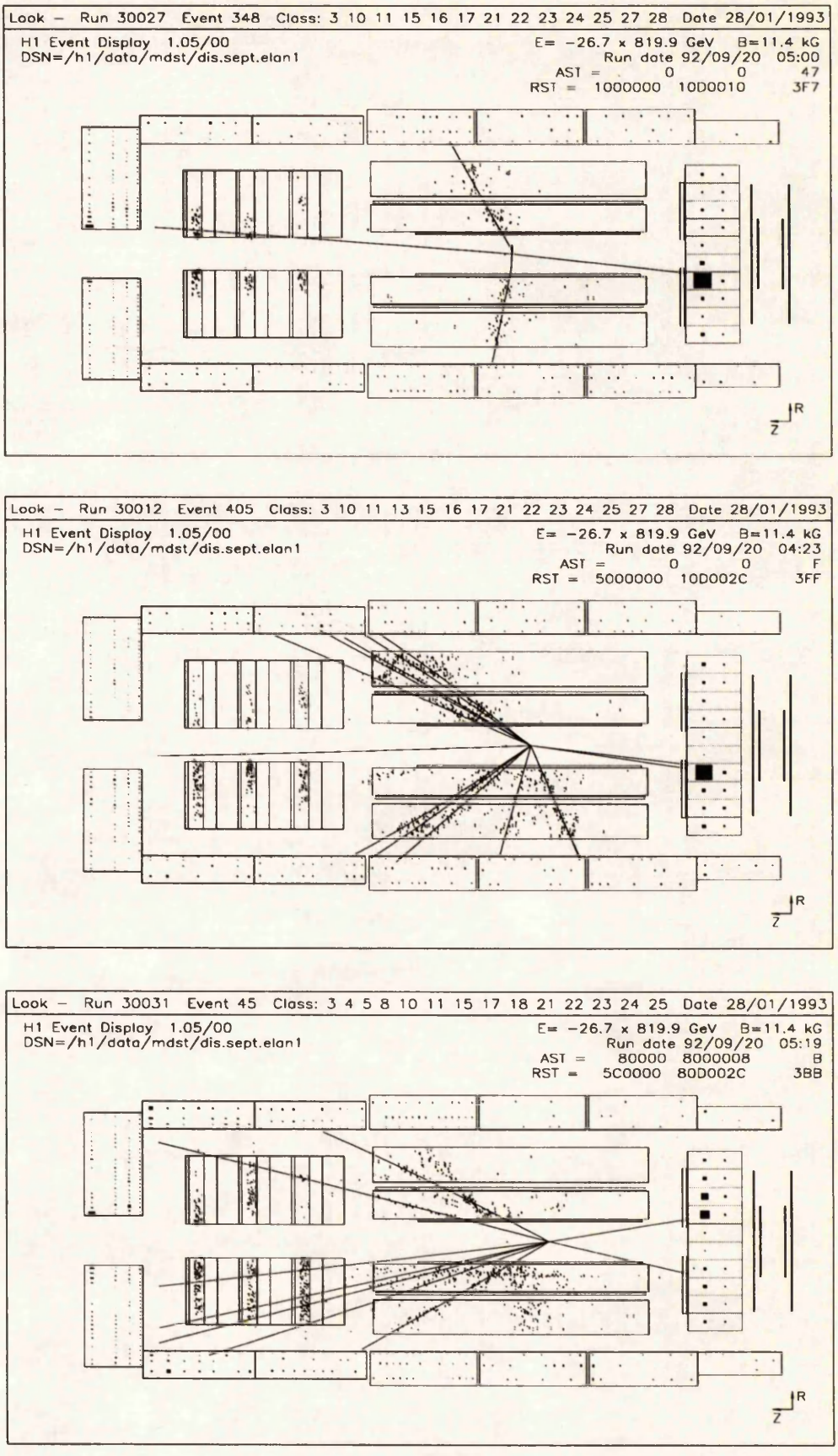


Figure 7.6: H1ED outputs

7.2 Simulating Photoproduction

Having established that photoproduction would constitute a source of false electrons which would reconstruct to high y (low E_{BCLR}) and low x – i.e. the very regions in which the sensitivity to different proton structure functions is at its best– it is a priority to ensure that a realistic Monte Carlo model for the process exists with which to develop some background rejection scheme.

The generator used hitherto for the production of figures such as 6.7, 6.9 was PYTHIA 5.6. PYTHIA as implemented at H1 was intended to generate the *hard* part of the total photoproduction background, that is the perturbative QCD part with the photon either resolved into a $q, \overline{q}$ pair or interacting as a point-like particle. Here, the photon has a partonic content described by the photon structure function which interacts with the proton's partons. Both the photon and proton structure functions are (in theory) selectable by the user. Throughout all chapters, the PYTHIA outputs here were created with the LAC1 [33] photon structure function parametrization and the Morfin-Tung B1 (MT-B1) [25] proton structure function.

It is vital to emphasize that since this study is not primarily one of photoproduction, its treatment of this topic may be somewhat less than exact, and here interest in it is exhausted once there exists a degree of confidence that the principal background to DIS can be simulated sufficiently accurately. It has already been mentioned that there is a significant degree of uncertainty in the photoproduction cross section (section 6.6) and indeed the scale of the hard processes modelled by PYTHIA is largely constrained by HERA data [77]. Any simulation is therefore unlikely to be accurate in its predictions to better than factors of two. Provided that any final model is physically reasonable in terms of current knowledge about the nature of these low Q^2 processes, and makes reasonable predictions on the behaviour of the data in at least some aspects, then this must (initially at least) be sufficient.

7.2.1 Initial Results

The first indications that the existing Monte Carlos for photoproduction were inadequate arose with the advent of real data in June 1992 and the

Sub-process	Cross-section (mB)
All included sub-processes	$267.9 \cdot 10^{-3}$
$qq \rightarrow qq$	$4.41 \cdot 10^{-3}$
$q\overline{q} \rightarrow q\overline{q}$	$3.00 \cdot 10^{-5}$
$q\overline{q} \rightarrow gg$	$6.07 \cdot 10^{-5}$
$qg \rightarrow qg$	$72.7 \cdot 10^{-3}$
$q\gamma \rightarrow qg$	$4.44 \cdot 10^{-4}$
$gg \rightarrow q\overline{q}$	$3.86 \cdot 10^{-3}$
$g\gamma \rightarrow q\overline{q}$	$1.18 \cdot 10^{-3}$
$gg \rightarrow gg$	$185 \cdot 10^{-3}$

Table 7.1: PYTHIA 5.6 generated kinematics

first spectra of the scattered electron energy. The physics simulated by the version of PYTHIA used for most DIS analyses at HERA is summarised in table 7.1. Here, a dataset of 30,000 events was generated.

The total cross-section for this file is thus 267.9 μb. Such a figure corresponds to the most extreme predictions for the *total* photoproduction (hard, soft, diffractive) cross-section. It was however never intended that this value be adopted as a figure to be directly employed in calculations. PYTHIA was deliberately run in a low Q^2 regime which was known to be unreasonable for a program written to model the hard, higher Q^2 processes. PYTHIA was believed to be a good simulator of the harder events, and it was hoped that it could be extended to model softer physics as well by lowering the minimum $\hat{P}_t^{min}$ (see below) used in the generation. This was valid if, and only if, the resulting distributions (of energy, momenta etc.) turned out to be reasonable. The cross-section that the Monte Carlo produced ceased to be believable once the program was operated outside of its designed physics range (although this fact was not always appreciated).

Figure 7.7 uses a log scale to demonstrate the scale of disagreement which exists between the PYTHIA dataset as generated and data. The plot shows the energy spectra of electrons selected as the maximum $E - P_z$ BCLR cluster, with the provisions that $E_{bclr} \geq 4$ GeV, a BPC hit exists within a radius of 15 cm, and that the ToF and tracking cuts outlined in section 7.1 are satisfied. Superimposed is the combined spectrum resulting from apply-

	x	y	Q^2
min	4×10^{-5}	5×10^{-4}	3 GeV^2
max	0.999	0.99	200 GeV^2
cross section			326.6 $\pm$ 1.6 nb

Table 7.2: DJANGO 10 KMRS B- generated kinematics

ing the same selection criteria to the output of DJANGO 10 (for DIS events) and PYTHIA, both having been processed by H1SIM and the same version of H1REC used to reconstruct the real data. The DJANGO file used is summarised in table 7.2. The B- parametrization of the KMRS group [22] was used for this Monte Carlo.

The two Monte Carlo outputs were then normalized (i.e. scaled to the same luminosity as each other), and the combined spectrum was then re-scaled so that the high energy DJANGO peak was the same height as that on the data plot. The kinematic peak at $\approx$ 26.7 GeV is one part of the spectrum which is certain to be well described by DJANGO and figures 6.9 (a), (b) demonstrate that the likelihood of the peak containing false electrons from PYTHIA is minimal, so matching the DIS heights at this point is certainly a valid exercise.

Referring to figure 7.7, it can be seen that the data and Monte Carlo begin to disagree badly below energies of $\approx$ 12 GeV. At 4 GeV, the PYTHIA prediction is 30 times too large. There is a clear need to tune the PYTHIA file to match the data, such that there is an acceptable level of agreement across the energy range. This is not simply a matter of applying an overall scale factor to the PYTHIA output, since this alone would clearly not suffice; the shape of the distributions (i.e. the gradients of figure 7.7 at low energy) differ considerably.

7.2.2 PYTHIA Weighting

PYTHIA assigns to each event a weight according to the value of $\hat{P}_t$ used in the generation, where $\hat{P}_t$ is the transverse momentum of the basic hard scattering process. PYTHIA also requires an input, non-zero value for

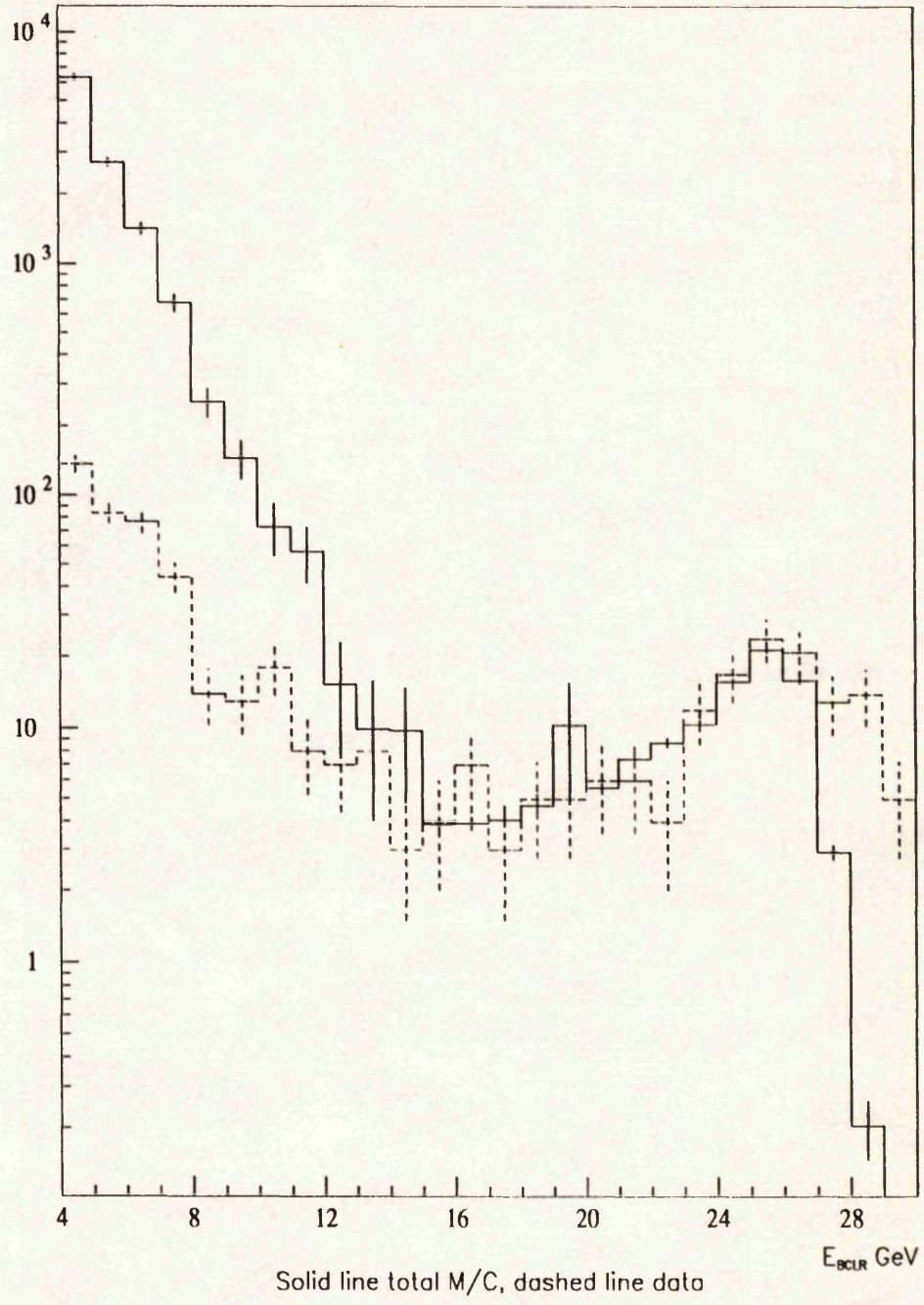


Figure 7.7: Monte Carlo, data e^- energy spectrum

$\hat{P}_t^{min}$ (also called simply P_t^{min}), the minimum allowed transverse momentum of the hard process, necessary for the perturbative calculation of the total jet cross-section. With a zero $\hat{P}_t^{min}$ the cross section would become infinite.

The program as used assigned weights according to

$$W = \frac{0.01}{A \cdot e^{B \cdot \hat{P}_t}}$$

so that a high weight, W, was associated with a low $\hat{P}_t$. The distribution of these weights for the original file is shown in figure 7.8 (a). This weighting was done because the cross section is dominated by low $\hat{P}_t$ events; in order to ensure that a reasonable number of higher $\hat{P}_t$s existed in the output, fewer of the low value events were actually generated, but those that were had a weight W $\geq$ 1 assigned to compensate. All events with a $\hat{P}_t \geq 4$ GeV/c have W = 1, hence the spike at that point. The $\hat{P}_t^{min}$ used for all PYTHIA files was 1 GeV/c.

One approach, once it was recognized that the resulting PYTHIA distributions were not good descriptions of the data, was to re-run the generator with a different value of $\hat{P}_t^{min}$, and to treat this as a free parameter which could be altered until a good match with data was obtained. There are objections to this approach; Hoeger [78] states that $\hat{P}_t^{min}$ is *not* a free parameter, but is entirely determined by the relative contribution of soft (Vector Dominance) photoproduction events. Certainly, if an attempt were made to model the whole range of photoproduction physics, combining Monte Carlos such as LUCVDM [79] (for low P_t vector dominance events) with PYTHIA, then the problem of providing a smooth transition between the programs, and fixing the total cross-section would constrain $\hat{P}_t^{min}$.

In fact, a slightly different strategy was adopted. Rather than re-running PYTHIA with a set of different $\hat{P}_t^{min}$ values (an extremely time consuming process—although large numbers of events may be generated quickly, the detector simulation and event reconstruction stages are lengthy [$\sim$ 2 s/event] and memory-expensive) the existing file was used and the *weight* distribution adjusted.

It was felt that PYTHIA was almost certainly a reliable generator of high P_t events ($P_t > 3$ GeV/c), and so it was intended that those events with

W = 1 should not have their weight altered, i.e. that they should remain with high P_t. The important task was to provide a better model of the low P_t events. The relatively large cross-section of the low P_t component ([80] reports that the relative cross-sections for non-diffractive soft and hard photoproduction are $\sigma_{ND}^{soft}/\sigma_{ND}^{hard} \approx 2.9$) means that only a fraction of these would be required to migrate into the DIS sample to produce a large background.

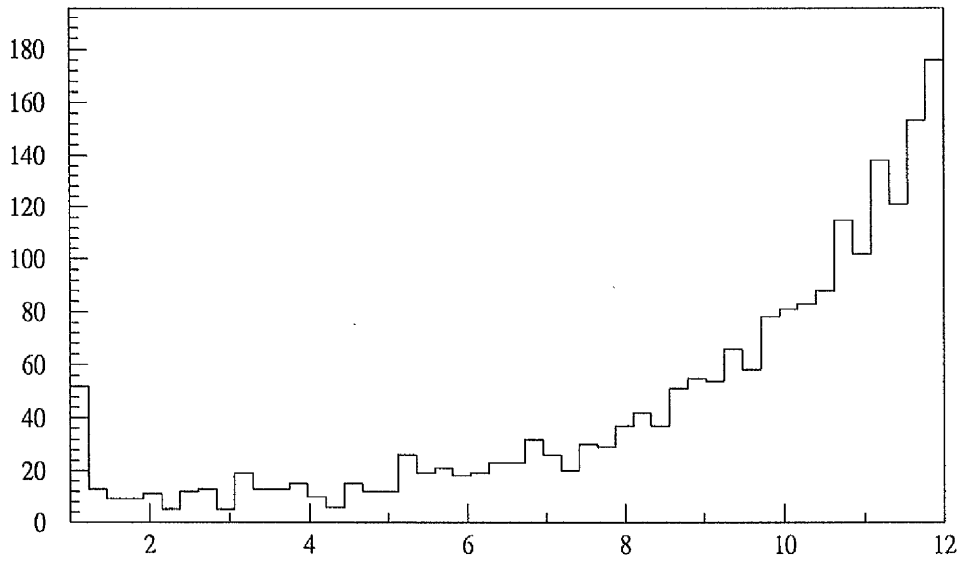
At this early stage of analysis it was thought unnecessary to pay attention to the detailed relative contribution of the various sub-processes, as listed in 7.1. The breakdown into direct and resolved photons, soft and hard VMD was not considered, as the only important factor was to ensure that the resulting spectra of energy, P_t etc. were in agreement with the data.

It was decided not to tune the weighting function to match the identified electron energy spectrum immediately, but to first use some independent experimental data instead. The data used as a 'target' was the single particle P_t spectrum from the OMEGA Photon collaboration experiment at CERN WA69 [81]. This experiment studied single charged- particle spectra arising from (tagged) photon, pion or kaon beams impinging upon a fixed hydrogen target. The experiment studied processes such as photon-gluon fusion ($\gamma g \rightarrow q\bar{q}$), QCD Compton scattering ($\gamma q \rightarrow gq$) and higher-twist reactions ($\gamma q \rightarrow$ Meson q), and investigated the hadron-like component of the photon down to very low P_ts. The P_t spectrum arising from photon beams in the energy range 110-170 GeV was used here (although this was in fact almost identical to the spectrum for data taken in the 70-90 GeV range).

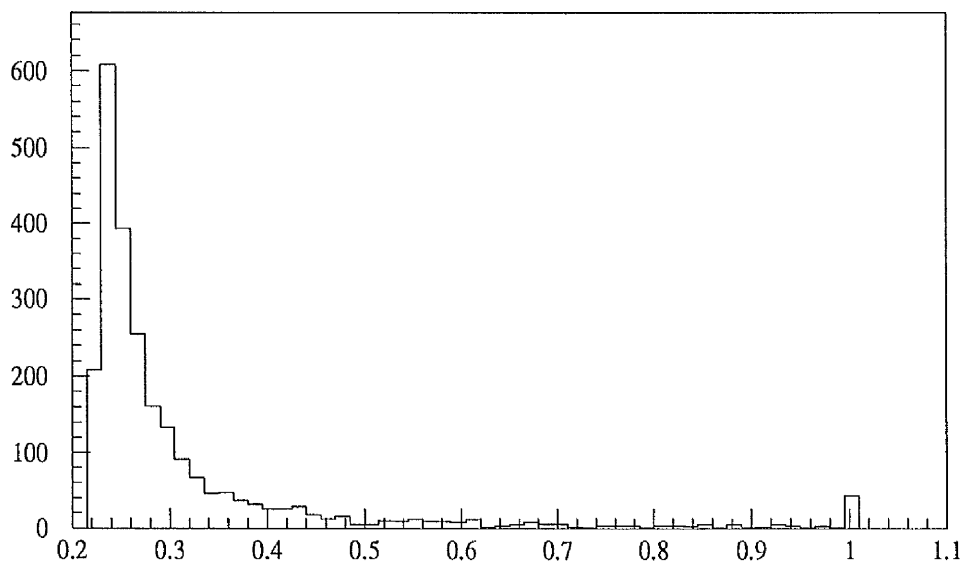
7.2.3 Comparisons with WA69

The results of this operation are shown in figures 7.9 (a)-(f). 7.9 (a) is a plot of the WA69 single particle P_t data. Plots (b)-(f) are of PYTHIA charged pion P_t, with $\approx 70,000$ entries from $\approx 35,000$ events. For each histogram a straight line has been fitted to the range $0.5 < P_t < 2.5$ GeV/c. All the distributions exhibit the feature that beyond $P_t \sim 2.0$ GeV/c they tail off into a flatter shape. The WA69 graph has two lines fitted to it, for $P_t < 2$ GeV/c, $P_t > 2$ GeV/c.

The decision was made to consider the low P_t part of each graph only



(a) Pythia original weights W



(b) Pythia modified weights F(w)

Figure 7.8: PYTHIA original and modified weight spectra

Weight	g, $F(P_t) = A.e^{g \cdot P_t}$	error
W	-5.01	0.03
$\frac{1}{W}$	-4.00	0.05
$F(W)$	-4.19	0.03
$W = 1$ *only*	-2.83	0.2
$W = 1$ *all*	-4.8	0.03
WA69 data	-4.1	0.9

Table 7.3: PYTHIA 5.6 / WA69 P_t spectrum slopes

because it is the low P_t events which are the major contributors to the cross section. Any effort spent matching the high P_t end is largely wasted because those events are few, and in any case will *mainly* be associated to events with high $\hat{P}_t$ (low weight) whereas the uncertainties with PYTHIA are focussed on its low $\hat{P}_t$ events.

The PYTHIA P_t graphs are shown for different weightings. Figure 7.9 (b) uses the weights W as generated, (c) uses the reciprocals $1/W$, (d) a function $F(W)$, (e) plots the spectra for only those events with (generated) weight $W = 1$, and (f) is obtained using equal weights for all events.

Comparing graphs (b) and (e) illustrates the point made above, that the high P_t tracks are largely from the low weight (W=1, $\hat{P}_t \geq 4$ GeV/c) events. The straight lines fitting the $P_t \leq 2$ GeV/c regions have the slopes and errors listed in table 7.3.

Comparing the slopes for the WA69 data and the various weight strategies, it is clear that the best fits are produced by the function F(W) and $1/W$. The function used was

$$F(W) = \frac{1}{W - (W-1) \cdot 0.7} = \frac{1}{0.3 \cdot W + 0.7}$$

which originated from the desire to have a form similar to $1/W$, but keep F(W)=1 if W=1 so that the weight of the high $\hat{P}_t$ entries was preserved. It was fine tuned so the resulting PYTHIA mis-identified electron spectrum matched H1 data better than the simple $1/W$, though the difference is not significant.

