ISBN 91-628-3636-6
UNFD6/(NFFL-7168) 1999

Different Methods for Ouark/Gluon Jet Classification on Real Data from the DELPHI Detector

Göran Tranströmer

DEPARTMENT OF PHYSICS, LUND UNIVERSITY
LUND 1999

Till min gode vän Jens
från närmaste kollegan
Göran

ISBN 91-628-3636-6
LUNFD6/(NFFL-7168) 1999

Different Methods for Ouark/Gluon Jet Classification on Real Data from the DELPHI Detector

Thesis Submitted for the Degree of
Doctor of Philosophy in Physics
by

Göran Tranströmer

Department of Particle Physics
Lund University
Professorsgatan 1
Box 118
S-221 00 Lund
Sweden

Elégie

J'ouvre la première porte.
C'est une grande chambre inondée de soleil.
Une lourde voiture passe dans la rue
et fait trembler la porcelaine.

J'ouvre la porte numéro deux.
Amis! Vous avez bu de l'ombre
pour vous rendre visibles.

Porte numéro trois. Une chambre d'hôtel étroite.
Avec vue sur une ruelle.
Une laterne qui étincelle sur l'asphalte.
Belles scories de l'expérience.

Tomas Tranströmer

Abstract

Different methods to separate quark jets from gluon jets have been investigated and tested on data from the DELPHI experiment. A test sample of gluon jets was selected from $b\bar{b}g$ threejet events where the two b-jets had been identified using a lifetime tag and a quark jet sample was obtained from $q\bar{q}\gamma$ events where the photon was required to have a high energy and to be well separated from the two jets.

Three types of tests were made.

Firstly, the jetenergy, which is the variable most frequently used for quark/gluon jet separation, was compared with methods based on the differences in the fragmentation of quark and gluon jets. It was found that the fragmentation based classification provides significantly better identification than the jet energy only in events where the jets all have approximately the same energy. In Monte Carlo generated symmetric $e^+e^- \rightarrow q\bar{q}g$ threejet events, where the jet energy does not provide any identification at all, the gluon jet was correctly assigned in 58% of the events. More importantly, however, is that the identification has been divided into two independent parts, the energy part and the fragmentation part.

Secondly, two different sets of fragmentation sensitive variables were tested. It was found that a slightly better identification could be achieved using information from all the particles of the jet rather than using only the leading ones.

Thirdly, three types of statistical discrimination methods were compared:

- a cut on a single fragmentation variable;
- a cut on the Fisher statistical discriminant calculated from one set of variables;
- a cut on the output from an *Artificial Neural Networks* (ANN) trained on different sets of variables.

The three types of classifiers gave about the same performance and one conclusion from this study was that the use of ANNs or Fisher statistical discrimination do not seem to improve the results significantly in quark/gluon jet separation on a jet to jet basis.

Contents

1 Introduction 2

2 Theory 5
- 2.1 One hundred years of Particles Physics . . . 5
  - 2.1.1 Only three particles . . . 5
  - 2.1.2 Hundreds of particles . . . 6
  - 2.1.3 A subatomic periodic table . . . 6
- 2.2 The Standard Model . . . 8
  - 2.2.1 Quantum Chromodynamics . . . 9
- 2.3 Grouping of particles . . . 12
  - 2.3.1 Matter constituents . . . 12
  - 2.3.2 Force carriers . . . 13
- 2.4 The fragmentation process . . . 14
- 2.5 The Z^0 decay into b-quarks . . . 17
- 2.6 B-particles . . . 17
- 2.7 Simulated events . . . 18
  - 2.7.1 Monte Carlo generation . . . 18
  - 2.7.2 Pertubative QCD . . . 18
  - 2.7.3 Non Pertubative QCD . . . 19
- 2.8 Tuning of fragmentation variables . . . 20

3 LEP and the DELPHI experiment 22
- 3.1 Why so large dimensions? . . . 22
  - 3.1.1 Radiation losses at LEP . . . 22
- 3.2 The accelerator complex . . . 24
  - 3.2.1 LEP, the Large Electron-Positron collider . . . 24
  - 3.2.2 LEP and LEP2 . . . 30
- 3.3 The DELPHI Detector . . . 30
  - 3.3.1 Tracking detectors . . . 31
  - 3.3.2 Calorimeters . . . 33
- 3.4 DAS, the data acquisition system . . . 34
  - 3.4.1 DSTANA . . . 34
  - 3.4.2 DELANA . . . 34
  - 3.4.3 DELSIM . . . 34
  - 3.4.4 SKELANA . . . 34

4 Multi-variable discriminant analysis 42
4.1 Linear discrimination analysis . . . 42
4.2 Non-linear discrimination analysis . . . 45
4.2.1 Artificial Neural Network models . . . 46
4.2.2 Feed-Forward Neural Networks . . . 47
4.2.3 Implementation or the learning phase . . . 50

5 Overview of the Main Analysis 52

6 Event selection and jet reconstruction 55
6.1 Data selection . . . 55
6.2 Particle selection . . . 55
6.3 Event selection . . . 56
6.4 Selection of threejet events . . . 56
6.5 Jet reconstruction . . . 56
6.6 Jet algorithms . . . 57

7 B-tagging 60
7.1 Lepton Tag . . . 60
7.2 Lifetime Tag . . . 61
7.3 The tagging method used in the analysis . . . 65
7.3.1 The event tagging . . . 65
7.3.2 The individual tagging of the jets in the event . . . 65

8 Quark/Gluon jet samples 68
8.1 Association of quarks/gluons to jets in $q\bar{q}g$ MC events . . . 68
8.2 Differencies between the true quark/gluon jet samples . . . 72
8.3 Selection of quark jets in $q\bar{q}\gamma$ events . . . 73
8.4 Selection of gluon jets in $b\bar{b}g$ events . . . 75

9 Jet classification using the jet energy 86
9.1 Definition of efficiency and purity . . . 86
9.2 Efficiency and purity on a jet to jet basis . . . 88
9.3 Efficiency/purity on an event basis . . . 88

10 Fragmentation variables 92
10.1 Fragmentation sensitive variables for gluon and quark jets . . . 92
10.1.1 Variables based on leading particles . . . 92
10.1.2 The Fodor moments . . . 94
10.2 The energy dependence of the fragmentation variables . . . 97

11 Jet classification using fragmentation 102
11.1 Using the Fisher discrimination method . . . 103
11.2 Using only one fragmentation variable . . . 105
11.3 Using artificial neural networks . . . 105
11.3.1 The training of an ANN . . . 108

11.3.2 JETNET 109
11.3.3 Description of the different ANNs used 110
11.4 Comparison of Monte Carlo data with real data 111
11.5 Performance of the classifiers used on Monte Carlo data 112
11.5.1 Purity on a jet to jet basis 115
11.5.2 Purity on an event basis 118
11.6 The classifiers used on experimental data 120

12 Applications of jet identification **130**
12.1 Investigation of the splitting of quark and gluon jets 130
12.2 Measurement of the triple gluon vertex 132
12.3 Scale dependence of the multiplicity 135

13 Conclusions **138**

A Geometry for Classification Purposes **144**

B The energy dependence of the longitudinal momenta **147**

C The energy dependence of the transverse momenta **150**

D The energy dependence of the Fodor momenta **153**

Chapter 1

Introduction

The theory of Quantum Chromo Dynamics, QCD, contains two fundamental types of particles, quarks and gluons. Quarks carry a single colour while gluons carry two colours and therefore quarks and gluons differ in their coupling strength to another gluon. Quarks and gluons themselves are not visible but after an interaction they hadronize into a jet of particles. Jets originating from quarks and gluons are expected to differ in energy spectrum, multiplicity and angular distributions and many previous studies [1], [2], [3], have confirmed that gluon jets have a softer energy spectrum, a higher multiplicity and a wider angular distribution than a quark jet. The energy spectra can be calculated to high precision from the matrix elements in pertubative QCD. The differences between the energy spectra of quark jets and gluon jets can essentially be explained by the fact that gluons are normally produced in an additional bremsstrahlung-like radiative process from a quark or another gluon and thus fragmentation differences are due to the different colour charge of the gluons compared to the quarks. Because the fragmentation takes place at a low energy scale, pertubative QCD is not applicable and the process must be studied with phenomenological models.

At LEP, the simplest topology involving a gluon jet occurs when a (hard) gluon is radiated from one of the quarks in a $q\bar{q}$ pair produced in a Z^0 decay. Hard here means energetic enough to create a shower of hadrons, a *jet* which is separated in space from the two jets created by the quarks. It is, in principle, possible to see if the original parton was a quark or a gluon. Quark jets and gluon jets are distinguishable, at least on a statistical basis, from differences in their energy spectra and their fragmentation properties and the problem studied in this work is how to best separate these two types of jets. If one can distinguish between quark jets and gluon jets the predictions of QCD can be tested and the fragmentation features can be better understood. A more specific example is five jet events at LEP2 coming from a WW-pair that have radiated a hard gluon. If the gluon jet could be identified the possibility to combine jets to the correct W-particles would be enhanced and the mass determination would be better.

Most of the interesting events at hadron colliders, for instance the Tevatron, are dominated by quark jets and the gluon jets are a background to these events. For instance $t\bar{t} \rightarrow W\ W\ b\bar{b}$ where the W-particles both decay by $W \rightarrow q\bar{q}$. The reverse is also true, i.e. if the elusive Higg's particle will be produced, its decay rate to two gluon jets (via a top loop) could, at least in principle, be seen if the overwhelming "background" of b-jets

could be overcome. The determination of the parton distribution of the proton would also be improved if the different jets could be well separated.

Normally, the jet energy provides a high degree of separation between quark jets and gluon jets. The intention in this thesis is to demonstrate how the quark/gluon jet separation using only the fragmentation (so far mostly studied at Monte Carlo level) performs on real data, collected during the period 1992-1994 by the DELPHI experiment at LEP. The quark/gluon jet separation using the fragmentation is a statistical problem which may be solved by using methods such as Fisher statistical discrimination or Artificial Neural Networks (ANNs). In this theses ANNs trained on different types of fragmentation variables and with different training methods are compared. For completeness, the Fisher statistical discrimination as well as the discriminative power of the single best input variable are compared with ANNs.

A short historical background, leading to QCD, is given in chapter 2. Some sixty years ago there was a simple picture with only three fundamental particles. Then "chaos" broke out with hundreds of elementary particles, some 40 years ago, but soon order was restored by the quark hypothesis and the invention of the Standard Model (SM) with its three families.

The accelerator LEP and the DELPHI detector, presented in chapter 3, are intimately connected, LEP provides an extremely clean initial condition and the full solid angle coverage of the DELPHI detector makes it possible to fully use the powerful constraint of energy-momentum conservation.

A short introduction to multi-dimensional methods, such as the Fisher statistical method and Artificial Neural Networks is given in chapter 4. An overview of the main analysis is given in chapter 5.

The event selection, which is more or less the same for all DELPHI analysis, is discussed in chapter 6. In the same chapter, different jet algorithms are presented. The LUCLUS algorithm was mainly used to reconstruct the threejet events for this analysis.

The different methods to identify or tag specific events are given in chapter 7. With the silicon vertex detectors, now installed in every LEP experiment, there is a possibility to select special events with a high purity as well as high efficiency. This will be used to find a pure sample of tagged b-quark jets that will be used to produce a sample of clean gluon jets from real data.

The different ways of obtaining, from the hadronic events, a quark sample as well as a gluon sample from the Monte Carlo is the content of chapter 8. It also reviews the anti-tagging method, which will be used to obtain the sample of real gluon jets from $b\bar{b}g$ events. The quark jet sample has been obtained from $q\bar{q}\gamma$ events. The same types of samples have been used in other studies measuring the multiplicity and particle contents of quark jets and gluon jets [1, 2, 3].

The must powerful variable to use when considering a quark/gluon jet classifier for the bulk of the events is the jet energy and it is discussed and evaluated in chapter 9. The performances of the various analysis methods can be compared to the corresponding samples obtained from detector simulated Monte Carlo data. Here is also a discussion on how to adequately present the performance of the classifiers in different jet topologies.

The intention of the thesis is to demonstrate how the quark/gluon jet separation using only the fragmentation performs on real data without using the jet energy which is

particular relevant for the subset of events which have similar jet energies. A discussion of the most suitable ways of doing this is discussed in chapter 10 and the different sets of fragmentation variables used as input are also presented in this chapter.

The selected fragmentation variables are then used in chapter 11 as input to the Fisher statistical discrimination method as well as to Artificial Neural Networks and their performances as quark/gluon jet classifier are evaluated. The traditional method of averaging over the entire jet sample is compared with the more useful measure of performance associated with a specific event type. Here is also a presentation of the results on the sample of DELPHI quark jets.

A brief account is given in chapter 12 of results of recently performed DELPHI analyses, where quark and gluon jets have been selected and identified in ways similar to those described in this thesis.

In the last chapter, conclusions are drawn of this study of jet identification.

Chapter 2

Theory

2.1 One hundred years of Particles Physics

Human knowledge is said to depend on our faculty to receive some raw material from the outside world and on our faculty to organise this material by means of certain concepts. Using this definition our knowledge of basic elementary particle physics has changed several times during the last century.

2.1.1 Only three particles

When Thompson discovered the negatively charged electron and its elementary nature about a hundred years ago, it became clear that *the atom* was not elementary. He also found the weight of the electrons to be about 2000 times less than the smallest atom, that of hydrogen. The discovery of the electron was reported to the British Royal Institution on 30 April 1897, so this date can be considered as the birthday of particle physics.

In the beginning of this century, Rutherford performed his scattering experiment with α-particles bombarding a very thin gold foil. The heavy, positively charged α-particles, which had been discovered some decade earlier by Becquerel and the Curie couple, penetrated the foil easily in most cases. Occasionally, however, they scattered into large angles, or even bounced backwards[1], which could be interpreted as if the mass and the charge of the atom was concentrated to a very small space in the centre, called the *nucleus*, about 10000 smaller then the size of the atom ($\sim 10^{-10}$ m).

In 1919, Rutherford returned to Cambridge to become director of the Cavendish laboratory where he had previously done his graduate work under J.J. Thomson. Here he discovered the proton from α-particle bombardment (from Po^{214}) on nitrogen, by observing the first nuclear reaction, $N^{14}(\alpha,\mathrm{p})O^{17}$.

When Chadwick, in 1932, found the neutron, the picture of particle physics was very clean. All ordinary matter is made of atoms, which are built by a nucleus of protons and neutrons (nucleons) surrounded by as many electrons as there are protons in order to make the atom neutral. For a while these particles were considered to be truly elementary, i.e. the number of elementary particles, at this time, was only three.

[1] At a rate of one in eight thousand. "As if you fired a 15-inch naval shell at a piece of tissue paper and the shell came right back and hit you" was Rutherford classical comment.

2.1.2 Hundreds of particles

Physicists knew that nuclei are difficult to break up, so what held the protons and the neutrons, i.e. the *nucleons*, together? In 1935, considering the size of nuclei, about 1 Fermi, Yukawa predicted a particle, which he called meson (meson = middle), which ought to be heavier than the electron ($m_e = 9,109 \cdot 10^{-31}$ kg) but lighter than the 1836 times heavier proton. The particle should be responsible for *the strong nuclear force.* Yukava drew a parallel with the well known electromagnetic force between charged particles. This force could be interpreted in such a way that one of the particles is emitting radiation which is absorbed by the other. According to the quantum theory this radiation is quantised and the force is originating from the exchange of photons. They cannot detach themselves from the charges, so they are said to be *virtual.* Photons are *force carriers* and are characterized by integer spin (in units of $\hbar$). Such particles are called *bosons.* Since the strong nuclear force is acting at a very short range, approximately 10^{-15} m, the energy (and therefore its mass) of the strong force carriers could be estimated according to Heisenberg's uncertainty principle, to be around 140 MeV ($M = 10^6, 1\ eV = 1.6021 \cdot 10^{-19}$ J).

Anderson, who in 1933 also discovered the first antiparticle, the positron, found in 1937 yet another new particle in the cosmic rays, the muon or μ, which was first mistakingly identified with the Yukawa particle. It was, however, a puzzle that this new particle interacted "weakly". The real Yukawa particle, the pion, π, was discovered in 1947 in balloon flight carrying emulsion from the interaction of cosmic protons with the upper atmosphere producing pions which sub-sequentially were observed to decay to Anderson's muon.

In the fifties an explosion of discoveries of new particles followed. Some of these were called strange particles, for example K, Λ, Σ and Ξ, because they were produced in pairs and decayed very slowly, in about 10^{-10} seconds and sometimes even longer. In order to explain why these particles lived so long, Gell-Mann proposed that they carried a fundamental property that he called *strangeness*, which had to be conserved in strong decays. Therefore the strange particles had to decay via the weak nuclear force, in which strangeness was not conserved and so managed to live longer. The clean particle physics picture from 1930 was now a "zoo" of particles, and the situation was very much like the one in chemistry, before the table of elements (the periodic system) was presented by Mendeleyev in 1869.

