Fermi National Accelerator Laboratory

FERMILAB-Pub-92/215-E

# The Use of Neural Networks in High Energy Physics

Bruce Denby

*Fermi National Accelerator Laboratory*
*P.O. Box 500, Batavia, Illinois 60510*

August 1992

Invited Review for *Neural Computation*

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CHO3000 with the United States Department of Energy

## Disclaimer

# THE USE OF NEURAL NETWORKS IN HIGH ENERGY PHYSICS*

BRUCE DENBY
*Fermi National Accelerator Laboratory*
*M.S. 318*
*Batavia, Illinois 60510 U.S.A.*
*denby@fnal.bitnet*

ABSTRACT

In the past few years a wide variety of applications of neural networks to pattern recognition in experimental high energy physics has appeared. The neural network solutions are in general of high quality, and, in a number of cases, are superior to those obtained using 'traditional' methods. But neural networks are of particular interest in high energy physics for another reason as well: much of the pattern recognition must be performed online, i.e., in a few microseconds or less. The inherent parallelism of neural network algorithms, and the ability to implement them as very fast hardware devices, may make them an ideal technology for this application.

## 1. Introduction

High energy physics (HEP) is the field which studies the basic constituents of matter and the fundamental forces through which they interact. Recently, high energy physicists have become interested in neural networks as HEP data analysis tools. It has been only a few years since the first investigations of neural networks for HEP were undertaken [Denby 1988, Peterson 1989], and much of today's work is still exploratory; however, the growth in applications to HEP is quite striking. At the Second International AIHEP Workshop [AIHEP 1992] at La Londe-les-Maures, France, in January, 1992, 25 applications of neural networks in high energy physics were presented. For comparison, at the first workshop in this series, in Lyon, France in March, 1990, there were only two such presentations.

In applications to date, neural networks have proven themselves to be more efficient classifiers than the simple cuts normally used in HEP, have allowed certain measurements to be made with smaller uncertainties due to their superior ability at function approximation, and have permitted analyses to be made even from heavily overlapping distributions due to their good approximation to Bayes probabilities. There have been some extremely interesting results using hardware neural networks: it appears possible to make rather sophisticated pattern analyses directly in the readout hardware of HEP experiments rather than in the standard, time consuming offline analysis.

### *1.1 HEP Accelerators*

HEP data is produced in experiments at the large accelerator centers worldwide as detailed in table I. Each site features a 'ring' in which opposing beams of particles are made to

---

* Invited review for *Neural Computation*

collide at one or more 'interaction regions' (figure 1).[1] In the collisions, daughter particles of many kinds are produced, and these are detected in arrays of particle detectors surrounding the interaction region (see figure 2). The data from these detectors constitute the HEP data sets from which physics results must be extracted.

| Accel. | Lab | Location | Beams | Energy | Period | Startup | Major Experiments |
|---|---|---|---|---|---|---|---|
| Tevatron | Fermilab | Batavia, IL | $p, \bar{p}$ | .9 x .9 TeV | 4 μs | 1986 | CDF, D0 |
| LEP | CERN | Geneva, Swit. | e+,e- | 50 x 50 GeV | 26 μs | 1988 | Delphi, Aleph,Opal,L3 |
| Hera | DESY | Hamburg, Ger. | e-, p | 27 x 820 GeV | 96 ns | 1992 | H1, Zeus |
| SLC | SLAC | Stanford, CA | e+,e- | 50 x 50 GeV | 7 ms | 1988 | SLD, Mark II |
| Tristan | KEK | Tsukuba, Jap. | e+,e- | 30 x 30 GeV | 5 μs | 1986 | Amy, Topaz, Venus |
| SSC | SSC | Ellis Cty., TX | p,p | 20 x 20 TeV | 16 ns | 1999 | SDC, GEM |
| LHC | CERN | Geneva, Swit. | p,p | 10 x 10 TeV | 16 ns | 1999 | under discussion |

Table I. Names and locations of the major world accelerator centers with the type and energy of beam used, time between collisions of particle bunches, first date of operation, and the names of the major experiments at the site. e- stands for electron, e+ for positron, p for proton, and $\bar{p}$ for antiproton. These particles are discussed in more detail in the following section. The unit of energy is the giga- or tera- electron volt (GeV or TeV), and time is measured in microseconds (μs) or nanoseconds (ns). The LHC and SSC are two large new machines scheduled to turn on before the end of the decade.

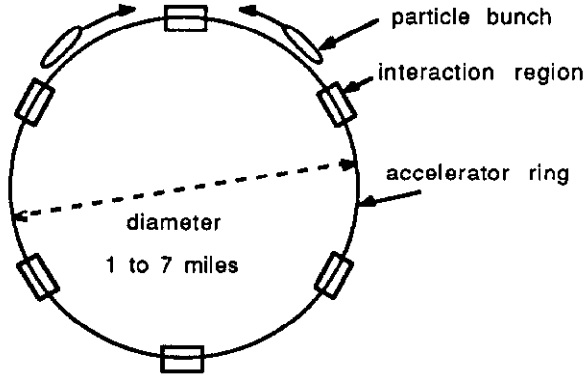


Figure 1. An accelerator with 6 interaction regions. Particles in bunches circulate in opposite directions, being brought together for collisions within the interaction regions. Normally only one or two particles within the bunches will actually collide during the crossing of two bunches. The Fermilab Tevatron has a diameter of about 1 mile. The SSC to be built in Texas will be about 7 times larger.

As more powerful particle accelerators are built, the accompanying experiments grow tremendously, both in physical size and in the demands they place upon their data readout systems. Figure 2, detailing the CDF (Collider Detector at Fermilab [CDF 1988]) experiment at Fermilab, gives an idea of the scale and complexity of the detectors used in a current experiment. Detectors at LHC and SSC will be larger again by a factor of two

[1]There are also experiments in which the extracted beam is directed onto a fixed target; for simplicity we shall not discuss these here.

or so. The volume of data produced in these detectors and the rate at which it must be analyzed are daunting. A typical experiment may record hundreds of thousands of individual detector channels, corresponding to about 1 million bits of information, for each collision, or 'event', as they are usually called, and it is not uncommon to record many millions of events during a data taking run. The particles within a beam are stored in 'bunches'. The rate of collisions varies considerably from machine to machine, and is determined by the spacing between the bunches stored in the machine, since typically only one or two particles will actually collide in each bunch crossing. In all cases, the rates are rather challenging from the standpoint of realtime processing: at the Tevatron, bunch crossings currently occur every 4 microseconds; at the SSC and LHC, they will occur about every 16 *nano*seconds. The growth in data set size and complexity, and the unprecedented data rates at today's and future colliders have been the major motivating factors in the search for more powerful data analysis tools for HEP.

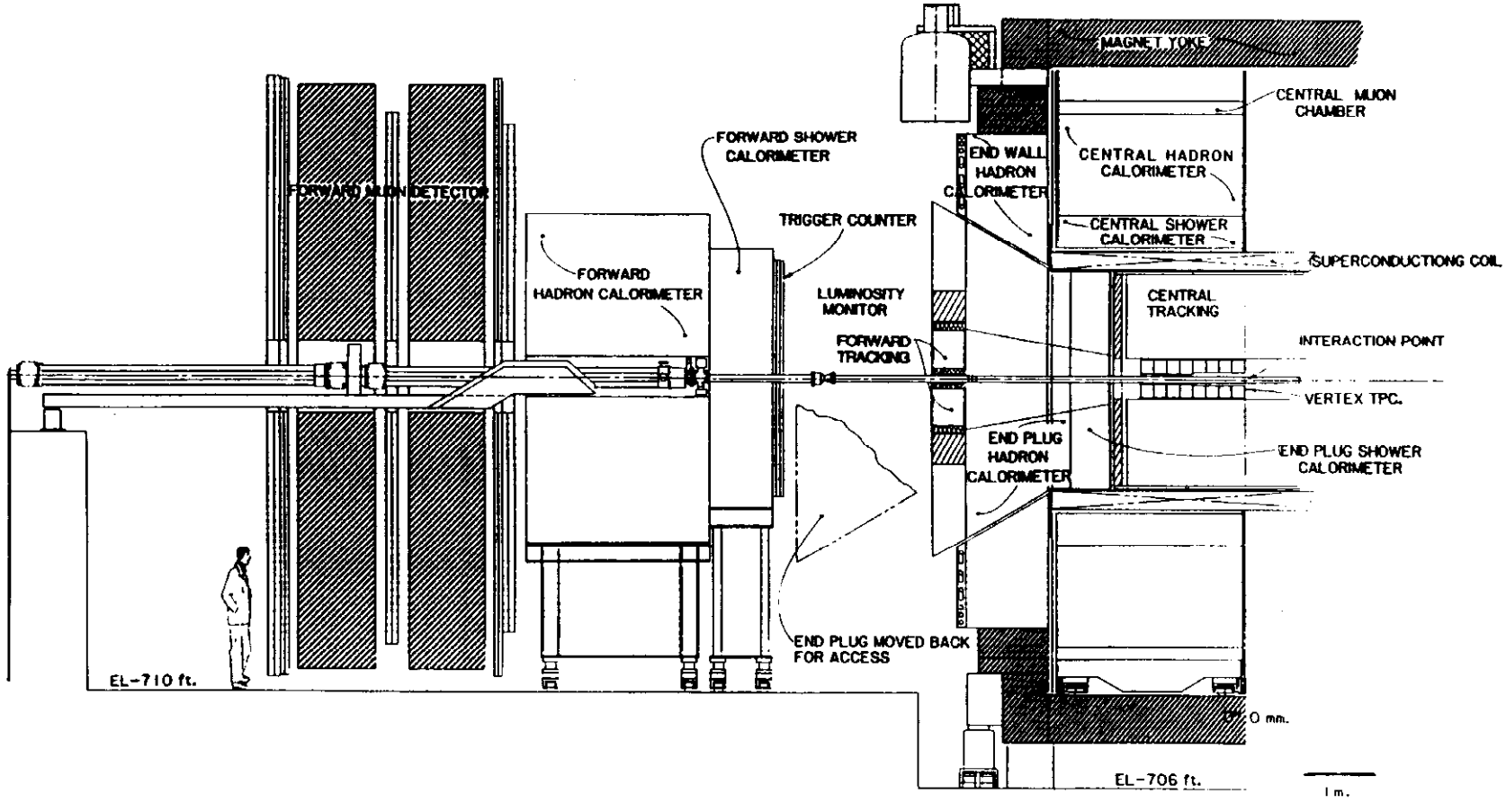


Figure 2. Elevation view of the CDF experiment at the Fermilab Tevatron. Only half of the apparatus is shown; it is symmetric about the point marked 'interaction point'. In the text, applications of neural networks to track reconstruction in a central tracking chamber; vertex finding in a vertex chamber; electron finding in an endplug calorimeter; and muon identification in a muon chamber are presented.

In the discussion of HEP neural network application which will follow, it will be necessary to have some familiarity with the terminology associated with high energy particle collisions and the detectors that record them. Sections 1.2 and 1.3 provide an introduction.

### *1.2 Areas of HEP Research - the Standard Model*

Much of current research in HEP is involved with the completion and verification of the so-called 'Standard Model' of particle physics. In this model, the basic constituents of matter are *quarks* and *leptons* as described in Table II. The constituents interact with each other by 'exchanging'[2] particles called 'bosons', as described in Table III. Particle interactions are described in more detail in the Appendix. Both quarks and leptons can interact via the *electroweak* force, carried by the W, Z and $\gamma$ bosons. This force combines the electric force, responsible for such phenomena as electricity and magnetism, with the weak force which is responsible for radioactivity. Quarks can also interact through the *strong* force, which is carried by bosons called *gluons*, usually represented as *g*. Individual quarks and gluons are not observable. The naturally occurring particles are either single leptons, or a 'composite' of two or more quarks as in table IV. A proton, for example, is composed of two 'u' quarks and a 'd' quark which are bound together by exchanging gluons. Composites containing quarks are also referred to as 'hadrons'. Leptons and hadrons interact differently in matter, as described in section 1.3.

| | type | symbol | name | charge | mass | comments |
|---|---|---|---|---|---|---|
| quarks | light quarks | $u,\bar{u}$ | up | +2/3,-2/3 | ~100 MeV | ordinary matter composed of up and down quarks |
| | | $d,\bar{d}$ | down | -1/3,+1/3 | ~100 MeV | |
| | | $s,\bar{s}$ | strange | -1/3,+1/3 | ~500 MeV | strange matter exists in stars |
| | heavy quarks | $c,\bar{c}$ | charm | +2/3,-2/3 | ~1.5 GeV | discovered in 1973 |
| | | $b,\bar{b}$ | bottom | -1/3,+1/3 | ~5 GeV | current area of study |
| | | $t,\bar{t}$ | top | +2/3,-2/3 | ~130 GeV? | not yet seen; much sought |
| leptons | charged leptons | $e^+, e^-$ | positron, electron | +1,-1 | 511 keV | causes chemical bonds |
| | | $\mu^+, \mu^-$ | muon | +1,-1 | 106 MeV | exist naturally in cosmic rays |
| | | $\tau^+, \tau^-$ | tau | +1,-1 | 1.8 GeV | first heavy lepton discovered |
| | neutral leptons | $\nu_e, \bar{\nu}_e$ | electron neutrino | 0 | 0? | not visible in detectors. except as 'missing' energy. masses thought to be zero |
| | | $\nu_\mu, \bar{\nu}_\mu$ | mu neutrino | 0 | 0? | |
| | | $\nu_\tau, \bar{\nu}_\tau$ | tau neutrino | 0 | 0? | |

Table II. Quarks and leptons and their properties, including mass and electric charge. The equivalence of matter and energy allows us to write masses in energy units of eV.

[2]The word, 'exchanging' is used figuratively. The true interaction is a quantum process which defies classical explanation. The exchange of a particle is represented by a line in a Feynman diagram as discussed in the Appendix.

| | force | symbol | name | charge | mass | comments |
|---|---|---|---|---|---|---|
| bosons | weak | $W^+W^-$ | $W^+$ $W^-$ | +1,-1 | 81 GeV | discovered at CERN in 1983 |
| | weak | Z | Z | 0 | 91 GeV | carriers of weak force |
| | elect. | $\gamma$ | photon | 0 | 0 | light is composed of photons |
| | strong | g | gluon | 0 | 0 | binds quarks in composites |

Table III. The force carrying bosons and their properties. The first column tells the type of interaction the boson mediates: weak, electromagnetic, or strong.

A collision between particles is, in the Standard Model theory, an interaction between two of the elementary constituents which they contain. For example, in a collision between a proton and an antiproton, the 'true' collision may be between a quark and an antiquark, between a quark or antiquark and a gluon, or between two gluons. When physicists examine the debris of such a collision, they are seeking information on the constituents and force carriers which are produced in the collision.