The weight spectrum F(W) is shown in figure 7.8 (b), and is heavily peaked at F(W)≈ 0.2, corresponding to low $\hat{P}_t$, with the smaller peak at F(W)=1 still representing events with $\hat{P}_t \geq 4$ GeV/c. Thus, the fit to experimental data has resulted in the relative importance (≡ weight, W) of the low P_t events being diminished. This is a pleasing result, because the quality of simulation must be superior at high P_t, where PYTHIA was designed to operate. The reliance of subsequent results upon PYTHIA's treatment (of uncertain reliability) of the softer physics, is lessened.

The cross-section for the new PYTHIA file can be estimated from

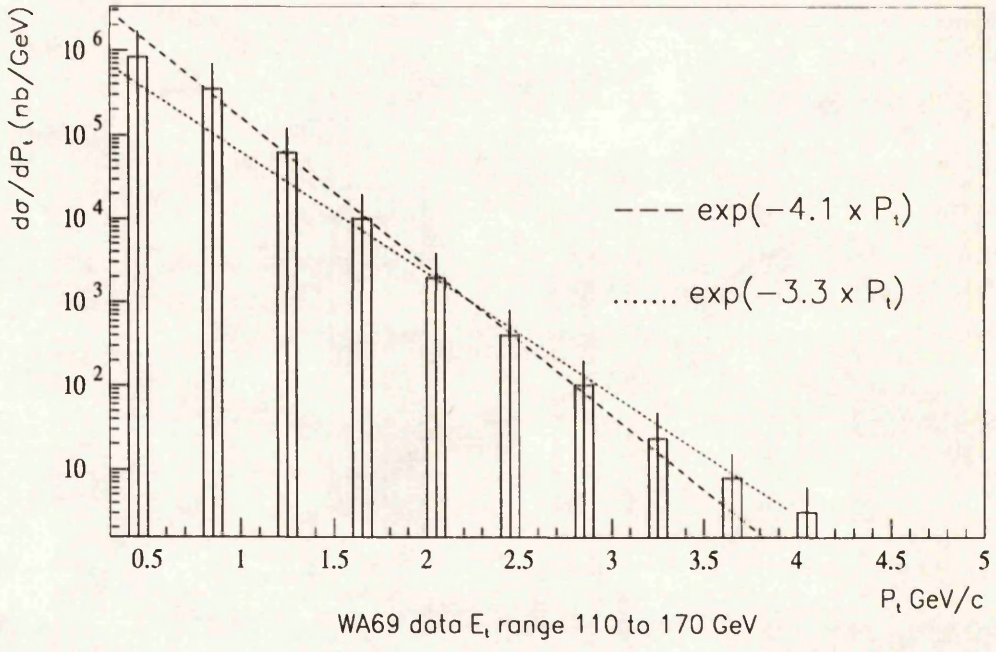
$$\sigma'_{PYTHIA} \approx \sigma_{PYTHIA} \cdot \frac{< F(W) >}{< W >}$$

where σ'_{PYTHIA} is the cross section of the re-weighted event file and $< W >$, $< F(W) >$ the means of the original and modified weight distributions. This gives a value of $\sigma'_{PYTHIA} \approx 8.6$ μb. This value is to be compared with 9.56 μb used by Besancon [43], and with the statement by Albert De Roeck [82] that in the HERA energy range, the hard photoproduction cross-section is expected to be $\sim$ 16 μb with an uncertainty of $\sim$ 100%.

It is this re-weighted PYTHIA which is used for the remainder of the analysis, on the basis that it has a P_t distribution closely following that measured by an independent experiment. The electron candidate energy spectrum which the modified PYTHIA and existing DJANGO files produce when combined is given in figure 7.12. Compared to 7.7 the benefits of the re-weighting are transparent. The Monte Carlo prediction is too large for the 4-5 GeV bin, and slightly too small for $E_{BCLR} \geq 6$ GeV, but a considerable improvement.

7.2.4 Further Work

At a late stage of the work described in chapters 6 and 7 data from the H1 experiment itself was made available which allowed the validity of the above operations with PYTHIA to be checked. The details of the work at HERA on photoproduction physics are to be found in [77]. Reproduced from [77] and [60] are figures 7.13 (a) and (b). Figure 7.13 (a) shows the ${P_t}^2$ spectrum for charged tracks in the central tracking chamber, i.e. $30° < \theta_{track} < 150°$ and $P_t > 0.3$ GeV/c as dots. Superimposed are the PYTHIA predictions for



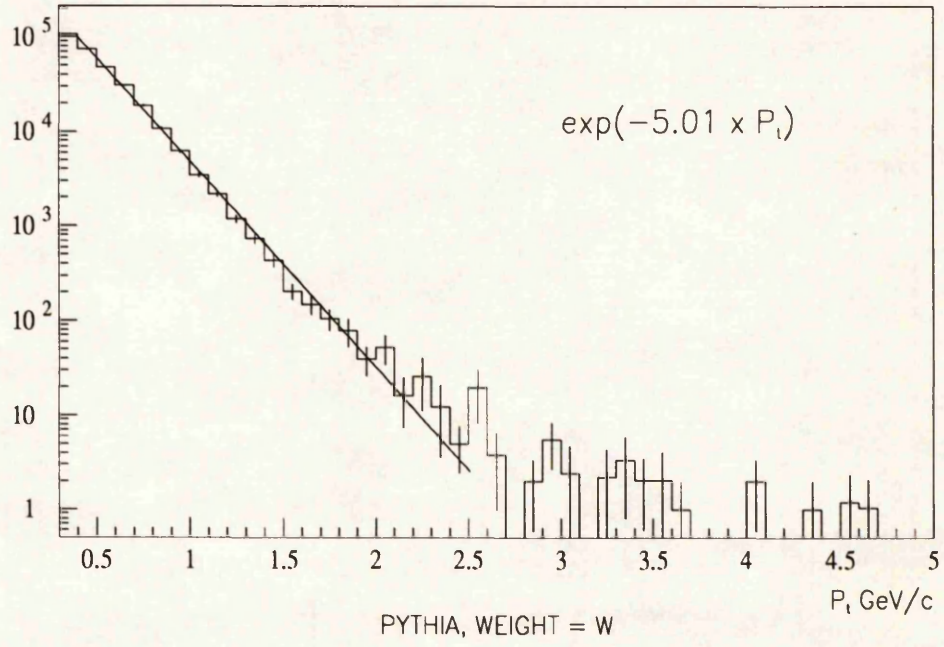


Figure 7.9: (a),(b) WA69 data and PYTHIA P_t spectra

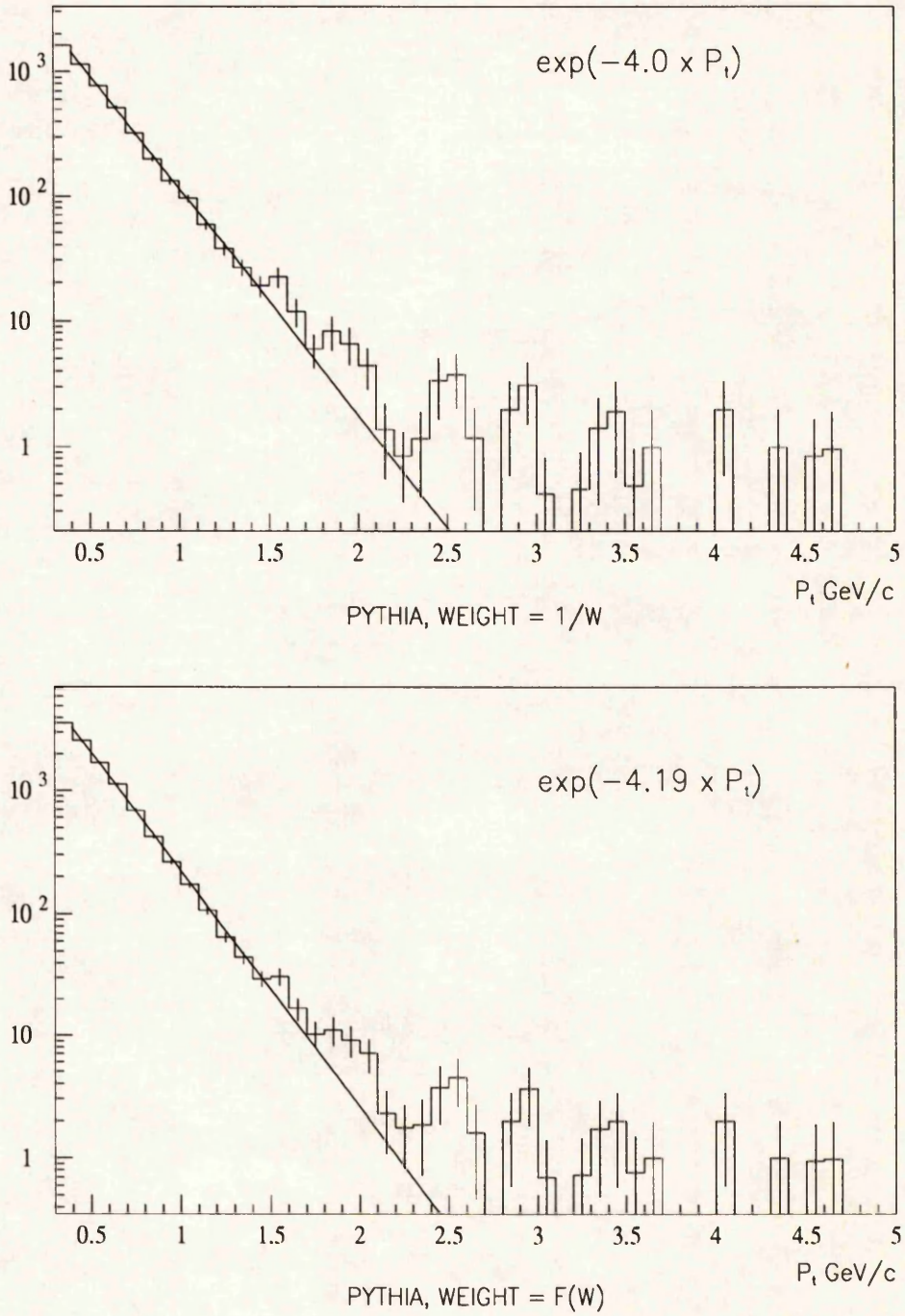


Figure 7.10: (c),(d) PYTHIA P_t spectra

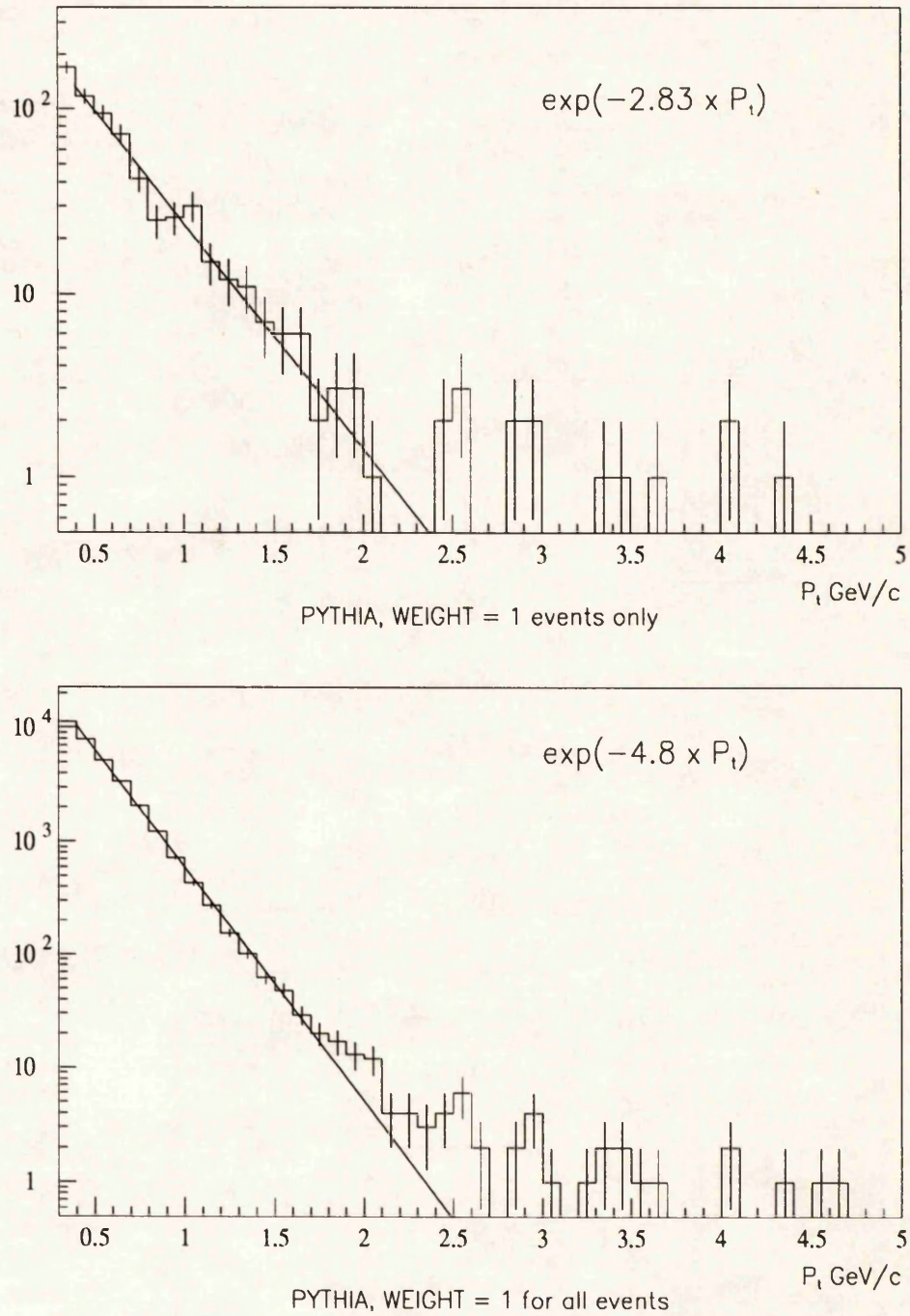


Figure 7.11: (e),(f) PYTHIA P_t spectra

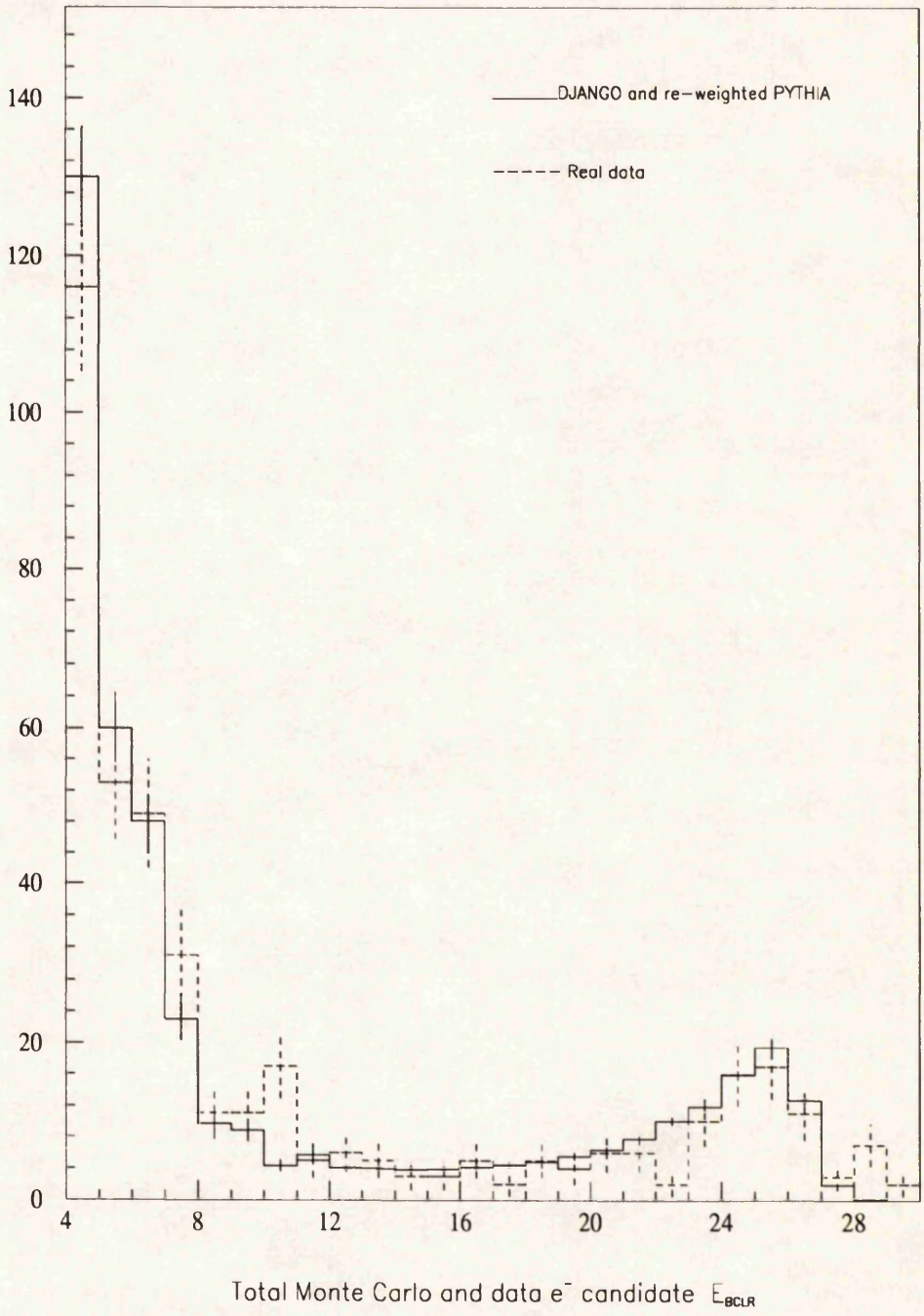


Figure 7.12: Reweighted PYTHIA, DJANGO and data energy spectra

the same distribution; the solid line was created using PYTHIA with a cut on $\hat{P}_t^{min}$ of 2.0 GeV/c, and the dashed line with a cut of 1.0 GeV, i.e. the original PYTHIA file. Not shown is a similar line for $\hat{P}_t^{min} > 2.5$ GeV/c, which lies slightly above the data points, (i.e. is less steep). This confirms that PYTHIA will accurately portray the data if generated with a cut $2 < \hat{P}_t^{min} < 2.5$ GeV/c. Figure 7.13 (b) shows the same P_t^2 spectrum for the original PYTHIA file with $\hat{P}_t^{min} = 2$ GeV/c and the re-weighted PYTHIA as used for this study. The pleasing result is that the WA69 scaling procedure did indeed produce an output which is equivalent to a $\hat{P}_t^{min}$ cut of ≈ 2 GeV/c, which has been subsequently shown to match the data well; but which used an independent set of results and did not require placing an explicit $\hat{P}_t^{min}$ cut, leading to a reduction in the number of PYTHIA events (leading to larger errors).

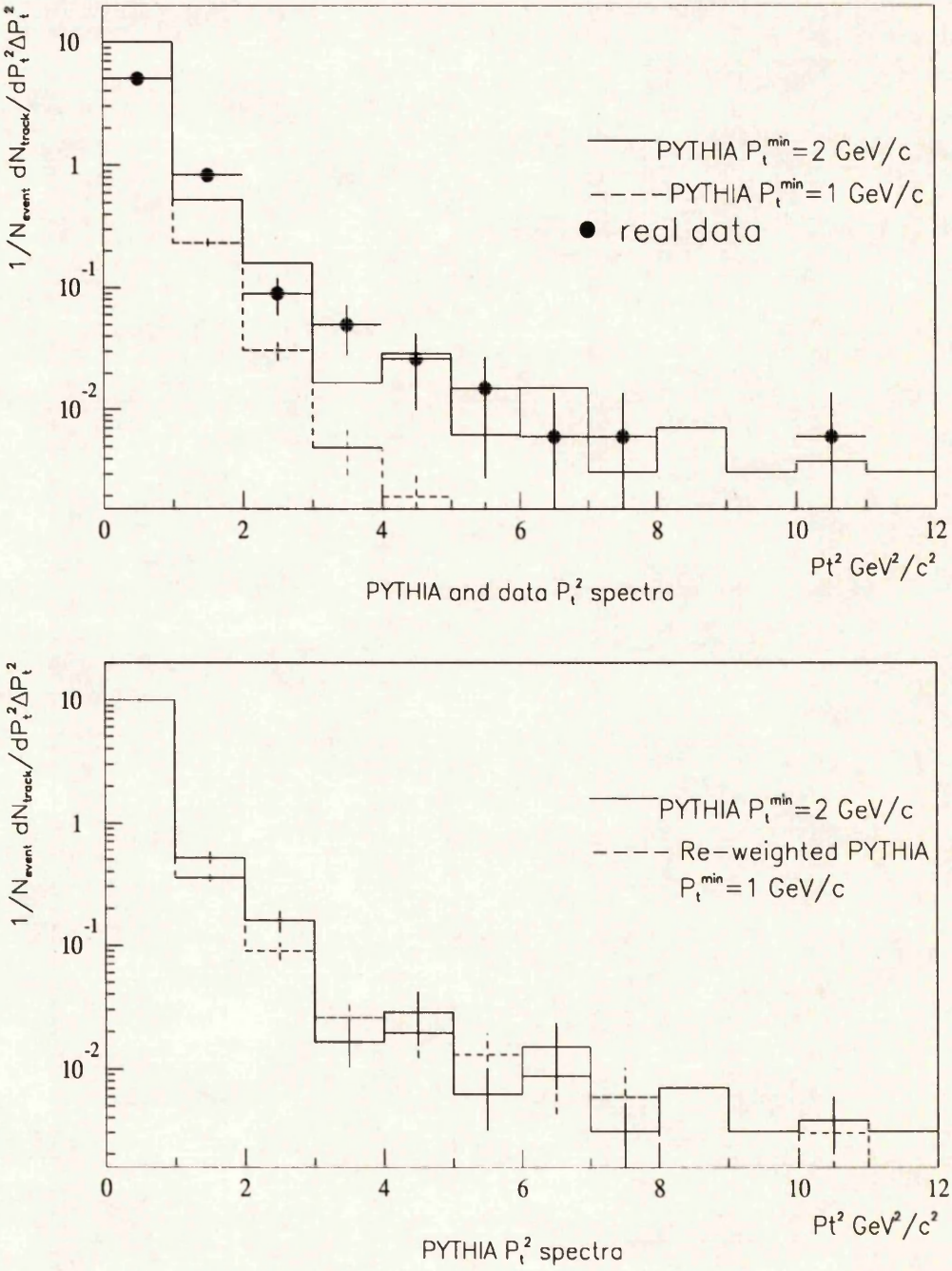


Figure 7.13: PYTHIA and Data P_t^2 spectra for single charged hadron tracks

7.3 Background Removal

Having developed a Monte Carlo that appeared to simulate the low energy photoproduction background to reasonable a level it was now necessary to supress these events relative to DIS. Once the background removal had progressed to such a stage where there were no PYTHIA events remaining, and the data and Monte Carlo DIS files agreed to within errors, this part of the analysis was regarded as complete.