2.1.3 A subatomic periodic table

Gell-Mann and Néeman constructed in 1961, the first successful subatomic periodic table by grouping the particles in families of specific geometrical patterns. These families contained sometimes eight and sometimes ten members. In 1962, Gell-Mann predicted a new particle, Ω^-, as a missing member in a group of ten. The question was whether the theory would be able to explain the truth underneath the geometrical patterns. Zweig and Gell-Mann showed that this classification was a natural fact if there were three basic constituents, later named *quarks.* Gell-Mann called them *up*, *down* and *strange* or just u, d and s. They must all have fractional charges, the up quark should have a charge of

$+2e/3$ and the down quark and the strange quark, a charge of $-e/3$. Fractional charges was a concept many physicists refused to accept. However, simple quark algorithms or rules could form all known hadrons (hadron = heavy, heavier than the proton), all baryons by putting three quarks together and all mesons by combining a quark and its antiquark. The proton was, for example, easily constructed of two up-quarks and one down-quark, giving the proton a positive "elementary" charge. The neutron was equally easy to built of one up-quark and two down-quarks, in that way being neutral. The quarks were at first accepted as some sort of a mathematical tool rather than real constituents.

A real problem appeared when the Ω^-, was discovered at Brookhaven's bubble chamber in 1963, see figure 2.1, since the model to construct the Ω^- violated a fundamental law of physics, the Pauli principle. The same argument applies to the particle Δ^{++}, consisting of three u-quarks and Δ^-, consisting of three d-quarks. This principle excludes

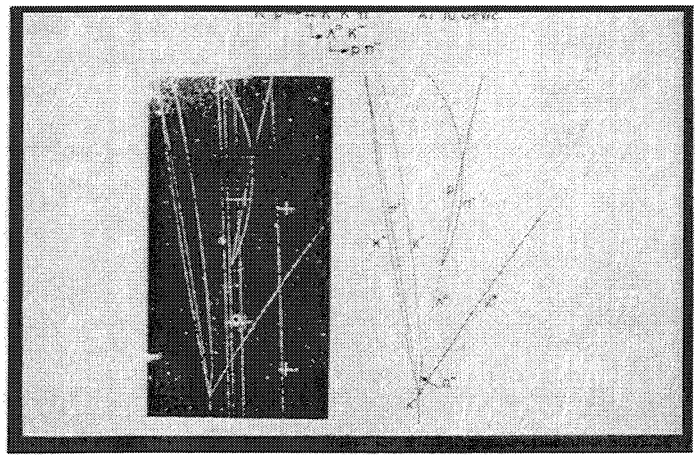

Figure 2.1: Production and decay of a negative omega particle, Ω^-, when a K^- has hit a proton, in the CERN two-metre liquid hydrogen bubble chamber.

the possibility that the three, identical, s-quarks, needed to construct the Ω^-, could be in the same state (i.e. have the same set of quantum numbers). It can be shown that the particle (having J=3/2 and L=0) must have the spins of three quarks aligned in the same direction and therefore could not exist inside the Ω^- simultaneously if they were identical, since a symmetric wave function could not possibly describe a fermion. The Pauli exclusion principle of quantum mechanics required an antisymmetric wave function. W. Greenberg solved this puzzle by introducing three distinguishable *colours* which restored the antisymmetry. Colour turned out to be the key to explaining how the quark model could be true. Now also the precise value of the anomalous magnetic moment of the proton and the neutron found an explanation in terms of the quark model with three colours. This had been a puzzle since the discovery in the beginning of the 1930s.

In order to bring some experimental proof into the quark idea, a rerun of Rutherford's classical experiment was performed at the Stanford Linear Accelerator Center (SLAC) in 1967, with its three kilometres long accelerator. In the SLAC-MIT experiment high energy electrons scattered inelastically against the nucleon. The results showed a much

higher rate of wide angle scattering than expected from previous models for the nucleon, revealing tiny charged bodies inside the nucleon. The first experimental evidence for the substructure of the nucleons required high momentum transfers. According to Bjorken and Feynman, the result of the experiment could be explained by assuming that the proton consisted of a few hard, pointlike scattering centres. In 1968 R. Feynman called those new objects partons, not yet ready to call them quarks. B.J. Bjorken worked out the detailed consequences of the quark-parton model for both electron and neutrino scattering.

In 1974, a neutrino beam experiment at Cern, when compared to the electron scattering experiment, showed that the main features of the data could be explained by the existence of the three valence quarks of the nucleon.

The very same year, 1974, a group at Stanford led by Richter and another group at Brookhaven led by Ting simultaneously discovered a new, heavy and long lived, totally unexpected, particle. The two research groups agreed about all details describing the particle except for the name. Therefore we are stuck with the name J/Ψ. Is was agreed that the fourth quark *charm*, or c, had been found in the J/Ψ particle. Since the particle was a quark-antiquark conglomerate the charm was hidden.

By 1974, the world of particle physics looked in a remarkably good shape, with four leptons (lepton = light), the electron, e^-, the electron neutrino, ν_e, the muon, μ^- and the muon neutrino, ν_μ, together with the four quarks up, down, strange and charm.

But one year later, the number of leptons suddenly jumped to six, because an even heavier brother than the muon to the electron, called tau, τ, was found, also at SLAC. Since electrons and muons had separated neutrino partners, a tau neutrino, ν_τ must also exist, although it has not yet been directly observed. To restore the equilibrium predicted for the sum of the electric charges for the leptons and quarks two extra quarks were needed. The fifth quark, *bottom*, or b, was found by Lederman at Fermilab in 1977. The particle, named upsilon, Υ, is a quark-antiquark conglomerate, the bottom quark was hidden, as was the case for the charm quark. In early LEP data there were strong, indirect, indications that the sixth quark, *top*, t, existed. The discovery was made at Fermilab in 1995 confirming the LEP estimate of the top mass (175 GeV). This completed the zoo of elementary building blocks of nature as we know it now!

2.2 The Standard Model

The theoretical picture that describes how the elementary particles are organised and how they interact with one another via the forces is called *the Standard Model*. This describes the quantum field theory that includes the theory of strong interactions (quantum chromodynamics or QCD) and the unified theory of weak and electromagnetic interactions (electroweak). The underlying principle, the spontaneously broken local gauge symmetries $SU(3)_C x SU(2)_L x U(1)_Y$ is nice and "simple". The model enables us to calculate, to a high precision, variables that can be observed and measured, like cross sections, branching ratios and lifetimes. Gravity [2] is, however, not incorporated in the model although it is

[2]Much of the problem is connected with the tensor nature of its predicted mediator of the gravitational force, the *graviton*, G, which is supposed to be a spin 2 particle. Since the gravitational force is very small compared to all other forces at present energies, it has a negligible effect on the measured data.

one of the fundamental interactions. The standard model accommodates all the quarks, leptons and bosons (except G); the gluons, the W^+, the W^-, the Z^0 and the photon. To produce massive particles the *Higg's mechanism* is needed, and at least one *Higg's particle* is predicted to be observable. (An event, where a Higg's particle and a Z^0 are produced, could look like the decay of a W^+W^--pair in figure 3.4). The standard model leaves some questions unanswered, e.g. the concept of charge and spin. Two of the most important puzzles are the questions of how to calculate the spectrum of masses and why there exist three generations of particles.

2.2.1 Quantum Chromodynamics

The strong nuclear force is due to colour exchange. The colour force is another example of a gauge theory, a theory using relativistic quantum mechanics and the principles of symmetry to summarise how the particles will interact. The theory, describing the interaction by gluons, postulates that each quark carries one of three types of *colour charge*, for convenience often called red, green and blue. If a red quark and an antired antiquark are combined, the resulting object, a meson, is colourless and the same result is obtained if a red, blue and green quark are combined giving a baryon. The deep inelastic scattering experiment at SLAC revealed that only half of the momentum of the nucleon was carried by the charged quarks. The rest must be due to the strong force carriers.

In $SU(3)_C$, a nonabelean, local, three-dimensional symmetry group, the quark colour force is transmitted by the eight possible types of bicolour charge carriers called gluons. Just as electrically charged particles interact by exchanging photons, in strong interactions colour-charged particles exchange colour-charged gluons. Since gluons have colour they interact among themselves, which is not the case with the electromagnetic force carrier, the photon. (Leptons, photons, and W and Z bosons have no colour charges, so they have no strong interaction.) One has experimentally found that one cannot isolate the colour charge. Therefore, quarks and gluons are *confined* into colour-neutral hadrons i.e. mesons or baryons. This confinement (binding) results from multiple exchanges of gluons among the colour-charged objects.

A great step forward came in 1973 when the concept of *asymptotic freedom* was worked out by D. Politzer, D. Gross and F. Wilczek. In high energy collisions, quarks and gluons are more or less independent particles, what is called asymptotic freedom and it also means that the colour coupling "constant" α_s decreases with increasing energy, α_s is said to be *running*. From then on a real calculable theory of strong interaction could be formulated. Asymptotic freedom allows pertubative calculation.

QCD predicts gluon emission by one of the quarks (bremsstrahlung) analogous to the photon emission by electrons in QED. The proposition was that the gluons should manifest themselves as jets not very different from these of quarks and that three-jet topologies should be observed. This was showed in PETRA experiments in 1979. Figure 2.2 shows a nice picture from DELPHI of a threejet event including a gluon jet. The corresponding Feynman diagram is seen in figure 2.3. Inside the hadron at long inter-quark distances, the strong force is reasonably. When the quarks are further separated, the potential energy is increasing linearly with distance and it is clear that it would take an infinite amount of (kinetic) energy to remove a quark from the hadron. Eventually, before that

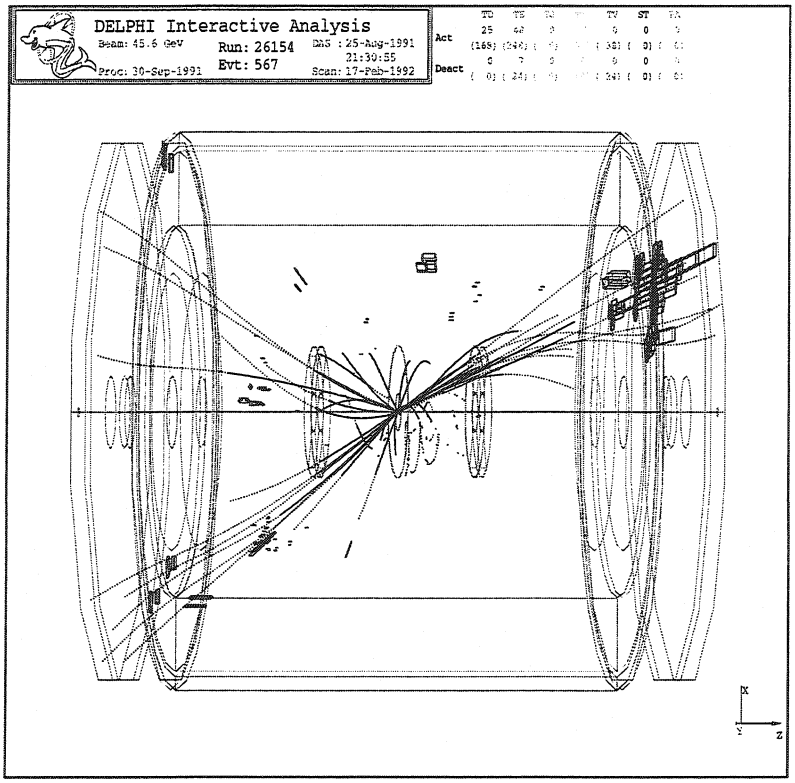

Figure 2.2: Z^0 decay into two quark jets and one gluon jet in the Delphi detector.

happens, the energy is converted into additional quark-antiquark pairs, see figure 2.4. The objects, finally emerging as a jet of particles, are colour-neutral combinations of baryons and mesons, see table 2.3 and table 2.4, respectively.

The α_s, depends on the energy transfer scale Q^2, where the different processes, including the self coupling of gluons and loop diagrams have to be considered, giving us the following "running" coupling constant,

$$\alpha_s(Q^2) = \frac{12\pi}{(11N_c - 2N_f) \cdot ln(\frac{Q^2}{\Lambda^2})}$$

where:

- N_f is the number of fermion families (i.e. three);
- N_c is the number of colours (i.e. three);
- Λ is the mass scale, the only free parameter of the theory.

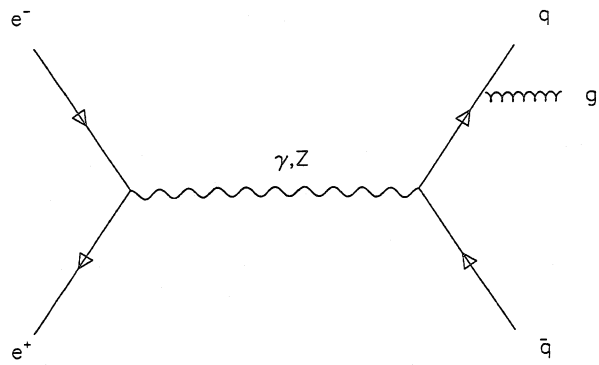

Figure 2.3: A Feynman diagram of a Z^0 decay into two quark jets and one gluon jet.

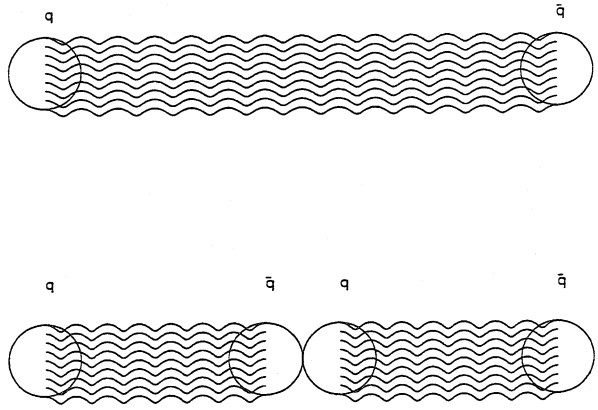

Figure 2.4: The $q\bar{q}$ string pulled far enough apart can create a new $q\bar{q}$ pair.

Λ is the reference energy that we choose depending on our energy scale. A measured $\alpha_s(Q^2)$ at some Q^2, will define the value of Λ, which can only be determined from experiment. A typical value for $\Lambda \sim 200MeV$ is well suited for the hadronic world. At large Q^2, the strong coupling constant is sufficiently small to treat the quarks and gluons as almost free particles but at smaller values of Q^2 this is not the case and it is obvious from equation 2.2.1 that when Q^2 is close to Λ^2, $\alpha_s(Q^2)$ will approach infinity i.e. α_s diverges. This is called *infrared slavery* and believed to be the origin of the *confinement* of colourless hadrons. Therefore perturbation theory fails at energies $\sim 1\ GeV$ where interactions become very strong and some other calculation technique, not based on perturbation expansion, is required.

2.3 Grouping of particles

The particles to be found in the Particle Data Tables are grouped according to the intrinsic spin, parity and charge conjugation. Fermions have half-integer spin and bosons have integer spin. They obey different statistics, Fermi-Dirac- and Bose-Einstein-statistics, respectively.

2.3.1 Matter constituents

The building blocks of matter are fermions, i.e. particles with half-integer spin. There are two classes of such particles, namely quarks and leptons. One of the first aims of the LEP experiments was to determine the total number of elementary fermions by studying the mass and the width of the Z^0 particle, the mediator of the weak force, with high precision from a scan at five accurately known beam energies around the peak. The width of the Z^0 allows a determination of the number of light (i.e. with a mass, $m_\nu < 45GeV$) neutrinos to three, see figure 2.5. Since the shape is fully determined by quantum field theory, by calculating the width of the Z^0, Γ_Z and using

$$\Gamma_Z = \sum_l \Gamma_{l^+l^-} + \sum_q \Gamma_{q\bar{q}} + \Gamma_{inv}$$

and getting Γ_Z from

$$\sigma_{peak} = \frac{12\pi}{M_Z^2} \cdot \frac{\Gamma_{l^+l^-} \cdot \Gamma_h}{\Gamma_Z^2}$$

the result will be

$$\frac{\Gamma_{inv}}{\Gamma_{l^+l^-}} = 5.957$$

and defining

$$N_\nu \equiv \frac{\Gamma_{inv}}{\Gamma_{\nu\bar{\nu}}}$$

measuring the fraction

$$\frac{\Gamma_{\nu\bar{\nu}}}{\Gamma_{l^+l^-}} = 1.991 \pm 0.001$$

will give as the best fit the value of the number of light neutrinos

$$N_\nu = 2.992 \pm 0.011$$

This makes it very plausible that the number of quark families is not more than three, hence there are probably six leptons, see table 2.1, and six quarks, see table 2.2.