Quarks and gluons emerging from a collision are not directly observable in the detector; they are said to 'fragment' into 'jets' containing many particles as they emerge from a collision. This process is discussed in more detail in the Appendix. Jets from quarks and from gluons are slightly different in their properties, as will be discussed in section 4.2.

| | type | symbol | quark content | charge | mass | comments |
|---|---|---|---|---|---|---|
| composites | proton | p | uud | +1 | 939 MeV | atomic nuclei made of protons and neutrons |
| | neutron | n | udd | 0 | 940 MeV | |
| | pion | $\pi$ | ud, uu + dd | +1,-1,0 | ~135 MeV | most commonly produced composites |
| | kaon | K | us, ds | +1,-1,0 | ~500 MeV | |

Table IV. The composites most commonly encountered in HEP detector systems.

The most 'fashionable' areas of research in HEP today are: the study of the production and decay properties of the 'heavy' (i.e., massive) quarks, c and b; the search for the heaviest quark, called 'top', or simply, 't', which is postulated but as yet undiscovered; studies of the bosons W and Z; the search for the Higgs particle (Table V), an essential but as yet unobserved element of the Standard Model believed to be the origin of the masses of all particles; and the study of the characteristics of jets.

| | symbol | name | charge | mass | comments |
|---|---|---|---|---|---|
| Higgs | $H^0$ | Higgs | 0 | ? | essential to theory. not yet seen. gives mass to particles. |

Table V. The Higgs particle.

### *1.3 HEP Measurement Tools*

Although there are quite a number of different types of measurement tools used in high energy physics, most can be classified as one of two main types, *tracking chambers* and *calorimeters*. Figure 3 shows a generic HEP detector system with a central tracking chamber and a vertex tracking chamber, calorimeter with sections called 'electromagnetic' and 'hadronic', muon shielding iron, followed by another set of tracking chambers called

muon chambers. The figure illustrates the behaviour of the detectors for the four most commonly encountered types of particles and for a jet.

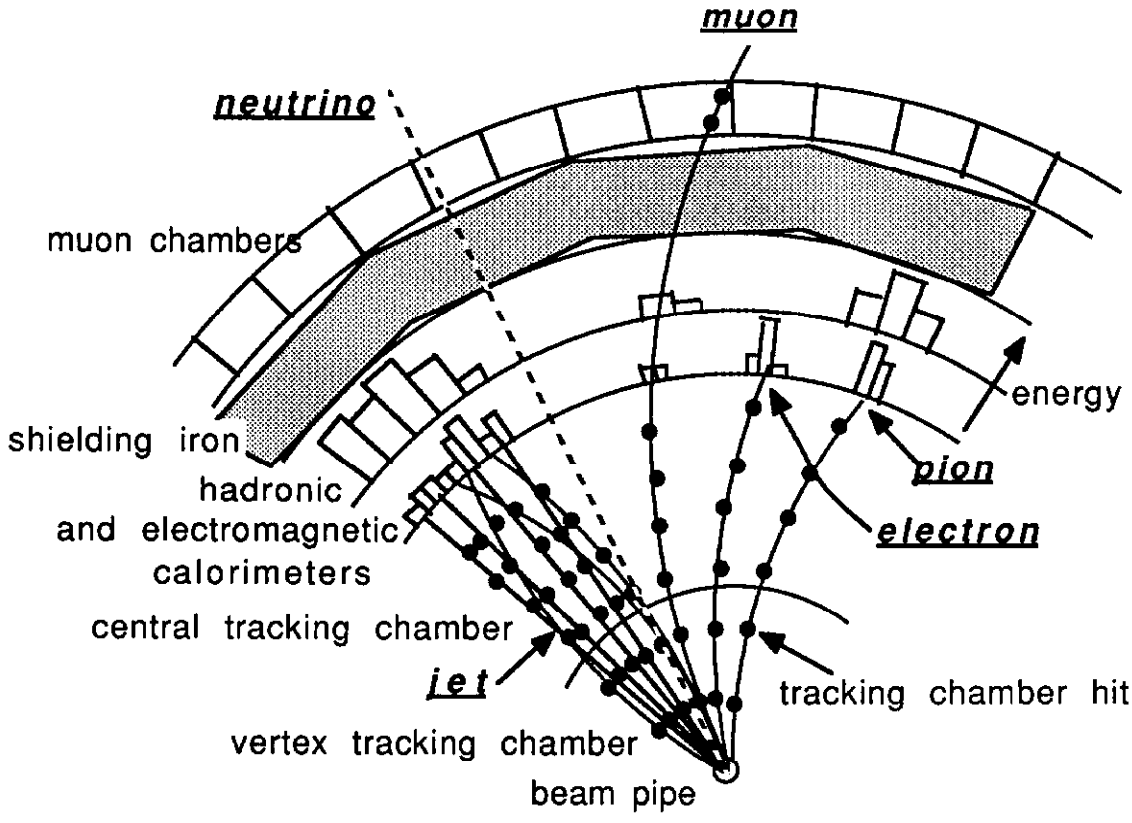


Figure 3. Behaviour of a muon, electron, pion, neutrino, and jet in a HEP detector system. The beam pipe is perpendicular to the plane of the page. The muon passes completely through the calorimeters, depositing only a small amount of energy in each section, and through the shielding iron, to be finally detected in the muon tracking chambers. The electron deposits all of its energy in a localized region of the electromagnetic calorimeter. The pion deposits its energy over a region of both electromagnetic and hadronic calorimeters. The jet is composed of many particles of different types, mostly pions, and deposits energy both in electromagnetic and hadronic sections of the calorimeter over a broad region. The neutrino does not interact at all and passes undetected through the apparatus.

Tracking chambers are used to detect the trajectories of electrically charged particles emerging from a collision. Usually the tracking chamber volume is within a magnetic field. This causes the path of the charged particle to curve, enabling a measurement of the momentum[3] of the particle. A knowledge of the momenta of all charged particles allows a complete study of the underlying dynamics of the collision to be made. When a charged particle passes through the chamber, gas molecules along its trajectory are ionized (there are also tracking chambers which do not use gas as an active medium, but we shall not discuss them here). High voltage wires spaced regularly throughout the tracking volume collect this ionization in the form of electrical pulses, which can then be passed on to the data acquisition system for analysis and reconstruction of the tracks. Position resolution finer than the wire spacing is obtained by using an electronic device to measure the time it

[3]The momentum P of a particle is defined as $P = Ev/c^2$ where E is its energy, v is its velocity, and c is the speed of light. For nonrelativistic particles $E \equiv mc^2$, where m is the mass, giving $p \equiv mv$.

takes for the ionization to drift to the wire. This is referred to as the 'drift time'. In figure 3 only wires closest to the trajectories, called 'hit' wires, are shown.

There are many different types of calorimeters but all have the same basic principle of operation. Calorimeters are normally built from many layers of metal interleaved with layers of a plastic or gas active medium. Quite the opposite of the tracking chambers, through which the particles pass uninterrupted, a calorimeter is designed to cause most particles incident upon it to interact and deposit all of their energy within its volume. The energy may be in the form of ionization or of light, but will ultimately be converted into an electrical impulse with a size proportional to the energy of the particle. Most calorimeters have two sections, called 'electromagnetic' and 'hadronic' of different composition. The electromagnetic section is designed to absorb almost all of the energy of the electromagnetically interacting particles, i.e., electrons and photons, while hadrons will deposit the largest fraction of their energy in the hadronic section. Calorimeters are usually highly segmented in order to give information on the spatial extent of the energy deposit from the particle, as shown in figure 3, where the energy in each cell is represented by the height of the tower drawn at each cell. Note that the segmentation in the electromagnetic section is twice as fine as in the hadronic section.

Calorimeters are particularly useful for identification of electrons. An electron will deposit almost all of its energy in a highly localized region of the electromagnetic calorimeter. By looking for a charged track which points at this localized region, and matching the calorimeter energy to the track momentum, an electron can be reliably identified.

Muons are charged particles which are capable of penetrating through great thicknesses of material with only minimal energy loss. For this reason, special muon tracking chambers are placed outside the calorimeter and a thickness of uninstrumented shielding iron in order to detect possible tracks from muons produced in a collision. The energy of other types of particles will be completely absorbed in the calorimeters and the shielding iron. The muon can be identified by measuring its momentum in the central tracking chamber and seeing if its projection through the calorimeter and iron matches well with a track 'stub' found in the muon chambers.

The detectors' response to pions, neutrinos, and jets is described in the caption of fig. 3.

### *1.4 Pattern Recognition in HEP - Standard Methods*
### *1.4.1 Introduction*

The only particles which are directly observable are those which have a natural lifetime long enough to allow them to be detected in the apparatus, i.e., photons, muons, electrons, and some of the low mass composites such as pions and kaons. Neutrinos normally leave no trace in the apparatus and are detectable only by their 'missing' energy. Most of the constituents produced in a collision quickly decay into these observable particles, or, in the case of quarks and gluons, fragment into jets containing many particles. The properties of the constituents must therefore be inferred from patterns in the 'visible' particles into which they decay or fragment.

Reconstructing an event involves two types of pattern recognition. The first, which we shall call *low level pattern recognition* consists of such things as finding tracks in the tracking chambers or identifying a candidate electron in the calorimeter (figure 3). The second type, which we shall call *physics process determination*, uses more sophisticated features, for example the angular distribution of the jets in the event, to try to identify the

underlying physics of the interaction which took place. Note that this nomenclature is not the same as typically found in classical pattern recognition, since classification, normally considered 'high-level', can occur both in our low-level and high-level pattern recognition[4] . In HEP, the distinction between high-level and low-level pattern recognition is based upon the complexity of the features used to perform the classification. Examples of the two types will be given in the sections to follow. We shall see that neural networks have found application to both.

In HEP it is also necessary to distinguish whether the pattern recognition is to be performed 'on-line', i.e., in real time, or 'off-line'. On-line pattern recognition is performed on the data before it is logged, in a part of the experiment referred to as the 'trigger'. Off-line pattern recognition is done with conventional computers operating on the data after it has been logged to permanent storage media. These two areas will be discussed in more detail below.

### *1.4.2 Triggering*

New HEP experiments study increasingly rare physical processes. The implications of this for data acquisition systems are best illustrated by an example. One of the main motivations for the construction of LHC and SSC is the search for the Higgs particle. The probability of producing a Higgs particle when two protons cross paths is so small that this would have to occur $10^{34}$ times per second[5] in order to produce a reasonable sample of detectable Higgs particles, say 1000, during a one-year run. The probability for other processes however, not involving the Higgs, is higher by a factor of about $10^{13}$. This implies that, during this one-year run, events containing background processes will be continuously produced at a rate of about 1 billion per second. It is neither desirable, nor feasible, to log all of these events to permanent storage media such as magnetic tape. On-line pattern recognition, called 'triggering', is required to reject background events and retain the rare interesting events. Although LHC and SSC represent an extreme case in high-rate HEP data acquisition, the problems are common to all HEP experiments.

---

[4] Segmentation of the data into events is performed trivially using timing information which correlates a block of data with the time of a particular bunch crossing.

[5] The $10^{34}$ per second is technically the accelerator 'luminosity' required to produce the 1000 Higgs particles. Luminosity is defined as the the square of the number of particles per bunch, times the number of bunches per beam, times the revolution frequency of the bunches within the ring, divided by the cross sectional area of the beams.

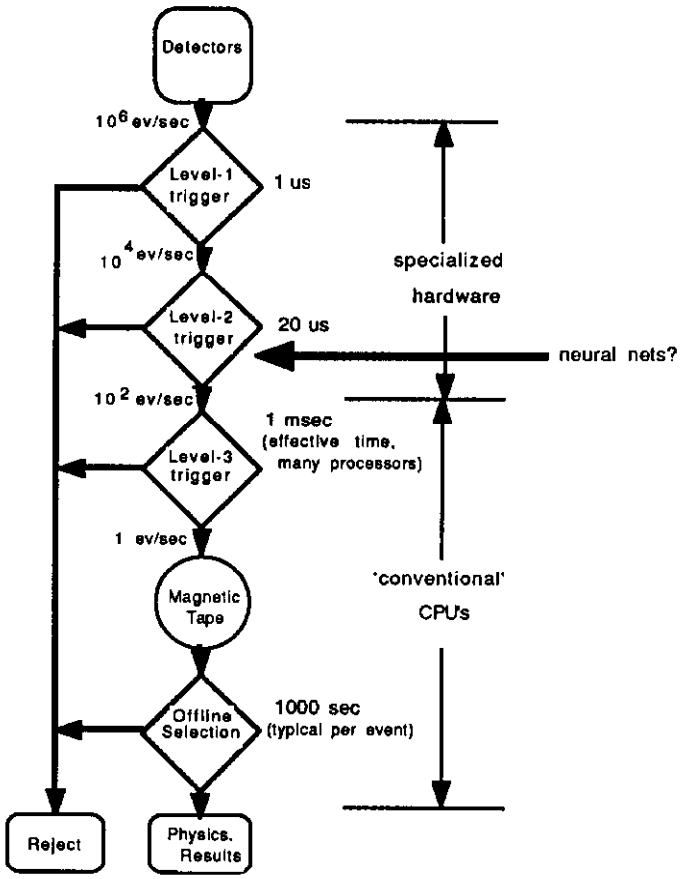


Figure 4. Generic HEP multilevel trigger system.

Figure 4 shows a typical multilevel HEP trigger system. The data from the detectors passes into the trigger as a stream of events, each containing all the detector data produced in a single collision. Each level of trigger rejects most of the events it receives and passes the remainder on to the higher level triggers. Levels 1 and 2 are typically implemented as fast specialized analog or digital hardware, while level-3 is a 'farm' of conventional processors. The processing times and event rates shown at each level are generic, but typical of those encountered at current proton-antiproton collider experiments such as the CDF experiment at Fermilab [CDF 1988]; rates will be one to two orders of magnitude higher at LHC and SSC.

In level-1, simple tests on global event information are performed, for example: (1) comparing to a threshold the summed transverse energy, $E_t = \Sigma_i E_i \sin\theta_i$, where $E_i$ is the energy in calorimeter cell i and $\theta_i$ is the angle with respect to the beam axis of a line from the collision point to the calorimeter cell; (2) looking for the presence of a charged track with transverse momentum, $P_t = P \sin\theta$, where P is the track momentum, above a threshold; (3) looking for the presence of one or more track segments in the muon chambers. The first and second cuts eliminate 'soft' interactions. Most interesting physics processes involve 'hard' scatters of two constituents in the beam particles, which produce particles at large angles to the beam direction and thus deposit in the calorimeter substantial energy transverse to the beam direction. 'Soft', glancing collisions of beam particles are much more copiously produced than hard scatters, and most must be

rejected. The third cut is useful since high $P_t$ muons are produced in many of the interesting processes currently under study, but are produced only with low probability in background processes. Level-1 triggers have a typical processing time of about 1 microsecond and reduce the rate due to backgrounds by about two orders of magnitude.