At all times during this operation, the final task — of attempting to distinguish between different structure function parametrizations — was held in mind. This required that events at low Q^2, small x be preserved as far as possible, because it is these events which form the tail of the scattered electron's energy spectrum. Section 2.5 showed how the ratio of the height of the kinematic peak to the tail height is a parameter which is sensitive to the underlying structure function, and so enough of the tail must still exist to enable this to be calculated.

7.3.1 Energy Cut

The above caveats ruled out what would certainly be the most effective anti-photoproduction cut of all, an energy cut at $E_{BCLR} \geq 12$ GeV. For studies entirely restricted to calibrating the BEMC by matching the kinematic peak height to the Monte Carlo or studying the resolution , this cut would be entirely sufficient. It was found to be impossible however to avoid one further direct energy cut (over the 4 GeV threshold for electron definition) and the sharp rise of the background peak at around 10 GeV suggested a value around that point, and the value of 7 GeV was eventually chosen. Although this is the first cut in the sequence described here, it must be realised that all the methods in this section were developed and tested in parallel; the choice of 7 GeV was less arbitrary than it may appear, and was made once it was established that all other techniques could not assist below this point, and that the bins 4-6 GeV were still heavily contaminated by PYTHIA events.

The immediate effect of the energy cut on the accessible x, y, Q^2 ranges can be observed from figure 7.14, where the cut was applied to DJANGO

DIS events in combination with the ELAN μDST cuts. It acts as a y cut as expected, limiting $y \leq 0.74$, but does not restrict the small x range too severely, removing $\approx 40\%$ of events at $x \leq 10^{-4}$. The only part of the hitherto accessible (x, Q^2) plane which is removed totally is the high y band.

7.3.2 P_z, P_t Cuts

The easiest quantities to study in an effort to reject the background are simple kinematic ones, in much the same way that the electron identification problem was approached. Figure 7.15 displays the P_t, P_z spectra for DJANGO and PYTHIA. In the high angle range $154° \leq \theta_e \leq 176°$ the P_z spectrum is in effect the energy spectrum reflected about the y-axis, and so yields no more useful cut possibilities. A cut on P_t at 2 GeV/c would seem reasonable, but is also correlated with the energy cut; 80% of those DJANGO events with $E_e < 12$ GeV have $P_t < 2$ GeV/c, and would be lost by such a cut.

7.3.3 E_Δ Cut

One possible quantity of interest is the amount of energy deposited in the innermost BEMC stacks. Photoproduction events consist of a spray of high angle hadrons and photons, and it was thought likely that much of their energy would be concentrated within a small radius around the beam pipe. This would however also apply to high y DIS events, for which the quark jet is at high angle. The angle of the quark can be expressed as

$$\theta_j = 2 \cdot \sin^{-1} \left(\frac{1}{1 + \frac{x.E_p.(1-y)}{y.E_l}} \right)^{\frac{1}{2}}$$

(where E_p is the initial proton energy = 820 GeV, E_l the initial electron energy = 26.7 GeV) and it is trivial to see that $\theta_j \rightarrow 180°$ as $y \rightarrow 1$ at fixed, low x. One attempt to eliminate background events was to cut on the total energy found in the BEMC stacks immediately adjacent to the beam pipe, at the risk of losing some high y, small x genuine DIS. The innermost ring of the BEMC consists of the triangle stacks, see fig 4.3, and the summed energy in them is referred to here as E_Δ. Low Q^2 events in which a near-kinematic

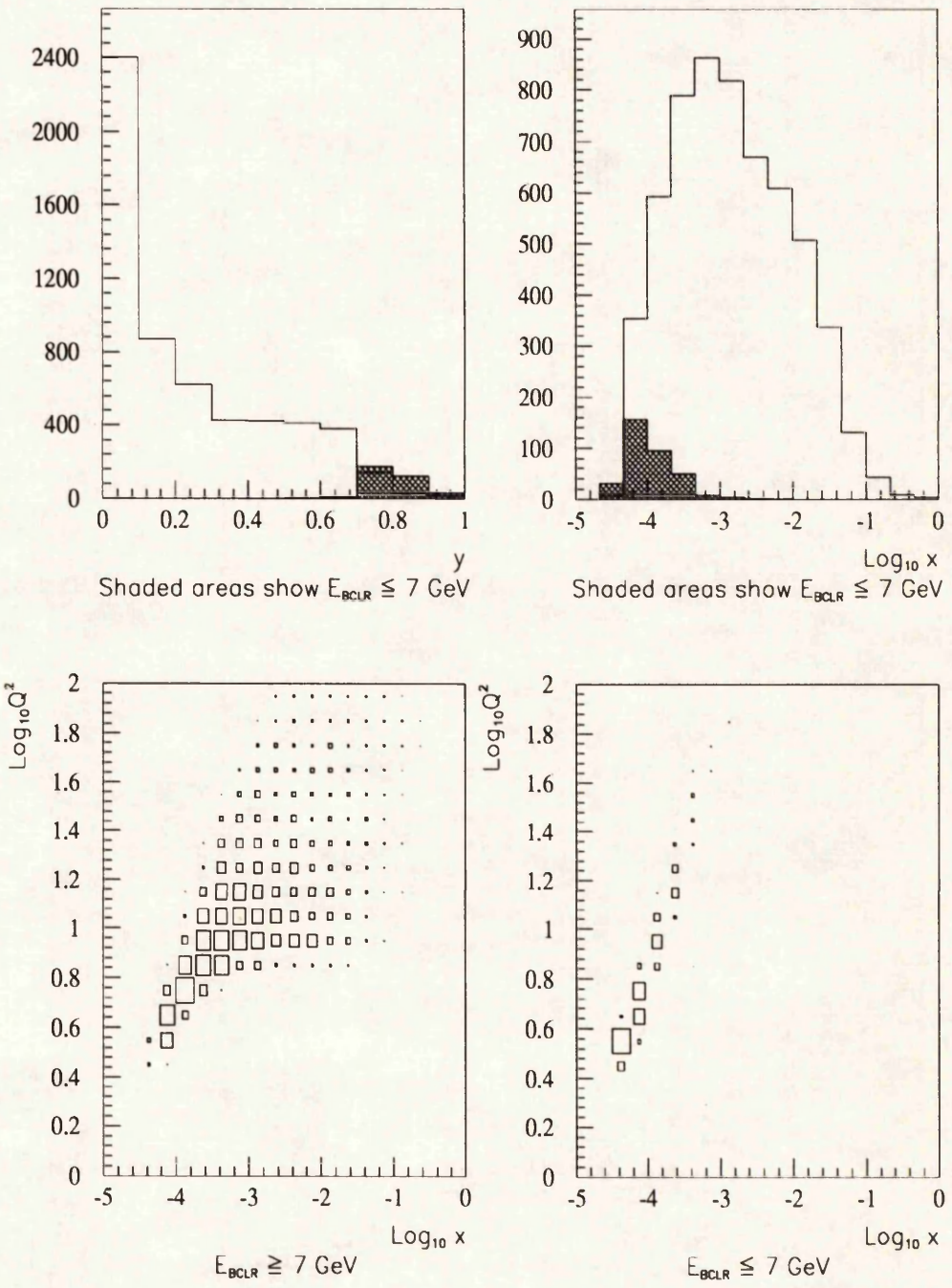


Figure 7.14: Effect on x, y of e^- energy cut at 7 GeV DJANGO DIS

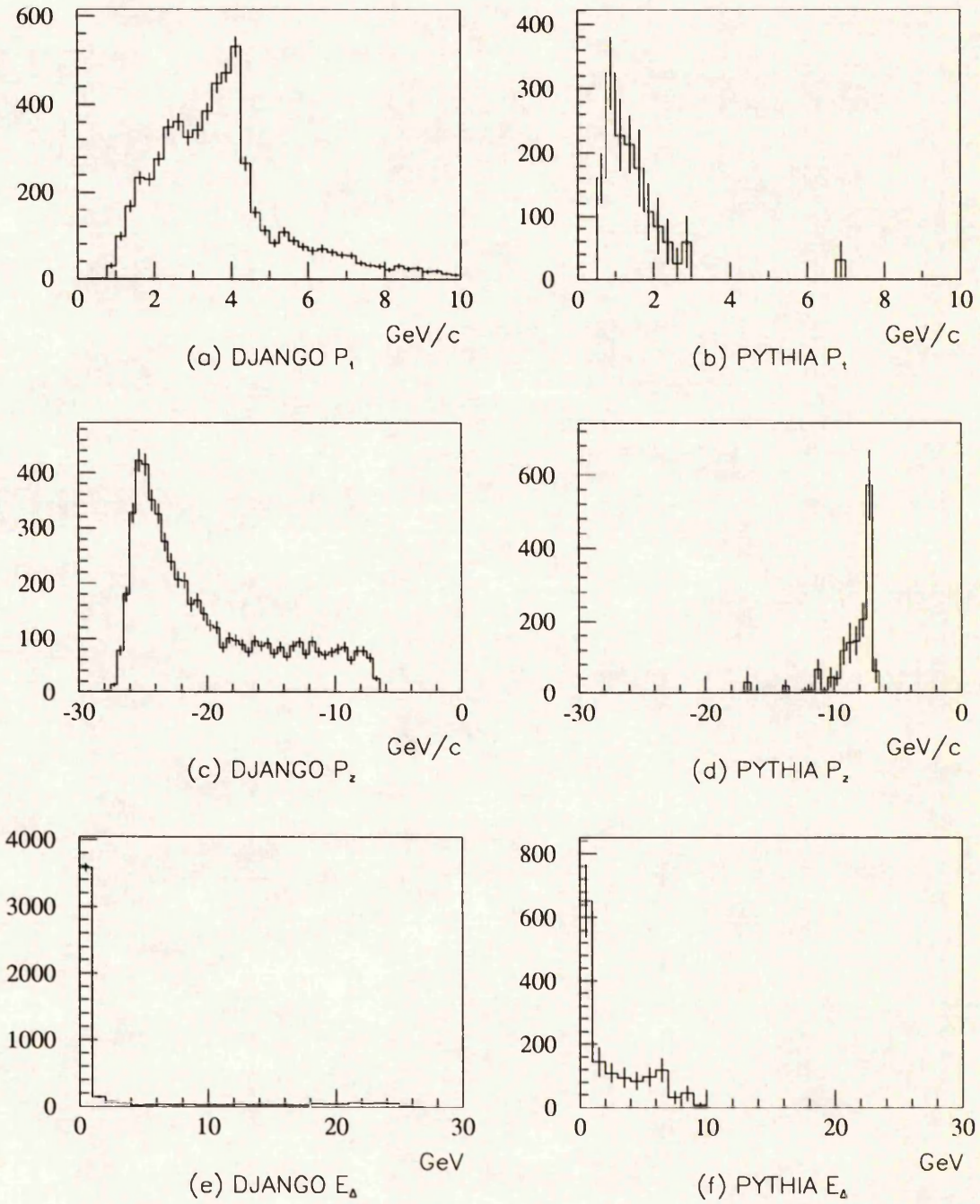


Figure 7.15: P_t, P_z and E_Δ spectra for Monte Carlo

peak electron was at $\theta_e \geq 171°$, within the triangle stacks, may also have been lost by such a cut. Figure 7.15 (e), (f) shows plots of E_Δ. PYTHIA events show a much greater tendency to deposit energies up to 10 GeV in the triangle stacks, with the long tail of the DJANGO distribution including low Q^2 kinematic peak events. A cut on some low value of E_Δ ($\approx$ 2 GeV) removes only $\approx$ 20% of DJANGO events with $E_e < 12$ GeV whilst allowing $\approx$ 60% of PYTHIA events to survive. (The quantity E_e/E_Δ exhibits a similar behaviour.) These figures are quite acceptable, and of potential use when combined with other cuts, but there is a good reason to avoid use of the inner BEMC stacks; they are the least-well understood components of the BEMC calorimeter in terms of their calibration and performance, and the part most susceptible to beamwall/beamgas backgrounds (see section 7.4.1).

7.3.4 E_{semi} cut

The next cut tried to make use of the differing topologies of DIS and photoproduction events. Figure 7.16 illustrates the principle on which it operated. The electron candidate and the hadronised quark jet are shown projected onto the circle of the BEMC. The BEMC was then split into 2 semicircles along the line perpendicular to the candidate's ϕ radius, such that one semicircle has the candidate inside, and one does not, labelled S_e and S respectively. The energy of all clusters inside S_e *except* that of the candidate was then summed, the result being E_{semi}. In figure 7.16, $E_{semi} = E_x$. For DIS events the electron was expected to be isolated from the hadrons, and so $E_{semi} \rightarrow 0$. For photoproduction events, the electron candidate is simply the $E - P_z$ maximum cluster within a dense shower of particles, and $E_{semi} \neq 0$.

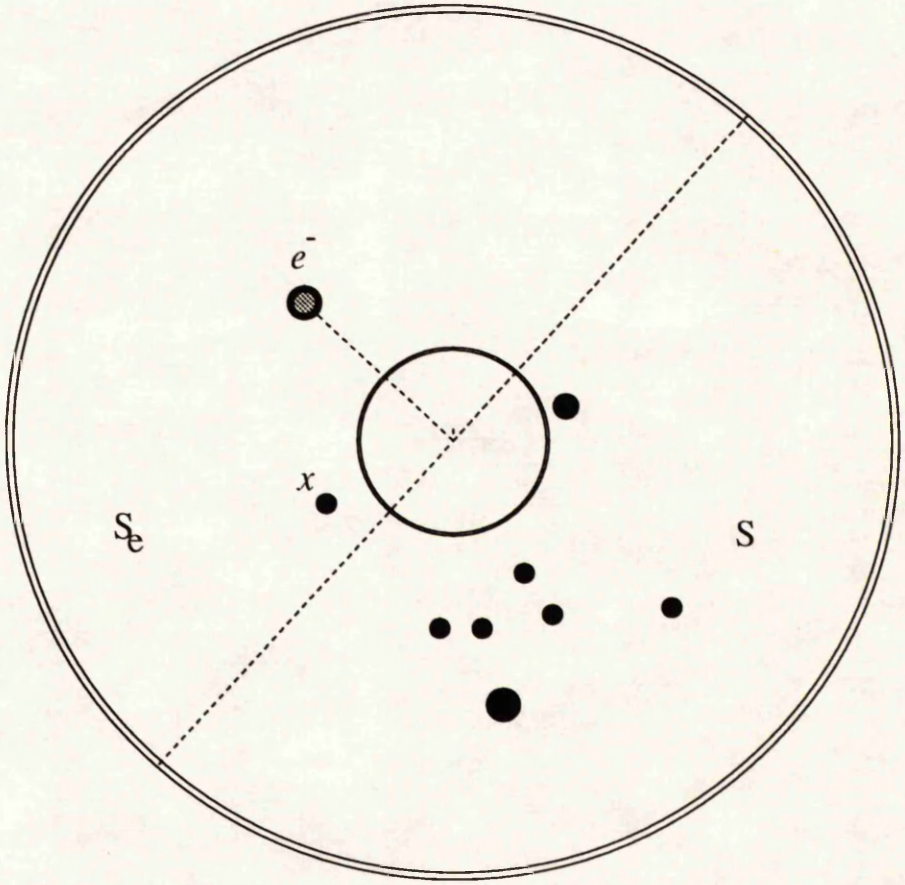
The likely performance of this could be estimated immediately, given that it would fail in the same manner as all the electron identification methods described in section 6.1; the problems arose when the quark jet and electron within DIS become close in angle, destroying the electron's isolation. This is confirmed by figure 7.17 (a)-(c), showing histograms of E_{semi} for DJANGO and PYTHIA files, and a plot of the average value of E_{semi} (for the DJANGO events) on a $\log x, y$ plane (using *generated* kinematics). As expected, the higher averages are for events at small x, high y. The same explanation applies

once more; at high y the current jet is at high angle, and at low Q^2 (i.e. at small x for a given y), the electron is in a similar angular range. The overlap of the two renders these DIS events indistinguishable from the photoproduction background. A cut on E_{semi} at $\approx$ 1 GeV would remove a larger fraction of background, but once again at the expense of some of the most interesting DIS events.

7.3.5 $\sum E - P_z$ Cut

The earlier look at P_z and P_t spectra showed that these offered no way forward, being too correlated with the electron energy, and $E - P_z$ for the electron candidate is similarly of no use. Figure 7.18 (a), (b) shows two more distributions, this time of the quantity $\sum E - P_z$, for the two Monte Carlos after the ELAN μDST selections. This is the sum over $E - P_z$ for *all* particles, over all angles. It differs from all previously studied entities in that it necessarily uses the liquid argon calorimeter for particles outside the BEMC angular region. The DIS spectrum is peaked at $\approx$ 53 GeV, twice the initial electron energy. Since $E - P_z = E \cdot (1 - \cos\theta)$, backwards travelling particles contribute an amount always greater than their energy, increasing with angle. DIS events with a high angle electron and a low P_t, forward travelling, spectator quark jet therefore have $\sum E - P_z \approx 2 \cdot E_e$, explaining the peak. Any detected hadrons will raise the sum, increasing with increasing quark jet angle. The loss of particles down the rear beam pipe depletes the sum by approximately twice their energy, the loss of the spectator quarks down the forward beam pipe does not affect $\sum E - P_z$ significantly. The graphs suggest that a cut of $\sum E - P_z > 40$ GeV is possible to disallow $> 50\%$ of photoproduction. In preference to this however, the cut was combined with another parameter implicitly suggested by both this discussion and the previous considerations of E_{semi}, namely the total energy in the BEMC. Figures 7.18 (c), (d) show the summed BEMC energies, $\sum E_{BCLR}$ for the same files, which are very similar in shape to the spectra of the candidate electrons (because the electron is invariably the highest energy particle, and dominates over all other BEMC clusters in magnitude).

The correlation between $\sum E - P_z$ and $\sum E_{BCLR}$ is demonstrated in



Plane view of BEMC face

$E_{semi} = E_x$

Figure 7.16: How E_{semi} works

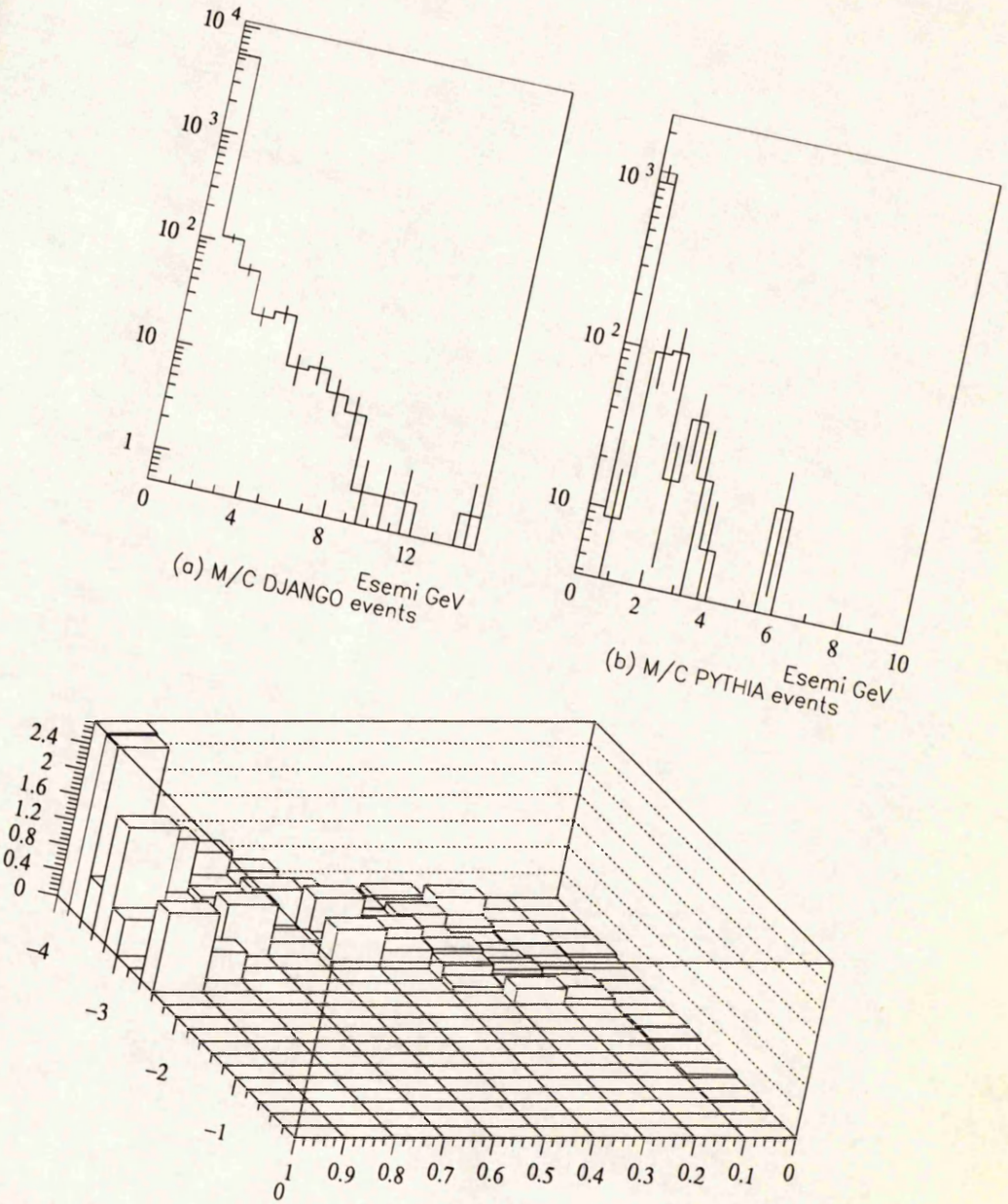


Figure 7.17: E_{semi} for Monte Carlo

fig 7.19. The box sizes are here representative of the number of events in a given bin. Referring to the DJANGO plot, the majority of events lie along or under the line $\sum E - P_z = 2 \cdot \sum E_{BCLR}$. This is easy to understand for the case where the electron is the only particle in the BEMC. It contributes an amount $\leq 2 \cdot E_e$ to $\sum E - P_z$, which is then raised by non-BEMC calorimeter hits. A few events ($\sim 1\%$ of the total) which lie above the line are those events where the electron was the only detected particle, and $\sum E - P_z = E_e - P_{z_e}$. The same plot for the PYTHIA events shows a band of entries with low total BEMC energy confined to a narrower range of $\sum E - P_z$.

A cut on a region of the graph was then placed such that events were rejected if

$$\sum E - P_z < 40 \text{ GeV} \quad \text{AND} \quad \sum E_{BCLR} < 13 \text{ GeV}.$$

The values were chosen to eliminate the highest density region of entries on the PYTHIA graph. Studying the final plot (c) which uses real data shows that $\approx 20\%$ of events will fail this cut.

The consequences introduced by this cut on the (x, Q^2) plane are demonstrated in fig 7.20 (a)-(d). The $\log Q^2, \log x$ plane for DJANGO and PYTHIA events (with both kinematic variables calculated from the electron) before and after the $\sum E_{BCLR} / \sum E - P_z$ cut are shown. It is obvious that no serious reduction in the number of DIS events occurs even at low Q^2 and small x, whilst over 50% of the photoproduction background is removed. In fact, 99% of DJANGO events with $E_e > 12$ GeV and 80% of those with $E_e < 12$ GeV pass, with 65% of the PYTHIA events failing. This is a better performance than the previously most promising cut (E_Δ), and does not rely on the triangle stacks.