Hadrons are made up of quarks bound together by gluons. The particles with three quark constituents form all the *baryons*, or *antibaryons*, see table 2.3 for some examples. Baryons range from the lightest states, i.e. proton and neutron, which are constituents

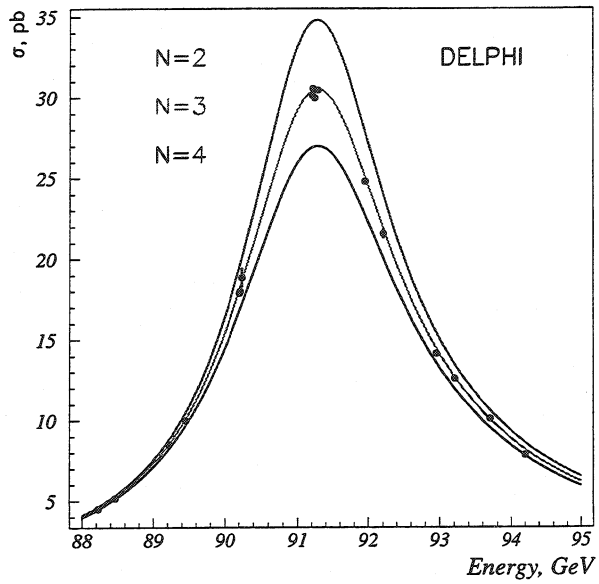

Figure 2.5: The width of the Z^0 allows a determination of the number of (light) neutrinos.

leptons	symbol	charge, e	mass, GeV/c^2
electron neutrino	ν_e	0	$< 5 \cdot 10^{-9}$
electron	e	-1	0.000511
muon neutrino	ν_μ	0	< 0.00017
muon	μ	-1	0.105658
tau neutrino	ν_τ	0	< 0.018
tau	τ	-1	1.777

Table 2.1: The leptons.

of the atomic nucleus, to the ever-growing families of "hyperons" and other baryonic "resonances" at higher energies, exited particle states. The particles with quark-antiquark constituents form all the *mesons*, (or *anti-mesons*), see table 2.4 for some examples.

2.3.2 Force carriers

All the force carriers are bosons. In some respect these particles are even more fundamental than the matter constituents, e.g. the gluons could by themselves form matter. However, nature did not chose this option. See table 2.5 for a specification of the force carriers.

quark flavour	symbol	charge, e	mass, GeV/c^2
down	u	-1/3	0.0015-0.005
up	d	+2/3	0.003-0.009
strange	s	-1/3	0.060-0.170
charm	c	+2/3	1.1-1.4
bottom	b	-1/3	4.1-4.4
top	t	+2/3	173.8±5.2

Table 2.2: The quark flavours.

baryon	quark content	charge, e	spin, $\hbar$	mass, GeV/c^2
proton, p	uud	+1	1/2	0.938
antiproton, $\bar{p}$	$\bar{u}\bar{u}\bar{d}$	-1	1/2	0.938
neutron, n	udd	0	1/2	0.940
Σ^+	uus	+1	1/2	1.189
Σ^0	uds	0	1/2	1.193
Σ^-	dds	-1	1/2	1.197
Δ^{++}	uuu	+2	3/2	1.232
Ξ^0	uss	0	1/2	1.315
Ξ^-	dss	-1	1/2	1.321
Ω^-	sss	-1	3/2	1.672
Λ_b^0	udb	0	1/2	5.641

Table 2.3: Some baryons.

2.4 The fragmentation process

The fragmentation (or hadronization) is the process in which quarks and gluons transform into observable hadrons. A phenomenological model for the fragmentation process can be seen in figure 2.6. The example is from LEP running as a Z^0-factory.

- Phase 1 (10^{-19} m)
  The e^+e^- pair produces a Z^0 at the resonance, where the contribution from the photon exchange channel can be neglected. Before the annihilation, the electron and the positron can have initial state radiation (ISR). The process is well described by the electroweak theory. When the Z^0 decays hadronically, a $q\bar{q}$ pair is created.

- Phase 2 (10^{-17} m)
  The $q\bar{q}$ pair can radiate gluons (primary branching) often called hard gluon *bremsstrahlung*. In the next branching a gluon can split into two gluons as well as into a $q\bar{q}$ pair according to QCD. This prediction of QCD, that gluons couple to themselves, have been proven experimentally in fourjet events and therefore the nature of QCD is non Abelian. These first steps of the whole process can be described by pertubative QCD, either based on exact matrix elements calculations to

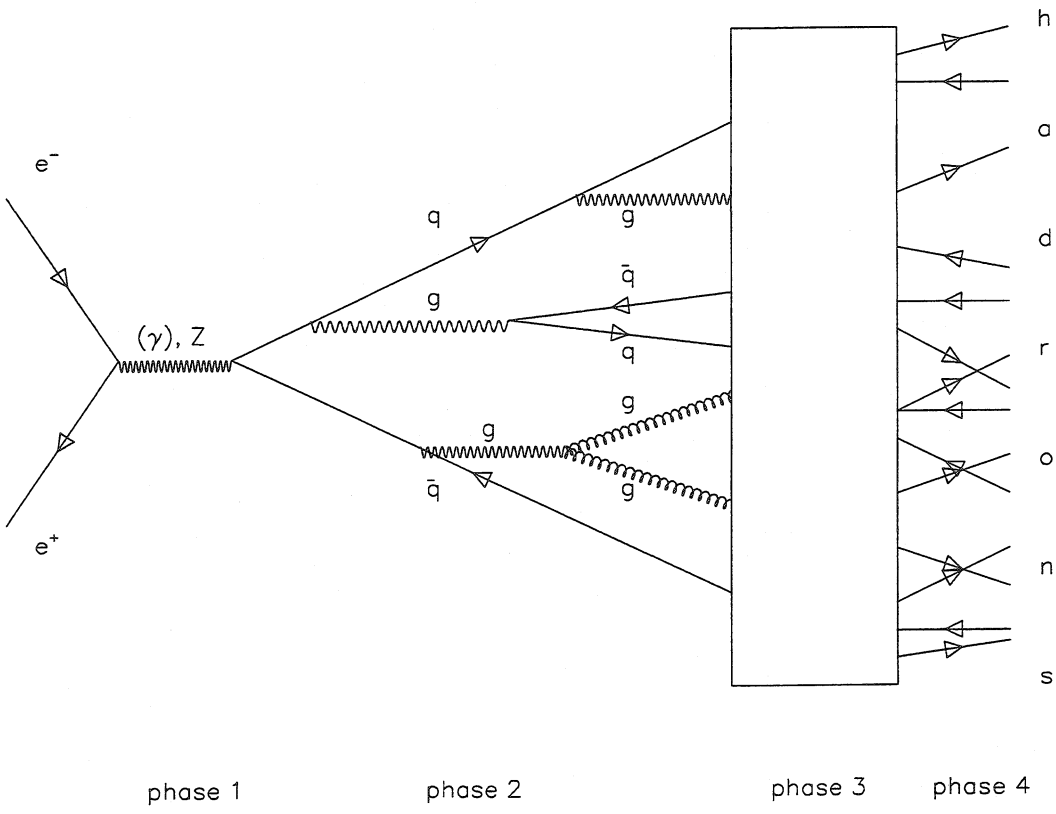

Figure 2.6: The hadronization process.

mesons	quark content	charge, e	spin, $\hbar$	mass, GeV/c^2
π^+	$u\bar{d}$	+1	0	0.140
π^-	$\bar{u}d$	-1	0	0.140
K^+	$u\bar{s}$	+1	0	0.494
K^0	$d\bar{s}$	0	0	0.498
$\bar{K}^0$	$s\bar{d}$	0	0	0.498
K^-	$\bar{u}s$	-1	0	0.494
D^0	$\bar{u}c$	0	0	1.865
$\bar{D}^0$	$\bar{c}u$	0	0	1.865
D^+	$\bar{d}c$	+1	0	1.869
D^-	$\bar{c}d$	-1	0	1.869
D_S^+	$\bar{s}c$	+1	0	1.969
D_S^-	$\bar{c}s$	-1	0	1.969
J/$\Psi(1S)$	$c\bar{c}$	0	1	3.097

Table 2.4: Some mesons.

force	force carrier	charge, e	spin, $\hbar$	mass, GeV/c^2
strong	g	0	1	0
elec. magn.	γ	0	1	0
weak	W^+, W^-, Z^0	$\pm$1,0	1	80.33 91.187
gravity	G	0	2	0

Table 2.5: The force carriers.

some order or parton shower models which are based on leading log approximations. Parton showers are describing how a single initial parton gives a bunch of partons in the final state and the process can be treated in a pertubative way, which has been showed above in equation 2.2.1. When Q^2 is less than Λ^2, the shower is no longer developing.

- Phase 3 (10^{-15} m)
  As said above, when Q^2 is less than Λ^2, it is no longer possible to treat the process by pertubative QCD and therefore phenomenological models have to be used. In the *fragmentation* models, the partons now regroup into confined colour neutral hadrons, forming jets of hadrons.

- Phase 4 ($> 10^{-15}$ m)
  In this last phase unstable hadrons, i.e. short lived resonances, will decay into relatively stable particles that will finally be seen by the various detectors.

2.5 The Z^0 decay into b-quarks

According to the standard model the Z^0 created in the e^+e^- annihilation rapidly decays into a fermion-antifermion pair with the following branching ratios:

- $Z^0 \rightarrow e\bar{e} \approx 3.367 \pm 0.0009$ %
- $Z^0 \rightarrow \mu\bar{\mu} \approx 3.367 \pm 0.0013$ %
- $Z^0 \rightarrow \tau\bar{\tau} \approx 3.367 \pm 0.0015$ %
- $Z^0 \rightarrow \nu\bar{\nu}$ (i.e. $\nu_e\bar{\nu_e}$ and $\nu_\mu\bar{\nu_\mu}$ and $\nu_\tau\bar{\nu_\tau}$) $\approx 20.01 \pm 0.16$ %
- $Z^0 \rightarrow$ hadrons $\approx 69.90 \pm 0.15$ %

In this study the interest is mainly focused on b-quarks and to some extent c-quarks and therefore it is interesting to know the fraction of the hadronic decays that will decay into these particles.

- $Z^0 \rightarrow c\bar{c} \approx 11.9 \pm 1.4$ %
- $Z^0 \rightarrow b\bar{b} \approx 15.45 \pm 0.21$ %

2.6 B-particles

The b-quark has a running mass, depending on the production mechanism. The value obtained from Υ-decay is around 4.1-4.4 GeV and is therefore more than twice as heavy as the c-quark and four times heavier than the familiar proton, see table 2.2. As we have seen in chapter 2, the b-quark can form B-mesons, together with lighter quarks. These particles will therefore have heavy masses and also *long* lifetimes, of the order of picoseconds [3]. A selection of B-mesons can be found in table 2.6.

B-mesons	quark content	charge, e	spin, $\hbar$	mass, GeV/c^2
B^+	$u\bar{b}$	+1	0	5.2789
B^-	$\bar{u}b$	-1	0	5.2789
B^0	$\bar{b}d$	0	0	5.2792
$\bar{B}^0$	$b\bar{d}$	0	0	5.2792
B_s^0	$s\bar{b}$	0	0	5.3693
$\bar{B}_s^0$	$b\bar{s}$	0	0	5.3693

Table 2.6: Some B-mesons.

[3] A picosecond, 10^{-12} seconds, seems like a very *short* time, but the time scale of strong interaction, in which the B-mesons are formed, is some 10^{-23} seconds.

2.7 Simulated events

Simulated events exist at three levels:

- generator level;
- detector level;
- reconstructed level.

The events are produced in an event generator, normally the JETSET program, giving events as detailed as could be observed by a perfect detector. The events are sent to a detector-simulator where all effects of the detector are taken into account. The output of the detector simulation should give the electrical signals as produced by the detector. In the next step the events are put into the same reconstruction program as for the real events.

2.7.1 Monte Carlo generation

A typical hadronic final state from a Z^0 decay contains about thirty stable particles, i.e. some twenty charged particles, mainly pions, and about ten neutral hadrons, mainly π^0's, decaying into photons. Monte Carlo event generator programs, e.g. [7] have been developed providing accurate detailed descriptions of complete hadronic events. The aim of the Monte Carlo is to simulate the physics built on a theory and the response of the physical event of the experiment. In order to use this type of simulation for physics analysis the reconstructed Monte Carlo events must include detector simulation in order to make a perfect "copy" of the detector in the computer. Usually it is difficult to incorporate all detector faults and inefficiencies in one step. If the Monte Carlo was completely correct, any distribution of experimental data compared to the corresponding Monte Carlo data, would be identical within statistical fluctuations.

2.7.2 Pertubative QCD

There exist two attempts within the framework of pertubative QCD in order to construct the final states of the originating partons.

- Matrix Elements (ME)
- Parton Shower (PS)

In both of these models one starts with a system of partons generated according to pertubative QCD.
In the Matrix Element model the pertubative state hadronizes directly according to the Lund string fragmentation model and the calculations are completely determined analytically and exact up to second order pertubative QCD.
The parton shower development is based on that the initial quark-antiquark pair evolves into a system of quarks and gluons according to the leading log (LL) approximation which will allow more than four partons in the final state (using the Altarelli-Parisi

formula). Successive branching of the type $q \rightarrow qg$, $g \rightarrow gg$ and $g \rightarrow q\bar{q}$ are repeated until the invariant mass off the resulting partons falls below a specified cut-off value of the order of $\approx 1\ GeV$. The final conversion of partons into colour neutral hadrons is simulated using the string fragmentation scheme.

2.7.3 Non Pertubative QCD

Several schemes have been developed to describe the quark and gluon fragmentation. The most important are:

- independent fragmentation;
- cluster fragmentation;
- string fragmentation.

The independent fragmentation approach of the hadronization process is almost synonymous with the work of Field and Feynman. This fragmentation model assumes that any system of partons can be described by a non coherent sum of individually fragmenting partons and the fragmentation axis of the jet is given by the direction of the originating parton. The best known Monte Carlo generator using independent fragmentation is the "colour blind" COJETS where the fast particles in the jet are used which will give an isotropic string fragmentation with a diquark-antidiquark. Typically a large cut-off value, $Q_0 \approx 3\ GeV$, is used which will result in a final state with only a few partons. COJETS uses fewer parameters than HERWIG (see next section) or JETSET but gives, on the other hand, a poorer fit when comparing Monte Carlo data with experimental data.

The best known Monte Carlo generator using cluster fragmentation is HERWIG which, in the final state will produce only quarks. Therefore the gluons remaining at the cut-off Q_0 are forced to split into (light) $q\bar{q}$ pairs. Then neighbouring quark-antiquark pairs form colour singlets with an effective mass. The fragmentation will allow a smallest mass $m_g \simeq 0.7$ GeV for the gluons, and $m_q \simeq 0.3$ GeV for the quarks clusters.

The physics generator mostly used in this work is the Parton Shower MC JETSET 7.4 [7], the Lund Monte Carlo, specially suitable for e^+e^- physics. The fragmentation is the Lund string scheme. The Lund fragmentation model is described in detail in [8], where the basic ideas are presented. A more modern approach is in [9] and reviews of fragmentation models can be found in [10] and [11]. In the string picture a production of $q\bar{q}$ pairs are causally disconnected but the gluon field between the quarks forms a string. A convenient way of describing the linear confinement is to use the picture of a massless relativistic string with a string constant k of the order of the appropriate energy scale vs. the transverse dimension of the string i.e. $k \sim 1GeV \sim 1fm$.
As described above, when the quarks are separated, the field energy will increase linearly with distance, neglecting a short-range 'Coulomb' term, and the string will break, when the energy density along the string is greater than the string value, thus converting

the energy into an additional quark-antiquark pair, see figure 2.4. The model also admits solutions that include localized energy-momentum carrying 'kinks' which have successfully been identified with gluons. In this way the geometry of the string can be seen as two string pieces attached to the gluon and the open ends to the quark, q, and the anti-quark,$\bar{q}$, respectively. This process will continue until all systems are so small that they are all leading to hadronic states. In the simplest case this will lead to a system with two "cones" of particles, each of them forming a *jet*. In describing the the string breaking, there is no preference for where along the string the fragmentation process starts. The *fragmentation function* will, since the fragmentation process should look the same from the q-side as well as from the $\bar{q}$-side, therefore have a left/right symmetry. This leads to the "Lund symmetric fragmentation function", which is very suitable for describing light quarks, u-, d- and s-quarks.

$$f(z) \sim \frac{1}{z}(1-z)^a \exp\frac{-b \cdot m_T^2}{z}$$

where m_T is the transverse mass of q and $\bar{q}$ and

$$z = \frac{(E+p_L)_{hadron}}{(E+p_L)_{quark}}$$

The fragmentation function will be retained by a hadron after the breakup of the string and can be modelled in many ways to suit different situations. Heavy quark fragmentation function has to be modified in comparison with the one for light quark. If for instance the string breaks very close to a heavy quark and therefore receiving a very small amount of mass, giving the result that not many particles can be created and therefore when describing c- and b-quarks the most popular choice for the fragmentation function, is the Peterson parameterisation [12],

$$f(z) \propto \frac{1}{z(1-\frac{1}{z}-\frac{\epsilon_Q}{1-z})^2}$$

where

$$\epsilon_Q = \frac{m_q^2 + p_{qT}^2}{m_Q^2 + p_{QT}^2}$$

2.8 Tuning of fragmentation variables

The fragmentation variables used in the Monte Carlo in DELPHI are modified slightly from the default values in the Jetset 7.4 Monte Carlo, see table 2.7, in order to best describe the hadronic decay of a Z^0. There are many possible choices for the fragmentation parameters, depending on the Monte Carlo one prefers to use.

- Parton Shower (PS) + String Fragmentation (SF)
- PS + SF + Peterson
- Ariadne (AR) + SF

- AR + SF + Peterson
- NN (Nlljet 2.0) + SF
- Herwig (HW) + SF(decays in Jetset)
- Matrix Element (ME) + SF + Peterson
- ME + Independent Fragmentation (IF) + Peterson
- PS + IF + Peterson

The slightly modified JETSET version used in our case is PPSPE1, that is PARTON SHOWER + STRING FRAGMENTATION + PETERSON, and it is best in describing the fragmentation function for heavy quarks, which will be used in this analysis. A selection of the most important parameter settings in this DELPHI MC compared to the default values in JETSET 7.4 are shown in table 2.7.