In the level-2 trigger, somewhat more sophisticated tests can be done, for example: (1) looking for a match between a high-$P_t$ track and an energy cluster in the electromagnetic calorimeter, indicating the presence of a candidate electron, or between a high-$P_t$ track and a track segment in the muon counters, indicating a candidate muon; (2) looking for the presence of localized clusters of energy in the calorimeter, which will correspond to jets, with $E_t$ above some threshold. Validating the presence of leptons and jets as in (1) and (2) above ensures that the event is more likely to have come from an interesting physics process. Ten to twenty microseconds are available for level-2 decisions.

Level-3 triggers are executed using algorithms written in standard high level computer codes running on a 'farm' of conventional processors which operate in parallel on separate events. As each event comes into level-3, it is immediately sent to an available processor. The processing done by level-3 can be quite sophisticated, in some cases being identical to the code used in offline analyses. Some of the typical analyses performed in level 3 are: 1) reconstruction of charged tracks; 2) accurate calculation of the position of the collision point in order to reject events too far from the detector center, to allow more accurate calculation of $E_t$ of calorimeter cells, and to detect multiple vertices; 3) high quality electron and muon identification using accurate $P_t$ measurements of the tracks; 4) imposition of isolation cuts, i.e., requiring that an electron or muon have very little energy surrounding it in the calorimeter; 5) formation of composite triggers, e.g. electron plus missing transverse energy plus one or more jets would be a good trigger for top quark production. Such calculations as these are too complicated to be performed in level-2. The time to process a single event in level-3 may be of the order of a second, however as there are many processors operating in parallel, the effective processing time is a few milliseconds per event.

### *1.4.3 Offline Reconstruction*

Offline reconstruction is the final event reconstruction in which all available information is processed using whatever data analysis techniques may be available. Normally all the data from a run will be processed in a single reconstruction pass in which data sets of special interest are created, e.g., one for the physics of b and c quarks; one for the search for the top quark; one for W and Z physics, etc. These are often analyzed many times over with ever more refined sets of selection cuts. Analysis usually proceeds with the definition of several feature variables upon which one dimensional cuts are placed. The use of likelihood techniques is also common.

The offline analysis does not have the same real time constraint as online reconstruction; however, the codes used to process high energy physics data are normally tens of thousands of lines long and require substantial computing resources in order to complete the processing in a reasonable amount of time. It is not uncommon for a complete offline reconstruction of a particular physics process to take one or two years.

## *2. The Need for Neural Networks*

In high energy physics, neural networks have been used both in real-time and offline applications. Most applications to date have used MLP's trained with backpropagation, although a few instances of the use of learning vector quantization (LVQ) and feature

maps have also appeared. Recurrent networks have been applied to the problem of charged track reconstruction as discussed in section 5.1.

For the offline applications, the advantage to HEP is the same as that for other fields: near optimal classification with a minimum of computational overhead. In the real-time applications, neural networks present an advantage because of their parallel architecture which allows for faster processing. We now discuss these two areas in more detail.

### *2.1 Neural Networks for Triggering*

It is interesting to note that some of the functions performed by standard level-1 and 2 triggers as discussed above, i.e., thresholding performed upon a linear combination of inputs, already resemble those performed by an artificial neuron. High energy physicists building fast trigger electronics have for decades been making use of electronic devices called 'discriminators' for performing this function. The idea of applying true neural network technology in HEP triggering, however, is quite new [Denby 1988, Denby 1990], and it is far from being accepted as a standard tool.

Neural networks are a natural choice for incorporation into triggering systems due to their speed of execution, made possible by their parallel architecture and the ability to implement this architecture in silicon. This processing speed can be extremely valuable in very high rate data acquisition systems. At present most projects to use neural networks in triggering foresee an application at level-2, since the processing times of existing neural network chips are of the order of a few microseconds and are thus too slow for level-1. As faster hardware becomes available, level-1 applications can also be envisioned.

Although trigger systems using conventional electronics can probably be made to handle the rates to be found at SSC and LHC, neural networks can make the triggers far more efficient and less costly by moving to level-2 the complex pattern recognition normally done in level-3. In section 3 we shall show some specific examples of this: accurate muon $P_t$ measurement in a few microseconds; application of an isolation cut at level-2; a possible scheme for determining the position of the collision point online, etc. This will reduce the requirements placed on the level-3 processor farm and significantly reduce the amount of data which must be recorded on tape for later analysis.

Another attractive feature of neural nets for triggering is their programmability. In the past, many level-2 triggers have been built as hard-wired special purpose electronic devices. To change the algorithm in such a device implies rebuilding it or re-wiring it. In a neural network, the algorithm can be changed simply by downloading a different set of weights, which will make neural network triggers much more flexible than their predecessors.

### *2.2 Offline Applications*

Historically, high energy physicists have eschewed 'complicated' data analyses in favor of simple one dimensional cuts. In HEP, such problems as incomplete understanding of detector response, and heavy dependence upon Monte Carlo models render the extraction of a final physics result from the experimental data an extremely difficult and time consuming task, sometimes requiring hundreds of man years of effort. There was a strong tendency to try to keep the analyses as simple as possible. However, over the years in HEP, considerable experience in detector construction techniques and in software generation has been gained, and detector simulation packages which model instrumental effects have become extremely sophisticated. Too, with the growth of collaboration size,

particular groups of researchers within an experiment have been able to devote themselves exclusively to data analysis problems.

The key to the value of neural networks in offline HEP analyses is in creating efficient cuts to retain events from rare physics processes while rejecting as many as possible of the background events. A further advantage is that neural networks may make possible certain analyses which previously were considered hopeless precisely because simple one dimensional cuts were known to be ineffective discriminators. An example of this is the classification of quark and gluon jets, which we shall discuss in section 4.

It has been argued that although a series of one dimensional cuts is less efficient than a multidimensional cut, this can be compensated for by taking more data. As the interesting physics processes to study become more rare, however, this reliance on increased statistics becomes impossible: it becomes necessary to extract as much information as possible from the data at hand.

### *2.3 The Problem of Training Data*

One of the major goals of HEP is to identify and characterise the properties of as yet unseen constituents in the standard model. This however presents a problem for classification schemes involving supervised learning since there is no existing real data containing these particles. It follows that Monte Carlo data must be generated according to some model. In some cases, there are a number of rather different models to choose from. Any classification based upon these models will therefore be biased towards the model chosen. This is of course a problem for any type of classifier, however a number of high energy physicists are concerned that it will be more difficult to understand model dependence using neural networks than using a simpler type of classifier. This is used as an argument against using neural networks in HEP analyses. Although it is true that model dependence in a nonlinear classifier is somewhat more difficult to characterize than in a linear classifier, the superior performance of nonlinear classifiers has led some researchers to expend the additional effort necessary to characterise the model dependence. This will be seen in some of the applications described in section 4.

This effect is particularly important in triggering. Events rejected by a trigger will not be recorded, and so can never be used to check what the trigger was doing. For this reason, there has been a tendency in the past to keep trigger cuts as simple as possible to facilitate understanding of the trigger efficiency. This this 'validation' problem is not important for triggers based upon low level pattern recognition such as track segment finding or electron identification since modern detector simulations can quite reliably simulate such simple entities as tracks and electrons. However, because of possible biases from model dependence, there is still work to be done in HEP to show convincingly that unbiased information can be extracted from data taken with triggers which select specific physics process, whether they use neural networks or more conventional technology.

## 3. Applications to Low Level Pattern Recognition

These applications, as well as those in later sections, are summarized in Table VI.

| | Problem | Training Set | Test Set | Network | Results and Comments |
|---|---|---|---|---|---|
| Low Level - Triggering | muon trigger test beam experiment Fermilab | Monte Carlo tracks | Real online data | 15-64-64 MLP ETANN | Fiftyfold improvement in position resolution over conventional trigger. To be applied to D0 expt. muon upgrade |
| | isolation and b trigger for CDF calorim. Fermilab | Real and Monte Carlo data | Real and Monte Carlo data | 50-1<br>50-4-1 (isol.)<br>50-10-1 (b)<br>MLP+ETANN | Simulation results show big improvements in trigger rejection power. Will take data soon. |
| | level-2 trig. for H1 expt. at HERA | Monte Carlo data | Monte Carlo data | 19-??-1 MLP+silicon | Proposed. Simulations show a 10 fold background rejection and 10 microsecond execution time, suitable for level-2. |
| | electron i.d. for LHC at CERN | Monte Carlo data | Monte Carlo data | 192-96-1 MLP<br>92-92-1 silicon | Simulation studies show good rejection. Prototype chip had propagation time of 15 ns, suitable for LHC or SSC |
| Low Level - Offline | Find primary vertex at E735 expt. Fermilab | Real collider data | Real collider data | 18-128-62 MLP | Overlapping 18 wire sections summed. 3 times better resolution than TOF. Finds multiple vertices naturally. |
| | Kink finding for charged particle tracks | Monte Carlo data | Monte Carlo data | 5-5-1<br>5-10-1 (par.)<br>14-7-1<br>14-14-1 (res.)<br>42-6-1 | Both parameter and residual neural net methods exceed performance of standard chi squared technique. Parameter method gives 20x speedup. |
| Physics Process Determination | Z decay probabilities into b,c, and (uds) | Monte Carlo data | Real data from Delphi expt. | 19-25-3 MLP one output node for each of b,c, (uds) | Z decay probabilities into b,c, and (uds) measured more accurately than with standard method. |
| | Quark/gluon discr. at CDF Fermilab | Monte Carlo data | Real data from CDF experiment | 8 - 6 - 1 MLP (feature map in new analys.) | Heavy overlap of quark/gluon distrib. First evidence for quark fraction increase with Et. |
| | B tagging | Monte Carlo data | Monte Carlo and real data | MLP<br>MLP + LVQ | Numerous references on b tagging, mostly at LEP. |
| Recurrent nets and track reconstr. | Track recons. with Denby-Peterson Net | hand wired | real data from Aleph expt. | fully connected recurrent Hopfield network | Neurons are links between hits. Links form tracks as system settles. Tested on Aleph expt. |
| | Deformable templates/ elastic arms | none | Monte Carlo data | dynamical system | Inspired by Denby-Peterson net and elastic net methods; not really a neural network. |

Table VI. Summary of the main HEP neural network applications covered in this paper.

### *3.1 Trigger Applications*

We will treat in this section only those trigger applications which have already been realized or have been seriously proposed. Some of the other low level pattern recognition applications which follow are also intended for triggering but are still just studies.

#### *3.1.1 First Real-time Application: Muon Trigger*

The first real-time application of a neural network in HEP was accomplished recently at Fermilab [Lindsey 1992].

##### *3.1.1.1 Conventional Method*

Identification of a muon with a transverse momentum $P_t$ above a threshold is a useful trigger for detecting decays of W's, Z's, and b quarks since each will decay about 10 percent of the time to a muon. The cut on $P_t$ is necessary since background processes produce many low $P_t$ muons. A measurement of the $P_t$ of a muon in the trigger requires a knowledge of the angle of the muon track at the muon chamber. Although, offline, the wire drift times can be used to calculate the track angle quite accurately, in the trigger, only the information on which wires were hit is available, resulting in an inaccurate measurement of $P_t$ in the trigger. It is therefore necessary to set the $P_t$ trigger threshold quite low in order to avoid discarding high $P_t$ events which have been poorly measured. This introduces a large amount of background.

##### *3.1.1.2 Test Beam Results*

In a simple test beam experiment at the Fermilab Tevatron, slopes and intercepts of muon tracks traversing a small prototype drift chamber were calculated accurately, in real-time, using a commercial VLSI neural network chip incorporated into the standard drift chamber data acquisition system. This was a test experiment carried out in an auxiliary particle beam; in a full scale collider experiment, the drift chamber would be duplicated many times over to cover an area of many square meters surrounding the other measuring devices, as in figure 3. The drift chamber sense wires signals appeared on Time to Voltage Converters (TVC's) which convert the drift time of the ionization to the wire into a voltage. The setup is shown in figure 5. The beam dump in the figure simulates the shielding iron of figure 3. The small circles in the drift chamber volume represent the wires and the small horizontal lines above and below represent the TVC values interpreted as a drift distance. Note that there is an ambiguity as to on which side of the wire the particle passed. The neural net must resolve this ambiguity.

The wires in figure 5 are paired vertically. For each of the three pairs, two signals are produced: a drift time and a latch indicating whether the lower or upper member of the pair was hit. The drift time signals had to be duplicated 4 times in order to achieve sufficient fanout for the analog neural net chip. These 12 signals were coupled with the three latch signals to form the 15 inputs to the neural network chip, configured as a MLP. Sixty four hidden units in a single layer were used. The output layer consisted of sixty four units divided in a group of 32 to encode slope and a group of 32 for intercept (this type of readout has been used in several previous studies of tracking with neural networks [Denby2 1990, Lindsey 1991, Lindsey2 1991]). Each output unit covers .625 centimeters in intercept or .05 radians in slope. The network was trained on 10000 tracks generated with a simple Monte Carlo, using gradient backpropagation. Target patterns consisted of gaussian histograms with means equal to the target slope and intercept and r.m.s. width of one bin. Architectures with fewer hidden units were also tried, but these

resulted in degraded performance. (In an analog hardware network such as this, extra hidden units may be needed simply to increase fanout.) The weights obtained were downloaded into an Intel Electronically Trainable Analog Neural Nework chip (ETANN) after performing emulation and chip-in-the-loop training using the Intel ETANN Development System [Intel 1991].

The intercept position resolution available using the conventional trigger technique, which does not make use of the drift times, is 5 centimeters. The neural network trigger was found to have a position resolution of 1.2 millimeters. This resolution is only about a factor of two worse than the best obtainable offline using the complete reconstruction algorithm, but is available in about 8 microseconds. The neural network result, as shown in figure 5, can be passed back to the readout motherboard for readout with the rest of the event information, without introducing dead time in the data acquisition system.

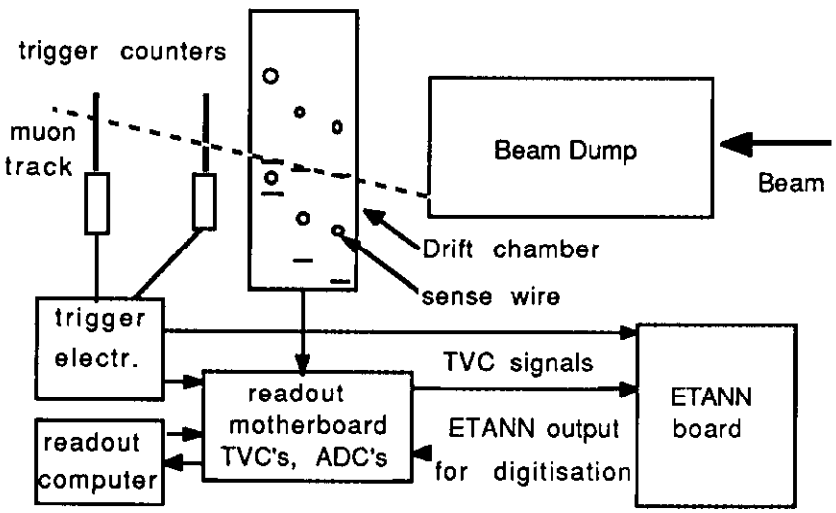


Figure 5. Setup for the drift chamber neural net trigger test.