If the PYTHIA model of photoproduction is to be useful, then it must not only describe the background well at the start of the analysis, but also after the application of any cut. (This must be particularly so for a simple kinematic cut which does not depend on H1SIM's ability to simulate more detector-specific details such as the performance of particular BEMC stacks or BPC resolutions.) The combined Monte Carlo and data spectra after the $\sum E_{BCLR} / \sum E - P_z$ cut are shown in figure 7.21, and should be compared with figure 7.12. This demonstrates that the re-weighted PYTHIA does indeed

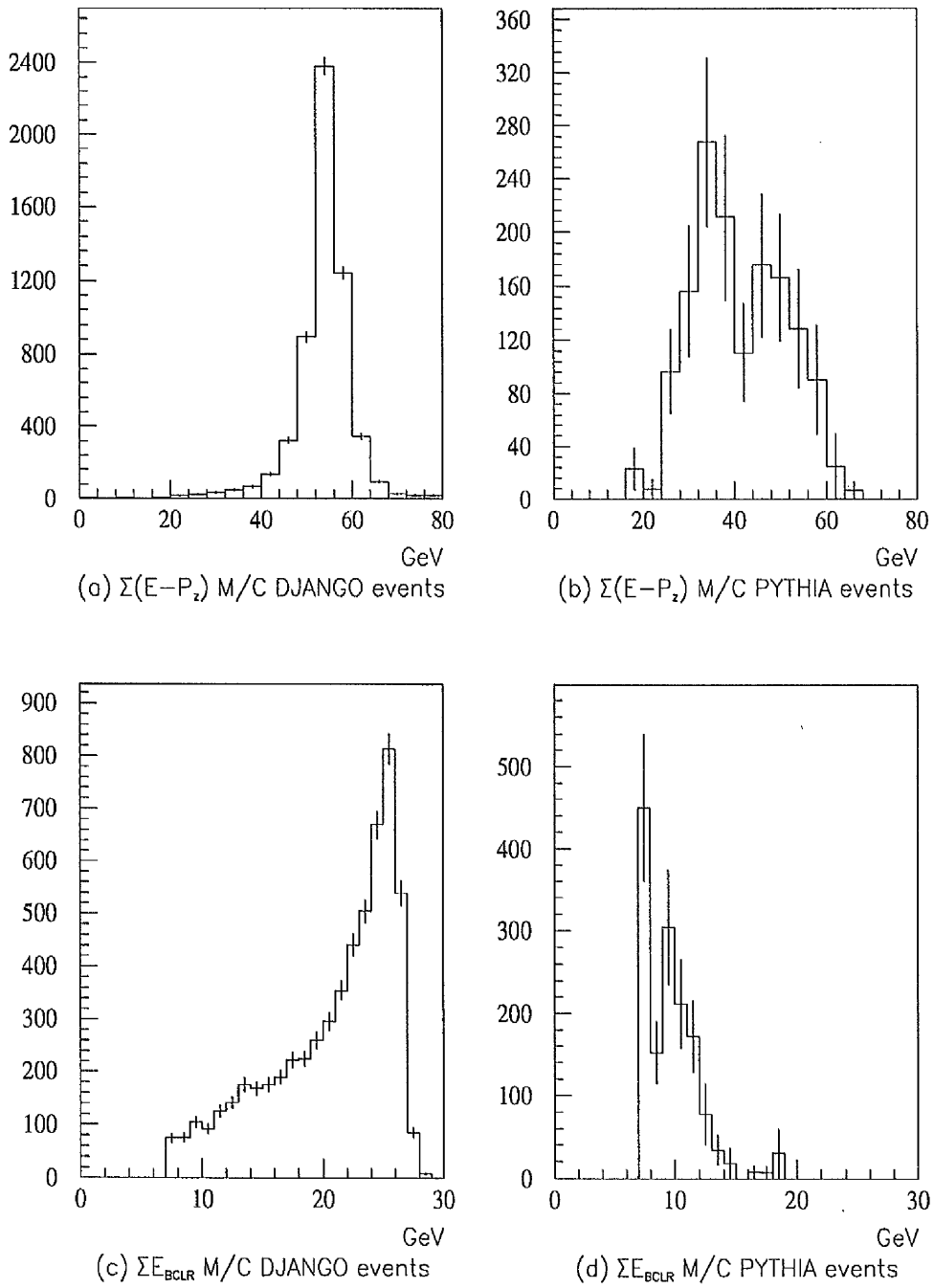
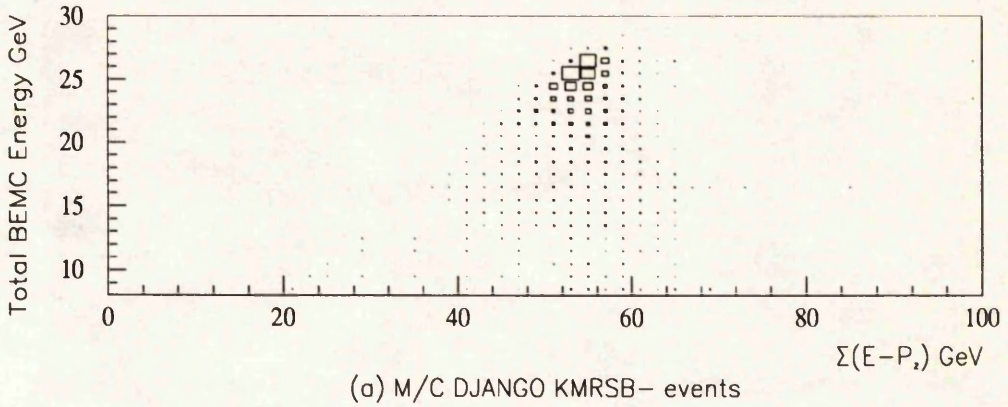
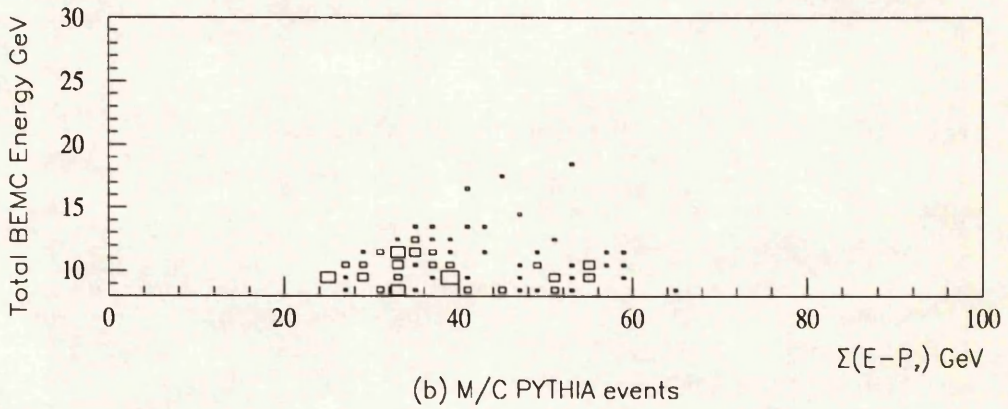


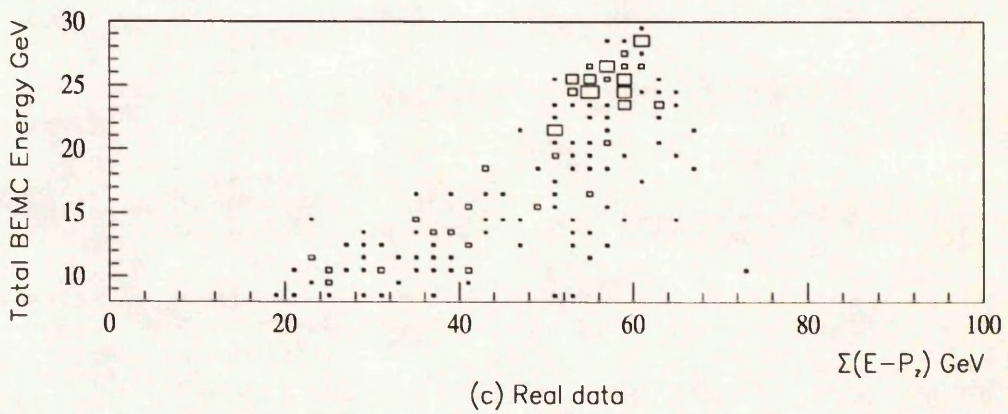
Figure 7.18: $\sum E - P_z, \sum E_{BCLR}$ for Monte Carlo



(a) M/C DJANGO KMRSB− events



(b) M/C PYTHIA events



(c) Real data

Figure 7.19: $\sum E - P_z / \sum E_{BCLR}$ correlation

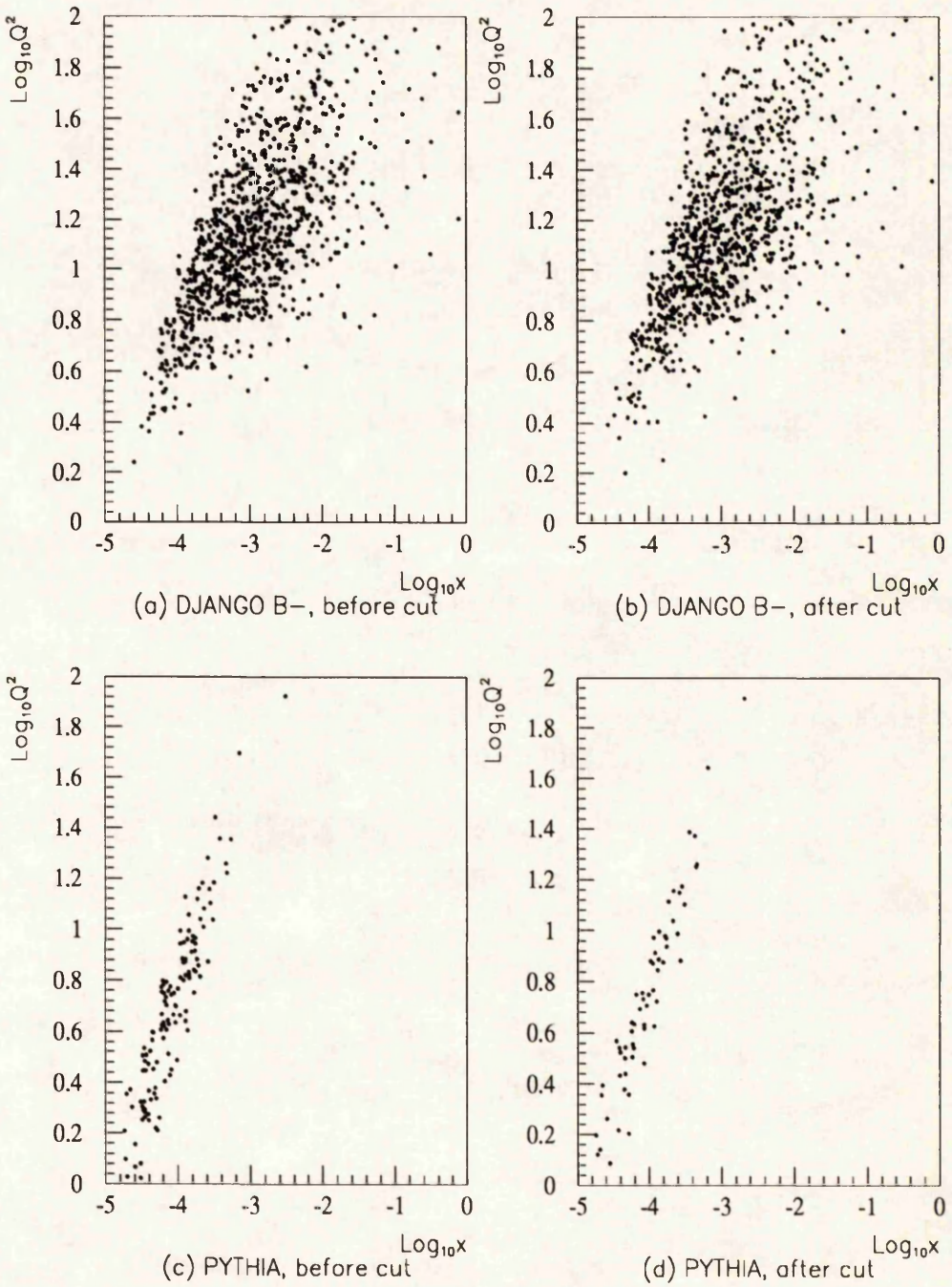


Figure 7.20: Effect on kinematic plane of $\sum E - P_z / \sum E_{BCLR}$ cut

continue to follow the true background quite accurately, and implies that its kinematic (if not its topological) description of the events is a good one, quite adequate for the purposes of this analysis.

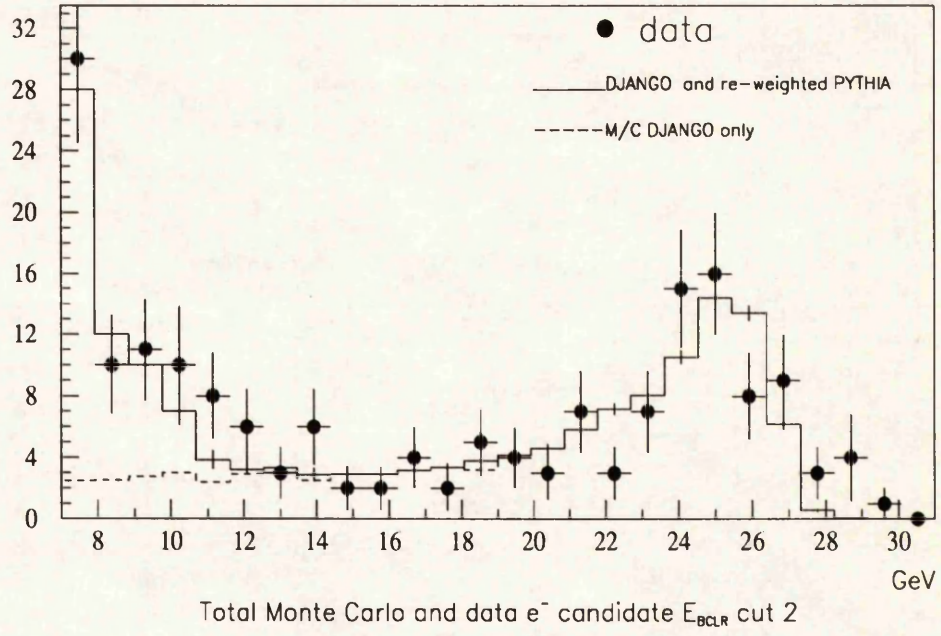


Figure 7.21: Data and Monte Carlo after $\sum E - P_z / \sum E_{BCLR}$ cut

A second approach to forming an anti-photoproduction cut using $\sum E - P_z$ can be followed by considering equations 2.8 and 2.18 for y_{e^-} and y_{jb}, i.e. the kinematic variable y evaluated from the electron alone or using the Jaquet-Blondel formulae and the hadrons only. From these it follows that

$$y_{jb} - y_{e^-} = \frac{\sum E - \sum P_z}{2 \cdot E_l} - 1$$

where E_l is the original electron energy, and the summation is over all observed particles. This equation demonstrates that the two methods for calculating y would give the same answer in a perfect detector, where all particles could be detected. Losses of hadrons down the rearward beampipe destroy the symmetry, since then $\sum E - \sum P_z \neq 2 \cdot E_l$. (This argument only applies to the non-radiative case; the treatment of detected radiated photons in even a perfect detector would still require care.) Figures 7.22 (a)-(c) show (on log scales) the correlation between y_{jb} and y_{e^-} for DJANGO, PYTHIA and real data after the μDST cuts. Concentrating on the DJANGO plot, the disagreement

between the y values from the two methods increases at very small y. This is expected from 2.6 where it was stated that the resolution of y_{e^-} becomes large at small y because of a $1/y$ term. The performance of y_{jb} may also be expected to deteriorate at very small y, where the quark jet is at low energy. The PYTHIA events are concentrated in a narrow band at high y_{e^-} independent of y_{jb}. For these events, the electron escaped down the beampipe, and the low energy photons/hadrons identified by mistake inevitably lead to y_{el} being high (since y is the fraction of energy lost by the original electron). Only a very slight correlation is then found with y_{jb}, as expected, because this is calculated from a meaningless sub-set of the remaining hadrons. The data in figure 7.22 (c) exhibits features of both DJANGO and PYTHIA. Superimposed are the lines $y_{e^-} = 0.6$ and $y_{jb} = y_{e^-}/2$. A cut rejecting events with $y_{jb} < y_{e^-}/2$ if $y_{e^-} < 0.6$ was used by the ELAN group on the basis of these results. The effect of this cut is very similar to the $\sum E - P_z / \sum E_{BCLR}$ method, but slightly looser; the resulting energy spectrum shows a clear rise at $E_e < 12$ GeV which is not present with the $\sum E - P_z / \sum E_{BCLR}$ cut.

7.3.6 EBDI

There are a number of other possible criteria which could be used in the task of background reduction. The first of these was a so-called estimator EBDI (Estimated BEMC DIspersion) calculated within the PHAN analysis code. EBDI is defined as

$$\sum_{stacks} \frac{|E_{stack} - E_{mean}|}{E_{mean} \cdot N_{stack}}$$

where: *stacks* or *cells* refer to all those BEMC scintillator stacks which are grouped together to form the energy cluster in question. E_{stack} is the energy of an individual stack, E_{mean} is the mean over the N_{stack}s involved.

EBDI is a measure of the homogeneity of the distribution of a cluster's energy between neighbouring BEMC stacks, and has been histogrammed in figures 7.23 (a)-(c) for the electron candidate. The explanation for the EBDI distribution shapes is complex. Using the above formula it can be trivially confirmed that for a cluster with just one cell, EBDI=0. EBDI then increases as the energy distribution amongst N cells becomes more *uneven*; for instance,

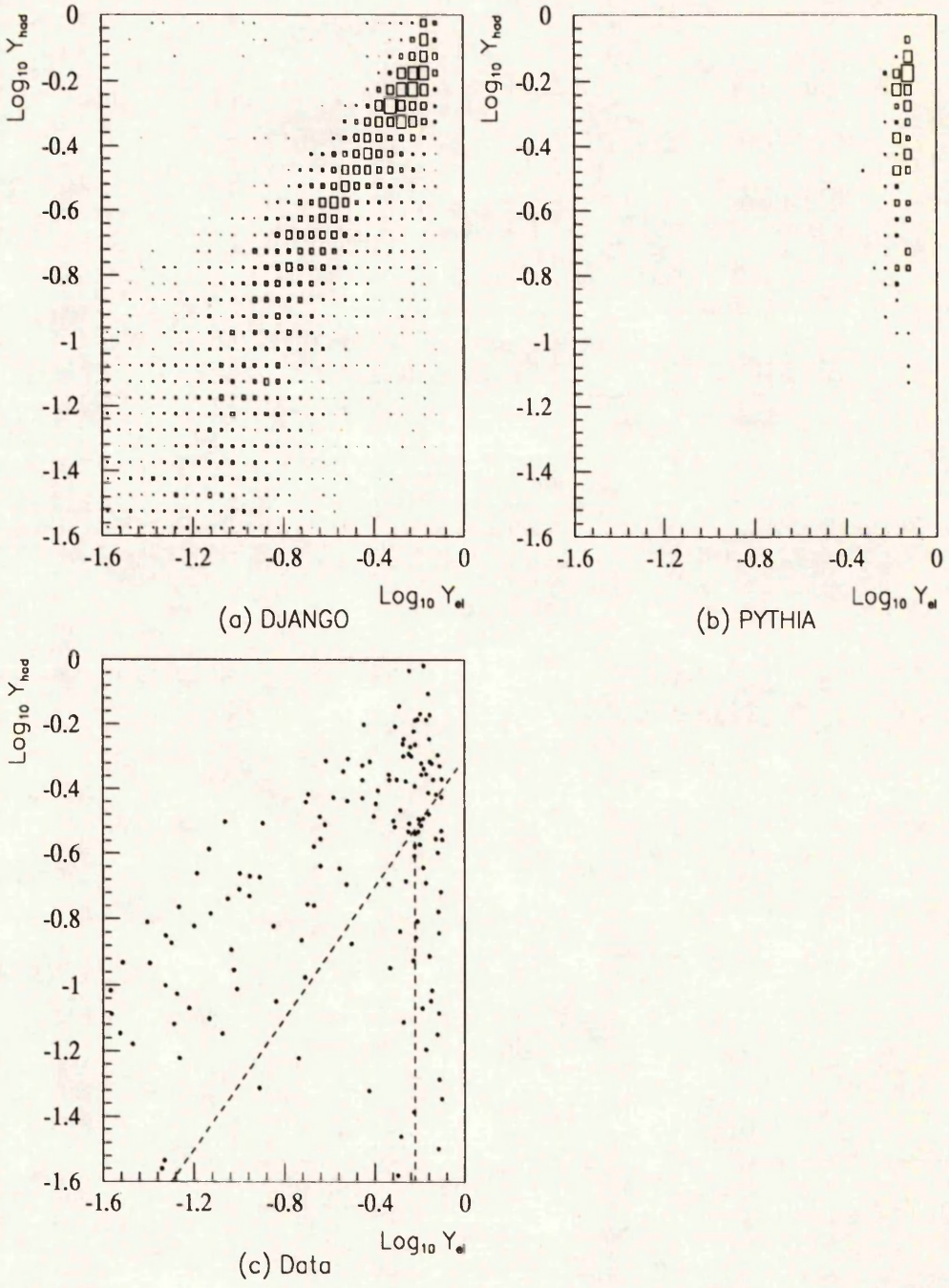


Figure 7.22: Correlation of y_{jb} with y_{e^-}.

a 15 GeV deposit in one cell with two 5 GeV neighbours results in EBDI = 0.533, but the same 25 GeV total spread as three cells at 20, 2.5, 2.5 GeV gives EBDI=0.933 . The number of cells grouped to form a cluster in practice is shown in figures 7.24 (a), (b). As can be seen, DJANGO e^- clusters almost always consist of nine cells, in fact grouped into a 3×3 grid with the hottest (maximum energy) cell at the centre, whilst PYTHIA events more frequently involve fewer cells. Some features of the DJANGO EBDI spectrum can be explained by considering the energy to be virtually all in one cell with 8 equal, small (but non-zero) neighbours. If E is the central energy, $\overline{E} = \frac{E}{9}$ and the upper limit is given by

$$EBDI = \frac{|E - \frac{E}{9}| + 8 \cdot |0 - \frac{E}{9}|}{9 \cdot \frac{E}{9}} \approx 1.77$$

If the 8 neighbouring energies each have energy A,

$$EBDI = \frac{16 \cdot (|E_{BCLR} - 9A|)}{9 \cdot E_{BCLR}}$$

and with $A \approx 0.15$ GeV, this gives rise to an EBDI shape starting at ≈ 1.2 GeV and rising sharply to a peak at around 1.7 GeV (with a mean value of 1.6 GeV), using the DJANGO E_{BCLR} values. It is supposed that a convolution of distributions derived from similar models, each with a slightly different method for sharing the energy – but all with the common feature of having one very dominant cell – would explain the overall EBDI shape exactly. Certainly the fact that DJANGO events give rise to one hot cell is confirmed by figure 7.24 (c), (d) which show the ratios $\frac{Energy\ of\ SECOND\ hottest\ stack}{Energy\ of\ hottest}$ and $\frac{Energy\ of\ THIRD\ hottest\ stack}{Energy\ of\ hottest}$ respectively. From these, it is plain that it is rare for the second most dominant cell to have even 10% of the maximum energy. The corresponding PYTHIA graphs 7.24 (e), (f) confirm what the broader, lower EBDI shape implied, namely that these events tend to share the energy more evenly.