Parameter	Parameter name	DELPHI MC	JETSET 7.4
Light flavours	MSTJ(11)	3	4
Max flavour number	MSTJ(45)	6	5
Transverse momentum σ_q	PARJ(21)	0.395 GeV	0.36 GeV
Lund a	PARJ(41)	0.20	0.30
Lund b	PARJ(42)	0.385 GeV^{-2}	0.58 GeV^{-2}
Peterson ϵ_a	PARJ(54)	-0.0474	-0.05
Peterson ϵ_b	PARJ(55)	-0.005	-0.005
Λ value	PARJ(81)	0.30 GeV	0.29 GeV
Cut value on m_{min}	PARJ(82)	1.74 GeV	1.0 GeV
Mass of Z^0	PARJ(123)	91.25 GeV	91.187 GeV
Width of Z^0	PARJ(124)	2.5 GeV	2.489 GeV
$sin^2(\theta_W)$	PARU(107)	0.22	0.232

Table 2.7: DELPHI tuned parameters, compared to default values in JETSET 7.4.

Chapter 3

LEP and the DELPHI experiment

3.1 Why so large dimensions?

The high energy accelerator enables the investigation of smaller and smaller distances, i.e. small compared to the dimensions of the proton. The famous formula of de Broglie, $\lambda = h/p$, gives the answer to the question why higher energy (or momentum) is needed to see small objects, so physicists use accelerators of higher and higher energies to probe deeper and deeper inside the particles and to create larger and larger masses.

When accelerated particles collide at a speed close to the speed of light, they create an energy density as high as that of the early stages of the big bang (some trillionths of a second old). For a symmetric collider, like LEP, the available energy in the center-of-mass system, E_{cms}, for creating new particles, is

$$E_{cms} = 2 \cdot E_{beam}$$

This available energy, E_{cms} goes into a virtual force mediator, which subsequently decays and yields e.g. jets of particles.

To observe and record the created particles, physicists are using sophisticated detectors and powerful computers to sort through and analyse the enormous amount of data, collected from these experiments. The search for new particles and phenomena can test the predictions from the theories supplied by theoretical physicists.

3.1.1 Radiation losses at LEP

Inside LEP, the electron and positron beams are tuned to produce Z^0 particles. Tuned, means that the energy has been selected for creating the the Z^0 boson which has a rest mass of 91.2 GeV.

The particles are being accelerated with the help of a large number of powerful RF-cavities, and magnets so that they bend correctly. Due to the opposite charge of electrons and positrons, they move in the opposite direction in the same vacuum chamber, and are brought to collide at four points along the circumference.

The problem with circular colliders (such as LEP) is, that when electrons are accelerated, they will radiate energy. The maximum electron energy that can be reached is

therefore limited by the available accelerating power that can be provided to compensate for the *synchrotron radiation*. The energy loss per turn, $-\Delta E_{beam}$, by synchrotron radiation, is given as,

$$-\Delta E_{beam} \approx \frac{q^2 \cdot \gamma^4}{3\epsilon_0 \cdot r} \approx 8.85 \cdot 10^{-8} \cdot \frac{{E_{beam}}^4}{r}\ eV$$

where:

- r is the radius of curvature (in meter);
- γ is $\frac{1}{\sqrt{1-v^2/c^2}}$;
- v is the velocity of the electron;
- E_{beam} is the energy of the beam (in MeV).

and the natural constants are:

- $q \approx 1.602\ 177 \cdot 10^{-19}$ As;
- $\epsilon_0 \approx 8.854\ 188 \cdot 10^{-12}$ As/Vm.

The synchrotron radiation at LEP with $E_{beam} \sim 45\ GeV$ and a radius of curvature $r \sim 3096\ m$ gives an energy loss of 120 MeV/turn. This gives a loss of power of 1.6 MW (The ohmic losses are 14 MW at $\sqrt{s} = E_{cms} = 100\ GeV$) which for LEP2 with an energy twice that of LEP is 16 times larger and the synchrotron radiation at LEP2 is $\sim$ 2 GeV/turn. These losses have to be compensated by RF cavities along the LEP ring.

At LEP each beam contained 8 bunches of particles in a so called Pretzels scheme, and measured (150 x 10 x 10000)μm, along the bending radius, perpendicular to the bending plane and along the beam direction, respectively, during 1994. Each bunch contains $\sim 2 \cdot 10^{11}$ particles and traverses LEP approximately every 89 μs, giving a luminous volume at the interaction point every 11 μs. The bunches are kept apart outside the instrumented interaction points by electrostatic separators. Typically, it took 30 minutes to 'fill' LEP with electrons and positrons, 30 minutes to 'ramp' the beams up to full energy and 'squeeze' to produce good luminosity and 15 minutes to correct orbits, cure backgrounds and produce stable colliding beams. Once set-up with high current beams the luminosity is dominated by beam-beam interference effects and careful tuning of the machine parameters can result in $\sim$ 15 hours of data taking by the experiments at luminosities[1] of $\sim 1.8 \cdot 10^{31} cm^{-2} s^{-1}$. When the beam energies are tuned to the Z^0 mass the hadronic (quarks) and visible leptonic (electrons, muons or taus) cross-sections are 30 nb[2] and 4.5 nb respectively, giving $\sim$ 15000 hadronic and $\sim$ 2250 leptonic events a day.

[1]The luminosity, L, is the product of the incident beam flux (particles/unit time) with mean target density (particles/unit area). If the cross section for a special event is σ, the product $L \cdot \sigma$, gives the corresponding event rate.

[2]The prefix n = 10^{-9}, 1 barn = $10^{-24}\ cm^2$

3.2 The accelerator complex

Studies of the design of the LEP machine started at CERN in 1976 and the first practical design was published in 1978. This machine had a cost-optimized energy of 70 GeV per beam and measured 22 km in circumference. Later a cheaper design for the main magnet system was developed and also a more economical system for the RF accelerating system using a storage cavity scheme. The accelerator complex at Cern is shown in figure 3.1, where the LEP ring is the last accelerator in a chain of five, each of which handles the same electrons and positrons generated on every pulse by the electron gun and the positron converter. The LEP injectors consist of two linacs of 200 MeV and 600 MeV followed by a 600 MeV Electron-Positron Accumulator (EPA), which injects into the CERN Proton Synchrotron (PS) operating as a 3.5 GeV e^+e^- synchrotron. The PS then injects into the CERN Super Proton Synchrotron (SPS), which operates as a 20 GeV electron-positron injector for LEP. The decision to use the two already existing CERN proton synchrotrons (the SPS and the PS) and all the infrastructure associated with them, resulted in significant economies both in cost and in time.

3.2.1 LEP, the Large Electron-Positron collider

Physics goals of LEP

The first aim was the production of the Z^0 at an energy of 90 GeV. Since these bosons can be produced singly, the LEP energy is initially about 45 GeV per beam, giving 90 GeV in the centre of mass system. The second aim was the production of pairs of the charged intermediate boson pairs (W^+W^-, each with a rest energy of about 80 GeV) at an energy of about 160 GeV. To explore the Z^0 region, only a part of the RF accelerating cavities and their power amplifiers were needed in the machine. To reach the LEP2 energy stage one had to install super-conducting cavities while all the other components of the LEP machine were designed for this higher energy.

LEP construction

Between 1983 and 1988 LEP was the largest civil-engineering undertaking in Europe. The 26.67 km main ring tunnel formed the most impressive part of this work, even though it represented less than half of the 1.4 million m^3 of material which had to be excavated for the project. The remainder of the underground work consisted of the four experimental caverns, 18 pits, 3 km of secondary tunnel, and some 60 chambers and alcoves. After an extensive campaign of test borings in and around the area proposed for the LEP tunnel it was decided to incline the plane of the tunnel by 1.4 degrees. This decision was made so as to ensure that all underground caverns and the main part of the tunnel would be located in solid rock while and, at the same time, limiting the maximum depth of any of the shaft to less than 150 m. The guidance of the tunnelling machines on their desired trajectory to a precision of about 1 cm and the alignment of the collider components within the LEP tunnel to a required short-range relative precision of less than about 0.1 mm on the scale of many kilometres is impressive. The basic reference for this work was provided by a geodesic network between the hills surrounding the site. The base lines of

up to 13 km length were measured with precisions of 10^{-7} by the use of a two-wavelength laser interferometer (Terrameter). Recently these measurements have been verified with excellent agreement by satellite observations using the NAVSAT satellite system. The first precise measurements with beams indicated that the LEP circumference was in fact more than twice as precise as predicted: better than 1 cm in 26.67 km.

Beam elements

The electromagnetic guide field system of LEP consists of dipoles, quadrupoles, sextupoles, horizontal and vertical dipole correctors, rotated quadrupoles, and finally electrostatic dipole deflectors. About 3/4 of the LEP circumference is occupied by 'standard cells'. Each of the eight arcs contains 31 of these standard cells, which are comprised of magnets in the following order:

- a defocusing quadrupole;
- a vertical orbit corrector;
- a group of six bending dipoles;
- a focusing sextupole;
- a focusing quadrupole;
- a horizontal orbit corrector;
- a second group of six bending dipoles;
- a defocusing sextupole.

The length of a standard cell is 79.11 m. The electrons and positrons are bent in a piecewise circular trajectory by the strength of the dipole magnets. As previously indicated, the bending field of these dipoles has been made unusually low (about 0.1 T) so as to increase the bending radius and thereby reduce the amount of synchrotron radiation. The low bending field allowed a novel design of the dipole magnet cores, with the 4 mm gaps between the steel laminations (1.5 mm thick) filled by mortar. Compared with classical steel cores, this produced a cost saving of around 40%. The quadrupole magnets, which produce fields linear to the transverse direction, act as magnetic lenses and focus the beam to be comfortably contained within the vacuum chamber. The alternating polarity of the quadrupoles in the standard cells produces alternating-gradient focusing or 'strong' focusing. The cell sextupoles produce a field which is quadratic in transverse displacement, and they are used to compensate the dependence of the focusing strength on the beam energy ('chromaticity'). The small horizontal and vertical correctors are individually powered so as to allow 'steering' of the beam through the centre of the LEP aperture.

Each experimental collision point in LEP is surrounded by a large solenoid magnet used for particle identification. The bunches of each beam must be tightly focused ('squeezed') to very small dimensions in the centre of these detectors in order to increase the luminosity or particle production rate. This is accomplished by a set of superconducting quadrupoles

with very strong field gradients that focus the transverse beam dimensions to about 10 μm and 150 μm in the vertical and horizontal planes respectively. The solenoid detector magnets produce another effect, since they cause the horizontal oscillations to be 'coupled' into the vertical plane and if this were uncompensated it would greatly increase the vertical beam size and cause a reduction in the luminosity. For this reason, rotated quadrupoles are installed around each solenoid to compensate this magnetic coupling. These quadrupoles are similar to conventional quadrupoles but rotated about their axis by 45°.

For the main dipole and quadrupole supplies, absolute accuracies down to 2 parts in 10^5 have been achieved with a resolution typically three times better. Each magnet has its own cooling circuit. For the majority, the cooling is provided by demineralized water circuits, which are connected to a total of 10 cooling towers with a capacity of 10 MW each. Some of the small correcting magnets are cooled by air, while the superconducting quadrupoles and the superconducting experimental solenoids are cooled by liquid helium at 4.2 K from the cryogenics installation.

The RF acceleration system consists of 128 four-cell copper cavities powered by sixteen 1 MW klystrons via a complex of waveguides. Each accelerating cavity is coupled to a spherical low-loss storage cavity in such a way that the electromagnetic power continuously oscillates between the two sets of cavities. The coupling is arranged so that the power is at its peak in the acceleration cavities at the instant of the passage of the beam bunches. In this way, the bunches receive the maximum possible accelerating gradient, but the power loss due to heating of the copper cavity walls is greatly reduced since the electromagnetic power spends half of its time in the very-low-loss storage cavities. The operating frequency is 352.21 MHz, which corresponds to 31320 times the revolution frequency of a beam circulating in LEP. The sinusoidal electric field that is generated in each accelerating cavity cell produces a 'potential well', inside which each particle of each bunch can perform stable oscillations with respect to the particle at the centre of the bunch, i.e. the synchronous particle. These oscillations are around the energy of the synchronous particle and, in azimuthal distance, ahead of or behind this particle. When represented in a phase-plane plot of normalized energy as a function of normalized distance, the trajectories form circles for the case of small oscillation amplitudes while the maximum stable oscillation possible inside the potential well forms a closed contour which, in accelerator jargon, is called an RF bucket.

The vacuum system

The two main components of the vacuum system are the vacuum chamber itself and the pumping system. Of the 27 km of LEP vacuum chamber, a length of about 22 km passes through the dipole and quadrupole magnets, and is subjected to the heating due to synchrotron radiation. Although this heating represents a mere 100 W/m at LEP, it rises to more than 2000 W/m at LEP2. Therefore the chambers need water-cooling channels and are constructed from aluminium because of its good thermal conductivity.

However, only about half the radiated power would be absorbed by the aluminium and the remainder would normally escape into the tunnel and produce such a high radiation dose that organic materials such as gaskets, cables or electronic components, would rapidly be destroyed. In addition, severe damage could result from the formation of ozone

and nitric oxides, which produce highly corrosive nitric acid in the presence of humid air. For these reasons, the aluminium chamber is covered with a lead cladding of a thickness varying between 3 and 8 mm, which greatly reduces the radiation that escapes into the tunnel during operation. These chambers are interconnected by bolted flanges with aluminium gaskets and stainless-steel bellows, which allow for minor misalignments during installation and thermal expansion during machine operation and high-temperature 'baking' of the chambers. Other types of chambers are used in special regions such as the injection, RF, electrostatic separators, and the detector regions. For the main part these are made of stainless steel except for the detector regions where, for reasons of transparency to particles, they are fabricated from beryllium or carbon-fibre composites.

In previous electron storage rings, the ultra high vacuum was normally produced by linear sputter-ion pumps operating in the field of the bending magnets. However, this proved impossible in LEP since the bending field strength is below the threshold for efficient operation of such pumps. Consequently, a novel type of ultra high-vacuum pumping system was required for the LEP storage ring. The solution adopted, for the first time in an accelerator, was the use of non-evaporating getter (NEG) strips, installed in pumping channels running parallel to the beam channel with pumping holes between the two. The NEG strip is 3 cm wide, and extends over 22 km. It is fabricated by coating constantan with a zirconium-aluminium alloy. The NEG material forms stable chemical compounds with the majority of the active gases and consequently the residual gas molecules inside the pumping channel simply 'stick' to the NEG ribbon. During long periods of pumping, the getter surface becomes progressively saturated and loses some of its pumping capacity. An essential operation is therefore reconditioning, which consists of heating the getter up to 400° C for about 15 minutes and results in the diffusion of O_2 and N_2 from the saturated surface layer into the bulk of the material. The very low static pressure of less than 10^{-11} torr requires initially very clean internal surfaces. This was achieved by careful chemical cleaning of all chambers, followed by storage under chemically inert conditions. After installation in the tunnel all chambers were 'baked out' at 150° C for 24 hours. The bake outs were performed by pumping super heated water at 150° C and at a pressure of $5 \cdot 10^5$ Pa into the cooling channels of the aluminium-lead chambers. Sputter-ion pumps, valves, gauges, all stainless steel chambers, and special equipment such as the electrostatic separators and the feedback system, were baked out by electrical heating elements and jackets.

Beam diagnostics

The LEP beam instrumentation system is used to observe the position, shape, or other relevant properties (such as polarization or electrical current) of the beam. The beam electrical current is measured in LEP as in other accelerators by current transformers placed around the vacuum chamber. These transformers are capable of measuring the current of a single injection or of a steady circulating beam. In the latter case the beam lifetime can be evaluated by accurate measurement of the current as a function of time. In order to position the beam accurately in the middle of the vacuum aperture, it is essential to measure the transverse beam positions at many azimuthal locations on the circumference.

In the case of LEP this 'closed orbit' is measured by 504 monitors fairly evenly distributed around the circumference. Each monitor consists of four electrostatic pick-up 'buttons' and these are positioned in housings, which are connected directly to the end faces of the quadrupole magnets. The electromagnetic field of the bunches induces voltages on each of the four buttons and when properly analysed, these can give an accurate measurement of the horizontal and vertical beam positions relative to the centre of the monitor. This system has been designed to be capable of measuring the positions of all eight individual bunches during more than 1,000 revolutions to an accuracy of better than a millimetre.

The betatron-tune value is defined as the number of transverse oscillations, around the closed orbit, made by the beam per revolution. The measurement and correction of this parameter is of paramount importance for the stability of a beam in a storage ring. Owing to non-linear resonances driven by magnetic imperfections in the guide field, there are many undesirable values for the horizontal and vertical betatron tunes as well as undesirable combinations of the two. In LEP the betatron-tune values (or the Q values) are measured by two Q meters, one for each plane. The Q-meter system consists of a magnetic beam 'shaker', which excites the beam, and of dedicated electrostatic pickups, which measure the phase and amplitude of the induced transverse oscillations as a function of the excitation frequency. Since the beam acts as a resonator of high quality, the amplitude of the response is at a maximum when the excitation frequency equals the natural resonant frequency of the beam, which gives the betatron-tune value.