### *3.1.1.3 Future Plans*

The drift chamber used in the above tests was a prototype of chambers which are currently installed in the D0 experiment at Fermilab [D0 1983]. A group on the D0 experiment is currently installing an ETANN chip on one of their chambers to take test data during the 1992 run [Haggerty 1992]. They also plan to incorporate the ETANN readout into the trigger of the upgraded D0 detector in the 1994 run of that experiment [Fortner 1992]. This will allow a more accurate determination of the muon $P_t$, which will allow the threshold to be lowered and significantly reduce the amount of background data recorded.

### *3.1.2 Test Case: the CDF Experiment*

Neural network trigger hardware is being installed for the 1992 run of the CDF experiment. We describe below the conventional CDF calorimeter trigger and the neural network improvements to it.

### *3.1.2.1 Conventional Techniques*

The trigger for the CDF experiment at Fermilab has been in operation since the first experimental run in 1987 [CDF 1988]. In this trigger, signals from the calorimeter cells appear as analog levels (i.e., voltages) at the ends of special 200 foot cables, where they are received by the trigger receiver boards. From this point on, the trigger can be thought of as operating on an array of voltages of size 24 (azimuthal angle) by 42 (pseudorapidity, related to polar angle) by 2 (electromagnetic/hadronic compartment), which represent the energies in the calorimeter. Analog processing is used in level-1 and level-2 for the cluster analysis, in which the total $E_T$ of the cluster, the number of towers in the cluster and the cluster width are computed. Once the cluster analysis is finished, additional digital processing is performed, operating upon the cluster quantities using the level-2 processors and special function modules; e.g., calculation of the ratio of the cluster's energies in the electromagnetic and hadronic calorimeters.

### *3.1.2.2 CDF Neural Network Triggers*

The existing CDF calorimeter trigger is very powerful, but is based upon the philosophy that clusters can be adequately described by their position, their width, the number of towers they contain, and the ratio of hadronic to electromagnetic energy they contain. Indeed, this information is adequate for a great many triggers. However, there are instances when a more sophisticated cluster analysis would be fruitful. A neural network trigger is currently being installed at the CDF experiment [Wu 1990]. For every cluster found by the cluster finder, the new trigger selects 5 by 5 trigger tower region of interest (in hadronic and in electromagnetic compartments) centered on the cluster and passes the 50 analog signals to analog neural network chips [Intel 1991]. The chips are programmed to execute three different cluster algorithms: (1) determine if the cluster could be an isolated photon in the central calorimeter; (2) determine if the cluster could be an isolated electron in the endplug[6] calorimeter; (3) determine if the cluster could have come from the semileptonic decay of a b quark[7]. None of these analyses would be possible using the existing calorimeter trigger without extensive hardware modifications.

We choose the isolated endplug electron trigger [Denby 1991] as simple illustrative example. There is a very high rate of clusters in the endplug which pass the conventional electron trigger but are in fact due not to electrons but to background processes. In the past, a high energy threshold was used in the endplug in order to reduce the rate from these false electrons. This, however, is undesirable since it rejects a significant number of real electrons along with the background. Electrons from the decay of a W are normally isolated in the calorimeter; i.e., have very little energy surrounding them. In 1992, an isolation requirement, implemented by a neural network, will be tried in the level 2 trigger to allow the same trigger rate but with a lower energy threshold. Normally such a cut would have been made in the level-3 trigger. The conventional level-2 trigger cannot implement this cut since it no longer has access to the individual tower energies after cluster finding.

---

[6]The endplug is a name given to calorimeters or other detectors which fit into the end openings of the cylindrical central detectors (figure 2).

[7]A semileptonic decay is one in which a quark decays to a lepton plus other particles. In a purely leptonic decay, the quark decays to a charged lepton and a neutrino.

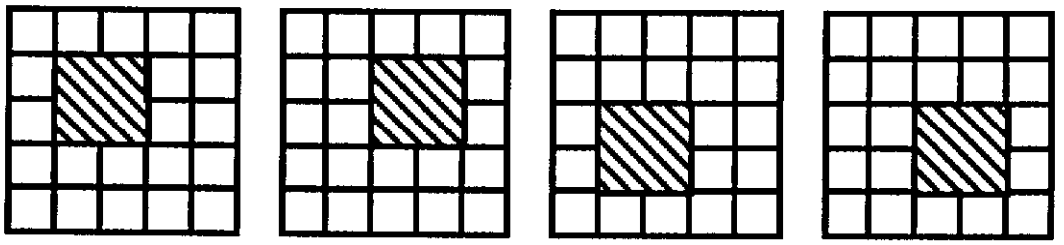

Figure 6. Isolation templates for plug electron trigger.

The neural net endplug isolation trigger operates upon 5 by 5 tower regions of the electromagnetic and hadronic calorimeters as shown in figure 6 (only the electromagnetic part is shown in the figure). The dark central region is meant to contain the electron, which normally produces a narrow cluster in one or two towers. Four templates are necessary since some of the electron's energy may spill over into 2 to 4 towers and since the center of the tower as found by the cluster finder may not perfectly center it in the 5 by 5 array in all cases. Each template will be represented as a hidden unit in the neural network, and each tower has a weight connecting it to one of these hidden units. Cells in the central region have a weight of F, and cells in the outer region have a weight of -1. Thus, the quantity presented to the hidden units, which are used as comparators, is

$$F * E_{inner} - E_{outer}$$

If this quantity is negative, the hidden unit will not 'fire': the energy outside the central region was greater than some fixed fraction of the central region energy and the cluster is thus not isolated. If the quantity is positive, the neuron fires, indicating an isolated cluster. If any of the templates fires, the cluster is isolated; i.e., the output unit simply sums up the outputs of the hidden units.

The value F = .16 was found to be optimum in the present application. (Since the network is very simple, and essentially 'hand wired', it was not necessary to train the network using, e.g., backgpropagation.) Using this value, in a simulation of the trigger operating on real data from a previous CDF run, it was possible to lower the energy threshold for endplug electrons from 23 GeV to 15 GeV, while reducing background by a factor of 4 and retaining 95% of electrons from the decay of W bosons. This will allow access to electrons of energies lower than were previously obtainable, which will be valuable for studying certain decay properties of W bosons.

The isolated central photon trigger operates in an analagous way, except that it operates in the central region of the calorimeter rather than the endplug, and in this case has only one template with a single tower in the central region of the 5 by 5 grid. This trigger will provide access to a class of physics events containing so called 'direct' photons, which tend to be isolated in the calorimeter. Without the isolation cut, the high rate of background limits the amount of good data which may be taken.

In the case of the semileptonic b-trigger [Wu 1990], a Monte Carlo program was used to generate events containing the semileptonic b jets and background events not containing b jets. The semileptonic b jets will contain an electron as well as other particles, while the background jets will not contain electrons. A full detector simulation was used in order to model as closely as possible any instrumental effects. A training set was made from 5 by 5 regions centered on the b jets extracted from the signal and background events. This was used to train a feed forward neural network with 50 inputs, one hidden layer of 10 units and a single output unit to discriminate between b jets and non-b jets. This is the

only one of the three CDF neural network triggers which uses a network trained with backpropagation. The other two are 'hand wired' nets. A simulation of the trigger showed a reduction of background of a factor of about 100 while retaining 30 percent efficiency for b's. The weights found in the simulation will be loaded into the neural network chip in order to allow online identification of the b-jets. It would be impossible to carry out a discrimination such as this using conventional computer hardware within the time limits of the level-2 trigger, i.e., about 20 microseconds.

The hardware for these triggers is being installed and should begin taking data soon. All three of the triggers will be implemented with identical hardware. It is remarkable that such different algorithms can be implemented with the same hardware simply by downloading different weights. Future modifications to any of the algorithms will also be easy because of the programmability of the neural net.

### *3.1.3 Other Trigger Applications*
#### *3.1.3.1 The H1 Experiment*

The Hera accelerator , which collides electrons upon protons, is just coming on line at the time of writing. The experiments H1 and Zeus there will study the momentum distribution of constitutents within the proton and measure the coupling strength of the gluon to the different quarks. At Hera, the rate of produced events due to background processes such as interaction of a beam particle with a residual gas molecule in the vacuum system is some $10^5$ larger than the rate due to physics processes of interest. In the H1 experiment, a 4 level trigger system is envisioned in order to reduce this high rate to a manageable level of about 100 Hz. Level 1 is a digital pipeline which reduces the rate by about a factor of 100. An additional reduction of a factor of 10 is required in level 2 in order to provide an acceptable rate into levels 3 and 4, which are implemented in software on conventional computers. The level 2 trigger must complete its processing within 20 microseconds. A hardware neural network has been proposed as a solution to this problem [Ribarics 1991, Ribarics 1992, Ribarics2 1992]. We describe the approach below.

In level 1, 16 simple trigger quantities, such as total summed energy, total summed transverse energy, total energy in the central region of the calorimeter, etc., are compared to thresholds. Level 1 however ignores correlations among the input variables. More sophisticated cuts will be made in level 2 by augmenting the level 1 quantities with additional information which becomes available after the level 1 decision time and feeding the resulting list of variables to a feed forward neural network. At present 19 input variables, including energy sums in subsets of the calorimeter, information on the vertex position, number of charged tracks, etc. are used. The neural network will use these 19 variables to determine whether the energy patterns in the event have come from an electron proton collision or from a beam-gas collision or other background.

The detailed architecture of the neural network is still under development, however typical results using Monte Carlo data with a MLP show retention of 98 percent of events from interesting physics processes and rejection of 90 percent of background events; i.e., the reduction factor of 10 is achieved while maintaining excellent efficiency. The algorithm is planned to be executed by a Siemens MA16 neural network chip [Siemens 1992], which should be able to finish processing in 10 microseconds, well within the allocated time.

#### *3.1.3.2 Trigger R&D at CERN*

Some of the research and development projects at CERN are investigating neural networks for triggering applications for the LHC accelerator. In one project, a type of detector called a 'transition radiation detector', TRD, was designed to tell electrons from pions in an online trigger [Hansen 1992]. The TRD will have 192 input wires, embedded in a special substrate, which sense the passage of the electron. The analog values from these wires will be fed into a MLP with 96 hidden units, and one output unit which signals whether or not an electron was present. In a simulation, the TRD rejected 92% of pions, and accepted 90% of electrons. These results were better than the 89% rejection, 90% percent acceptance obtained with a more traditional analysis. Ultimately the neural network will be implemented in silicon with fixed weights. A prototype chip has already been built which has 32 input units and 32 hidden units. The propagation time through the chip is 15 nanoseconds; thus, the processing is sufficiently fast for incorporation into a first level trigger for LHC or SSC.

A group at the Dutch lab NIKHEF is investigating a calorimetry based neural network trigger for the LHC accelerator as part of a research collaboration at CERN [Vermeulen 1992]. The approach is similar to the CDF trigger in that it will perform simple pattern matching upon energy patterns in local regions of the calorimeter. This is a two year pilot project which will compare the neural net solution to other techniques. The exact hardware implementation is still under development but will probably use a fast digital signal processor to implement the neural network algorithm.

### *3.2 Other Low Level Pattern Recognition Applications*
#### *3.2.1 Track Segment and Vertex Finding*

This discussion is from [Lindsey 1991] in which data from a proton antiproton collider experiment were fed to a MLP trained to find the primary vertex of the event, based upon drift times in the z-chamber, a drift chamber with three layers of wires placed near the beam pipe. The primary vertex[8] is the point from which the tracks in the event emanate, and marks the location of the collision. Figure 7 shows the hits in the chamber for a typical event; here, only the hit wires are shown, not the drift times. The hits appear to emerge from a point on or near the beam line.

The vertex position in collider experiments is normally not available online. This would, however, be very useful since it could be used to improve trigger calculations which assume a nominal vertex position at the center of the apparatus, and to flag or reject events which contain multiple interactions (i.e., more than one primary vertex). Vertex calculations are normally not performed until the offline analysis. A cross check of the offline analysis is provided by the time-of-flight (TOF) system, which crudely measures the vertex position using timing information.

---

[8]This vertex is related to but not technically the same as the vertices discussed in the Appendix in connection with Feynman diagrams. The discussion here is of vertices which are physically discernible in the apparatus. The vertices in a Feynman diagram are mathematical entities.

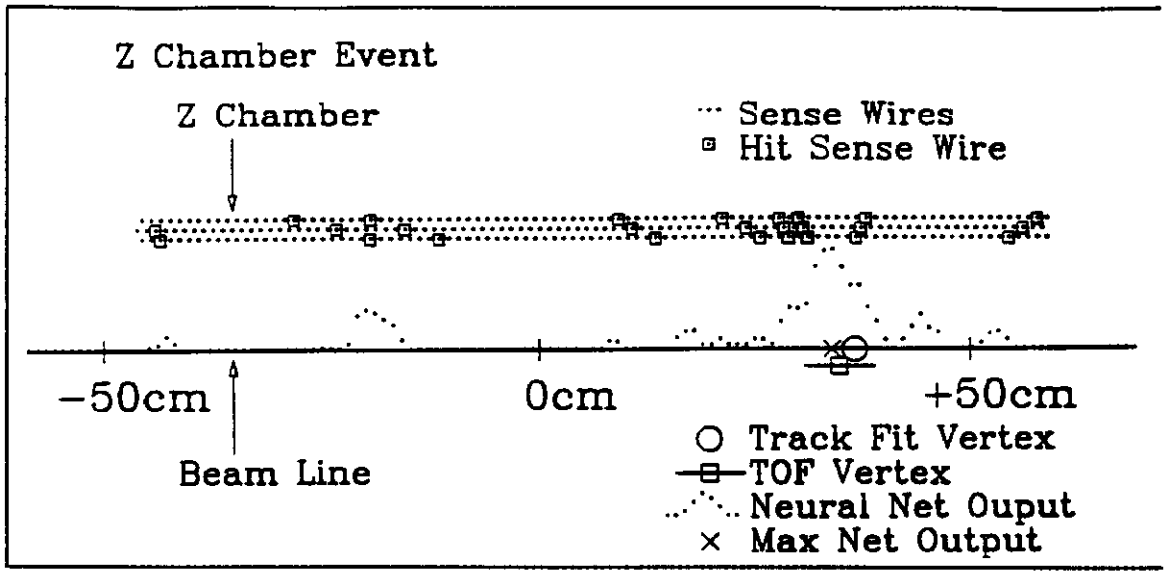


Figure 7. A typical proton antiproton collision viewed in the z-chamber of the E735 experiment.