This difference is presumably because the false electrons found by PYTHIA are overlaps of different primary particles, e.g. charged pions and photons/neutral pions. This gives rise to broader clusters spread over a number of BEMC stacks. On the basis of these results, a cut EBDI > 1.3 was added. EBDI shows no correlation with electron energy or angle or any of the kinematic variables x, y, Q^2, which makes it particularly useful. A cut on the

ratio of first / third most significant cell energies would have been equivalently good, and equally uncorrelated with energy and angle, but EBDI was an easier quantity to access within the analysis code.

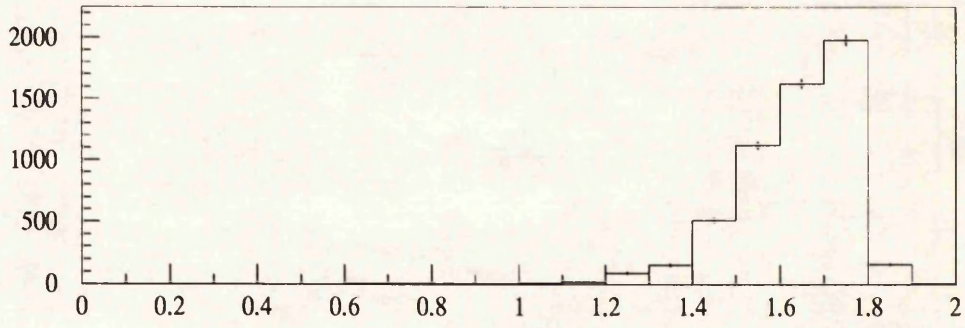
7.3.7 Other Cuts

The official ELAN BPC requirement discussed in section 7.1.4 was tightened so that the required radial distance between an electron candidate BEMC cluster and a BPC hit point was $\leq$ 8 cm. The more advanced form of the ToF selection as detailed was implemented, and the insistence upon at least one track (forward or central) was maintained.

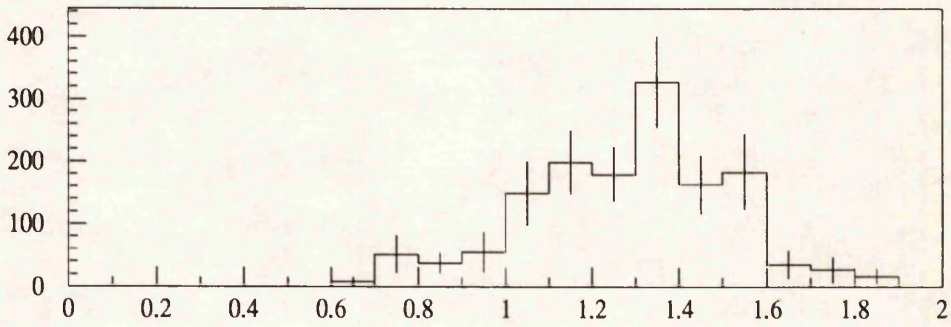
7.3.8 Effect of the Cuts

Not one of the cuts discussed above is on its own sufficient to remove the background, i.e. PYTHIA events still remain in the combined Monte Carlo sample. The decision was made to apply both the two most effective anti-photoproduction cuts, i.e. $\sum E_{BCLR}/\sum E - P_z$ and EBDI with the low energy cut on E_{BCLR}. No improvement was found to result by combining with E_{Δ}, E_{Δ}/E_{BCLR} or E_{semi}. Used in conjunction with the μDST selections discussed above, this combination resulted in a severe reduction of the data sample from 559 to 120 events.

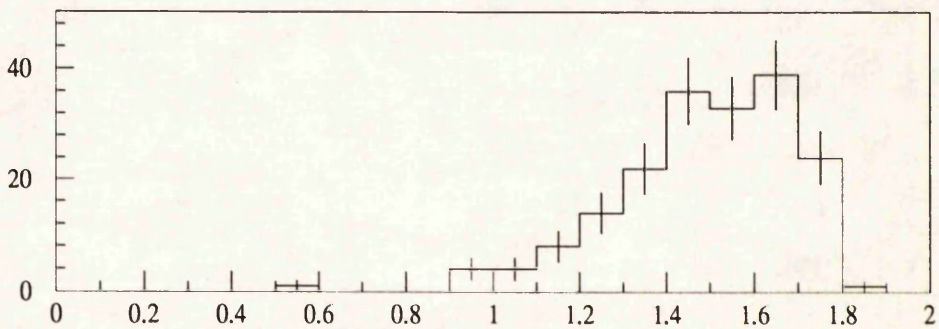
The effect of all the preceeding cuts is illustrated in figures 7.25 (a), (b) and 7.26 (a), (b). The first graphs are of the electron energy spectrum, with the total Monte Carlo, i.e. the combined PYTHIA and DJANGO, superimposed (dashed line). The Monte Carlo graph has been deliberately scaled so that the height of the 25-26 GeV bin matches that of the data. Also superimposed is the DJANGO component alone (dotted). The errors on the Monte Carlo are much smaller than those on the data across the spectrum, and so the data errors ($= \sqrt{N}$, where N is the bin content) alone have been depicted. Figure 7.25 (b) is the same as (a), but with one difference; here the superimposed Monte Carlo graphs have been normalised to correspond to a luminosity $\mathcal{L}$ = 1.3 nb^{-1}, the measured luminosity of the summer 1992 data-taking period. The difference in scaling factors between the (a) and (b) amounts to 1.27.



(a) EBDI M/C DJANGO events



(b) EBDI M/C PYTHIA events



(c) EBDI Real data

Figure 7.23: Estimator EBDI spectra

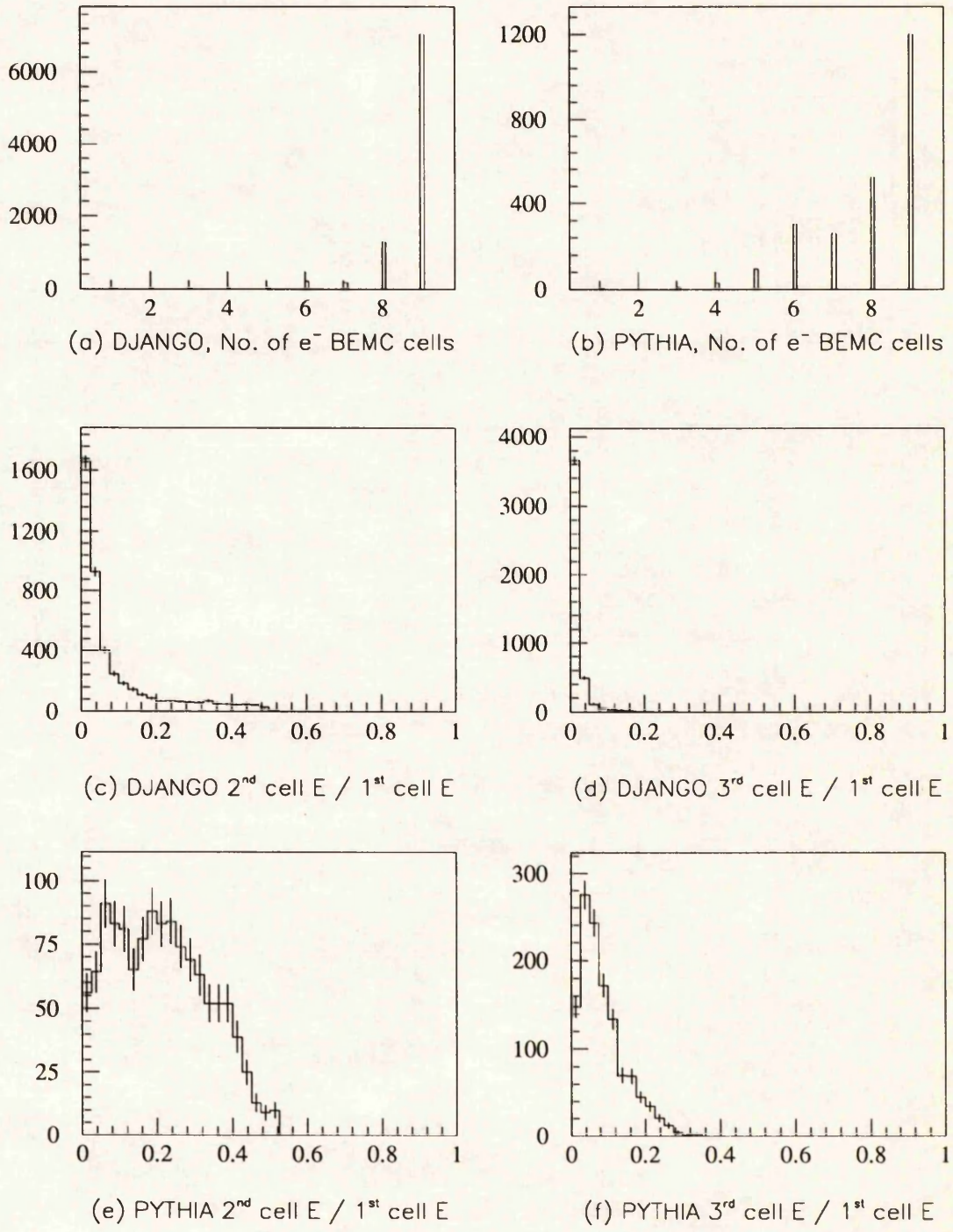
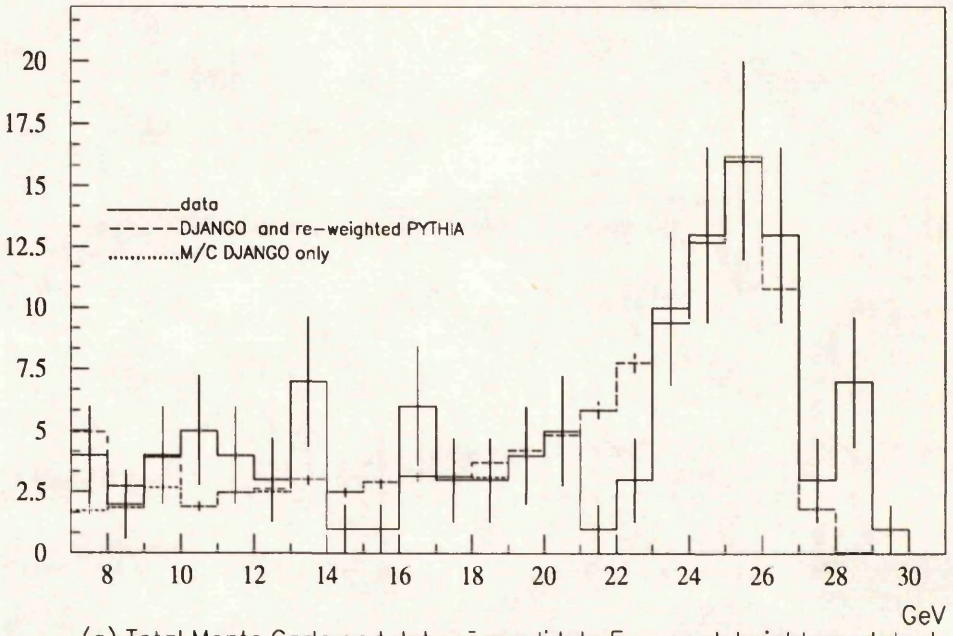
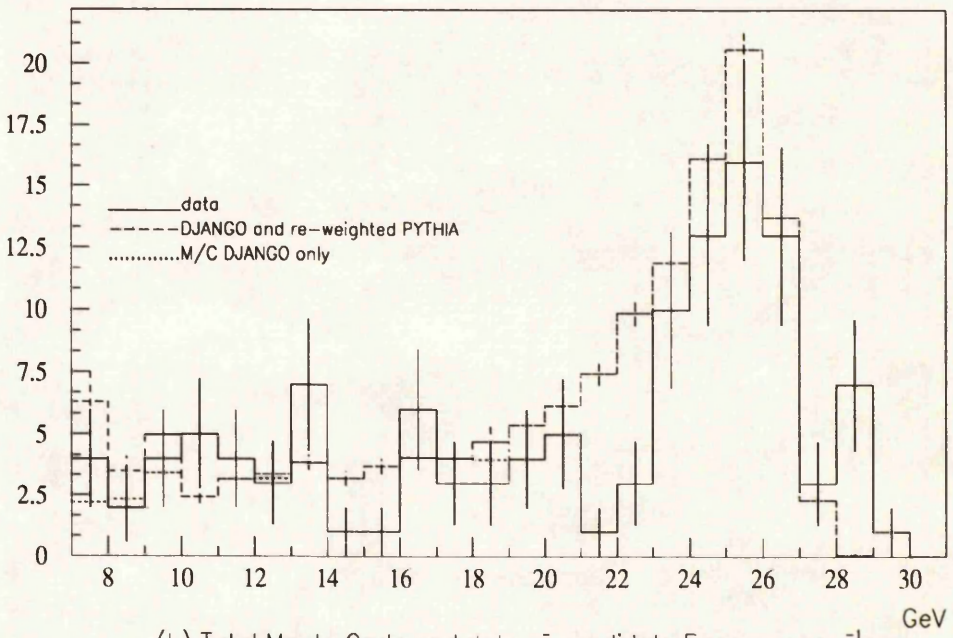


Figure 7.24: No. of cells, Ratio of $2^{nd}, 3^{rd}$ largest cell energies to largest



(a) Total Monte Carlo and data e^- candidate E_{BCLR}, peak heights matched



(b) Total Monte Carlo and data e^- candidate $E_{BCLR,\ coled\ to\ L=1.3\ nb^{-1}}$

Figure 7.25: Electron energy spectra after all cuts

Figures 7.26 (a), (b) are histograms of the log Q^2 and x values, with the combined Monte Carlo values depicted as a bold line, the data values as points. The Monte Carlo has again been scaled to the nominal data luminosity value. Both the x and Q^2 spectra were obtained from a sub-set of the data, neglecting those events with no z vertex (i.e. a value $z = 0$ was assigned because the reconstruction code was unable to find the genuine vertex), some 5% of the total sample, since the error on θ_e for these events is necessarily larger. These events were however left in the energy spectrum. Data at $x \leq 10^{-4}$ is clearly present in the remaining sample, which demonstrates HERA's value in opening up a new regime in low x physics. The disagreement between data and Monte Carlo becomes apparent at $Q^2 \leq 25$ GeV^2 and $x \geq 10^{-2}$. The x values were reconstructed from the electron alone, but two methods were used for Q^2; from the electron only, i.e. using equation 2.1, and the two-angle method alluded to in section 2.6. In this latter case, the relevant equation is

$$Q^2(\theta_e, \gamma) = {E_e}^2 \cdot \frac{(\sin\theta_e + \sin\gamma - \sin(\theta_e + \gamma))(1 + \cos\theta_e)}{\sin\gamma}$$

where θ_e is the electron angle and γ is the hadron angle. This in turn is defined in terms of the Jaquet-Blondel variables,

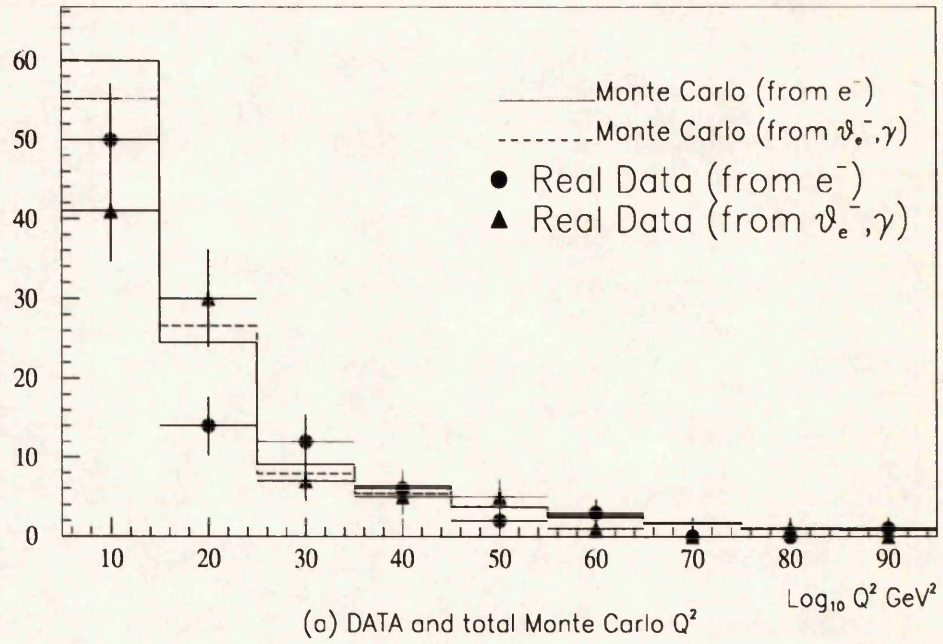
$$\cos\gamma = \frac{Q^2_{jb}(1 - y_{jb}) - 4{E_l}^2{y_{jb}}^2}{Q^2_{jb}(1 - y_{jb}) + 4{E_l}^2{y_{jb}}^2}.$$

with E_l the initial electron energy. This second method was used as a double-check of the results. For both data and Monte Carlo, the two methods agree well above $Q^2 = 35$ GeV^2. Below this, the two methods diverge, more so for the data than the Monte Carlo. The diagram produced by Bernardi et al., and reproduced as figure 2.8 illustrated the regions in which the different reconstruction methods were most effective. This shows that the two-angle method is not suitable below $Q^2 = 10$ GeV^2, and so may explain the large difference in the lowest Q^2 bins between the two data points. The two-angle method can of course be used to derive $x(\theta_e, \gamma)$ and $y(\theta_e, \gamma)$, and indeed the scattered electron's energy $E_e(\theta_e, \gamma)$. These are not illustrated because the agreement is very poor, both for DJANGO and real data. This is a result of being in the low Q^2, small x region of the kinematic plane, where the current jet angle tends towards 180°. An event-by-event study shows that in events where the hadrons were strongly backwards pointing, and suffered

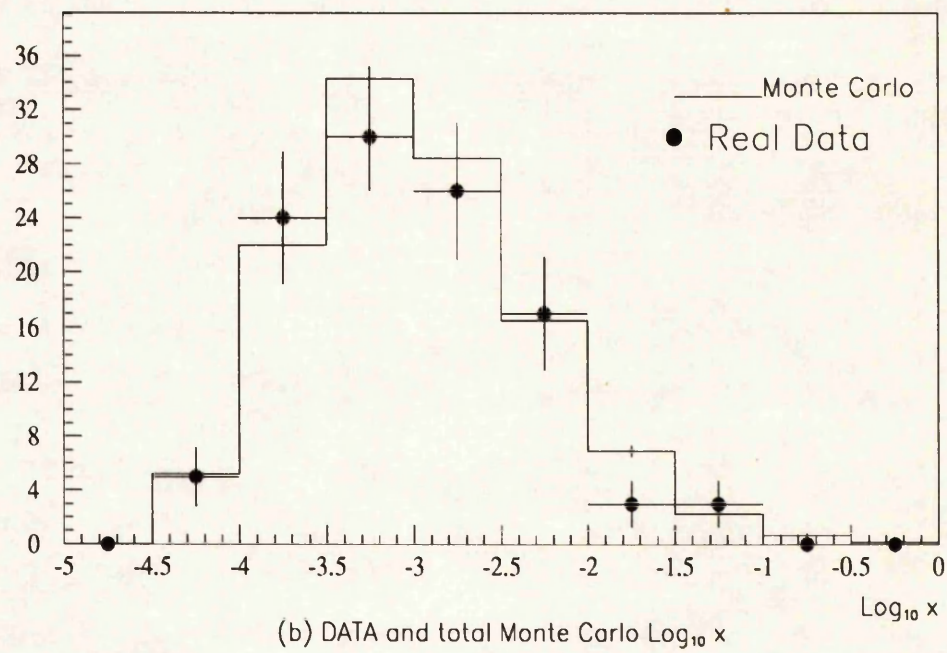
losses in the beampipe, the angle γ is very severely mis-calculated. In effect, the presence of small quantities of hadronic energy at lower angles (in the BEMC or Lar calorimeter) 'pulls' γ towards 90°. Referring again to the plots of quark/electron energy/angle (figures 2.3, 2.4) demonstrates the outcome of this; at a fixed θ_e, pulling $\gamma = \theta_j \rightarrow 90°$ results in $y \rightarrow 0$. This effect is observed in the plot of $E_e(\theta_e, \gamma)$, where events with genuinely low E_e ($\equiv$ low x, low Q^2) are migrated into the kinematic peak ($\equiv y \rightarrow 0$). It may then be asked why $Q^2(\theta_e, \gamma)$ produces such reasonable results. The answer is that Q^2 is so strongly dependent upon the electron angle, that the errors introduced into γ have only limited effect.

The above exercise provides a good demonstration of the problems that may be caused by an unwise choice of reconstruction scheme, and the way in which the choice is affected by the location of the data on the kinematic plane. Figure 2.8 confirms that the low x, low Q^2 region in which the current data sits is best reconstructed from the electron alone.

Returning to the energy spectra, several features are worthy of note. There is now a complete absence of any low energy peak, with just the 7-8 GeV Monte Carlo bin indicating a rise. The tiny contribution made by the few remaining PYTHIA events is evidenced by the merging of the dotted and dashed lines at 10 GeV. Only in the lowest energy bin is the data more consistent with DJANGO+PYTHIA than with the DJANGO prediction alone, only in the bins 7 - 10 GeV does the Monte Carlo indicate any possibility of PYTHIA events still existing. The PYTHIA events have larger errors on them than the DJANGO events. The events are weighted, so the error on a bin with N entries is not $\sqrt{N}$ but $\sqrt{\sum_i^N F(W)_i^2}$, where $F(W)_i$ is the (new) weight of the i^{th} entry. Using this, the remaining PYTHIA events constitute $6 \pm 3\%$ of the total Monte Carlo. At almost every bin the Monte Carlo and data agree to within the data errors, the obvious exceptions being the 21-23 GeV and 28-29 GeV bins. The cuts which are responsible for the gap at 21-23 GeV are the EBDI cut and (in particular) the insistence upon at least one track. Although tempting to restore these events by loosening these requirements, it is difficult to justify the inclusion of events which have no tracks whatsoever, especially when (by visual inspection of the tracker display using the H1ED program) the electron candidate cluster is at $\theta \leq 170°$ i.e. any track from a reasonable



(a) DATA and total Monte Carlo Q^2



(b) DATA and total Monte Carlo Log_{10} x

Figure 7.26: x, y spectra after all cuts

vertex to the cluster must have passed through the central trackers.