Synchrotron radiation

As previously described, when charged particles are bent in a circular trajectory they radiate photons. Consequently, the beams can be 'seen' by measuring this flux in the UV frequency range. Four UV monitors are used in LEP to measure the transverse dimensions of both beams at two different locations. The UV range is preferred because it produces a sharper image on the digital TV camera. The images are transmitted to the control room to give a real-time view of the beam, while the digital signals are processed to provide numerical values for the beam sizes. The UV synchrotron radiation monitors cannot give the absolute beam dimensions, instead for calibration a *wire scanner monitor* is used, which is a 37 μm diameter carbon fibre that traverses the beam at 0.5 m/s and thereby creates photons which are detected outside the vacuum chamber. The beam profile is given by the density of the detected photons, plotted as a function of the position of the flying wire, which can be measured to an accuracy of 10 μm. The synchrotron light signal can also be used to measure the length of the bunch with a precision of a picosecond.

The synchrotron radiation results in another problem since it gives a background originating from the high energy spectrum of the photon emissions. In order to reduce this background, collimators are installed around each experimental point. Each of the collimators consists of movable jaws of tungsten and copper, which can intercept and absorb the high-energy photons very close to the beam. The system of collimators has proved invaluable in LEP and has resulted in low background conditions in the detectors.

Colliding beams

Under certain circumstances it is essential that the beams of electrons and positrons do not collide. This is particularly true at injection energy, where the electromagnetic fields associated with each bunch would destroy the opposing bunch long before a sufficient number of particles could be accumulated. If electrons and positrons, travelling in opposite directions are subjected to the same transverse electrostatic field, they will be displaced in opposite directions, thereby avoiding collisions and greatly reducing the beam-beam effect. In LEP this has been achieved by equipping each of the eight possible collision points (four bunches could make eight collision points) with four electrostatic separators, each of which is 4 m long and produces a vertical electric field of 2.5 MV/m between the plates, which are separated by 11 cm. This produces a separation between the bunches of electrons and positrons of more than 40 standard deviations of the vertical beam size. The separators are powered in all eight possible collision points during injection, accumulation, and energy ramping. Some time before physics data taking starts the separators in the experimental points are switched off to allow collisions. At higher energies in LEP however, the bunches may not necessarily collide perfectly 'head-on', even if the separators are off. For this reason the separators have been equipped with a vernier adjustment, which allows vertical steering of the beam positions in the collision points.

The duration of a typical operation to fill LEP with particles for a physics run is 12 hours. During this time each of the 10^{12} particles in the beams will have traversed the complete 26.67 km of the LEP vacuum chamber about 500 million times. In order to minimize particle losses due to collisions with residual gas molecules, the whole vacuum chamber must be pumped down to very low pressures. The achieved static pressure for LEP is $8 \cdot 10^{-10}$ torr whereas in the presence of beam the pressure rises to about 10^{-9} torr. This pressure rise is due to gas desorption from the inner vacuum-chamber wall, provoked by the synchrotron radiation of the circulating beam, and has had a profound influence on the design of the LEP vacuum system.

Controlling LEP

Almost every single LEP component and piece of equipment must be remotely controllable from the main control room by means of the LEP computer control system, which consists of more than 160 computers and microprocessors distributed over 24 underground areas and 24 surface buildings. Communication between the computers and microprocessors is provided by a data network and a synchronization timing system. Many of the design choices for this system have been dictated by the size and topology of the project. For example, the prohibitive cost of laying many dedicated cables around the 27 km circumference led to the decision to replace cables by a Time -Division Multiplex (TDM) system, which allows many communication channels on a single cable. The network is composed of two logical levels. The upper level consists of the central consoles and servers, which are situated in the control room, and the lower level consists of local consoles and process computers. This network is based on the Token Passing Ring principle using TCP/IP as the high level protocol. The synchronization of control and data taking is a very important aspect of accelerator control. Two important examples of this are the energy

ramping process, where all 750 power converters and the RF system must be controlled in perfect synchronization and the acquisition, on the same machine revolution, of the closed-orbit data coming from the electrostatic pick-ups.

3.2.2 LEP and LEP2

From August 1989 till the end of 1995, LEP provided each LEP experiment with about 6 million Z^0-decays. This gave about 4.3 million multihadronic Z^0-decays with an average of 21 charged tracks and 21 photons per event. Figure 3.3 shows the integrated luminosity during the years 1992–1995, which summed over the whole period 1989–1995 gave a total integrated luminosity of 170 pb^{-1}.

Since the end of 1995, LEP has moved on from the Z^0 and entered its second phase. Its energy is being doubled to allow the study of the W^+, W^- pairs, the charged counterparts of the Z^0, thus opening a new domain of investigations and tests of the standard model, the search for new particles, particularly the Higgs boson and/or super symmetric particles. The first pair of W particles, produced by LEP2, was detected by the DELPHI experiment, on July 10, 1996, see figure 3.4.

The steps in energy taken so far are:

- 130-136 GeV operation in November 1995 (6 pb^{-1} [3] ; of data taken)
- 161 GeV operation in July and August 1996 (10 pb^{-1} of data taken);
- 172 GeV operation in October and November 1996 (10 pb^{-1} of data taken);
- 183 GeV in 1997 (54 pb^{-1} of data taken);
- 188 GeV in 1998 (150.5 pb^{-1} of data taken);
- 200 GeV is scheduled for the year 1999.

To provide these higher energies, 288 super conducting cavities with a designed gradient of 6 MV/m, are to be installed. A drawing of one of the superconducting cavities can be seen in figure 3.5 and a picture of the same cavity in figure 3.6. As mentioned before the energy loss due to the synchrotron radiation, at LEP2 is about 2.0 GeV/turn. If one wants to increase the energy even more, this value will increase rapidly with energy and brings a limit to how high in energy it is meaningful to go. For 1999/2000 additional cooling will be provided to push the LEP's cavities to 7 MV/m (more than 16% that they were designed for) which will allow the energy of the colliding beams to approach, and maybe go beyond, 200 GeV.

3.3 The DELPHI Detector

DELPHI is a DEtector with Lepton Photon and Hadron Identification. It is one of four main detectors at LEP. DELPHI was designed as a general purpose detector with emphasis

[4]The unit pb^{-1} is a way of expressing the integrated luminosity, $\mathcal{L}$, which is $\int_t L \cdot dt$ where t is the accumulated time of the measurement of data taking.

on particle identification. The general features of the detector can be seen in figure 3.7. DELPHI is a large structure consisting of several detectors and each of them are specified to perform a special task, such as energy measurement, particle trajectory measurements or being a trigger for other measurements. Some of the detectors are specialized to detect a certain kind of particles while the identification of others rely on the combined results of more than one detector.

DELPHI consists of one cylindrical section and two end-caps, with an overall length of 10 meters and a diameter of the same size. The total weight is around 3500 tons.

The collisions between electrons and positrons, going in opposite directions, take place inside the beampipe, which is made of beryllium-carbon inside the detector. The products from this annihilation fly almost isotropically, and if they are charged they will be guided by the strong magnetic field and their trajectories will be bent so their momentum can be measured. The magnetic field of 1.23 T is produced by a huge super conducting solenoid (once the largest in the world). DELPHI has high precision and 'granularity' over almost the full solid angle and it has the specific ability, using the Ring Imaging CHerenkov technique, to differentiate between all the various secondary charged particles. Design and construction of the DELPHI detector took some seven years. The present collaboration consists of about 550 physicists from 56 participating universities and institutes in 22 countries. Data have been taken for the last 8 years.

In the following, the subdetectors that are interesting for the following analysis are presented in some detail. For a more complete treatment of all the subdetectors in DELPHI, see [13].

3.3.1 Tracking detectors

There are several *tracking* detectors throughout DELPHI that will provide measurements of the coordinates of *charged* particles in the $r\phi$-plane, transverse to the beam direction, as well as the z-coordinate, along the beam direction. These are the vertex detector (VD), the inner detector (ID), the time projection chamber (TPC) and the outer detector (OD).

The vertex detector, VD

The vertex detector consists of three concentric and overlapping layers of silicon microstrip detectors which allows a very precise tracking, principally in order to detect very short lived, charmed and beauty particles, by extrapolating the tracks back towards the interaction point. The precision in the $r\phi$-plane is about 8 μm. The drawing of the vertex detector is seen in figure 3.8 and a picture of the same detector is seen in figure 3.9.

The inner detector, ID

The inner detector is a cylindrical jet chamber providing 24 coordinates in the $r\phi$-plane. The outer cylinder consist of five layers of multiwire proportional chambers (MWPC), each equipped with 192 anode wires parallel to the beam axis and 192 cathode strips forming full circles at constant z thus giving both the z and the $r\phi$-coordinates. The resolution is of the order of 60 μm in $r\phi$ and about 1.5 mrad in the ϕ-angle.

The time projection chamber, TPC

The TPC detector is the principal tracking device for momentum measurements and works as a drift chamber. In an ordinary proportional chamber anode wires are stringed up between two cathode planes. The space is filled with a gas that is ionized when a high energetic incident particle enters the tube. The electrons from the ionization drifts towards the anode, and in the high field close to the anode wires they ionize a large number of atoms leading to an avalanche of electrons. The number of electrons in the avalanche is proportional to the number of primary ion pairs. Measurement of the time it takes for the primary ionization to drift to the wires where the avalanche is created gives a precise measurement of the location of the particle. This requires a reference trigger in time in order to know when the particle entered the chamber. The TPC is of a somewhat different type, see figure 3.12. The electric field is in the same direction as the primary beam axis, and the anode wires are placed in the two end caps. These are divided into 6 sectors each with 16 circular pad rows in the cathode plane to give the coordinate along the wire. A sector is seen in figure 3.11.

In this way the TPC provides a three dimensional picture of the ionization pattern of the primary charged particle. In the TPC the electrons have drift distances of over one meter which means that the bunches of electrons would have long time to spread out by diffusion. This would be true if it was not for the magnetic field that is aligned with the electric field, holding the electrons together in curly-paths. The first meter of the track (except for the very first part covered by the ID and the VD) can in this way be detected and described by means of 16 points in the endcap of the TPC. Combined with the drift time, that gives the position along the beam axis, this gives a full three-dimensional information of the track, see figure 3.12. The TPC also measure dE/dx, the energy loss of the particle by pulse height measurements which combined with the particle momentum, given by the bending of the track, can be used to calculate the particle mass.

The TPC is a cylinder of 2x1.3 m, see figure 3.12, situated between the radii 0.29 m and 1.22 m. The two drift volumes are separated by a HT (20 kV) plate producing an electric field of 15 kV/m. A charged particle crossing the TPC produces by ionisation around 70 electrons per cm of gaz (80% Ar 20% CH_4). Under the action of the electric field these primary electrons drift in the direction of the proportional chambers (6 at each end of the TPC). Each of 12 TPC sectors have 16 rows of pads, allowing the reconstruction of 16 space points per track (for a non looping track). In each sectors there is 1680 pads. In front of the pads plates there is 3 grid levels and the one closest to the pad (the anode at a HV of 1430 V) contains 192 sensitive wires, separated by 4 mm per sector, performing the dE/dx (ionisation) measurement. There are 22464 electronic channels of the TPC, 2304 for the wires and 20160 for the pads. The precision of the measurement of the $r\phi$ coordinate is 250 μm and the accuracy of the z measurement is 900 μm.

The outer detector, OD

The outer detector is completing the tracking in the barrel region. It is composed of 24 planks of 5 layers x 32 columns of 4.7 meters long drift tubes operating in the limited streamer mode (section 1.65x1.65 cm^2). The electrons produced by a charged particle crossing a tube drift toward the wire (HT of 4.4 kV) and give a signal with short rise time

and high amplitude (typically 3 ns and 80 mV). Situated at 2 meters from the beam axis, it improves the precision of the momenta of charged particles measured by the TPC.

The precision over the measure of the $r\phi$ coordinate is 100 μm per track and the accuracy of the z-measurement is of the order of 4.4 cm.

3.3.2 Calorimeters

Calorimeters are used to measure the energy and the position of of the particles by total absorption. The calorimeter is different from most other detectors in that it can detect both neutral and charged particles. The neutrals are detected by detection of the secondary charged particles they produce when they are observed. Calorimeters are basically of two sorts, electromagnetic and hadronic. At Delphi the electromagnetic calorimeter is accomplished in the barrel region by the *high density projection chamber.*

The high density projection chamber, HPC

The high density projection chamber is the barrel electromagnetic calorimeter of 18 radiation lengths with a dynamical range from MIPs to a 50 GeV shower and it is installed as a cylindrical layer outside the outer detector and mounted on the inside of the solenoid. It consists of 144 independent modules, arranged in 6 rings of 24 modules each, and covers the polar angles θ from 40° to 140°. Each HPC module is a box filled with 41 layers of lead separated by 8 mm gas gaps (filled with 80% Ar and 20% methane). It is a gas sampling calorimeter and uses the time projection principle to measure the three dimensional distribution of charges induced by electromagnetic showers in the lead and the ionization of the gas. The charges drift to one end, where they are collected by a proportional chamber with pad readout (like in a TPC). The granularity of the HPC is 1° in ϕ and 4 mm in z.

The hadron calorimeter HCAL

The hadron calorimeter is a sampling gas detector incorporated in the magnet yoke, the barrel part covering polar angles between $42.6° - 137.4°$ and two end-caps between polar angles of $11.2° - 48.5°$ and $131.5° - 168.8°$. The barrel, is constructed of 24 modules with 20 layers of limited streamer tubes inserted into 18 mm slots between 50 mm iron plates in each module. The modularity of the end-caps is similar to the barrel with a sampling depth of 19 layers.

The detectors are wire chambers which consist of a plastic cathode forming 8 cells of 0.9 cm x 0.9 cm with one anode wire (80μm Cu-Be) in each. The inner surface of the cathode cells is coated with a poorly conductive graphite varnish. The copper clad readout boards are segmented into pads which pick up the streamer charges. Pads are shaped to form projective towers pointing to the intersection point. In the barrel 5 adjacent pads in the radial direction are combined into a tower. Each tower covers an angular region of $\Delta\phi = 3.75°$ and $\Delta\theta = 2.96°$ in the barrel, and $\Delta\theta = 2.62°$ in the end-caps. The dimensions of a typical tower in the barrel are 25x25x35 cm and there are 19008 towers in the hadron calorimeter.

The hadron calorimeter contains by far the largest volume of gas in DELPHI, and operates with relatively low iso-butane content: Ar/CO_2/iso-butane 10/60/30%. With

this gas mixture an average charge of 10 pC is induced on external readout boards with an HV of 4kV.

3.4 DAS, the data acquisition system

All DELPHI data are written in a format which allows the members of the collaboration to transport data from one computing architecture to another without having to convert the data from one format to another. The DELPHI data sets produced by the program DSTANA for use in the analysis is called short DST data, which is a compact data set, allowing faster analysis and it is mainly used by the hadronic analysis teams in DELPHI.

3.4.1 DSTANA

DSTANA stands for DST ANAlysis and is the main DELPHI software program for raw data event processing of *real* events collected by the DELPHI online system or *simulated* events produced by DELSIM. It is the code used for the short DST production and also the code needed to perform a physical analysis. The DSTANA library includes, for example, the following packages:

- track extrapolation, track refit and vertex fit;
- jet reconstruction packages;
- tagging of B-events and B-jets.

3.4.2 DELANA

The main DELPHI software for event reconstruction.

3.4.3 DELSIM

The simulated data were generated by a Monte Carlo program with full DELPHI detector response simulation called DELSIM which is the standard DELPHI simulation software package into which different event generator can be plugged. Each detector group provides its individual code to coexist with the entire DELPHI Detector set up. The DELPHI Detector Description and Calibration Database is needed to track through detector modules and to guarantee that events are reconstructed like real events. The output of the raw data structure from simulation events can be either with a perfect detector or with run conditions as for real data.

3.4.4 SKELANA

There exists a user friendly analysis program called *SKELANA*, [14]. SKELANA is a program for physics analysis. When working in the SKELANA frame work the DSTANA software can be considered as a black box with the physics information being available without having to take care of the technical details. SKELANA can run part of the

DSTANA software (e. g. to perform particle identification or btagging). The interface between the user and SKELANA is done through flags asking SKELANA to provide a given kind of information and SKELANA gives access to the data through the flags set by the user. The program reads the data and tracks and fills the *common blocks* which contains the required physics information thereby making them accessible for the analysis. The results of the analysis are typically stored in a HBOOK [15] file and visualised in *histograms*, accessed through the program Physics Analysis Workshop, or *PAW* [16]. To make a further analysis of the data, the results are sometimes stored in *ntuples*. The ntuple is an array of several variables, making it possible to handle the data in a very interactive way by plotting various histograms with different cuts on the chosen variables and much of the analysis is done in this way.