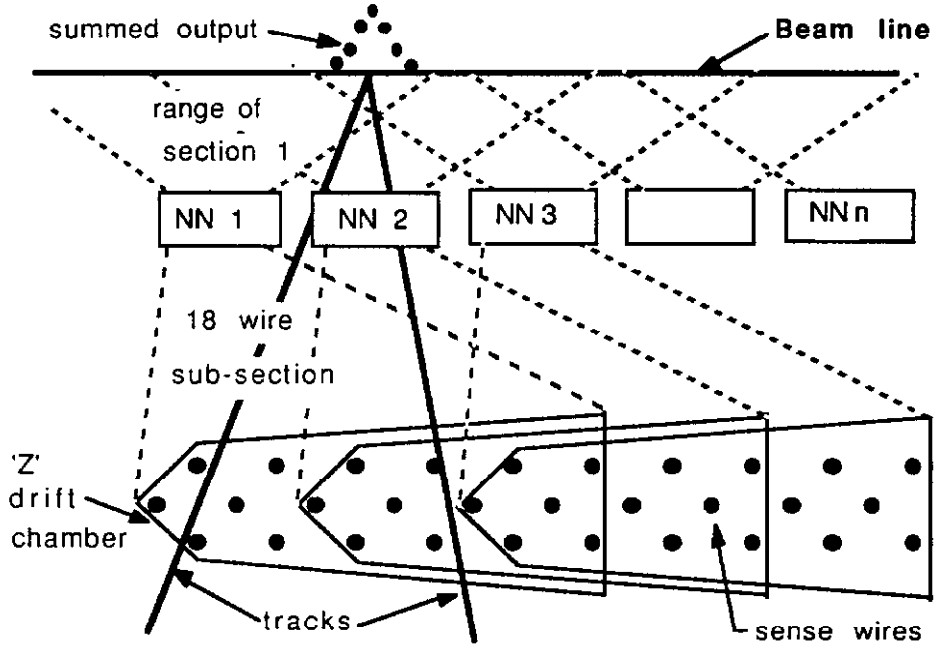


Figure 8. The z-chamber sits near the beam pipe to detect outgoing charged particles whose trajectories can be used to determine the vertex position. It is divided into 18 wire subsections each with its own MLP, whose outputs are summed to give a distribution whose peak indicates the most probable vertex position.

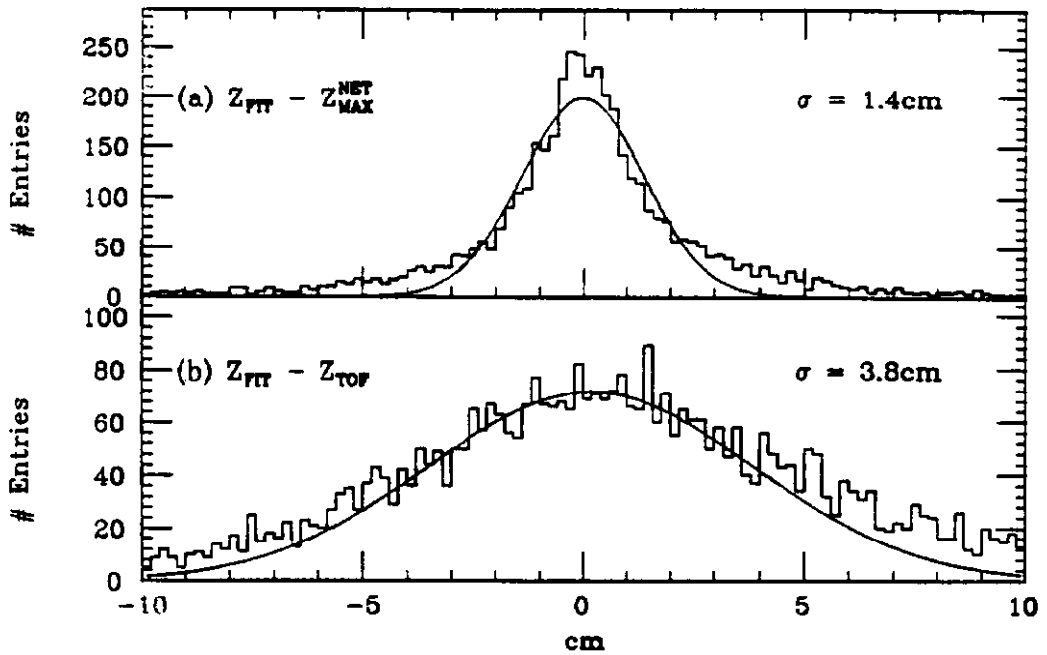


Figure 9. a) difference between Zvertex as measured by the neural net and by the standard offline program, in centimeters. b) difference between Zvertex as measured by TOF counters and standard offline program. The neural net resolution is much better.

The 288 sense wires of the chamber were broken up into 18 wire subsections (3 layers of 6 wires each) for processing by the network. The sets of 18 drift times became inputs to identical MLP's each with a single hidden layer of 128 units. Each output layer had 62 units, 60 representing 1.0 centimeter bins from -30 cm to 30 cm. and 2 'overflow' units. The 18 input subnetworks were trained to represent the vertex position by a Gaussian histogram in the output units, which gives good vertex position resolution with relatively few output units. Training was done using real data recorded in a previous run of the E735 experiment at Fermilab [E735 1991]. Targets were obtained using the Z position of the vertex calculated using the standard offline algorithm. The 18 wire subsections were chosen so as to overlap in order not to miss tracks which may span subsections. The outputs of the subnets are then simply added. This is illustrated in figure 8.

Figure 9 compares the distribution of $Z_{offline}$-$Z_{NN}$ to that of $Z_{offline}$-$Z_{TOF}$, where Z is the position along the direction of the beam particles. The neural network Z resolution is about 3 times better than TOF, and its performance can probably be even further improved by using additional wire layers in the chamber. TOF is currently analyzed offline. It might be possible to implement it online, but its resolution can probably not be improved because it is a technology which has already been pushed to its limits. Also, the TOF technique cannot handle cases of multiple vertices. The neural net treats these in a natural way: each vertex appears as a bump in the summed net output.

### *3.2.2 Kink Recognition*

A high energy pion or kaon will sometimes decay in a tracking chamber volume into a muon and a neutrino. The neutrino is neutral and is not seen in the tracking chamber. The muon is charged and is seen, however has a different momentum from the original particle. The result is a track which appears to have a 'kink' in it (figure 10).

In this work [Stimpfl 1991, Stimpfl 1992] , simulated pion tracks of 3, 5, and 10 GeV momentum were generated and transported through a chamber modelled upon that of the Aleph experiment at CERN. A detailed detector simulation was used to model noise hits and other instrumental effects. Two approaches were tried. In the first, helical track segments are fit to the hit positions in an inner region, 1, and an outer region, 2 (figure 10). The 5 helix parameters[9] in the two regions are then used as input to a MLP which tells whether or not this track is due to a decay. In the second approach, a single fit is done to the track across both regions, and the residuals of the fit are used as input to the neural network. There will be 42 residuals, one for each measurement along the trajectory. As a variant to this second approach, groups of three residuals were averaged to give 14 residuals as input to the network.

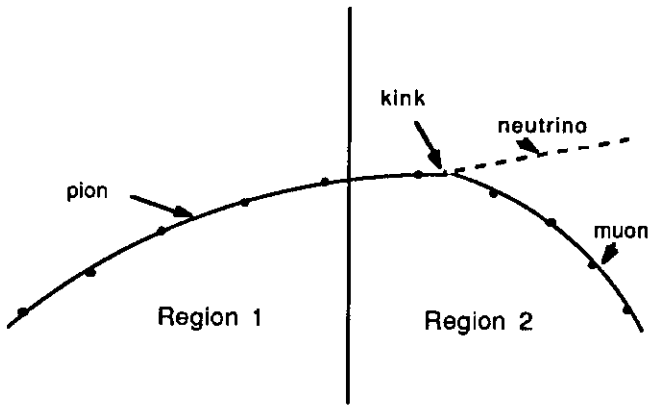


Figure 10. A pion decays to a muon and a neutrino to produce what appears as a track with a 'kink'. The kink is recognized by comparing found track parameters in region 1 and region 2

The results, are summarized in table VII, which also shows the network architectures tried. Also given in the table is the result obtained with the standard method for kink identification, called the analytical $\chi^2$ method, in which again the track is fit in two regions and a $\chi^2$ is calculated from the helix parameters in the two regions to determine the probability of the non-kink hypothesis. Both of the neural net methods are found to have higher efficiency than the standard chi-squared method. The neural network residual method is about 20 times faster to calculate than the analytical $\chi^2$ method, assuming that the residuals are already available from the standard track fit.

[9] The helix parameters are the z position of the vertex, the polar and azimuthal angles of the axis of the helix, the radius, and the pitch.

| Method | 3 Gev | 5 GeV | 10 GeV |
|---|---|---|---|
| 5 - 5 - 1 (par) | 78.9 | 67.0 | 53.5 |
| 5 - 10 - 1 (par) | 79.1 | 67.2 | 53.6 |
| 14 - 7 - 1 (res) | 79.9 | 65.5 | 51.5 |
| 14 - 14 - 1 (res) | 80.5 | 65.7 | 53.9 |
| 42 - 6 - 1 (res) | 80.3 | 67.7 | 54.5 |
| analytical $\chi^2$ | 76.0 | 62.0 | 40.2 |

Table VII. Efficiencies (in percent) for correctly identifying kinks (defined in text) in pion tracks of 3, 5, and 10 GeV momentum. Two MLP architectures were tried for the case of track parameters as net inputs, and three for the case of fit residuals as net inputs. The results for the standard method, analytical $\chi^2$, are also given.

### *3.2.3 Other Applications*

A variety of other applications of neural networks to low level pattern recognition in high energy physics have appeared, which we mention only briefly. The interested reader may consult the references. In an application to a Cherenkov[10] detector, MLP's were used to find a set of dots forming a ring pattern in a noisy image [Altherr 1992, deGroot 1992]. In another hardware application [Haggerty 1992], a discrete component hardware MLP was used to measure, in real time, the position of a muon track in a tracking chamber using charges induced on electrodes placed below the sense wire. MLP's have been used to perform electron/pion discrimination in a calorimeter [Garlatti Costa 1992, Teykal 1992] and identification of heavy quarks using the presence of multiple vertices in a vertex tracking chamber [Gupta 1991, Denby 1992]. Applications to charged track reconstruction will be discussed in section 5.

## 4. Physics Process Determination

### *4.1 B Tagging*

Numerous groups have used neural networks for identifying reactions containing b quarks. This is usually referred to as 'b tagging'. Typically this has been done at the four experiments at the LEP electron positron colliders [Proriol 1991, Proriol 1992, Bortolotto 1991, deGroot2 1991, Gottschalk 1991, Bellantoni 1991, Seidel 1992, Branchini 1992, Brandl 1992], although some work with simulated jets at proton antiproton colliders has also been reported [Denby 1990]. B tagging is of considerable interest since the properties of many particles containing b quarks have to date not been well studied. In the LEP work, the approach is typically to choose an ensemble of feature variables which describe the spatial distribution of energy within each jet and of the event as a whole. Additional information such as that from vertex tracking chambers may also be included. We choose as an example of this type of study the analysis performed by members of the Delphi experiment which extends the analysis to charm quarks and undifferentiated light

[10] A Cherenkov detector measures the mass of certain types of particles using the light the particle produces in passing through a transparent medium.

quarks in order to extract the decay probabilities into these quarks of the Z boson. This analysis is described in the next section.

### *4.2 Decay Probabilities of the Z*

The neutral boson Z can decay into any constituent plus its anti-constituent, e.g., electron plus positron, u quark plus $\bar{u}$ quark, etc. The standard model dictates the types of interactions which the constituents can undergo, but the relative strengths of the various interactions must be verified experimentally. A group from the DELPHI collaboration (one of the 4 major experiments at the LEP accelerator at CERN) has recently used a feed forward neural network to classify decays of the Z into three classes: $c\bar{c}$ pairs; $b\bar{b}$ pairs; or light quark (u,d, or s) -antiquark pairs [Cosmo 1992, De Angelis 1992, Eerola 1992, see also Bortolotto 1991]. This classification has permitted a measurement of the probabilities of the Z to decay into these particles to be made with higher precision than was previously possible.

The probability of the Z to decay into the leptons electron, muon, and tau has been well established. That measurement is 'easy' to make since these particles are relatively easy to identify in the apparatus. The case of the decay of the Z into quarks is considerabley more difficult since the final state quarks fragment immediately into jets. The problem then becomes deducing the type of quark involved in the decay from the properties of the jets themselves and from their distribution within the apparatus.

The standard technique for distinguishing heavy quarks from light quarks is through their so-called semileptonic decays, in which a particle containing a heavy quark decays to a lepton plus other particles. This technique has two disadvantages: 1) semileptonic decays account for only 20 percent of heavy quark decays; therefore with this technique it will be more difficult to obtain a sample large enough to assure small statistical errors; 2) in a semileptonic decay a neutrino is also emitted; these escape detection, making it impossible to completely reconstruct the event, leading to uncertainty in quark species in some cases. A technique which allows the use of all types of heavy quark decays is thus desirable.

In the DELPHI work, 19 jet and event-shape variables were created as inputs to an MLP. The variables describe the spatial distribution of energy in the jets and in the event as a whole, various kinematical combinations of the momenta of the particles in the jets, as well as information about the presence of leptons in the event. An exact description of the 19 variables is not very illuminating to the non-specialist; the interested reader is referred to the original works. The network architecture chosen had 25 hidden units and 3 output units to encode the three classes.

The training data for the network was generated with a standard physics Monte Carlo program and a program which simulates the response of the DELPHI apparatus to particle collisions. A total of 6000 training events were used. An independent set of 200,000 events was used for testing the network.

The trained network was then used to determine the relative fractions of b, c, and light quark decays in a sample of 123,475 real events from the DELPHI experiment. To do this, a 2-dimensional representation of the network output was devised as follows. The values of the 3 output nodes were normalized to sum to 1. Each event can then be represented as a point within an equilateral triangle where the perpendicular distances of

the point to the sides of the triangle represent the values of the output nodes. This type of representation is referred to as a Dalitz plot. Figure 11 shows the distribution within this plane of Monte Carlo events for b, c, and light quark decays, as well as for the real data. The fractions were obtained by fitting the real data distribution to a linear combination of the Monte Carlo distributions for the three classes:

$$R(u,v) = (1 - F_c - F_b)\, a_1(u,v) + F_c a_2(u,v) + F_b a_3(u,v)$$

where u,v are the variables defining the plane, R is the distribution of the real data, $F_c$ and $F_b$ are the fractions of decays containing c and b quarks, respectively, and $a_1$, $a_2$, and $a_3$ are the distributions of the Monte Carlo data. The results of the fit are:

$$F_c = .158 \pm .007_{stat} \pm .030_{param} \pm .008_{model}$$

$$F_b = .212 \pm .004_{stat} \pm .005_{param} \pm .011_{model}$$

where the first error is due to statistics, the second to an incomplete knowledge of certain parameters in the Monte Carlos, and the third to the dependence of the result upon which Monte Carlo model is used. Note that explicit reference is made to a model dependence of the result, a problem peculiar to high energy physics as discussed in section 2.3. For comparison, the best result to date for $F_b$ [Abreu 1992] using semileptonic decays is

$$F_b = .215 \pm .017_{stat+systematic}$$

where the systematic error contains effects due to parameter and model dependence. For the charm quarks, the best result to date [Abreu 1990] is obtained by identifying a characteristic low energy pion from the decay of a particle containing a charm quark. The result is

$$F_c = .162 \pm .030_{stat} \pm .050_{syst.}$$

The result using the neural network has a smaller uncertainty in both cases.