What is pleasing about all these graphs is that a succession of severe cuts using a wide variety of detector features (ToF, tracking, BEMC response etc.) have apparently eliminated $\approx 94\%$ of photoproduction events, allowing important features of the remaining data to be described by simulated DIS events alone. Data at $x \leq 10^{-4}$ is clearly present in the remaining sample.

The absence of almost all PYTHIA events and the agreement (to within the large errors on the data) between DJANGO and data means that the task of background reduction must now be considered finished, at least for the current sample, because there are insufficient simulated events remaining on which to test any additional schemes.

7.3.9 The Electron Tagger

The electron tagger arm of the luminosity monitor, described in section 4.13 is sensitive to electrons in the energy range $12 \leq E_{e^-} \leq 25$ GeV. Although designed to detect electrons from the process $ep \rightarrow ep\gamma$, it can of course serve to tag electrons from photoproduction events, provided that they are in the correct energy window. An event with a tagged electron and no tagged photon, accompanied by hadronic deposits could form a condition for rejecting events as photoproduction. Such methods will almost certainly form an important part of future stategies for background removal. The tagger was not used more extensively in this analysis because of uncertainties in its acceptance for low Q^2 processes. The claimed average acceptance $\overline{A}$ for events with $6 < E_e < 24$ GeV, $3 \times 10^{-8} < Q^2 < 10^{-2}$ GeV^2 is 31% [38]. Studying the low energy DIS tail (of the μDST), shows that in the 4-5 GeV bin, there are 136 events, the vast majority of which are assumed to be photoproduction. There are 46 events with electron tagger energies between 5 and 22 GeV, which means the acceptance corrected prediction is 148; such good agreement is not however found at higher electron energies, with the number of tagged electrons only accounting for $\leq 30\%$ of the observed background. The two distributions, BEMC electron candidate energy and electron tagger energy are plotted against one another in figure 7.27 for the uncut μDST. Of the 120 data events surviving the described cuts, no events with energy deposits in

the electron tagger (or in the photon tagger) were found. Whilst this cannot be offered as evidence against the existence of photoproduction events, the presence of tagged electrons could have indicated to the contrary.

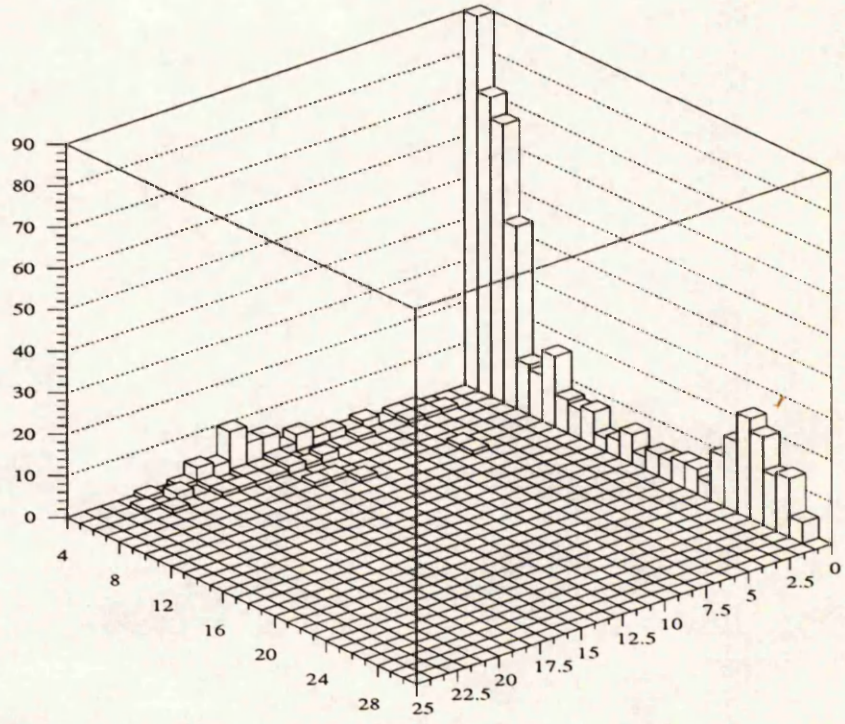


Figure 7.27: BEMC e^- energy plotted against electron tagger energy

7.4 Further Analysis

7.4.1 Peak Corrections

One of the obvious differences between the data and simulated energy spectra in figure 7.25 is the width of the kinematic peak. The data shows a tail extending to 30 GeV, whilst the DJANGO drops off sharply at 27 GeV. The existence of events at $E_e > E_l = 26.7$ GeV is allowed physically. Using the equations 2.6 - 2.1, at a fixed θ_e an electron may have an energy E_e from 0 to a maximum of

$$E_{e_{max}} = \frac{E_l}{\sin^2 \frac{\theta_e}{2} + \frac{E_l}{E_p} \cos^2 \frac{\theta_e}{2}}$$

corresponding to the minimum allowed y value, with E_p, E_l the initial proton and electron energies. For the lowest angle accepted by the BEMC, 152°, $E_{e_{max}} = 28.3$ GeV. The higher energy entries are therefore either the result of non-DIS physics, or of the inevitable smearing of energies due to detector effects. The first case can be ignored since it is hard to conceive of a suitable source creating such high energy electrons. It is assumed therefore that the sharp cut-off in the DJANGO spectrum is because the modelling of the BEMC response by H1SIM is too ideal, and it is necessary to correct for this if any even approximate comparison of spectra is to be successful.

There are several sources of BEMC energy smearing and corrections. A study of electromagnetic shower development and the performance of sampling calorimeters [53] shows how the primary contribution to the resolution comes from fluctuations in the energy loss and fluctuations in the number of photo-electrons captured by the photomultiplier tubes. These processes are governed by Poisson statistics and it can be shown that the resolution of a detector like the BEMC should go as $\sigma_E/E = const/\sqrt{E_e}$. This is what has been assumed so far, with the constant having been determined to be 10% [84], and is used as the sampling resolution in the H1SIM code. This figure was determined from beam tests, referred to in section 4.5.1. The noise in a typical BEMC electron cluster was estimated at 450 MeV based on a noise/stack measurement of 150 MeV. The construction of the BEMC means that there may be energy losses in the wavelength shifter bars surrounding each scintillator stack. These 'crack' corrections act to worsen the resolution, as does the leakage of energy out of the back of the BEMC. An additional term arises from loss of energy in the dead material in front of the BEMC, assessed to be about 1 electromagnetic radiation length (X_0) in average thickness. If, as here, data and simulations are compared at the level of BCLR energies, then dead-material corrections need to be added to *both* data and Monte Carlo results in order to achieve the correct physical energy. Detailed Monte Carlo work by Peppel [76] found the dead material and leakage corrections amount to $2.7\% \pm 1.2\%$ when combined. Other studies have revealed that the energy leakage is a function of stack shape; for the quadratic stacks it is assessed to contribute $< 1\%$ to the total resolution, but certain (trapeziod, triangular) stacks may have leakage terms as high as 10% [85]. The crack correction term may also show a shape

dependence. Another term in the resolution comes from the uncertainty in the relative calibration of neighbouring stacks, (the so-called stack to stack intercalibration), estimated to be 4%.

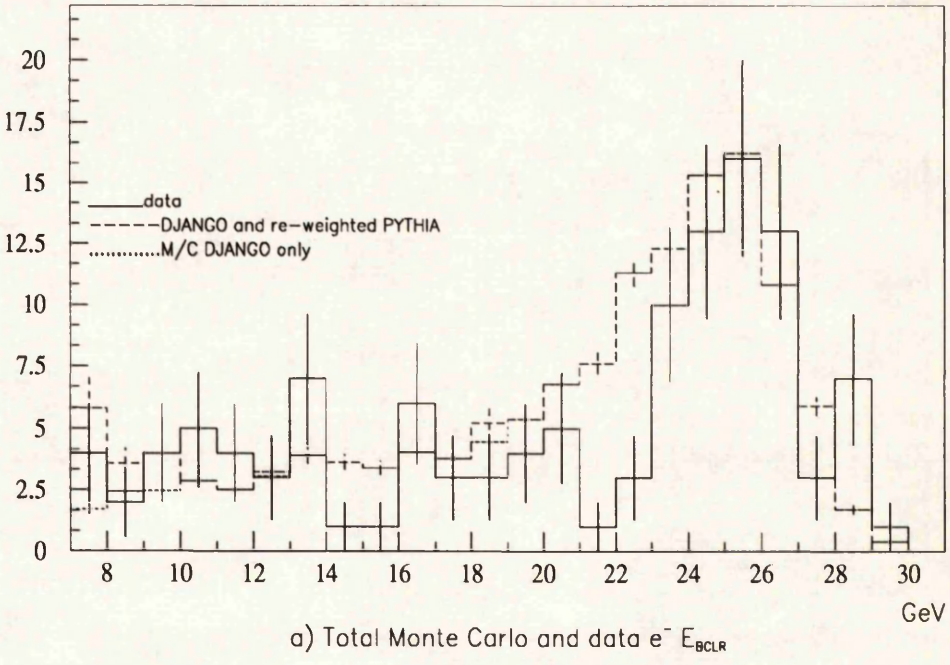
The above factors lead to the following form of the overall BEMC resolution, ignoring the uncertain crack-correction and leakage terms

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{\sigma_{noise}}{E}\right)^2 + \left(\frac{\sigma_{sample}}{\sqrt{E}}\right)^2 + (\sigma_{const})^2}$$

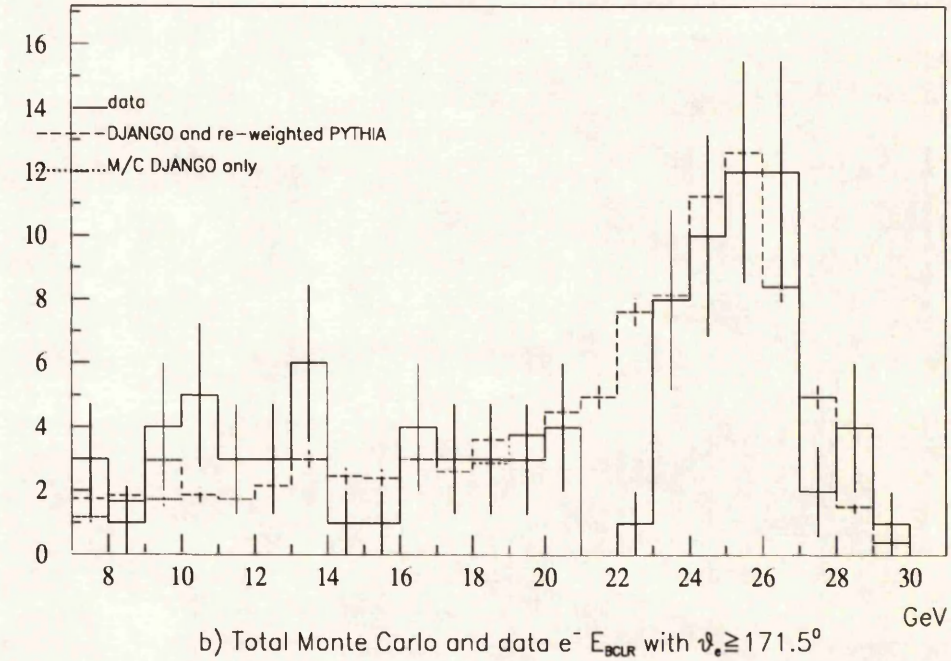
where $\sigma_{noise} = 0.45$ GeV, $\sigma_{sample} = 0.1$ GeV$^{\frac{1}{2}}$, $\sigma_{const} = 0.04$ [57]. For an electron energy at the expected peak value of $E_l = 26.7$ GeV this amounts to $\sigma_E/E = 24\%/\sqrt{E}$, with the constant term dominating.

The H1SIM detector model used the noise and sampling terms in the BEMC resolution, but not the stack-to-stack factor. The DJANGO files were therefore re-processed with this additional 4% term included into the BCLR energy resolution. The result of this is shown in figure 7.28 (a) in which the re-processed DJANGO file used in 7.25 is superimposed on the data spectrum. The DJANGO file was normalised to the data luminosity $\mathcal{L} = 1.3$ nb^{-1}, and two improvements are now to be seen; the Monte Carlo peak extends to the same 30 GeV limit as the data, and the peak bin heights match. Previously, the scale factor which was needed to achieve this coincidence in peak height was not the same as the luminosity normalisation factor (the difference between the two scalings of 30% being much larger than the error on the luminosity measurement at 7%), so this improvement is significant. The location of the peak in the data and Monte Carlo is the 25-26 GeV bin, (although the size of the error bars on the data spectrum must be considered) rather than at 26-27 GeV. The shift is compatible with the estimates of $\approx 2 - 3\%$ for the dead-material correction needed to turn EBCLR energies into real values. The correction was not applied here as it is not necessary for a comparison between data and simulation output.

More precise studies of the peak width, BEMC calibration and resolution details are beyond the scope of this work, and require a greater amount of data in order to be performed with the necessary precision. These exercises will become increasingly important as the HERA luminosity increases. Here, it is merely sufficient that the broad characteristics of the spectrum can be justified in terms of current knowledge of the BEMC performance.



a) Total Monte Carlo and data e^- E_{BCLR}



b) Total Monte Carlo and data e^- E_{BCLR} with $\vartheta_e \geq 171.5^\circ$

Figure 7.28: Monte Carlo and data e^- spectra with new Monte Carlo resolution

	x	y	Q^2
min	$5 \cdot 10^{-}5$	$5 \cdot 10^{-}5$	3 GeV^2
max	0.999	0.99	200 GeV^2
cross section		247.5 $\pm$ 0.9 nb	

Table 7.4: DJANGO 10 KMRS B0 generated kinematics

7.4.2 High Angle Cut

All the DIS Monte Carlo events so far used were created using the DJANGO 10 generator with the KMRS B- structure function parametrization. A second file was made with the same generator but using the B0 parametrization of [22] instead. A graph of F_2 against x for this version is flatter than for the B- as already discussed in section 2.4. The details of the resulting file are given in table 7.4.

It was noted that this file exhibited an apparently quite different reconstructed electron angle θ_e distribution when compared with its B- counterpart, demonstrated in figure 7.29 (a), (b). The lack of double peak of the B- spectrum was not however the result of its different structure function, but because of a pre-selection which it was discovered had been applied to it; the B- file had been cut prior to its processing through H1SIM so that the generated electron angle $\theta_{sim} \leq 174°$. (The upper limit to the B0 file was $\theta_{sim} \leq 176°$.) This was trivially confirmed by applying a similar cut to the B0 output and producing a single-peaked θ_e shape (7.29 (c)) with a shoulder at higher angles, exactly as for the B- events. It was decided on the basis of the θ distributions to add a cut removing events with $\theta_e \geq 171.5°$. (The seemingly odd value of 171.5° was chosen over 171°, because the latter would have reduced the number of data events to just 57.) This was done for two reasons; the cut on the *reconstructed* angle renders the data and all Monte Carlos (including the ones introduced later) directly comparable. In the case of the B0 dataset, 97% of events with electrons generated beyond 174° reconstruct to angles $\geq 171.5°$, and so are eliminated. In this way, the (small) changes that may have been introduced into the shapes of different E_e spectra due to their being generated from slightly different areas of the kinematic plane are removed. The second

reason was to restrict the comparisons to a region of high acceptance. The θ spectra demonstrate the competition which exists between the DIS cross section which rises sharply with angle and the acceptance which drops. The acceptance of the detector is low and uncertain in this high angular region; a very slight difference between the software's model of the trackers, the beam pipe, the BEMC geometry etc. and reality (or even between the versions of H1SIM used for different runs) could have a large effect on the high angle acceptance and result in many more DIS events entering the final sample, so it is wise to limit the study to a safer part of the detector. The high angle cut has other advantages. It rejects $\approx 70\%$ of events with the electron candidate in the BEMC inner triangle stacks (which start at 171.2°), itself a reasonable operation; the triangle stacks have resolutions and leakage corrections which are less well understood than the remainder, and receive the greatest rate of background events, an estimated 77 kHz, *cf* 93 kHz for the range $150° - 170°$ [84]. Secondly, the high θ_e cut acts as a powerful anti-photoproduction selection, because it is in effect a low Q^2 cut, restricting $Q^2 \geq 15.5$ GeV^2 for most of the x range. Although it would have been prefered to use only the selections developed above, and have more statistics, the θ_e cut does mean that the chance of photoproduction events remaining in the sample is now negligible. Thus, there is considerable justification for eliminating these large angle events even without the need to compensate for the hidden B- Monte Carlo cut. Figure 7.28 (b) shows the KMRS B- and data energy spectrum after the $\theta_e \leq 171.5°$ cut, with the Monte Carlo again scaled to the nominal 1.3 nb^{-1} luminosity. The data file now contained 95 events, reduced from 120.

The use of the μDST selections, EBDI, $\sum E - P_z / \sum E_{BCLR}$ and angle cuts has, examining 7.28 (b) resulted in an almost complete removal of the photoproduction background. The remaining PYTHIA events form $2 \pm 2\%$ of the total Monte Carlo entries. On this basis, it is considered safe to compare the data with DJANGO outputs alone.

It is worth considering at this stage that the PYTHIA file was reweighted not only to produce a good match with the low energy tail of the original data, but also to agree with the WA69 measurements for low P_t particles. (Indeed, the former was achieved upon managing the latter; F(W) was 'fine tuned' to match the data, but its basic form was found by comparing with

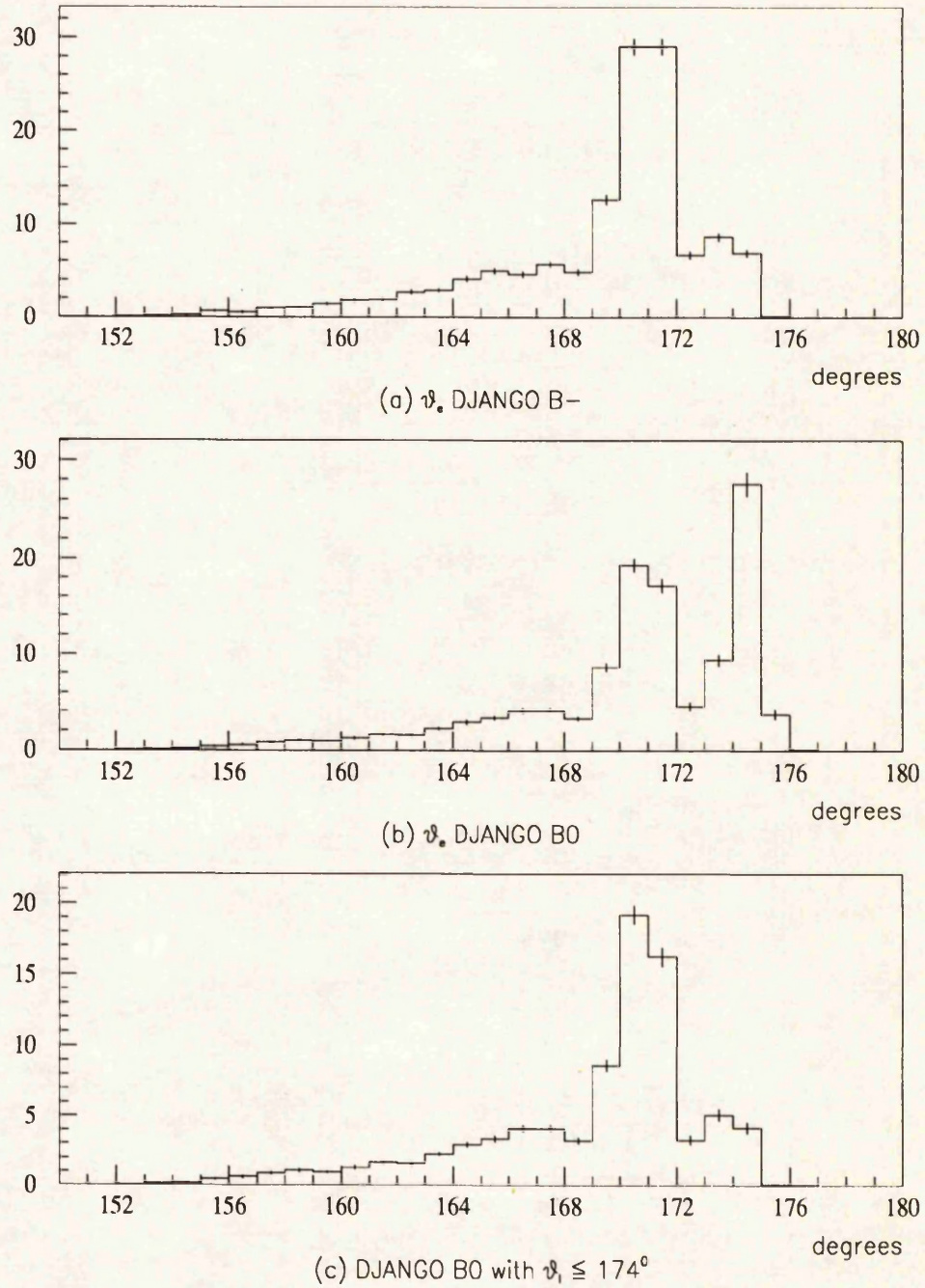


Figure 7.29: Reconstructed θ_e for DJANGO KMRS B-, B0

Sample	$\sum E - P_z / \sum E_{BCLR}$	EBDI	Θ_e	≥ 1 Track	BPC
Data $7 \leq E_e \leq 15$ GeV	81	39	15	-	11
Data $15 \leq E_e \leq 31$ GeV	-	10	31	36	-
DJANGO	3	5	38	27	2
PYTHIA	91	53	75	-	85

Table 7.5: Cut efficiencies for each data set

the WA69 results). Despite the effort made to correct the *low* P_t spectrum, more confidence could be placed in the performance of PYTHIA if it could be shown that it was the hard scattering events (high $\hat{P}_t$) which had constituted the dangerous background, since it was those events which the generator was originally written to simulate. This was partially confirmed by looking at the $\hat{P}_t$ values for those events which survived each level of cuts. Their mean value in the original re-weighted PYTHIA file was 2.2 GeV/c, rising to 3.1 GeV/c after the $\sum E - P_z / \sum E_{BCLR}$ cut, and was 7.5 GeV/c for the few events remaining before the θ_e cut. This provides some evidence that the dangerous events – those which survived the simple μDST cuts – belonged to the range of $\hat{P}_t$ where PYTHIA is believed to work well.

To conclude the work on removing the background, table 7.5 gives the relative effectiveness of various cuts. The 'efficiency' of a given selection rule has been defined here as the percentage increase in the number of events in a sample which results when the cut is removed from the total set of cuts. For instance, if the final selection rule is A . AND. B . AND. C, and there are 100 events in the sample, then if A . AND. B alone result in 200 events, cut C has a 100% efficiency.

The efficiencies for the data set are split into those for the lowest scattered electron energies and those for the mid-range/peak values. The $7 \leq E_e \leq 15$ GeV set show features in common with the PYTHIA values, i.e. the great effect that specifically anti- photoproduction cuts like the EBDI estimator and $\sum E - P_z / \sum E_{BCLR}$ have. The cuts have very similar effects on the higher energies of the data and DJANGO files, with the angle cut and track cut removing the greatest number of events.