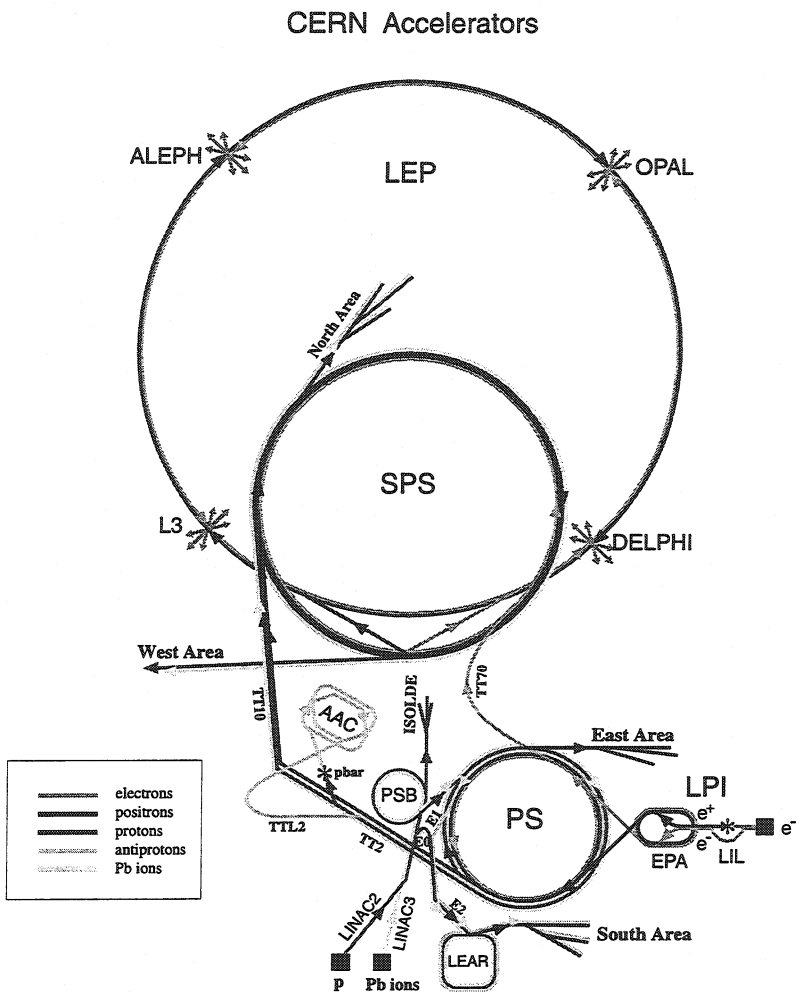

Figure 3.1: The different accelerators at CERN.

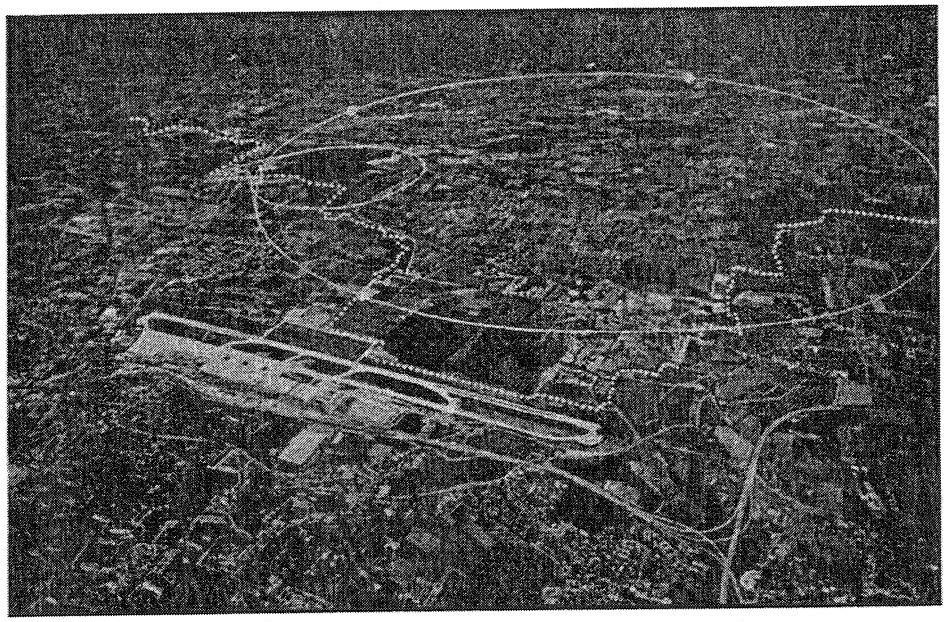

Figure 3.2: The LEP accelerator at CERN in an aerial view.

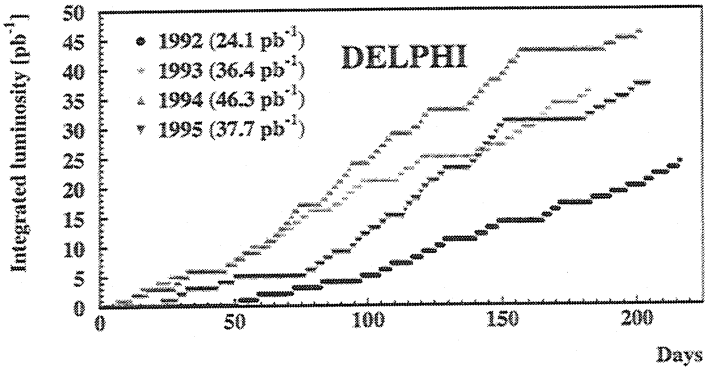

Figure 3.3: The integrated luminosity during 1992–1995.

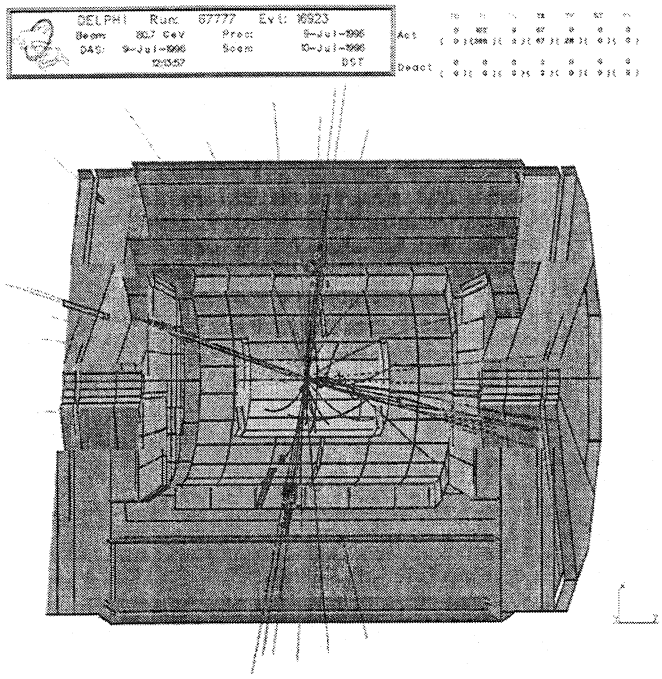

Figure 3.4: The first pair of W^+, W^- particles, produced by LEP2, was detected by the DELPHI experiment on July 10, 1996.

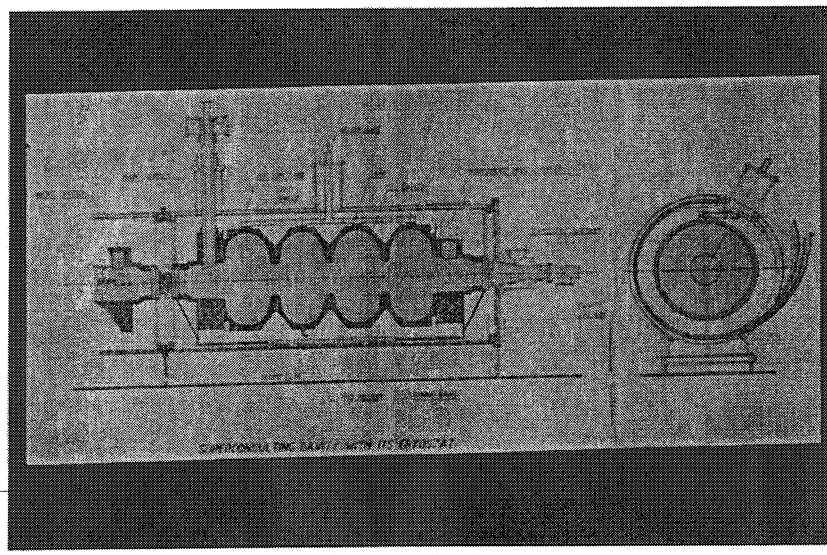

Figure 3.5: A drawing of one of the superconducting cavities.

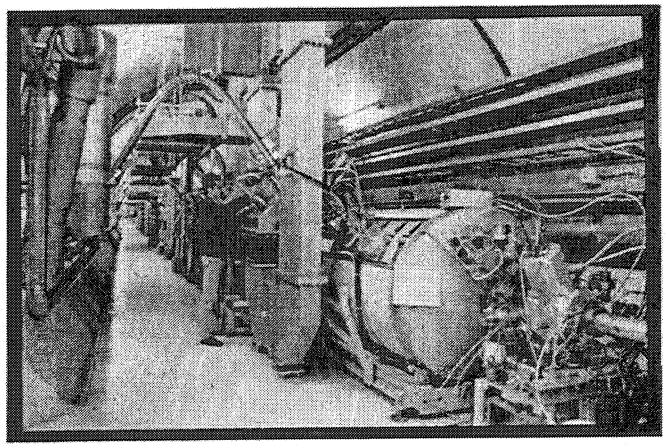

Figure 3.6: A cavity for LEP2 in preparation.

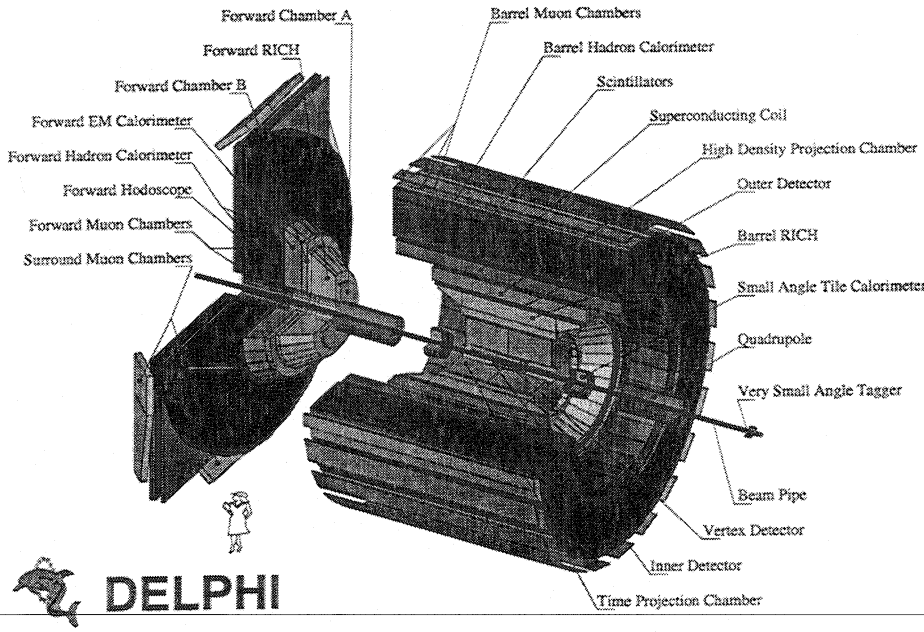

Figure 3.7: The Delphi detector.

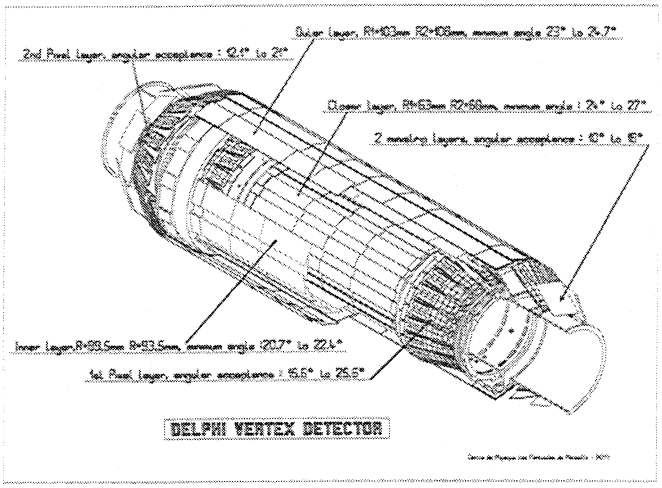

Figure 3.8: A drawing of the vertex detector.

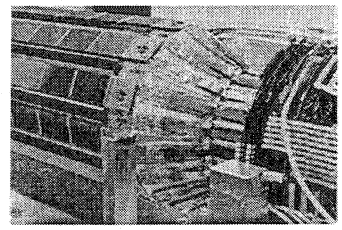

Figure 3.9: The vertex detector.

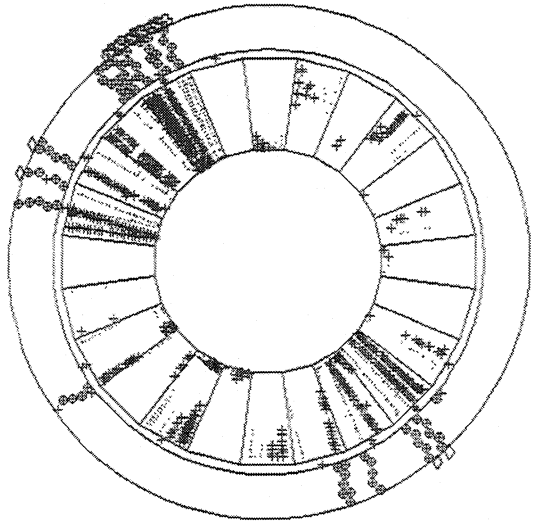

Figure 3.10: An event seen by the inner detector.

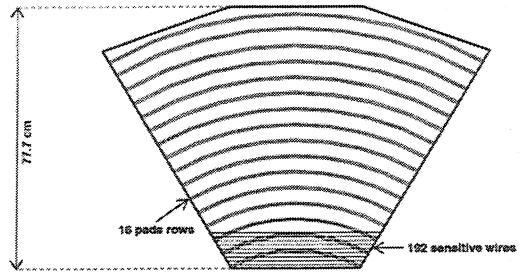

Figure 3.11: A TPC sector with some of the 192 sense wires indicated.

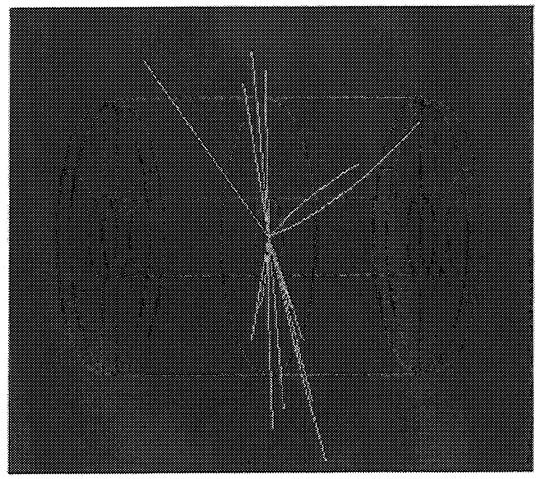

Figure 3.12: The TPC detector.

Chapter 4

Multi-variable discriminant analysis

Multi-variable data analysis has been used in high energy physics for more then twenty years. The principle is to analyse the events in the multi-space, spanned by the relevant observables, and to maximize the use of the given information. In the multi-space, that may be of a large dimension, the hope is that events coming from different elementary processes will populate different regions in space with none or little overlap between them. This separation clearly depends on the choice of the set of variables describing the space. A linear method, Fisher's Linear Discriminant, [17], is used to project d-dimensional data onto an appropriate line, thereby reducing the data to a single measurement. This is one version of a problem in multiple discriminant analysis where data are projected from R^d into R^q, where typically $q << d$. After the discussion about linear methods, non-linear methods, using artificial neural network (ANN) are presented with examples taken from [18], since ANNs are suitable in analysis of specific physics processes like event classification which depend upon a very large number of variables. These discrimination analyses will be used later for an application on the gluon/quark classification problem.