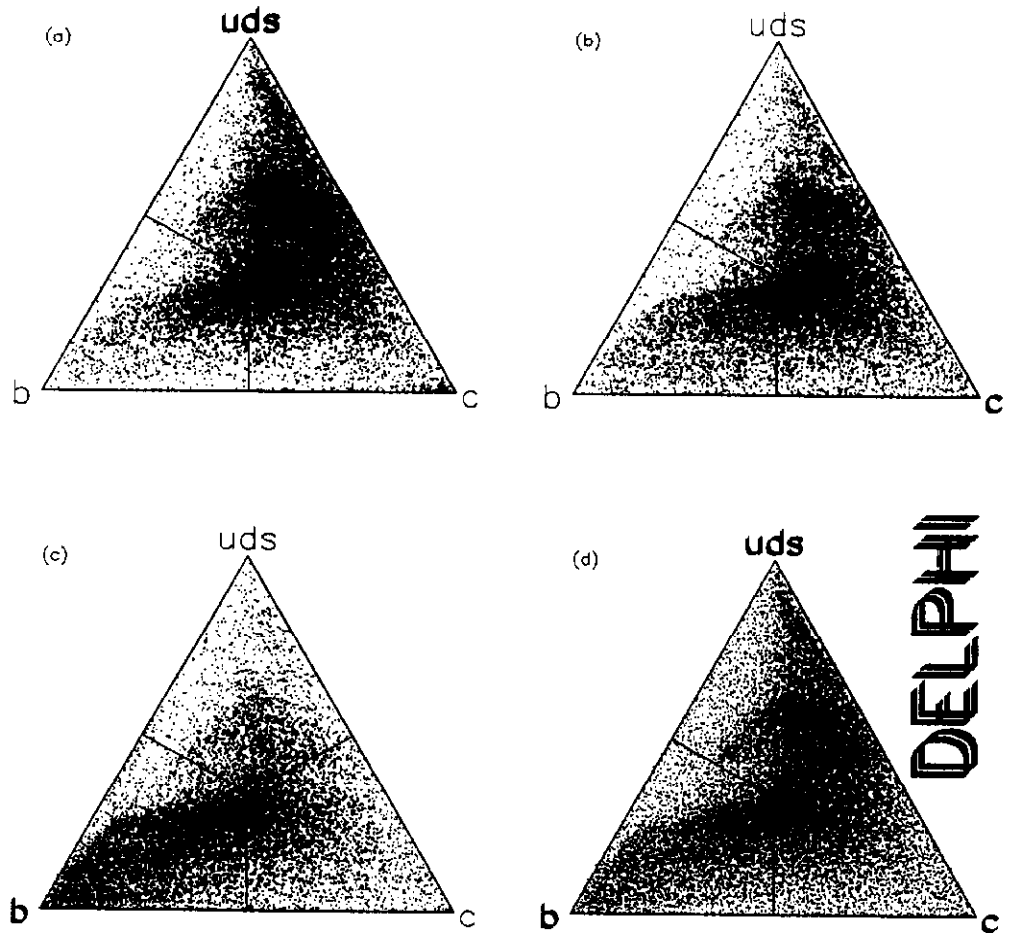


Figure 11 Dalitz plots used to measure the relative fractions of b, c, and light (uds) quarks in the decays of the $Z^0$. The activation of the network output node corresponding to each class is represented as the perpendicular distance from the side of the triangle opposite the corner labelled with that class. The outputs of the three nodes always sum to 1. a), b), and c) show the distribution of network outputs for Monte Carlo (uds), c, and b quarks respectively. d) shows the distribution for real data from Delphi. To extract the fractions of b,c, and (uds), the distribution in d) is fit as a linear combination of the distributions of a), b), and c), where the coefficients in the linear combination are the desired fractions.

### *4.3 Quark/Gluon Separation*

The ability to distinguish quark jets from gluon jets is clearly very desirable. The W and Z decay 80 percent of the time to two quarks, but normally these decays are unusable since it is not possible to distinguish these jets from the more copiously produced gluon jets. Furthermore, the most probable decay mode of the much sought top quark is into three quark jets, but this channel has long been considered unusable due to high backgrounds from multi-gluon final states. The ability to verify three quark jets would dramatically reduce the background. Distinguishing quark jets from gluon jets has been thought by many high energy physicists to be impossible due to the high degree of similarity between the two types of jets.

Separation of quark and gluon jets using neural networks has been treated in a number of references [Lonnblad 1990, Lonnblad 1991, Bhat 1990, Csabai 1991, Baer 1991,

Barbagli 1992]. These results have been almost exclusively based upon data generated by Monte Carlo. Recently a new result from the Fermilab Tevatron collider has appeared [Bianchin 1992, Bianchin2 1992] which for the first time appears to give evidence of quark and gluon components in real jets produced in proton antiproton collisions. In the Fermilab result, jets identified in the apparatus are represented by a set of 8 feature variables which describe the spatial distribution of energy within the jets, e.g., the amount of energy contained within each of three concentric cones centered on the centroid of the jet, the r.m.s. width of the jet, etc. A backpropagation neural network with these 8 variables as inputs was trained to separate quark jets from gluon jets based upon examples generated by Monte Carlo. It is necessary to use Monte Carlo since pure samples of quarks and gluons do not exist. The real data will always contain a mixture of quark and gluon jets, and in fact the relative ratio of quarks and gluons in various kinematical regions is one of the sought after results. For this reason, this problem too will suffer from the fact that the results will depend upon which model of quark and gluon fragmentation has been used.

There is considerable overlap of the two classes in all of the feature variables, and none is adequate to provide a useful classification of the jets. Figures 12 a) and b) shows the output of the trained neural network on independent test samples of Monte Carlo quarks and gluons. The quark and gluon distributions overlap substantially: quark and gluon jets are indeed very similar! However the separation achieved is useful because quark or gluon enriched samples can now be produced by placing cuts on the output of the neural network.

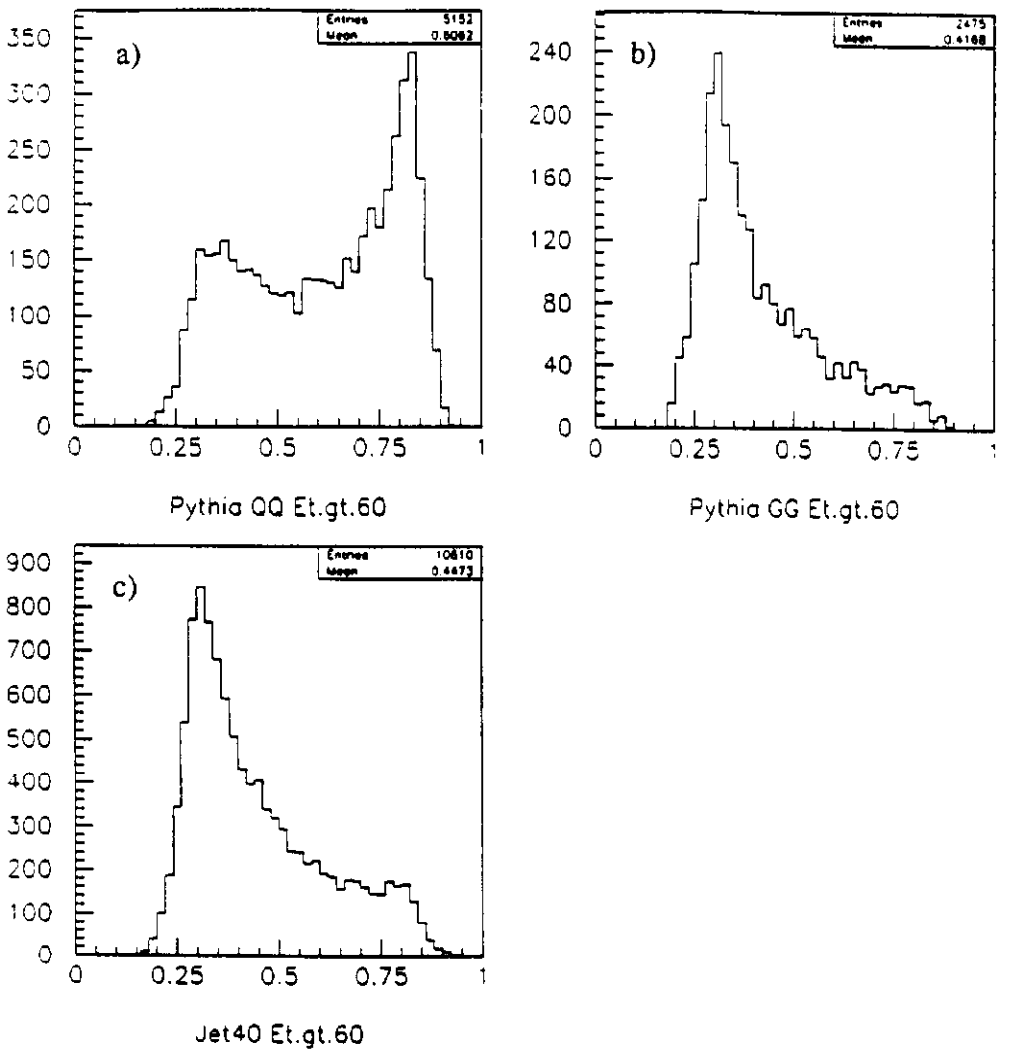


Figure 12. Output of 8-6-1 MLP for a) Monte Carlo (the Pythia Monte Carlo was used in these studies) quarks; b) Monte Carlo gluons; c) Real data from the CDF experiment (labelled 'Jet 40'). All the jets are required to have Et greater than 60 GeV. The real data appears to be predominantly gluon like with a small admixture of quarks, as expected from theory.

A study was made of the efficiency of the network, defined as the fraction of quark jets with network output above 0.5, as a function of the number of nodes in the hidden layer. Performance did not improve beyond the results with two hidden units, and in fact a simple perceptron (no hidden units) was only a few percentage points worse using this measure. However, the network output distributions for the zero and two hidden unit cases was much more gaussian in shape and did not include any regions in which the quark to gluon ratio was high. Such regions may prove valuable for placing cuts which enrich quark to gluon ratio at the price of reduced quark efficiency. For this reason, the results from the 6 hidden unit network were retained for the final analysis. The maximum efficiency achieved on the Monte Carlo data was 70 percent.

Figure 12c) shows the result of applying the trained net to a sample of real data from the CDF experiment. The real data distribution appears to be predominantly gluon-like with a non-zero admixture of quarks, which is consistent with the result expected on theoretical grounds for events in the kinematical regime in which the data was taken. A fit to the real data as a linear combination of the Monte Carlo quark and gluon distributions gives a good $\chi^2$, but because of model dependence and some subtleties in the Monte Carlo programs, it has not yet been possible to extract the exact quark fraction from this distribution in an unambiguous way. However, the results are encouraging and work is continuing.

More recently, another analysis was performed [Bianchin2 1992] in which a feature map was trained on a sample of mixed Monte Carlo quarks and gluons and then used to identify quarks and gluons in an independent sample. A somewhat higher efficiency, about 72 percent was obtained. The feature map trained on Monte Carlo is also being applied to the real data, and, conversely, a feature map trained on real data is being applied to labelled Monte Carlo data. Training using only real data is very attractive since it avoids the problem of model dependence, although it may be necessary to use the Monte Carlo data to label the nodes in the topological map. These analyses are still in progress.

### *4.4 Additional Physics Process Applications*

The use of learning vector quantization and topological maps is relatively new in HEP. An interesting application of topological maps appears in [Lonnblad3 1991] in which a map is used to discover the b,c, and light quark classes in a sample of mixed Monte Carlo data. A similar application is being attempted for data at the Tevatron [Bianchin2 1992]. LVQ has been used for b tagging [Proriol 1991, Proriol 1992] and discrimination of $t\bar{t}$ events from background [Odorico 1991]. Other MLP offline applications include: resonance searches[11] [Alexopoulos 1991]; calculation of the total mass of the particles in an event [Lonnblad2 1991]; determination of the charge of the initial quark which produced a jet [Varela 1991]; and identification of jet cascades with muons [Los 1992].

## 5. Neural Nets and Charged Track Reconstruction

### *5.1 Tracking with Recurrent Nets*

Recurrent networks have been used in HEP for track reconstruction, using an algorithm developed by Denby and independently by Peterson [Denby 1988, Peterson 1989, Stimpfl 1990, Denby3 1990, Barbagli 1992]. In this application a neuron is defined to be a directed link between two hits in a tracking detector. The approach resembles qualitatively the encoding used by Hopfield [Hopfield 1986] for solving the Traveling Salesman Problem with a recurrent net. The weight connecting two neurons i and j is determined by the angle $\theta_{ij}$ between them, (figure 13):

$$w_{ij} = A \cos^n\theta_{ij}/l_i l_j$$

where $l_i$ and $l_j$ are the lengths of the neurons (i.e., distance between hits), if i and j do not both point into or out of the same point, and $w_{ij} = -B$ if i and j are head to head or tail to

[11] A resonance is a bound state of two or more particles and appears as a peak in a mass distribution.

tail. An energy function is defined, $E = -1/2 \Sigma w_{ij} o_i o_j$, where $o_i$ is the output of neuron i. The energy function will be smallest when the angles between close together neurons sharing points are small. This favors neurons lying along smooth trajectories such as those of particles moving in a magnetic field. The constraint term -B ensures a unique direction to the tracks to avoid a degeneracy which prevents settling of the network. The evolution of the system is obtained by iteratively solving the update equations:

$$\tau \, du_i/dt = \Sigma_j \; w_{ij} o_i - u_i \; ; \; o_i = \text{sigmoid}(u_i).$$

On each iteration, dt is kept much smaller than $\tau$, the time constant of the system.

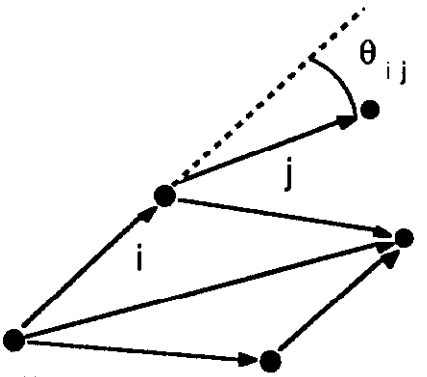


Figure 13. Neuron links in the Denby-Peterson Net

This method has been used on real data at the ALEPH experiment at LEP [Stimpfl 1990]. Figure 14 shows r-phi (i.e., looking down the beam line) and r-z (side) views of an event in which a Z boson decays to hadrons, with all links defined before network evolution (left side of figure), and the event after settling of the network, with tracks found (right side). The efficiency is as good as the conventional track reconstruction program but the neural net algorithm is somewhat faster. In this work, a study was made of execution time for the neural net and conventional algorithms as a function of track multiplicity (number of charged tracks in the event). The advantage of the neural algorithm is shown to increase with multiplicity. Although this type of algorithm has not yet been accepted as a standard track recognition algorithm, it may prove to be important in the future when track multiplicities will be larger.