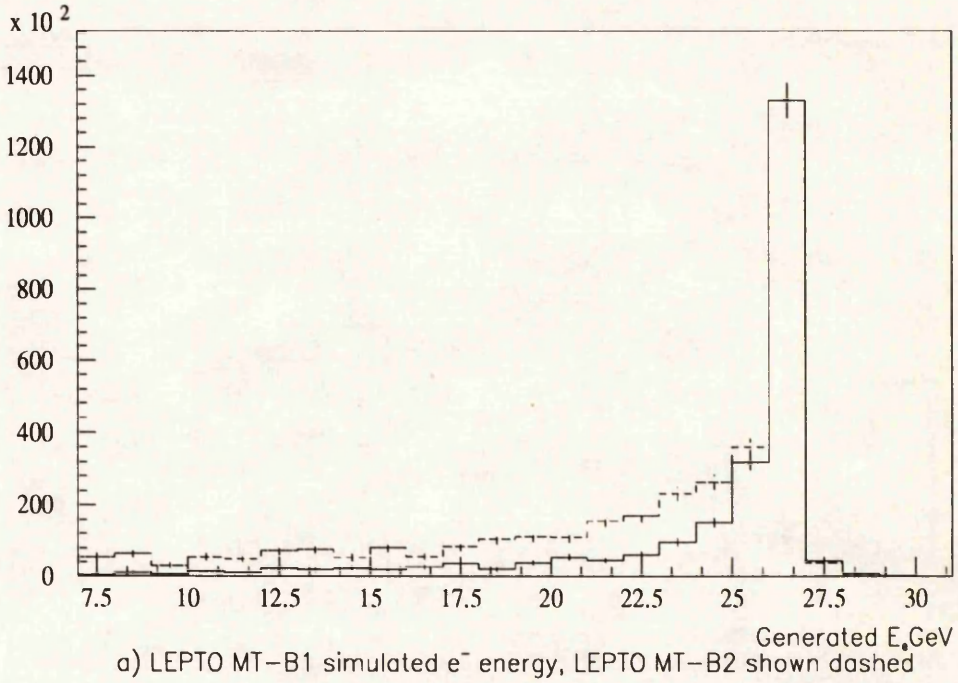
7.4.3 Other Structure Functions

Both the KMRS F_2 parametrisations so far used behave similarly at low x, as figure 2.6 confirms. The obvious choice of a third structure function parametrization for comparison was that of Morfin-Tung B2 [25], the most extreme of those illustrated. No such fully reconstructed Monte Carlo was available, and so a rather oblique way was used to provide a MT-B2 energy spectrum. LEPTO generator outputs of DIS events using the MT-B1 and MT-B2 forms existed, but these had not been processed through H1SIM/H1REC, and the constraints placed upon computing time prevented this step from being performed. Another DJANGO generator file created with the MT-B1 structure function was available however, and this had been passed through the entire reconstruction chain. The aim was to have a fully reconstructed set of events generated with DJANGO (for comparison with the KMRS models, also simulated using this program) using MT-B2 and this was now estimated from a combination of the above files. The energy spectra of the generated electron from each of the LEPTO files was plotted and the ratio of the MT-B2/MT-B1 heights calculated on a bin-by-bin basis, effectively a function R of the generated energy. For each DJANGO MT-B2 event with generated electron energy $E_{e_{gen}}$ and reconstructed energy $E_{e_{rec}}$, the spectrum of $E_{e_{rec}}$ was then plotted, with every event weighted by an amount $R(E_{e_{gen}})$. In effect, knowledge of how the *generated* energies for the two structure models compare has been used to re-form the MT-B1 *reconstructed* energies as if they were MT-B2.

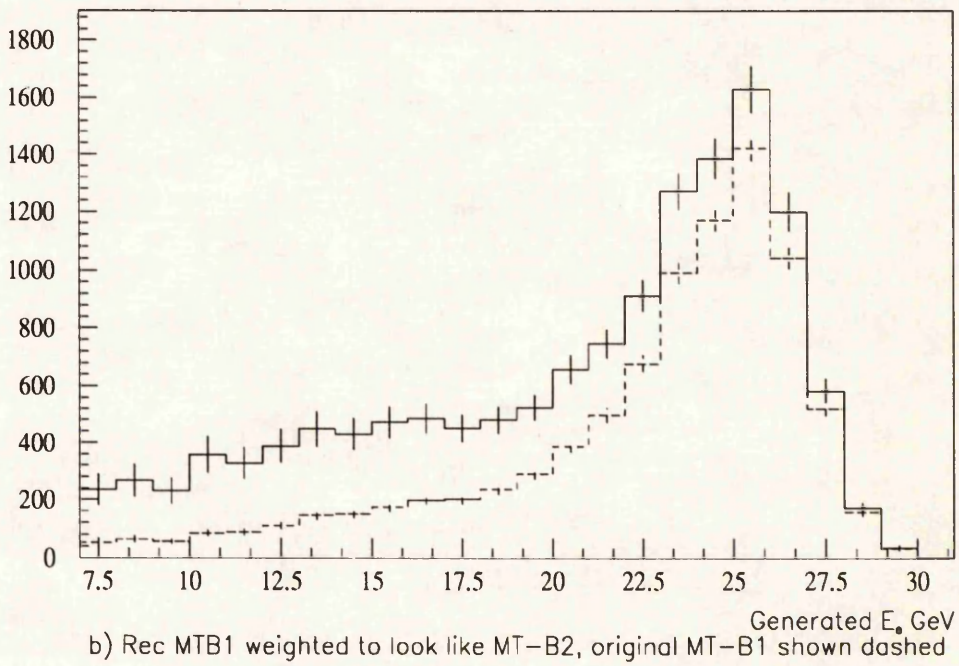
This process is illustrated in figures 7.30 (a), (b). These show the original LEPTO spectra, the MT-B1 E_e spectrum and the resulting, estimated MT-B2 reconstructed spectrum.

It was confirmed with this file that $> 95\%$ of events with $\theta_{gen} > 174°$ were removed by the chosen θ_e cut.

To extend the variety of F_2 models, two more DJANGO files were used, with the D- and D0 versions of the MRS parametrization [24]. Approximately 5000 events were run, their generated kinematic details given in 7.6



a) LEPTO MT−B1 simulated e^- energy, LEPTO MT−B2 shown dashed



b) Rec MTB1 weighted to look like MT−B2, original MT−B1 shown dashed

Figure 7.30: Simulating the DJANGO MT-B2 spectrum

	x	y	Q^2
min	$3.5 \cdot 10^{-}5$	$1 \cdot 10^{-}4$	3 GeV^2
max	0.999	0.99	200 GeV^2
cross section MRS D-			479 $\pm$ 2 nb
cross section MRS D0			326 $\pm$ 0.9 nb

Table 7.6: DJANGO 10 KMRS D-, D0 generated kinematics

7.4.4 Results: Comparing the Spectra

Once a set of Monte Carlos had been generated (with suitable cuts applied to restrict their generated kinematics to within the same limits), and the photoproduction background was found to be consistent with zero, the comparison of energy spectra could proceed. The low statistics did not allow for any sophisticated shape comparisons to be performed. The only parameters available were the relative heights of the tail ($E_e < 15$ GeV) and peak, and the slope of the tail. There remained the question of how to scale the Monte Carlo files to the data; previously the KMRS B- spectrum was normalised to the known data luminosity, but for a comparison of the spectral shapes this was not necessarily the best procedure. Two alternatives presented themselves, to scale the simulator output so that the peak-bin heights matched, or to scale so that the total number of events was equal, i.e. to match the area under the plots. These had the advantage that they were independent of any errors in the luminosity value or the calculated Monte Carlo cross-sections, and so were adopted. The only important factor here is the shape of the spectrum, and since the Monte Carlos and data are all supposed to be pure DIS samples, scaling in this manner is a legitimate exercise.

The results using the peak-height matching technique are shown in figures 7.31 (a)-(e). The data E_e spectrum is drawn in solid line and the DJANGO Monte Carlo superimposed. In all cases the errors on the data dominate due to the much lower number of events. The MT-B2 spectrum has errors on each bin $\approx 2\times$ those of the other Monte Carlos, arising from the division of the two LEPTO spectra and the combination of these errors with those of the original MT-B1 entries. All six Monte Carlos describe the data

Structure Function parametrization	$7 \leq E_e < 15$ GeV	$23 \leq E_e < 31$ GeV
KMRS B-	1.37	1.03
KMRS B0	1.74	1.07
MRS D-	1.17	0.82
MRS D0	1.44	1.35
MT-B2	1.27	0.83
MT-B1	1.72	2.05

Table 7.7: χ^2 per degree of freedom after scaling each E_e spectrum to same area as data.

Structure Function parametrization	$7 \leq E_e < 15$ GeV	$23 \leq E_e < 31$ GeV
KMRS B-	1.35	0.98
KMRS B0	1.85	0.72
MRS D-	1.31	0.86
MRS D0	1.49	1.06
MT-B2	1.30	0.85
MT-B1	2.03	0.85

Table 7.8: χ^2 per degree of freedom after scaling each E_e spectrum to match peak height of data.

well, and the differences between them are minor when the error bars of the data plot are considered. The distinction between structure functions with a steep and flat progression in x is best illustrated by comparing 7.31 (e) and (f), for the MT-B2 and MT-B1 models. The quality of match between the data and MT-B2 at energies $E_e < 17$ GeV is a direct result of the low x F_2 behaviour. A necessary consequence of forcing the areas to be equal is that the peak is too large if the tail is too small, and so the B0 and D0 and MT-B1 files are the worst fitting in both important regions. Re-scaling so that the peak-bin heights match exacerbates the very low tail of these files.

Three χ^2 fits were performed between each of the DJANGO files and the data E_e spectra. The fits were conducted over the ranges $7 \leq E_e < 15$,

$15 \leq E_e < 23$ and $23 < E_e < 31$ GeV, the first and last of these corresponding to the tail and peak regions respectively. The middle region posed a problem in that the 21-22 GeV and 22-23 GeV bins with 0 and 1 events in, $cf \approx 8$ for the Monte Carlos dominated the χ^2 result. If these bins were dropped from the calculation, then χ^2 per degree of freedom $\ll 1$ in every case because of the very good agreement and large data errors. Thus, results for the middle section give no information on the quality of overall fit, and have been dropped. This section of the spectrum, with a lower than expected population, is the one which demonstrated the most sensitivity (excluding that part which was supposed to be the photoproduction component) to the track and EBDI cuts. It is possible that a greater understanding of the acceptance of these cuts, and the manner in which they are simulated by the H1SIM code will enable a better fit to be obtained.

The χ^2 values are shown in tables 7.7 and 7.8, for all parametrizations in two energy regions and for both scaling methods. (The values given in every case is the χ^2 per degree of freedom, with 7 degrees; throughout this chapter the term χ^2 is used to mean χ^2/d.o.f.). The luminosity scaling results, whilst differing in individual values, displayed the same trends as those listed.

With statistics as low as 1 event per bin, it would seem sensible to re-bin all the spectra into bins (at least) 2 GeV wide, but rather than do this the 1 GeV bin width was preserved and an additional test was applied to supplement the χ^2 results. (The set of χ^2 results with bins 2 or 3 GeV wide does not in fact differ greatly from the above values (and not to an extent that contradicts the conclusions drawn from them). The only difference is that the χ^2 values in the upper energy band all tend to lie below 0.8, because the level of disagreement becomes small compared to the errors. The narrower bins were preferred also because they preserve the differing slopes of the Monte Carlo tails better.)

The extra test applied was the Run test, described by Barlow in [86]. In this, the number of consecutive sequences ('runs') of bins in the data which lie above or below the Monte Carlo values is found. Given the total number of bins which lie above/below the Monte Carlo prediction, the mean number of expected runs and its variance can be calculated. Comparing this with the actual number of runs yields a σ value which can be used as a test of the

Structure Function parametrization	$7 \leq E_e < 15$ GeV	$23 \leq E_e < 31$ GeV
KMRS B-	0.0 σ	0.21 σ
KMRS B0	1.7 σ	0.0 σ
MRS D-	1.02 σ	0.0 σ
MRS D0	1.7 σ	0.21 σ
MT-B2	1.02 σ	0.21 σ
MT-B1	1.7 σ	0.21 σ

Table 7.9: σ values after applying the Run test

quality of agreement. Although less powerful than χ^2, the Run test can be used to reject fits which have acceptable χ^2s (perhaps because of overestimated errors) but are still poor.

The results of the Run test are shown in table 7.9. They were found to be virtually identical for all three methods of scaling (i.e., matching areas, peak heights, luminosities).

The similarity of many of the Run test values is not surprising; it does not look at the magnitude of the difference between two bins, only the sign, and the same patterns frequently re-occur. Values of 0.0 result when the mean and measured number of runs coincide.

The χ^2 values from the area-fixing show that the best fitting structure function parametrizations for the $7 \leq E_e < 15$ GeV range are the MRS D- and MT-B2. The next best fit is KMRS B-. The flatter (in x) MRS D0, KMRS B0 and MT-B1 constitute consistently worse matches, the MRS D0 being the least bad. This latter parametrization is, significantly, the steepest of the set of three flatter forms. The fits for the peak are all good, which is to expected, especially in the case when the Monte Carlos were scaled to match the peak bin height. This pattern is repeated exactly in the peak-fixing results, with the three steep forms (MRS D-, MT-B2 and KMRS B-) giving χ^2s on average 40% lower than their flatter counterparts.

These findings are to a large extent confirmed by the Run test. KMRS B- and MT-B2 again fit considerably better to the energy tail than the other four. The 0.0 value is produced because, with 6 data bins higher than the

Monte Carlo and 2 below, the expected number of runs of consecutively 'above' or 'below' (= 4) is that found in practice, as can be confirmed from figure 7.31 (a).

The acceptance calculated from the Monte Carlo files allowed the data cross-section to be estimated. The cross-section was formed using

$$\sigma_{data} = \frac{N_{final}}{\mathcal{L} \cdot a}$$

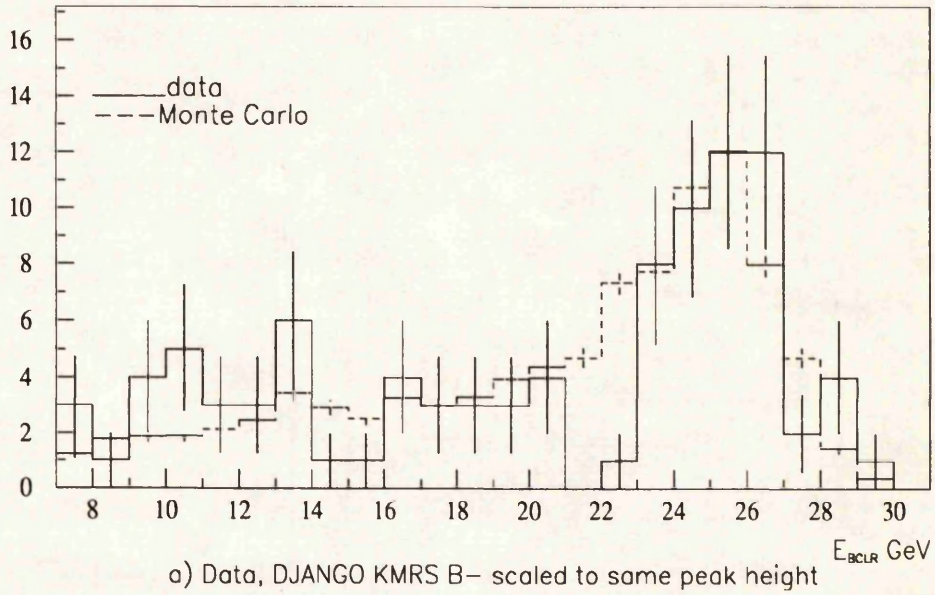
where N_{final} is the number of events in the final data sample, a the acceptance calculated from the Monte Carlo, and $\mathcal{L}$ the luminosity.

Restricting the kinematic domain to within the limits $Q^2 > 5$ GeV^2, $y < 0.7$, $\theta_e < 174°$, (i.e. before the high angle cut) the acceptance was found to vary by $\approx 13\%$ between the six different Monte Carlos used. The statistical error is assumed to be that on the number of entries in the samples ($\approx 10\%$ for the data and $\approx 1\%$ for the Monte Carlo), the systematic errors assumed are: error on the luminosity measurement (7%), error on acceptance value (13%), photoproduction contamination in remaining data and contamination from beamwall/beamgas events (7%). These final background values are an estimate based on the $\leq 10\%$ level of non-ep events assumed for the original μDST according to [76]. The tighter ToF cut and track requirement should reduce this by a factor of ~ 3, and this is combined with the percentage of remaining photoproduction events, 6%, as explained in section 7.3.8.

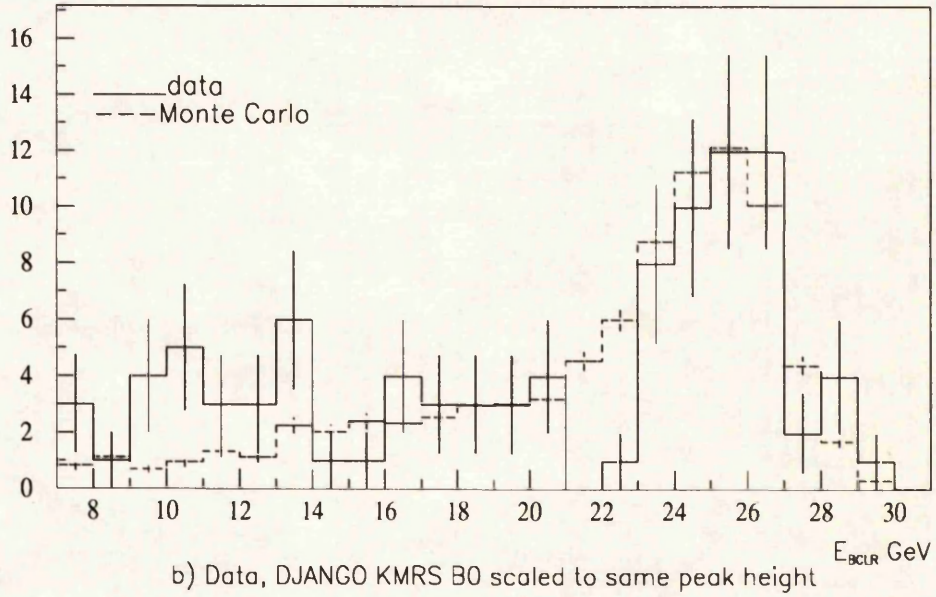
Over the stated limits, the cross-section was found to be 128 ± 13 (stat) $\pm$ 21 (syst) nb. This should be compared with the generated Monte Carlo cross-sections in the same kinematic region of 143 nb (MRS D-), 101 nb (KMRS B-), 98 nb (MRS D0) and 75 nb (KMRS B0). This result emphasizes that there does not exist any real possibility with these poor statistics and uncertain acceptances of making definite statements about the best structure function fits.

7.4.5 Discussion

There is a clear indication from the χ^2 results that flatter structure function parametrizations (KMRS B0, MRS D0, MT-B1) do not agree with the observed data as well as their steeper counterparts (KMRS B-, MRS D-,