4.1 Linear discrimination analysis

An often used linear discrimination analysis method is called the Fisher statistical discrimination method and it is used in natural and social sciences. Originally it was used on taxonomic problems of statistical nature. The technique is illustrated with a modified example from [19]. Consider a scatter plot of the variable x_1, considered to be the signal and the variable x_2, considered to be the background. A hypothetical distribution is shown in figure 4.1. The individual distributions, x_1 and x_2 overlap and do not provide a very efficient discrimination. However, a variable $F(x_1, x_2)$ can be constructed through a linear combination, $F(x_1, x_2) = f_1 \cdot x_1 + f_2 \cdot x_2$. It is evident from the figure that the two parameters f_1 and f_2 can be determined in such a way that the two classes (signal and background) are entirely separated, in this ideal situation, by a cut on $F(x_1, x_2)$. The scale of the variables is very important. For instance by compressing the axis for of one variable (i.e. x_1) the distribution of the two variables will approach one another, thus losing discrimination power. The best result is obtained by using variables renormalized

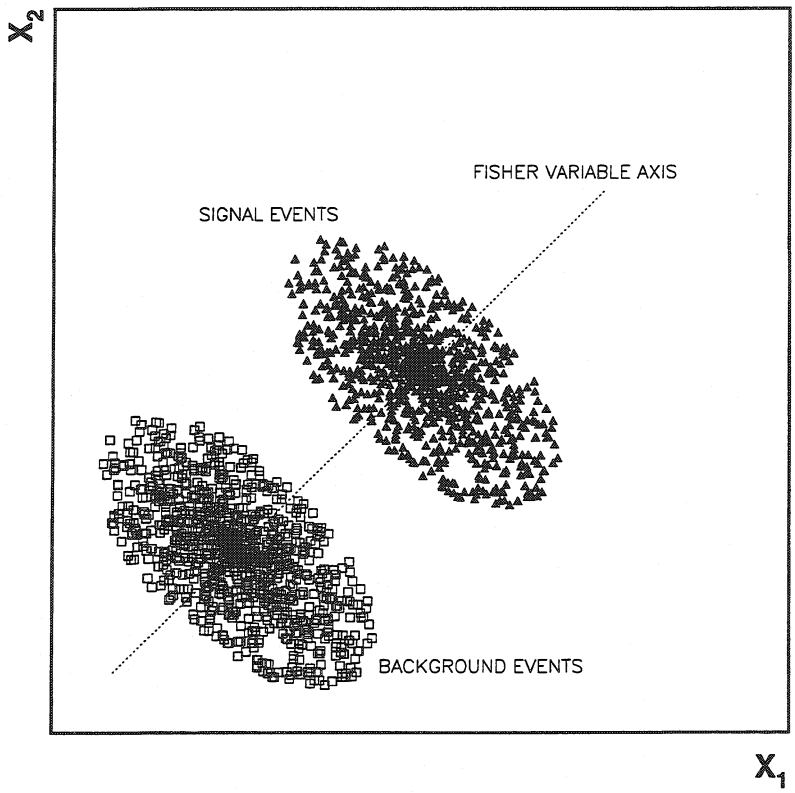

Figure 4.1: An example of signal/background classification using Fisher statistical discrimination. The two variables x_1 and x_2 do not individually provide a good separation between the two classes, whereas in the Fisher variable ($F(x_1, x_2) = f_1 \cdot x_1 + f_2 \cdot x_2$) the classes are entirely separated.

to their variances, σ_1, for the variable x_1,

$$\sigma_1 = \sqrt{< x_1^2 > - < x_1 >^2}$$

and similarly for x_2. The described two-dimensional linear transformation can be showed to be valid in a more rigorous mathematical approach and this is treated in appendix A and it can be generalised to an arbitrary number of dimensions. The variable $F(\vec{x}) = F(x_1, x_2, ..., x_n)$ is called the Fisher variable. A systematic procedure to calculate the optimal values of the parameters $f_1, f_2, ..., f_n$ is provided by the Fisher method.

The Fisher approach is based on the *projection* of d-dimensional data onto a line. The object of the method is to find the direction of this line to make these projections well separated by *class*. The output of the analysis is a probability that an event belongs to a class and the probability is calculated using a measure between the events and the average value of the class. Generalising the method outlined in the sections above and realising that the problem has only two classes, the following discrimination variables can be defined in the R^n space:

- x_i^{signal}, is the *signal* variable;
- x_i^{bg}, is the *background* variable.

Where the choice of the set, x_i^{signal}, $i = 1, n$, is one of the most crucial decisions to make for the analysis.

The covariant matrix A_{ij}^{signal} for the signal class, is defined as

$$A_{ij}^{signal} = 1/N_{signal} \cdot \sum^{N_{signal}} (x_i^{signal} - < x_i >^{signal}) \cdot (x_j^{signal} - < x_j >^{signal})$$

where:

- N_{signal} is the number of signal events;
- $< x_i >^{signal}$ is the average value of the signal variable.

The covariant matrix A_{ij}^{bg} for the background class, is created in a similar way.

$$A_{ij}^{bg} = 1/N_{bg} \cdot \sum^{N_{bg}} (x_i^{bg} - < x_i >^{bg}) \cdot (x_j^{bg} - < x_j >^{bg})$$

where:

- N_{bg} is the number of background events;
- $< x_i >^{bg}$ is the average value of the background variable.

The matrix A_{ij} is constructed with the help the matrices A_{ij}^{signal} and A_{ij}^{bg}, defined above

$$A_{ij} = \frac{A_{ij}^{signal} + A_{ij}^{bg}}{2}$$

The inverse of the matrix A_{ij} is written A_{ij}^{-1} and it fulfils the matrix multiplication rules,

$$A_{ij} \cdot A_{ij}^{-1} = A_{ij}^{-1} \cdot A_{ij} = I$$

where I is the *identity* matrix. Following the scheme, the inverse of the matrix A_{ij} is needed in order to form the Fisher coefficients.

The Fisher coefficients, f_i, are created in the following way

$$f_i = \sum_{j=1}^{N} A_{ij}^{-1} \cdot (x_j^{signal} - x_j^{bg})$$

where:

- i is the number of the discriminating variable, i = 1,2, ...,n, where n is the dimension of the multi-variable space, R^n;
- N is the number of events.

The last coefficient f_0 (the offset from the origin) is defined as

$$f_0 = -\frac{1}{2} \sum_{j=1}^{n} f_j \cdot (x_j^{signal} - x_j^{bg})$$

The Fisher variable is formed by

$$F(\vec{x}) = f_0 + f_1 \cdot x_1 + f_2 \cdot x_2 + ... + f_n \cdot x_n$$

where finally, the the projection of n-dimensional data has been reduced to a scalar number $F(\vec{x})$.

4.2 Non-linear discrimination analysis

For many centuries, man has developed machines to perform all cumbersome and tedious tasks so that we can enjoy a better life. The era of machine started with the discovery of simple machines such as the wheel and today scientists and engineers are trying to develop intelligent machines. Artificial Neural Networks are examples of such machines having a great potential to improve in suitable areas of problems. The human brain is, for example, excellent in fast recognising a photograph as one describing Albert Einstein but usually poorer in adding up the square roots of a million numbers. So only when very simple arithmetic operations are required, will the computer out match the human brain. The human brain is composed of some 20 billion nerve cells (there is in fact a gender[1] difference) called *neurons*, with about 10^6 inputs and around 10^{14} connections.

[1]Danish researchers say that they have found that men, on average, have about 4 billion more brain cells than women. But they have not figured out what men do with them. A Copenhagen Municipal Hospital neurologist (Bente Pakkenberg) leader of the research project, has told Danish radio that the conclusions came from an examination of the brains in 94 cadavers of people aged 20 to 90. The average number of brain cells in males was 23 billion, while the female average was about 19 billion. Asked what the males might be doing with the surplus, Pakkenberg said: "Right now it is a mystery. The knowledge we already have shows men are not smarter than women." Copenhagen, Denmark (AP).

The first mathematical model of a human brain neuron was introduced by McCulloch and Pitts, in 1943, see figure 4.2. This Artificial Neuron is a non-linear threshold device with many degrees of freedom in the input. The transfer (threshold) function, $g(\sigma)$, see table 4.1, is a step-function, and θ is the threshold value. First, the weighted input of the signals are summed in Σ and the sum is passed on to the threshold function (activity function) g that will give an output if the sum is greater than the threshold value θ otherwise will give zero output.

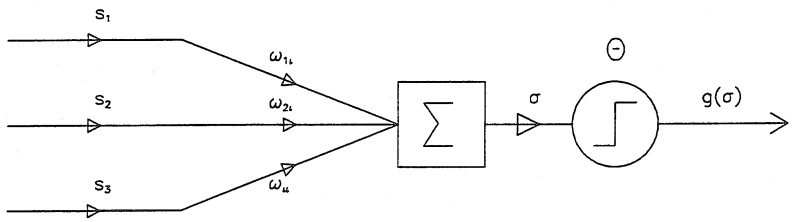

Figure 4.2: The McCulloch-Pitts neuron.

$$\sigma = \sum_i s_i \cdot \omega_i$$

Even if this neuron model is very simple it has a substantial potential and if it is provided

$g(\sigma)$	
$\sigma < \theta$	$\sigma > \theta$
0	1

Table 4.1: The threshold function, $g(\sigma)$.

with appropriate weights and thresholds it can perform basic logic operations.

4.2.1 Artificial Neural Network models

Artificial Neural Networks (ANNs) are mathematical models that are inspired by the functioning of the neurons and their connections in biological systems. They have given rise to a branch of research called neural computing, being used or tried out in many disciplines. The basic concept is based on two simple concepts:

- the topology of nodes and connections between them;
- and the transfer functions relating input and output of each node.

An ANN consists of neurons and connectors. To the connectors are associated weights which constitute the internal representation (the memory) of the ANN. The task of a node is to map nonlinearly (using a transfer/activation function), the weighted total output from the preceding layer and output the result to the neurons in the next layer.

An optimization procedure, based on the minimization of an error function will change the weights.

A node receives input data through its input connections, performs a very simple operation on these (doing a weighted sum and then using a threshold function), then passes the result on to its output connections, to be used in the next coming nodes or as a final output.

If the threshold function is a step function, a feed-forward network of more than one layer performs multiple piecewise linear transformations and the decision boundaries will be hyperplanes. However, for more complex threshold (transfer/activation)-functions, *sigmoid functions*, the interpretation is more complicated.

ANNs are often used as a way of optimizing a classification procedure or pattern recognition. ANNs also usually have more input than output nodes and they may in that way be considered to perform a dimensionality reduction on input data, in a more general way than principal component analysis. A further possibility of interpretation of the outputs of the ANNs is that of probabilities.

4.2.2 Feed-Forward Neural Networks

This class of ANN have their units organized in layers with the following characteristics:

- a hierarchical scheme;
- no connections between the nodes belonging to the same layer. The nodes can only be activated with a node from an immediate subsequent layer;
- an output layer can have more than one node;
- the activating function is continuously differentiable (typically a sigmoid);
- the learning is supervised.

To understand how it works one can start with the simplest Feed-Forward Neural Networks, *the perceptron*. The perceptron, was constructed in 1969, and can perform logic operations like AND, NOT and OR. This perceptron in figure 4.3 acts as an AND function

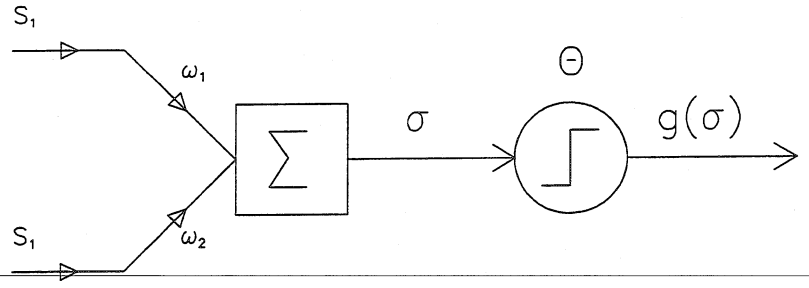

Figure 4.3: A simple perceptron.

with (for example) the following settings:

- $\omega_1 = 3/4$;
- $\omega_2 = 3/4$;
- $\theta = 1$.

However, the XOR function was not possible to construct with a perceptron and for quite some time artificial neural network were out. The renaissance for ANN came in the end of the 1980s, when help was to be found in the *hidden* layers forming a multi-layer perceptron, MLP.

The XOR function, that was impossible to construct with a single perceptron, is easily constructed with the help of the simplest possible hidden layer, see for example figure 4.4, with the following settings:

- $\omega_1 = 2$;
- $\omega_2 = 3/4$;
- $\omega_3 = 3/4$;
- $\omega_4 = 2$;
- $\omega_5 = -4$;
- $\theta = 1$.

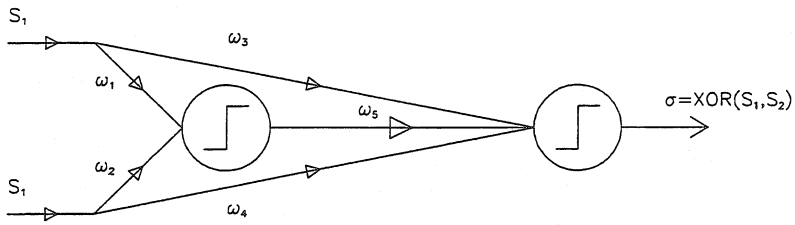

Figure 4.4: An XOR with one hidden "layer" (only one node).

MLPs with one hidden layer behaves essentially like a conventional discriminant analysis method, which projects the data on a subspace in R^n so as to optimise some criterion. It can be showed that a MLP with one hidden layer and linear units trained with a quadratic cost function, performs a discriminant analysis in its first layer of weight, having provided the number of hidden units equals the dimensions of the discriminant analysis space, so one can not do better than a MLP in such a case. A typical ANN, with one hidden layer, is shown in figure 4.5.

The learning rule mostly used for a MLP is called the back-propagating algorithm and it was suggested in 1986 [20]. From an arbitrary starting point, this rule allows one to find a set of weights via the following iterative process:

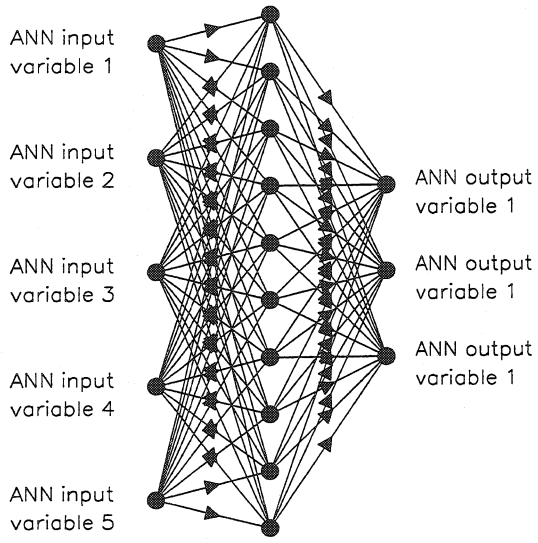

Figure 4.5: An ANN with five input nodes, one hidden layer with ten nodes and three output nodes.

- calculate the error between the desired output and the actual output from a node in the output layer;
- change each weight belonging to that node in proportion with the partial derivative of the error for that node in respect to that weight;
- use the updated weights in the output layer to update the weights in the layer closest to the output layer;
- continue to update the weights, layer by layer, going backwards through the net.

In the chosen structure, the node makes a weighted sum of the output values from all nodes in the previous layer using the sigmoid function as a temperature T. The output o_i is therefore

$$o_i = g_T(\sum_j \omega_{ij} o_j)$$

where the sum is taken over all nodes in the previous layer. In the training procedure the values of the weights are fixed by looking at the difference in the error function E

$$E = \frac{1}{2}\sum^{i}(o_i - x_i)^2$$

between the obtained output calculated by the neural network x_i from neuron i and the desired (true) target value o_i from output neuron i for the same input pattern.

The function is minimized by updating the weights by an amount computed from the error function by the gradient descent method [20].

$$(\Delta\omega_{ij})_{new} = \alpha \cdot (\Delta\omega_{ij})_{old} - \eta \cdot \frac{\partial E}{\partial \omega_{ij}}$$

where the gradient is used to update the changes of the weights in order to reach a minimum for E. The method is controlled by:

- the "learning strength parameter" η;
- the "momentum" α.

The artificial neural network program, JETNET [21], has been used throughout this analysis. The Lund JETNET program version 3.0, is a Fortran implementation of the standard gradient descent back-propagating algorithm for the updating of the weights as described earlier, which was used for the neural network simulations. More details about JETNET and the meaning of the parameters η and α can be found in [21].

4.2.3 Implementation or the learning phase

Many classification and pattern recognition problems can be expressed in terms of ANNs. One of their important properties is that they can be trained, i.e. they can be given training samples of events of different classes, and by learning algorithms, adjust the weights associated to all input connections until some overall function is maximized which characterizes the quality of the decision mechanism.

Impressive results can be achieved on classification problems of a small size, for instance, character recognition, where ANNs can learn to a good performance level without more input than the training samples. An optimization of the choice of input data and of the network topology is usually left to trial and error. Since the input data should ideally describe the event exhaustively, this might be interpreted as to use all possible variables as input as long as they have any relevance to the problem. However, too large and sometimes inadequate neural networks can be avoided by trying to reduce and transform the variables of the training sample into a fewer number or into new variables, before submitting them to a ANN training algorithm.

A trial-and-error approach is often used as the initial choice of weights in order to be able to start the learning process. The solidity is demonstrated by showing that different starting values are converging towards the similar results. Training can also be a continuous process, were the network weights are updated regularly by new training samples. This could be the case, for example if the training samples are not available from the start, so the network has to learn from the data.

The *learning* of an ANN is a minimization process in which the weights are changed so that the output of the neural network for a given input is as close as possible to the target output. The *training sample* consists of a relatively large set of input sequences together with their associated target output. The learning process of a real problem could look like this:

1. generate a training sample;

2. decide on a network architecture;

3. replace the step-function by a sigmoid function;

4. Adjust the weights and make the networks to learn the training sample.

The Heavy-side step-function is therefore replaced by a smoother function, *sigmoid function*, g_T, which has finite derivatives, see figure 4.6 where the T parameter, can be interpreted as a "temperature", deciding the "steepness" of the function.