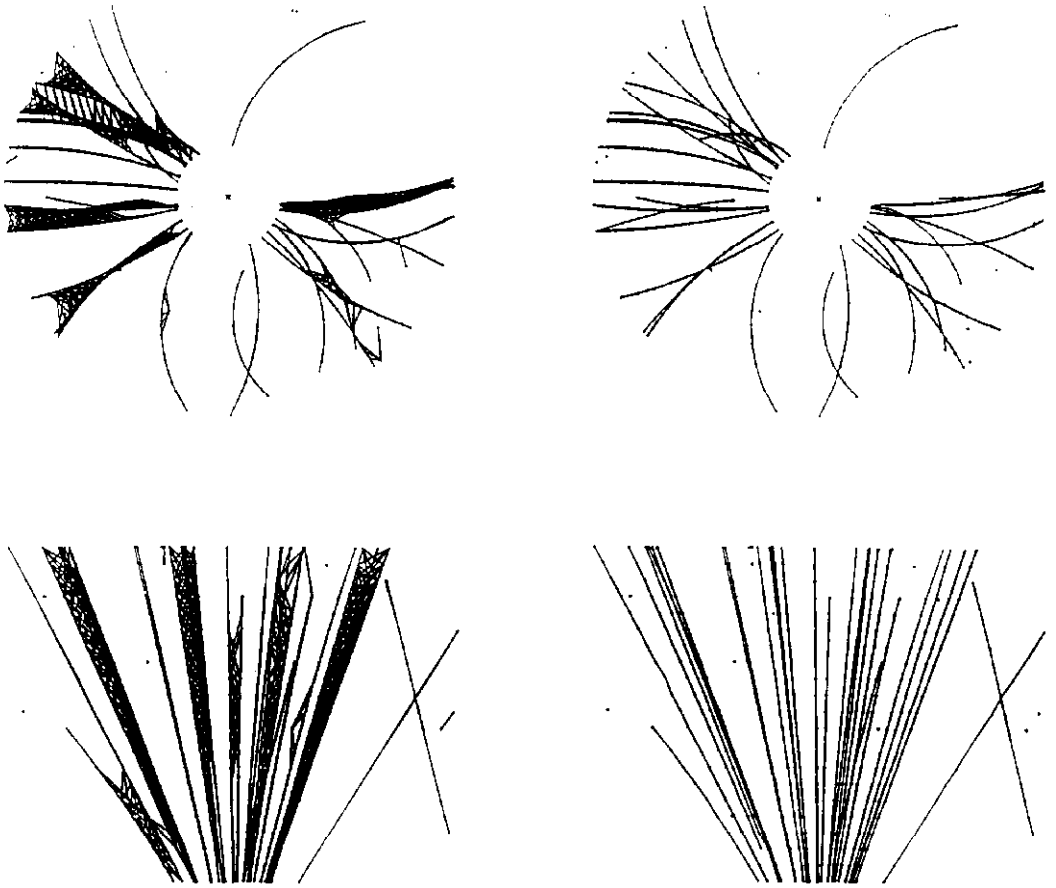

Figure 14. Charged track reconstruction on real data in the Aleph central tracking chamber, using a recurrent neural network algorithm. In the top figures the beam pipe is perpendicular to the plane of the page, in the bottom figures, horizontal. The left hand frames show the neuron links before evolution; at right are the found tracks at the end of evolution.

There is not a straightforward way to implement this algorithm in the fast hardware that would be needed to make it applicable at the trigger level, since the number of neurons and weights is high, and the weights must be recalculated for each event. In addition, the algorithm does not take advantage of all the available information, such as that tracks in a uniform magnetic field are known to be nearly perfect helices. This makes the algorithm more susceptible to noise since it will be less able to reject noise hits which happen to lie near the tracks.

### *5.2 Elastic Tracking*

Improvement to the neural tracking are the so-called elastic tracking [Gyulassy 1991] or deformable templates [Ohlsson 1991] approaches. In these approaches, a track is a helical object which settles into a shape which best fits the hits. The helix can be thought of as electrically charged and attracted to the hits which have opposite charge. Although these algorithms map the tracking problem onto dynamical systems, and are at least in principle parallelizable, they have lost some of the 'neural' flavor of the original Denby-Peterson net. Nonetheless, the efficiency and robustness to noise of the elastic methods

are excellent. One interesting study [Gyulassy 1991] compared the robustness to noise of the standard method, the Denby-Peterson net, and the elastic tracking method. The standard method of track reconstruction is called the 'roadfinder' since it starts with two nearby hits and then searches for additional hits on a 'road' in the direction of the segment joining them. Figure 15 from this study shows the efficacy of each method as a function of number of tracks. All data have 20 percent noise and 3 percent error on position measurement. The roadfinder breaks down between 5-10 tracks, the Denby-Peterson net at 10-15 tracks, but the elastic tracking always finds the correct answer in this study.

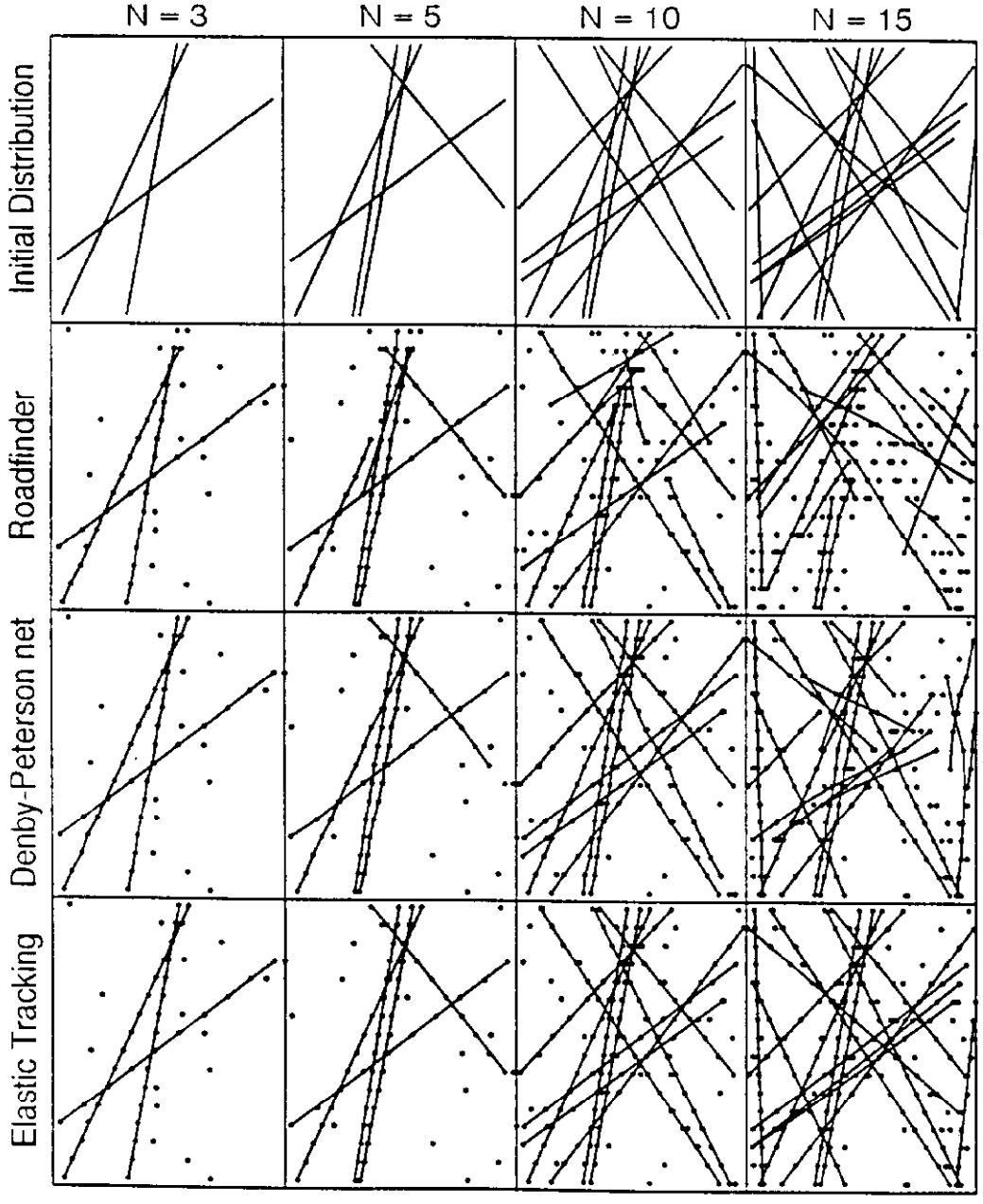


Figure 15. Comparison of track reconstruction performance for the standard method (roadfinder), Denby-Peterson net, and elastic tracking.

## 6. Conclusion

Five years ago there was no explicit mention of neural network techniques in HEP literature. A current bibliography of applications in HEP includes almost a hundred papers and reports. Much of the work is still exploratory and uses only the simplest techniques such as the MLP trained with backpropagation, although some interesting results using learning vector quantization and feature maps have also appeared.

In HEP, historically, data analysis has been done using simple one dimensional cuts; consequently the HEP community at large has yet to fully accept neural network techniques as standard tools. Nevertheless the neural network methods are beginning to show their worth. The decay probabilities of the Z boson into b,c and light quarks has been measured with higher precision than ever before using a technique based upon a MLP. A neural network technique has given higher kink finding efficiency and faster execution speed than the standard method. Results consistent with identification of quark and gluon components in jets produced at a proton antiproton collider have appeared for the first time using a feed forward neural network. Recurrent networks have provided a faster way of performing charged track reconstruction.

One of the most exciting promises of neural network technology is in the realm of triggering for HEP. One test has already been completed: a VLSI neural network used in the data acquistion system of a drift chamber has provided, in only a few microseconds, track intercept resolution 50 times more accurate than that previously obtainable online. Neural network triggers for three large collider experiments are either currently being installed or have been proposed for future installation. The neural network triggers will permit experiments to reject more background events earlier in the data stream, resulting in more efficient and cost effective data acquisition systems and enormously reducing data storage requirements.

It is intellectually quite stimulating to witness a marriage between such seemingly disparate domains as high energy physics and neural networks. Given the growth of applications and their success to date, HEP may turn out to be one of the driving forces in the integration of neural networks into science as data analysis tools.

## 7. Appendix: Particle Interactions

The constituents are arranged into doublets consisting of a lepton and its neutrino or of a quark and its conjugate quark (quark-conjugate pairs have a unit net charge) as in table VIII.

| | particles | antiparticles | interactions |
|---|---|---|---|
| leptons | $\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}$ | $\begin{pmatrix}\bar{\nu}_e\\ e^+\end{pmatrix}\begin{pmatrix}\bar{\nu}_\mu\\ \mu^+\end{pmatrix}\begin{pmatrix}\bar{\nu}_\tau\\ \tau^+\end{pmatrix}$ | W, Z<br>$\gamma$ (except $\nu$) |
| quarks | $\begin{pmatrix}u\\ \bar{d}\end{pmatrix}\begin{pmatrix}c\\ \bar{s}\end{pmatrix}\begin{pmatrix}t\\ \bar{b}\end{pmatrix}$ | $\begin{pmatrix}\bar{u}\\ d\end{pmatrix}\begin{pmatrix}\bar{c}\\ s\end{pmatrix}\begin{pmatrix}\bar{t}\\ b\end{pmatrix}$ | W, Z, $\gamma$, g |

Table VIII. Lepton and quark doublets and the interactions in which they participate.

Interactions of the members of the doublets are represented by three legged 'vertices'. Different symbols are used for the legs depending on the type of particle represented by the leg. Only certain vertices are allowed. In allowed vertices the legs must be:

1) a quark or lepton and its antiparticle plus a neutral boson,

2) the two members of a quark or lepton doublet plus a charged boson, or

3) three gluons

as shown in figure 16. Additional legal vertices may be generated by changing a leg from incoming to outgoing and replacing the particle the leg represents by its antiparticle.

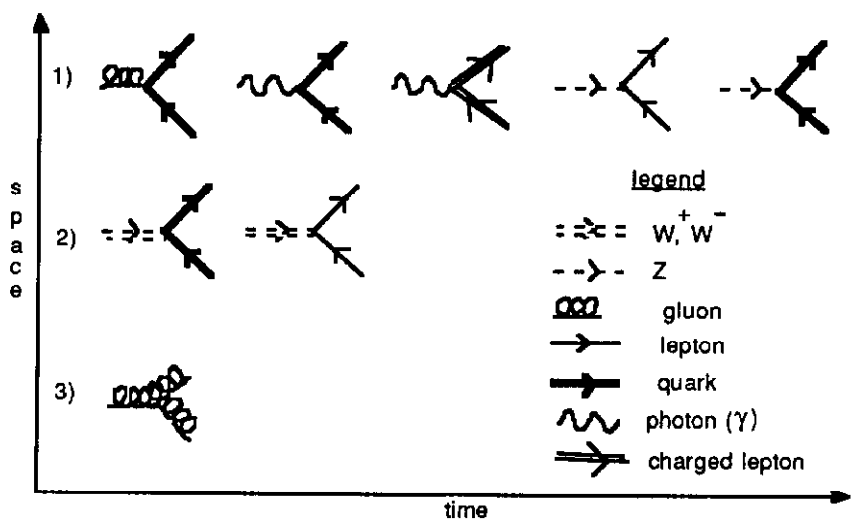


Figure 16. Allowed interaction vertices

The elementary vertices can be 'hooked together' to represent various physical processes such as collisions of constituents and decays. Some examples are given in figure 17. These so-called Feynman diagrams are drawn with strict rules, which we shall not discuss here, which allow the direct transcription of formulae which can be used to calculate the probability of the process represented in the diagram. The left hand side of such a diagram is called the 'inital state' and the right hand side the 'final state' since time progress from left to right.