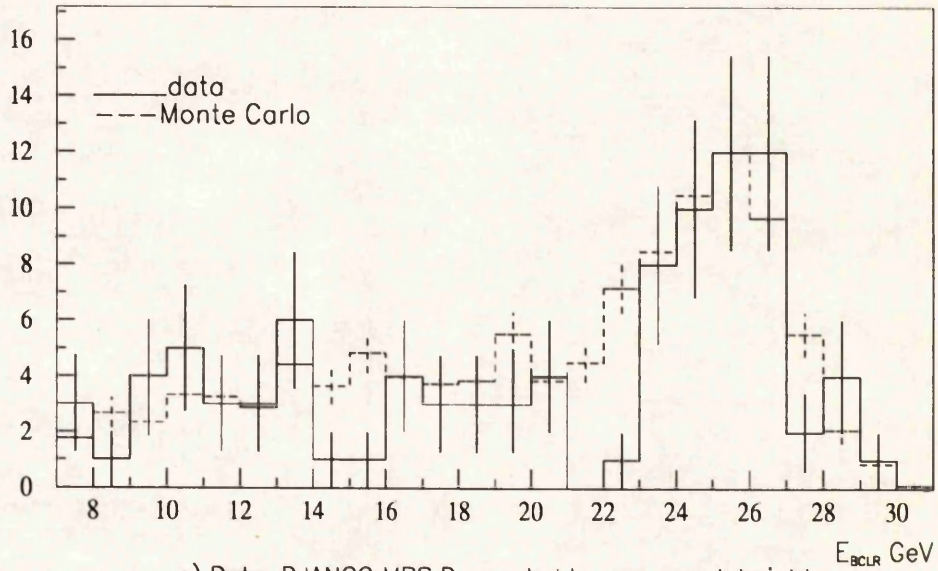


a) Data, DJANGO KMRS B− scaled to same peak height

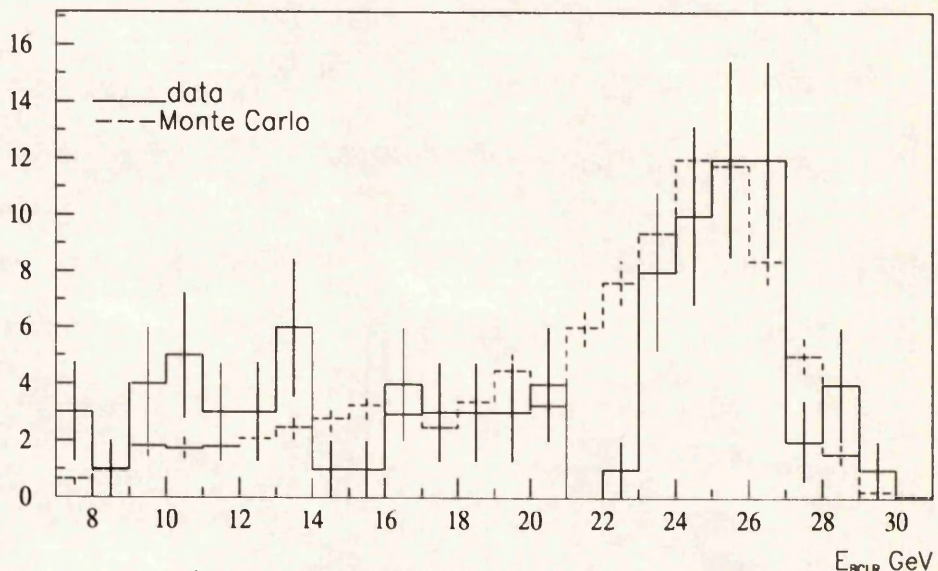


b) Data, DJANGO KMRS B0 scaled to same peak height

Figure 7.31: Data, DJANGO KMRS B-, B0 after all cuts and $\theta_e \leq 171.5°$

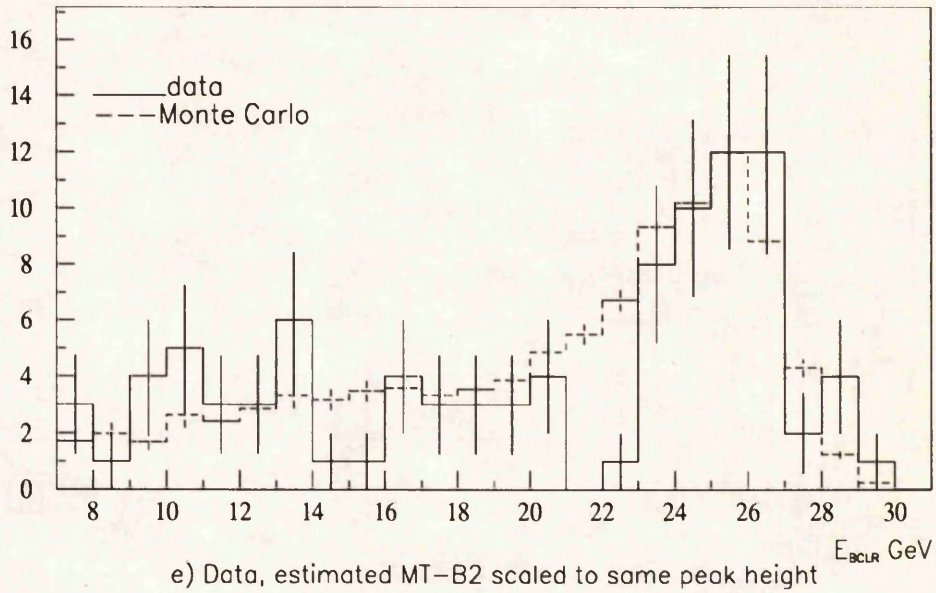


c) Data, DJANGO MRS D− scaled to same peak height

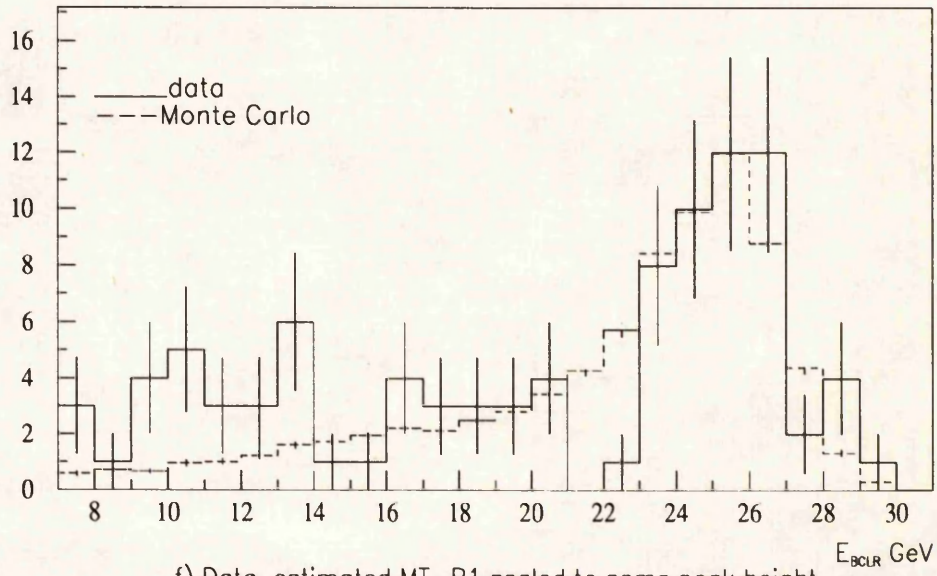


d) Data, DJANGO MRS D0 scaled to same peak height

Figure 7.32: Data, DJANGO MRS D-, D0 after all cuts and $\theta_e \leq 171.5°$



e) Data, estimated MT−B2 scaled to same peak height



f) Data, estimated MT−B1 scaled to same peak height

Figure 7.33: Data, estimated DJANGO MT-B2, MT-B1 after all cuts and $\theta_e \leq 171.5°$

MT-B2). (By 'flat' etc., is meant their relative steepness at low x over a narrow band of Q^2, see figure 7.31.) The three flat parametrizations consistently show worse agreement than the three steep ones. The dominant reason for the parametrizations' varying x dependences is the form of their 'starting' (low Q^2) gluon and quark densities, as described in section 2.3. These results provide weak evidence that starting distributions of the form $xG(x,{Q_0}^2) \sim C(x)x^{-1/2}$ are correct, since it is this form which characterise the B- and D- fits. It must be remembered that the MT-B2 scattered electron spectrum was obtained in a less-than-ideal manner, using the outputs from two different Monte Carlos, and so the results obtained from it are subject to a systematic error not included in the above analysis. The rejection of the flat MT-B1 parametrization is fully consistent with the other results, but the apparently good match with MT-B2 is probably not as significant as the KMRS B- and MRS D- results. These results are independent of the Monte Carlo cross-sections and experimental luminosity measurement. The temptation to use these results to make any more specific statement must be avoided. The χ^2 values alone would perhaps indicate that the MRS D- fit is the best, if one accepts that the quality of agreement can be based on the low energy tail alone; but this would be to go far beyond what is reasonable given the (very) low statistics, the unexplained holes in the data spectrum, and the small range of χ^2s produced. It would be decidedly premature to reject any of the models at all. All that can be stated is that there appears to be a trend in the data which is more compatible with steep structure parametrizations, but does not allow any to be ruled out. Certainly there is no hope whatsoever of choosing between different shadowing scenarios (by which steep input distributions can result in relatively flat forms of F_2 at low x) with this level of data.

The real success of this analysis lies not in the preliminary choice of structure function fits it makes possible, but that it has shown how the earliest data from the H1 experiment can be recorded, DIS events isolated from background, and results broadly in agreement with a range of theoretical models obtained.

Chapter 8

Summary and Conclusions

The main points of the work in this study may be summarised. It has be shown that a very simple kinematic method for selecting the scattered DIS electron could offer high acceptance rates at $y \leq 0.7$ and $Q^2 \geq 10$ GeV^2; the danger posed by high rate, low Q^2 photoproduction to the low energy ($y \geq 0.7$) part of the electron spectrum was demonstrated. The work done to combine detector features – the ToF scintillator, BPC wire chamber and tracking detectors – to form an algorithm that could separate DIS from photoproduction and non ep backgrounds was discussed. In this way, a sample of data collected over the first physics run period in the H1 detector at HERA was formed. The use and modification (using results from the WA69 experiment) of the PYTHIA generator to simulate the observed DIS contamination then led to an extended study of various techniques for removing the background. The presence of events in a new, low x regime at $x \leq 10^{-4}$ was confirmed by reconstructing the x spectrum from the scattered electron's energy and angle. Finally, the remaining 95 events were compared with DIS Monte Carlos incorporating different structure function parametrizations. From this, it was shown that the data has a trend towards forms steeper in x at $x \leq 10^{-2}$ and fixed Q^2, such as MRS D- and KMRS B-.

All the analysis here has been performed very much in a 'first day' manner, as an attempt to demonstrate that new physics could be obtained from the HERA experiment with only limited data, and with a detector (and software) in many respects far from complete. The next steps and areas for improvement

it suggests are clear. The BEMC calorimeter, so crucial for detecting DIS low Q^2 electrons must be studied in great detail. The spatial and angular resolutions it offers, the poorely understood energy resolution/calibration and the behaviour of the different component stacks impose severe limits on the accuracy of electron detection and background rejection. In part, this will improve with increasing luminosity, and eventually hardware upgrades will considerably improve the quality of physics results. An increased study of the acceptances of the cuts used, and the development of more sophisticated background rejection schemes will help improve the quality of any subsequent analyses. In the meantime, as more data is accumulated, the structure function program at H1 will continue. Even moderately improved statistics could enhance the conclusions reached here, and the ability to study how the cross-section (and hence structure functions) behaves at low x across several orders of magnitude in Q^2 would be valuable. Data collected will help new $F_2, xG(x, Q^2)$ parametrizations to be formed, and the existence and extent of proposed shadowing terms to be determined.

Bibliography

[1] B. Wiik, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.1 (1991). ed. W. Buchmüller, G. Ingelman.

[2] J. Dainton, *Physics at HERA*, Proceedings of the Annual conference of the HEPP Group of the Institute of Physics, April 2–3, p.79 Sheffield University (1990), ed. F. Combley.

[3] L. Maiani, Proceedings of the Workshop for Experimentation at HERA, NIKHEF Amsterdam, June 9–11 p. 3 (1983).

[4] J. Feltesse, *HERA, The New Frontier*, Proceedings of the Nineteeth SLAC Summer Institute on Particle Physics, (SLAC report 398) August 5–16, p. 155 (1992), ed. J. Hawthorne.

[5] EMC Collaboration, Nucl. Phys. B259 p.189 (1985).

[6] NMC Collaboration, CERN preprint PPE-92-124 (1992), to appear in Phys. Lett. B.

[7] J. Blümlein, M. Klein, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.101 (1991). ed. W. Buchmüller, G. Ingelman.

[8] R. Feynman, Acta Physica Polonica 24 p.697 (1963).

[9] R. Roberts, Proceedings of the School for Young High Energy Physicists, Rutherford Appleton Laboratory Sept. 1-14, ed. R. Barlow (1990)

[10] D. Perkins, *Introduction to High Energy Physics*, Addison-Wesley Publishing (1987).

[11] I. Aitchison, A. Hey, *Gauge Theories in Particle Physics (second ed.)*, Adam Hilger (1989).

[12] F. Halzen, A. Martin, *Quarks and Leptons* John Wiley & Sons (1984)

[13] J. Friedman, H. Kendall, Ann. Rev. Nucl. Sci. 22 p. 203 (1972).

[14] G. Altarelli, G. Parisi, Nucl. Phys. B126 p.298 (1977).

[15] E. Kuraev, L. Lipatov, V. Fadin, Phys. Lett. B60 p. 50 (1975).

[16] T. Hansel-Kozanecka, *Experimental Studies on QCD*, Proceedings of the Nineteenth SLAC Summer Institute on Particle Physics, (SLAC report 398) August 5–16, p.105 (1992), ed. J. Hawthorne.

[17] E. Reya, Dortmund University preprint DO-TH 91/09

[18] L. Gribov, E. Levin, M. Ryskin, Nucl. Phys. B188 p.555 (1981).

[19] J. Bartels et al., Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.203 (1991). ed. W. Buchmüller, G. Ingelman.

[20] R. Peccei, *High Energy Lepton Hadron Scattering as a Probe of QCD*, Proceedings of the Nineteeth SLAC Summer Institute on Particle Physics, (SLAC report 398) August 5–16, p. 259 (1992), ed. J. Hawthorne.

[21] BCDMS Collaboration, Phys. Lett. B223 p.485 (1989).

[22] J. Kwiecinski, A. Martin, W. Stirling, R.G. Roberts, Phys. Rev. D42 p.3645 (1990).

[23] CCFR Collaboration: S.R. Mishra presentation of preliminary wide band data at Lepton-Photon Symposium, Geneva (1991).

[24] A. Martin, W. Stirling, R. Roberts, Rutherford Appleton Laboratories preprint RAL-92-021 (1992).

[25] J. Morfin, W. Tung, FERMILAB preprint FERMILAB-Pub-90/74 (1990), published in Z. für Phys. C52 p.13 (1991).

[26] E. Eichten, I. Hinchcliff, K. Lane, C. Quigg, Rev. Mod. Phys. 56 p.579 (1984) & erratum 58 (1986).

[27] J. Bartels, K. Charcula, J. Feltesse, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.193 (1991). ed. W. Buchmüller, G. Ingelman.

[28] K. C. Hoeger, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.43 (1991). ed. W. Buchmüller, G. Ingelman.

[29] F. Jaquet, A. Blondel, Proceedings of the Conference on an *ep* Facility for Europe, Hamburg, (1979), ed. U. Amaldi.

[30] S. Bentvelsen J. Engelen, P. Kooijman, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.23 (1991). ed. W. Buchmüller, G. Ingelman.

[31] G. Bernardi, W. Hildesheim, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.79 (1991). ed. W. Buchmüller, G. Ingelman.

[32] M. Drees, K. Grassie, Z. für Phys. C28 p.451 (1985).

[33] H. Abramowicz, K. Charcula, A. Levy, Phys. Lett. B269, p.458 (1991).

[34] T. Sjöstrand, *PYTHIA at HERA*, Proceedings of the HERA Workshop Vol.3 Hamburg, Oct. 29–30 p.1405 (1991). ed. W. Buchmüller, G. Ingelman.

[35] S. de Jong, J. Engelen, Proceedings of the HERA Workshop Vol.2 Hamburg, Oct. 12–14 p.533 (1987). ed. R. Peccei.

[36] K. Goulianos, Phys. Rep. 101, No. 3, p. 169 (1983).

[37] A. Levy, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.481 (1991). ed. W. Buchmüller, G. Ingelman.

[38] S. Levonian, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.499 (1991). ed. W. Buchmüller, G. Ingelman.

[39] A. Ali, D. Wyler, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.669 (1991). ed. W. Buchmüller, G. Ingelman.

[40] R. van Woudenberg et al., Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.739 (1991). ed. W. Buchmüller, G. Ingelman.

[41] G. Ingelman, G. Schuler, *AROMA, A Monte Carlo Generator for Heavy Flavour Events*, Proceedings of the HERA Workshop Vol.3 Hamburg, Oct. 29–30 p.1346 (1991). ed. W. Buchmüller, G. Ingelman.

[42] G. Ingelman, *LEPTO 6.1 The Lund Monte Carlo for Deep Ineleastic Lepton Nucleon Scattering*, Proceedings of the HERA Workshop p.1366 (1991). ed. W. Buchmüller, G. Ingelman.

[43] M. Besancon, E. Evrard, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.181 (1991). ed. W. Buchmüller, G. Ingelman.

[44] The H1 Collaboration (T. Ahmed et al.), *Technical Proposal for the H1 Detector*, DESY, Hamburg, (1986).

[45] The BEMC Group, *The H1 Backward Electromagnetic Calorimeter*, H1 internal report, DESY, Hamburg, H1-08/92-233 (1992)

[46] A. Kwiatowski, H. Spiesberger, H. Möhring, Z. für Phys. C50 p.166 (1990).

[47] J. Kripfganz, H. Möhring, H. Spiesberger, Z. für Phys. C49 p.500 (1990).

[48] H. Spiesberger et al., Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.798 (1991). ed. W. Buchmüller, G. Ingelman.

[49] D. Duke, J. Owens, Phys. Rev. D30 p.49 (1984).

[50] M. Krasny, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.850 (1991). ed. W. Buchmüller, G. Ingelman.

[51] G. A. Schuler, H. Spiesberger, *DJANGO 1.0 – The Interface for Event Generators LEPTO and HERACLES*, Proceedings of the HERA Workshop Vol. 3 Hamburg, Oct. 29–30 p.1419 (1991). ed. W. Buchmüller, G. Ingelman.

[52] A. Kwiatowski, H. Spiesberger, H. Möhring, *HERACLES – An Event Generator for ep Interactions at HERA Including Radiative Processes*, Proceedings of the HERA Workshop Vol. 3 Hamburg, Oct. 29–30 p.1419 (1991). ed. W. Buchmüller, G. Ingelman.

[53] R. Fernow, *Introduction to Experimental Particle Physics*, Cambridge University Press (1986).

[54] H1 Calorimeter Group, *Study of Software Compensation for Single Particles and Jets in the H1 calorimeter*, contributed paper to the XXV Int. Conference on High Energy Physics, Singapore (1990).

[55] G. Cozzika, *The H1 Detector*, contributed paper to the Third Int. Conference on Calorimetry in High Energy Physics, Sept.29–Oct.2, Corpus Christi, Texas (1992).

[56] H. Kuhn, B. Naroska, *Trigger Studies for low x, Q^2 Events*, H1 internal report, DESY, Hamburg, H1-03/90-130 (1990).

[57] The H1 Collaboration (T. Ahmed et al.), *Observation of Deep Inelastic Scattering at Low x*, DESY preprint DESY-92-164 (1992) submitted to Phys. Lett. B (1992).

[58] J. Bán et al., *The BEMC Single Electron Trigger*, H1 internal report, DESY, Hamburg, H1-07/92-235 (1992).

[59] E. Vogel, thesis, Aachen University (1989).

[60] F. Eisele, Invited talk, Proceedings of the 26th International Conference on High Energy Physics, Dallas, Texas, August (1992).

[61] J. Heatherington, *ToF – The Time of Flight Device for H1*, Diploma report, Queen Mary and Westfield College, University of London (1992).

[62] E. Elsen, *The H1 Trigger and Data Acquisition System*, to be submited to Nuclear Instruments and Methods (1992).

[63] W. Haynes, *Experiences at HERA with the H1 Data Acquisition System*, DESY preprint DESY 92-129 (1992).

[64] S. Murray, *Excited Electron Production at HERA*, thesis, University of Manchester (1991).

[65] S. Levonian, U. Berthon, T. Naumann, (ed) *H1LOOK, General Purpose H1 Event Display* DESY central IBM, 'HERA01.H1LOOK.CMZ' deck 'MANUAL' (1992).

[66] J. Meyer (Ed.), *H1SIM manual*, DESY central IBM, 'HERA01.H1SIM' (1992).

[67] R. Brun et al., *GEANT Detector Simulation Program*, CERN Program Library.

[68] P. Schleper, *H1PSI Version Guide*, DESY central IBM, 'HERA01.H1PSI' (1992).

[69] G. Bernardi (Ed.). *H1REC Guide*, DESY central IBM, 'HERA01 .H1REC .JOBS' (1992).

[70] T. Sjöstrand, *The Lund Monte Carlo for Hadronic Processes – JETSET* LU TP 85-10; T. Sjöstrand and M. Bengtsson, JETSET version 6.3 (updates) Comp. Phys. Comm 43 p.367 (1987).

[71] M. Bengtsson, T. Sjöstrand, Z. für Phys. C37 p.465 (1988).

[72] A. F. Zarnecki H. Abramowicz, M. Krzyzanowski, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.121 (1991). ed. W. Buchmüller, G. Ingelman.

[73] J. P. Phillips *Kinematic Identification of Electrons in Neutral Current DIS events at HERA*, Diploma report, University of Manchester (1992).

[74] G. A. Schuler, Proceedings of the HERA Workshop Vol.1 Hamburg, Oct. 29–30 p.461 (1991). ed. W. Buchmüller, G. Ingelman.

[75] H. Hufnagel, *L5 Event Classification,* DESY central IBM, 'HERA01 .H1PHAN .CMZ' deck 'ECDOC' (1992).

[76] A. De Roeck, private communication to the ELAN group (1992).

[77] The H1 Collaboration (T. Ahmed et al.), *Hard Scattering in γp Interactions*, DESY preprint DESY 92-142 (1992).

[78] K. C. Hoeger, private communication (1992).

[79] H1 interface to program to the LUCVDM part of the LUCIFER package of G. Ingelman, A. Weigend, Comp. Phys. Comm. 46 p.241 (1987).

[80] The H1 Collaboration (T. Ahmed et al.), *Total Photoproduction Cross-section Measurement at Hera Energies*, DESY preprint DESY 92-160 (1992).

[81] The OMEGA Photon Collaboration, *Inclusive Photoproduction of Single Charged Particles at high P_t*, Z. für Phys. C43 p.63 (1989).

[82] A. De Roeck, private communication (1992).

[83] G. Ingelman, A. Weigend, Comp. Phys. Comm. 46 p.241 (1987).

[84] M. Fleischer, *Electronic Calibration of the Backward Electromagnetic Calorimeter*, H1 internal report, DESY, Hamburg, H1-07/92-256 (1992).

[85] J. Ferenci, *Electron Measurement and Trigger in the BEMC*, talk to Panel A group, DESY (1992).

[86] R. Barlow, *Statistics*, John Wiley and Sons, (1989)

Acknowledgements

> In England, more than in any other country, science is felt rather than thought...A defect of the English is their almost complete lack of systematic thinking...Science to them is a number of successful raids into the unknown. *John Desmond Bernal 1901–1971*

Well, maybe...

It is customary for students to adopt an ironic or humerous tone when writing this section of their theses, sentiments with which the author normally feels quite at home; the cynical and funny bits will have to be delayed for a few lines however, while the truthful and necessary bits are dealt with. Chief amongst these is my debt–and thanks–to my supervisor *Dr John Ellison.* Few other academics could ever have shown such dedication, energy, interest, patience or humour toward their students, especially during those periods when the students themselves displayed none of these things in their own work or lives. John never refused to help, and no problem was too dull or frustrating or difficult or time-consuming or jargon-riddled that he was unwilling or unable to help tackle it. John was also kinder and more approachable than he had any need to be, and to him I dedicate my final quotation:

> When I find myself in the company of scientists, I feel like a shabby curate who has strayed by mistake into a drawing room full of dukes. *Wystan Hugh Auden, 1907–1973*

To single out other members of the Manchester staff seems unfair, because my sincere thanks and respect goes to all of them. For their abilities in erradicating the administrative and financial problems that can beset students, and for their great humour, I will though mention *Dr Mike Ibbotson* and *Dr Fred Loebinger*; their common sense and friendliness has often served as a shelter in the drizzle of frustration and boredom which constitutes a good fraction of the students' end of academic life.

Thanks are due also to *Dr Christoph Hoeger*, who must have been driven to near despair by my stupid questions, and by my stupider answers to his sensible questions, but never showed it.

The rest of my year–*Paul Sutton, Robert McGowan and Graham Bahan*–deserve more thanks than can be expressed in fermented vegetable form and

drunk in a human lifetime, so they'll have to make do with this instead. *Paul* deserves my thanks in particular, because he's cleverer than any of us, and because he has helped me, housed me, humoured me and been the loyalest of friends. He has a degree of nouse which few give him credit for, probably because they've met him, and is truly a man of the world. *Which* world remains as yet undiscovered. Paul, thank you. *Robert* in particular too, for being the other cleverest person in the group, the best person with whom to go insane in a small tunnel tent, and for his attitude to life, dung, and the pedal abnormalities of South Sea Islanders. *Graham* was perhaps the most normal of any of us, but never to such an extent that he ceased to be a good friend.

Why no-one ever just told me to shut-up when they heard me ranting at my terminal I'll never know. They didn't though, so my thanks to those who read the manuals – *John Banks, Julian Phillips, Gareth Richards, Stuart Barnett* – which meant I didn't have to.

DESY convinced me that you don't *have* to have sandals and a bad hair cut to be a physicist, but that you probably will; three people whose hair remained excitingly bouffant, and whose stockinged feet never once peeped out from between the neo-leather slats of a cork-lined Dr Scholls were Dr Simon Lumsdon, Dr Sam Murray and Christine Milner. They are three of the best people I've met, and made my year in Hamburg what it was. Despite that, thanks. And another mention to *Sam*, to reciprocate her sentiments on the sort of person with whom one has to share offices.

The tear-ducts are about to run dry; brief mentions and sincere thanks to Lindsay Wolmersley, Andy Mehta, Phill Biddulph, Chris Hilton, Martin West, Doug Gillespie, Chris Barton, Chris Barham, Gareth Pierce, Harry 'shabby' Walton, and Zena Everitt (rooms to let).