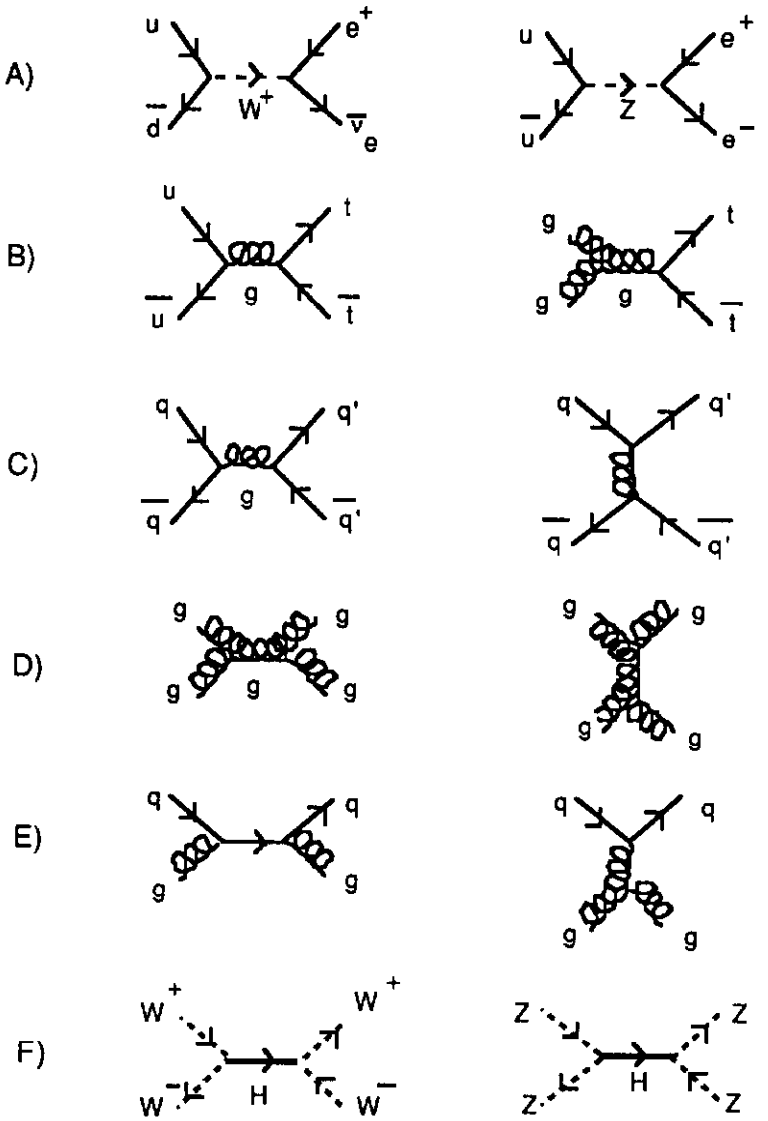


Figure 17. Processes shown are: A) quark production of weak bosons decaying to leptons; B) quark and gluon production of $t\bar{t}$ via a gluon; C) quark quark scattering; D) gluon gluon scattering; E) quark gluon scattering; F) weak boson production of Higgs decaying to weak bosons.

The diagrams shown are processes which occur at the level of the constituents. The final state may be considerably modified before any of the particles reach the detectors. This is because quarks and gluons are not visible as free particles. Rather, many additional

vertices will attach themselves to the final state particles with high probability. These outgoing quarks and gluons will bind into composites, resulting in a multiparticle jet of outgoing particles. This is illustrated in figure 18.

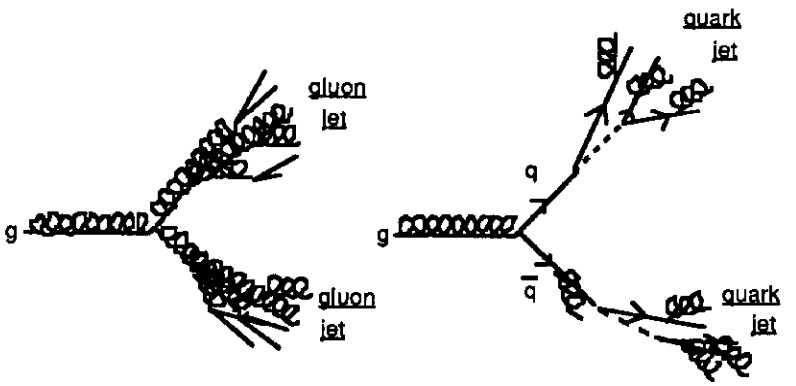


Figure 18. Example of how the final state quarks or gluons can evolve into multiparticle jets by the spontaneous attachment of additional vertices.

## 8. References

Abreu 1990: P. Abreu et al., *Phys. Lett.*, **252B** (1990) 140.

Abreu 1992: P. Abreu et al., "Measurement of the Partial Width of the $Z^0$ into $b\bar{b}$ Final States using their Semileptonic Decays", CERN - PPE/92-89, submitted to *Zeit. Phys.* C.

AIHEP 1992: *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Alexopoulos 1991. T. Alexopoulos, "Resonance Searches using a Neural Network Technique", talk at DPF 91, Vancouver, Canada, August 1991, submitted to proceedings, also T. Alexopoulos, Ph.D. Thesis, University of Wisconsin, unpublished.

Altherr 1992: T. Altherr et al., "Cerenkov Ring Recognition using Adaptable and non-Adaptable Networks", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Baer 1991: H. Baer, D. Karatas, G. Giudice, "Snagging the Top Quark with a Neural Network", FSU HEP 911130, Florida State University, Tallahassee, Nov. 1991.

Barbagli 1991: G. Barbagli, G. D'Agostini, D. Monaldi, "Quark/Gluon Separation in the Photoproduction Region with a Neural Network Algorithm", Universita di Roma 'La Sapienza', Internal note N.992, Feb. 1992.

Bellantoni 1991: L. Bellantoni, J.S. Conway, J.E. Jacobsen, Y.B. Pan, Sau Lan Wu, "Using Neural Networks with Jet Shapes to Identify b Jets in e+e- Interactions", CERN-PPE/91-80, 24 May 1991, submitted to *Nucl. Inst. & Meth.*

Bhat 1990: P. Bhat, L. Lonnblad, K. Meier, K. Sugano, "Using Neural Networks to Identify Jets in Hadron Hadron Collisions", *Proc. of the 1990 Summer Study on High Energy Physics - Research Directions for the Decade,* Snowmass, Colorado, June 25 - July 13, 1990.

Bianchin 1992: S. Bianchin, M. Denardi, B. Denby, M. Dickson, G. Pauletta, L. Santi, and N. Wainer, "Classification of Jets from PPbar Collisions at Tevatron Energies", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific,, and CDF Internal Note 1706.

Bianchin2 1992: S. Bianchin, M. Dall'Agata, M. De Nardi, G. Pauletta, L. Santi, B. Denby, N. Wainer, M. Dickson, "Jet Classification at CDF", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Bortolotto 1991: C. Bortolotto et al., "A Measurement of the Partial Hadronic Widths of the $Z^0$ Using Neural Networks", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Branchini 1992: P. Branchini, M. Ciuchini, and P. Del Giudice, "B Tagging with Neural Networks: An Alternative use of Single Particle Information for Discriminating Jet Events", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific, and INFN-ISS 92/1.

Brandl 1992: B. Brandl et al., "Tagging of Z Decays into Heavy Quarks in the Aleph Detector using Multivariate Analysis Methods: Neural Networks, Discriminant Analysis, Clustering", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

CDF 1988: The Collider Detector at Fermilab, a compilation of articles reprinted from *Nucl. Inst. Meth.*- A, North Holland, Amsterdam, 1988.

Cosmo 1992: G. Cosmo, A. De Angelis, N. De Groot, P. Del Giudice, P. Eerola, J. Kalkkinen, L. Lyons, M. Los, E. Torassa, E. Vallazza, Delphi Collaboration, "Classification of the Hadronic Decays of the $Z^o$ into b and c Quark Pairs using a Neural Network", submitted to the XXVI International Conference on High Energy Physics, Dallas, TX, U.S.A., August 5-12, 1992.

Csabai 1991: I. Csabai, F. Czako, Z. Fodor, "Combined Neural Network-QCD Classifier for Quark and Gluon Jet Separation", CERN Preprint CERN-TH.6038/91 and Eotvos University (Budapest) Institute for Theoretical Physics preprint ITP-Rep. Budapest 483, March, 1991.

De Angelis 1992: A. De Angelis, "Heavy Flavour Identification in Delphi", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

deGroot 1991: N. De Groot and M. Los, "B-Tagging in Delphi with a Feed-Forward Neural Network", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Denby 1988: B. Denby, *Computer Physics Communications,* **49** (1988) 429. Also, B. Denby "Neural Network and Cellular Automota Algorithms", Florida State University preprint FSU-SCRI-88-141, June, 1988. Tallahassee, Florida.

Denby 1990: B. Denby et al., *IEEE Trans. Nucl. Sci.* **37** No. 2 (1990) 248.

Denby2 1990: B. Denby, E. Lessner, and C.S. Lindsey, *Proc. 1990 Conf. on Computing in High Energy Physics,* Santa Fe, NM, (1990) AIP Conf. Proc. **209** 211.

Denby3 1990: B. Denby and S. Linn, *Computer Physics Communications* **56** (1990) 293.

Denby 1991: Denby et al., CDF Internal Note 1538, "Proposal for a Level-2 Isolated Plug Electron Trigger for the 1991/1992 Run", CDF Collaboration, Fermi National Accelerator Laboratory, Batavia, Illinois.

Denby 1992: B. Denby, "Quark Flavor Sensitivity of the Mammalian Cortex: Theoretical Foundations", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

D0 1983: D0 Design Report, Fermilab, Dec. 1983, and D0 Upgrade, Fermilab P-823, April 1991.

E735 1991: F. Turkot et al., *Nucl. Phys.* **A525** (1991) 165-170.

Eerola 1992: P. Eerola, "Classification of the Hadronic Decays of the $Z^0$ into b and c Quark Pairs using a Neural Network", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Fortner 1992: M. Fortner, "Analog Neural Networks in an Upgraded Muon Trigger for the D0 Detector", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Garlatti Costa 1992: P. Garlatti Costa, A. De Angelis, L. Lanceri, L. Santi, C. Vignaduzzo, E. Zoppolato, "A Neural Network for e/p Classification in a Calorimeter", INFN Sezione di Trieste, Italy, technical note INFN/AE-92/14, 27 April 1992.

Gottschalk 1991: T. D. Gottschalk and R. Nolty, "Identification of Physics Processes Using Neural Network Classifiers", Caltech Report CALT-68-1680, 1991.

Gupta 1991: L. Gupta, A. Upadhye, B. Denby, and S.R. Amendolia, *Pattern Recognition.* **25** (1992) 413.

Gyulassy 1991: M. Gyulassy and M. Harlander, *Computer Physics Commmunications* **66** (1991) 31.

Haggerty 1992: Herman Haggerty, Fermilab, private communication.

Hansen 1992: J.R. Hansen, "The Need for Neural Networks at LHC and SSC", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Hopfield 1986: J. Hopfield and D.W. Tank, *Science* 233 (1986) 625.

Intel 1991: Intel 80170NX Electrically Trainable Analog Neural Network, Intel Corporation, Santa Clara, California.

Lindsey 1991: C.S. Lindsey and B. Denby, *Nucl. Inst. & Meth.* **A302** (1991) 217.

Lindsey2 1991: C. S. Lindsey, "Tracking and Vertex Finding in Drift Chambers with Neural Networks", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Lindsey 1992: C.S. Lindsey, B. Denby, H. Haggerty, K. Johns, *Nucl. Inst. & Meth.* **A317** (1992) 346-356.

Lindsey2 1992: C.S. Lindsey, "Drift Chamber Tracking with a VLSI Neural Network", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Lonnblad 1990: L. Lonnblad, C. Peterson, and T. Rognvaldsson, *Phys. Rev. Letters* **65** (1990) 1321.

Lonnblad 1991: L. Lonnblad, C. Peterson, and T. Rognvaldsson, *Nucl. Physics* **B349** (1991) 675.

Lonnblad2 1991: L. Lonnblad, C. Peterson, and T. Rognvaldsson, "Mass Reconstruction with a Neural Network", Lund University preprint LU TP 91-25, October 1991, submitted to *Physics Letters B.*

Lonnblad3 1991: L. Lonnblad, C. Peterson, H. Pi, and T. Rognvaldsson, "Self Organizing Networks for Extracting Jet Features", Lund University preprint LU TP 91-4, March 1991, submitted to *Computer Physics Communications.*

Los 1992: M. Los, "Using a Neural Network for Classifying Jet Cascades with a Muon", *Proceedings of the Second Workshop, Neural Networks: from Biology to High*

*Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Odorico 1991: A. Cherubini and R. Odorico, "Identification by Neural Networks and Statistical Discrimination of New Physics Events at High Energy Colliders", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Ohlsson 1991: M. Ohlsson, C. Peterson, and A. Yuille, "Track Finding with Deformable Templates - The Elastic Arms Approach", Lund University Preprint LU TP 91-27, November 1991, Lund, Sweden, submitted to *Computer Physics Communications.*

Peterson 1989: C. Peterson, *Nucl. Inst. & Meth.*, **A279** (1989) 537.

Proriol 1991: J. Proriol et al., "Tagging B Quark Events in Aleph with Neural Networks", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Proriol 1992: J. Proriol, "Tagging B Quark Events in e+e- Colliders with Neural Networks. Comparisons of Different Sets of Variables and Different Methods", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Ribarics 1991: P. Ribarics et al., "Neural Network Trigger in the H1 Experiment", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Ribarics 1992: "Neural Network Level 2 Trigger in the H1 Experiment",to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Ribarics2 1992: "Neural Network Trigger in the H1 Experiment",*Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Seidel 1992: F. Seidel et al., "Extensive Studies on a Neural Networks for b Tagging and comparisons with a Classical Method", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Siemens 1991: U. Ramacher et al., "Design of a First Generation Neurocomputer", in *VLSI Design of Neural Networks,* eds. U. Ramacher and U. Rückert, Kluwer Academic Publishers, 1991.

Stimpfl 1990: G. Stimpfl-Abele and L. Garrido "Fast Track Finding with Neural Nets", UAB-LFAE 90-66, submitted to *Computer Physics Communications*, 1990.

Stimpfl 1991: G. Stimpfl-Abele and Lluis Garrido, "Recognition of Decays of Charged Tracks with Neural Network Techniques", Universite Blaise Pascal preprint, Clermont-Ferrand, France, submitted to *Computer Physics Communications*, May 1991.

Stimpfl 1992: "Neural Nets for Kink Finding", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Teykal 1992: H. Teykal, "Using Neural Networks for the Identification of Electrons and Pions in a Calorimeter for High Energy Physics", *Proceedings of the Second Workshop, Neural Networks: from Biology to High Energy Physics,* Elba International Physics Center, Isola d'Elba, Italy, 18-26 June, 1992, to appear in *International Journal of Neural Systems,* World Scientific.

Varela 1991: P. Silva and J. Varela, "Identification of the Quark Jet Charge Using Neural Networks", in *Proc. of the workshop Neural Networks: From Biology to High Energy Physics*, Elba International Physics Center, Isola d'Elba, Italy, June 5-14, 1991, ETS Editrice, Pisa.

Vermeulen 1992: J. Vermeulen, "A Study of the Feasibility of Using Neural Networks for Second Level Triggering at LHC", to appear in the *Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence, and Expert Systems for Nuclear and High Energy Physics,* La Londe les Maures, France, January 1992, World Scientific.

Wu 1990: D. Wu et al., CDF Internal Note 1310, "A Pattern Recognition Level-2 B Trigger at CDF in 1991", CDF Collaboration, Fermi National Accelerator Laboratory, Batavia, Illinois; and private communication.