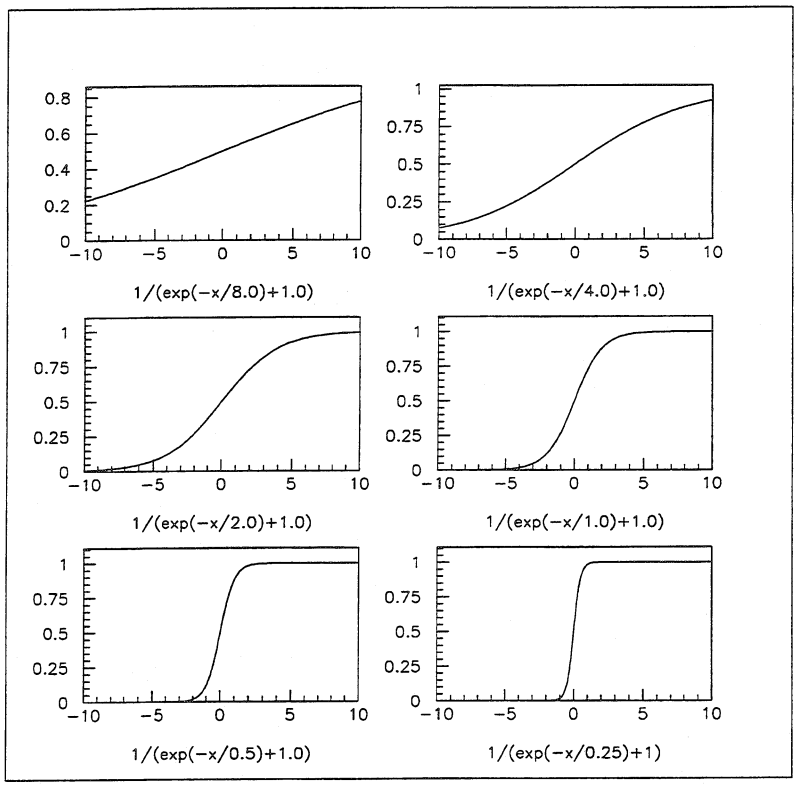

Figure 4.6: A sigmoid function, $g_T(x) = \frac{1}{1+\exp(-x/T)}$ for six different values (8.0, 4.0, 2.0, 1.0, 0.5 and 0.25) of T.

Chapter 5

Overview of the Main Analysis

The purpose of the work described in this thesis is to test different types of quark/gluon jet classifiers based on either the jet energy or fragmentation differences. A flow-chart of the analysis is shown in figure 5.1.

In the main analysis the quark/gluon jet classifiers are tested on samples of gluon and quark jets. In order to obtain these gluon and quark jet samples, which are the inputs for the analysis, some quality cuts are applied in the event selection. These selection criteria, which differ slightly for charged and neutral particles are discussed in chapter 6, where the cuts on hadronic events and the selection of threejet events are also discussed. In the same chapter, different jet algorithms are described. The LUCLUS algorithm has mostly been used but other algorithms are used when the performances of the different quark/gluon jet classifiers are studied.

Vertex detectors makes it possible to identify, or tag, $b\bar{b}$ events with high purity and high efficiency. Different b-tagging methods are presented and discussed in chapter 7.

In chapter 8, the selection of the three different jet samples is presented and these samples are:

- Hadronic events from Monte Carlo simulations where the jets of have been associated to quark or gluon jets using the simulation history
- Hadronic events from Monte Carlo where
  - the quark jets have been taken from $q\bar{q}\gamma$ events
  - the gluon jets have been anti-tagged as the remaining non b-jet in $b\bar{b}g$ events
- Hadronic events from real DELPHI data where
  - the quark jets have been taken from $q\bar{q}\gamma$ events
  - the gluon jets have been anti-tagged as the remaining jet in $b\bar{b}g$ events

---

In the first selection of these threejet event samples, associations of quarks and gluons to jets in $q\bar{q}g$ Monte Carlo events are made. By using the simulation history in hadronic threejet events it is in principle an easy task to identify quark and gluon jets. The detailed description of this procedure is a part of chapter 8.

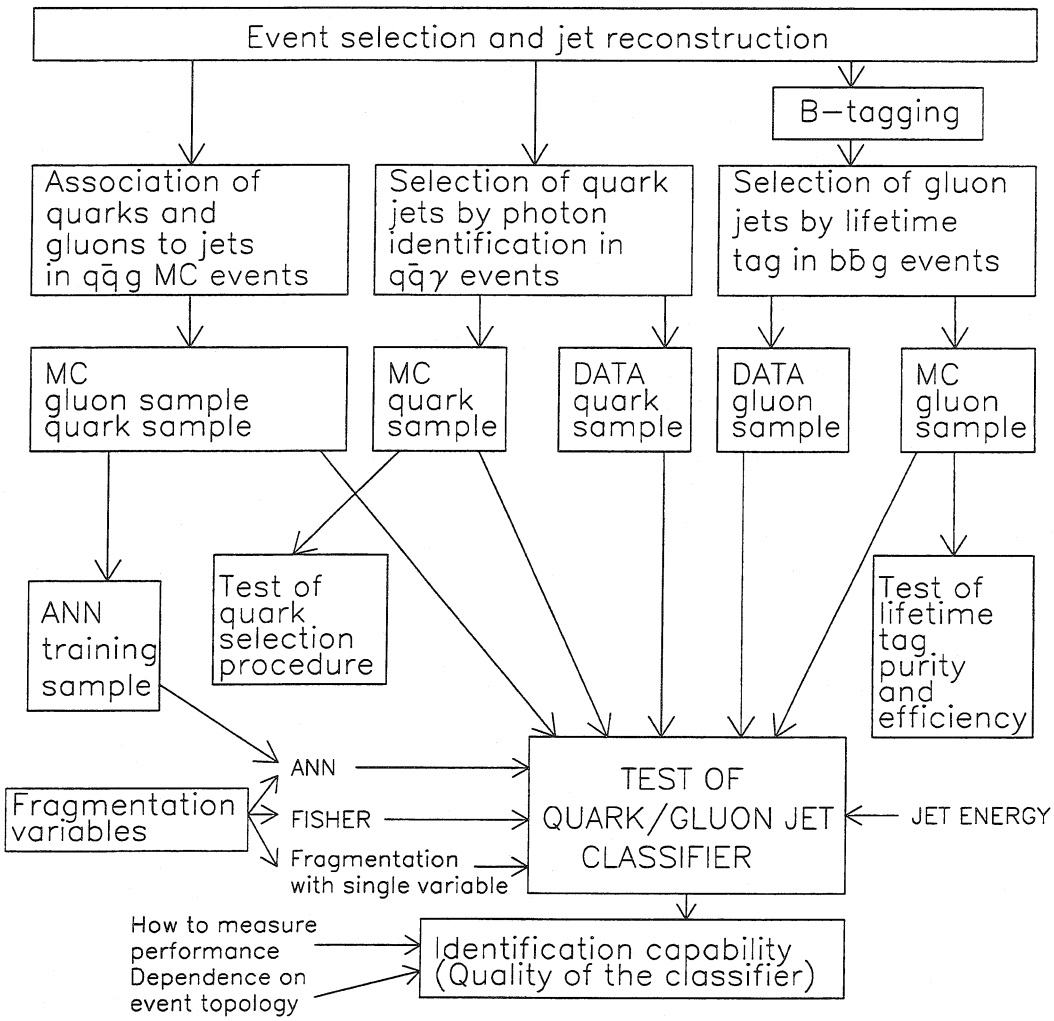

Figure 5.1: The main analysis.

Pure quark jet samples in DELPHI data are easy to obtain since twojet events contain only quark jets. These are, however, not suitable for the present analysis since these quark jets are mono-energetic and it is impossible to obtain gluon jets of a comparable energy. One way of getting quark jets at lower energy is to study quark jets in $q\bar{q}\gamma$ events and the selection of such a sample is described in chapter 8. The Monte Carlo quark jet sample, selected in the same way as the real data sample, is used to measure the purity and the efficiency of the method of the quark jet selection.

The selection of gluon jets by lifetime tag in $b\bar{b}g$ events is used to produce the third event sample. One of the ways of tagging events, described in chapter 7, is the b-lifetime tagging, which is used to identify b-quark jets. They are then used to produce a sample of anti-tagged gluon jets. Similarly selected samples have been used in other DELPHI studies measuring the multiplicity and particle contents of gluon jets, see [1, 2, 3]. A Monte Carlo gluon jet sample, produced with the same selection criteria as the real data sample, is used to measure the purity and the efficiency of the gluon selection and this is also presented in chapter 8.

The most powerful variable to use as a quark/gluon jet classifier is the jet energy and this is discussed and evaluated in chapter 9. Here is also a discussion on how to correctly present the performance of the classifier for different jet topologies.

Since one of the main objectives in this work is to separate the classification done by jet energy from that of fragmentation, a discussion of the most suitable ways of doing this, is found in chapter 10, where the different sets of fragmentation variables used as input to the quark/gluon jet classifiers are also presented.

The two training samples used as an input for the quark/gluon jet classifier are taken from the associated quark and gluon jet samples in chapter 8. These selected sets of fragmentation variables are then used as input to three different quarks/gluon classification methods:

- Fisher statistical discrimination method;
- A single variable discrimination method;
- Artificial Neural Networks.

Chapter 11 is dedicated to the problem of measuring the performances of the quark/gluon jet classifiers. There is also a section in which the Monte Carlo samples are used to estimate the purity of the classifiers using the traditional method of averaging over the entire jet sample. This is then compared to a more useful measure of performance which is associated with a specific event type. The chapter ends with a presentation of the results of the classifiers when tested on experimental DELPHI data.

Chapter 6

Event selection and jet reconstruction

6.1 Data selection

The data used was collected by the DELPHI detector in 1992–1994[1] around the Z^0 resonance which gave some 2.7 million hadronic events from e^+e^- annihilations. After the DELPHI standard hadronic cuts and the cuts on threejet events in table 6.2, a total of 741548 threejet events were selected in the Delphi data.

A slightly larger sample of events has been generated with the Lund Parton shower Monte Carlo JETSET 7.4 program [7], with parameters taken from the DELPHI tuning [22]. In total 793009 threejet events in Monte Carlo data were left after the cuts.

The jets were clustered using the LUCLUS jet-algorithm, but two other jet-algorithms, JADE and DURHAM, will be used in a later chapter and for convenience they are all described below.

6.2 Particle selection

The charged particles are measured in the barrel region by the set of cylindrical tracking detectors (VD, ID, TPC and OD) in the solenoid magnetic field of 1.23 T. In the reconstruction, a track is only accepted if the impact parameter at closest approach in the $r\phi$-plane is less than 5 cm·rad and less than 10 cm in the z-plane (the beam direction). A minimum track length of 0.50 m in the TPC, which is the main tracking device, was also required.

Charged particles are required to have a momenta between 0.1 GeV and 45 GeV and neutral particles between 0.5 GeV and 45 GeV. Only particles within the angular range $25° \leq \theta \leq 155°$ (where θ is the angle to the beam direction) are used.

[1] More specificly the real data events were taken from the short DST-tapes: short94_b2, short93_c1 and short92_d2 and the simulated events from sh_qqps_k94_2l_b1, sh_qqps_b93_1g_c1 and sh_qqps_b92_2g_d2.

6.3 Event selection

The standard Delphi cuts were used to select hadronic events. An event was accepted if the total visible energy was larger than 15 GeV and a minimum value of 3 GeV of visible energy deposited in each of the hemispheres, E^{+}_{hem} i.e. $\cos\theta > 0$ and E^{-}_{hem}, i.e. $\cos\theta < 0$ was also demanded. Furthermore, for all events, the polar angle of the thrust axis had to be in the interval 40° to 140° and the event was required to contain a particle multiplicity $N_{charged} + N_{neutral} \geq 8$.

6.4 Selection of threejet events

The selection criteria for individual jets in threejet events are given in table 6.1. Because

Cuts on individual jets	
multiplicity in jets	$N_{charged} + N_{neutral} \geq 4$
measured energy in jets	$E_{jet} \geq 5 \cdot GeV$

Table 6.1: Selection criteria for individual jets.

of momentum conservation, threejet events are coplanar. In order to enhance the contributions of three well defined jets, cuts are imposed on the angular sum of the three angles and a cut on the minimum angle between every pair of jets as summarised in table 6.2.

Cuts on three jet events	
planar events	$\theta_{12} + \theta_{13} + \theta_{23} \geq 357°$
minimum jet to jet angle	$min(\theta_{12}, \theta_{13}, \theta_{23}) \geq 35°$

Table 6.2: Cuts on three jet events.

6.5 Jet reconstruction

A particle jet is a cluster of particles with correlated directions which form a "cone". The different jet algorithms used in the data analysis are based on different ways of combining particles to jets. The algorithms used in e^+e^- experiment are *clustering* algorithms. A typical jet algorithm, e.g. JADE or DURHAM, distributes particles in sets of an invariant mass smaller than a given cut off parameter, y_{cut} ($\sim$ mass or $\sim$ energy), which is chosen in advance. It is required that a particle belongs to one jet only and that no other particle can be added to a given jet without exceeding this parameter. One usually introduces a dimensionless parameter y_{ij},

$$y_{ij} = M^2_{ij}/E^2_{vis}$$

that measures "the distance" between two particles, where:

- M_{ij} is $\sim$ mass or $\sim$ energy of the particle pair (i,j);
- E_{vis} is the visible (measured) centre-of-mass energy (i.e. s).

The particle pair (i,j) having the smallest y_{ij} are combined to a pseudo-particle with index k, $(i,j) \rightarrow k$ if $y_{ij} > y_{cut}$. In every step, the pseudo-particle gets the properties given by the sum of the properties of the two pseudo-particles that have been joined together. This is iterated until all particle pairs satisfy $y_{ij} > y_{cut}$. Depending upon how the cut parameter is chosen, the algorithm gives a number of jets as as result. If a higher y_{cut} is chosen it will mean that particles with a greater distance in phase space are joined together, giving a smaller number of jets as a result. The smaller the resolution parameter y_{ij}, the smaller the mass of the jet. If the jet mass is large, fluctuations in the jet evolution (like emission of soft or collinear gluons) will not affect its structure. On the other hand, small fluctuations for a light jet, can easily change its structure and the overall jet multiplicity.

6.6 Jet algorithms

From the discussion in the previous section, it is clear that even within the same algorithm the definition of a jet is not fixed. It therefore exist many different algorithms, which suit different analysis purposes and which are favoured by different people.

In $e^+e^- \rightarrow q\bar{q}g$ threejet events it is in principle possible to calculate the jet energies from the jet-jet angles given the kinematic constraints imposed by energy and momentum conservation. Assuming massless kinematics, the calculated jet energy, E_i^{calc}, for jet number i, can be expressed as,

$$E_i^{calc} = \frac{\sin\theta_{jk}}{sin\theta_{ij} + sin\theta_{ik} + sin\theta_{jk}} \cdot \sqrt{s}$$

where:

- θ_{jk} is the angle between jet number j and jet number k;
- s is the squared centre of mass energy;
- i, j and k are cyclic.

It is well known that the calculated jet energy corresponds better to the original parton energy than the measured jet energy, see for instance [23]. However, the latter is more general (i.e. not bound to a specific topology) and in the case of jets initiated by b-quarks, corrections have to be made due to the mass effect and therefore the measured energy has been used in this study.

As the main jet finder, the LUCLUS algorithm provided in JETSET, [7] with the separation parameter d_{join} set to 3.5 GeV/c (LUCLUS 3.5) has been used. To check the sensitivity of the quark/gluon jet classification with respect to jet-clustering, the following algorithms and separation cuts have also been tried,

LUCLUS with $d_{join} = 5.0\ GeV/c$, (LUCLUS 5.0)

JADE with $y_{cut} = 0.015$

DURHAM with $y_{cut} = 0.015$

A common feature of the jet algorithms are that they *order* the jets according to their energy. It is therefore always true in this analysis that

$$E_{JET1} > E_{JET2} > E_{JET3}$$

LUCLUS

The LUCLUS algorithm is based on the momentum of the particle. At the start of the algorithm, every particle in the final state is considered to be one cluster. The distance measure between two clusters, d_{ij}, is defined as,

$$d_{ij}^2 = \frac{4 \cdot |\vec{p_i}|^2 \cdot |\vec{p_j}|^2 \cdot \sin^2(\theta_{ij}/2)}{(|\vec{p_i}| + |\vec{p_j}|)^2}$$

where:

- θ_{ij} is the angle between particle i and particle j;
- $\vec{p_i}$ is the momentum vector of particle i;
- $\vec{p_j}$ is the momentum vector of particle j.

The algorithm uses binary joining, which means that the two nearest clusters are joined into one cluster, if their distance, d_{ij}, is smaller than a preset value, d_{join}. If the two particles or jets are merged, new jet axes are calculated and the particles are reassigned to the nearest jet. Then the new configuration of clusters is investigated and again the two nearest clusters are joined into one cluster if the joining condition is fulfilled, until all clusters are separated by a distance larger than d_{join}. The now remaining clusters are forming the jets. In this approach every single particle belongs to one, and only one, jet. The jet reconstruction depends on one single parameter with a definite physical meaning of a transverse momentum. If well separated jets is needed, a rather large d_{join} should be chosen, while a small value of d_{join} would allow the separation of jets very close to each other. The default value of d_{join} is 2.5 GeV/c, which is good for lower energies but for LEP energies, a higher value is often required. Therefore the selection of $Z^0 \rightarrow b\bar{b}g$ events used a d_{joint} of 3.5 GeV/c.

JADE

In the JADE algorithm, the jet resolution parameter, y_{ij}, is the measure of closeness for all pairs of particles in the events and it is defined as,

$$y_{ij} = \frac{2 \cdot E_i \cdot E_j \cdot (1 - \cos\theta_{ij})}{E_{visible}^2}$$

where:

- θ_{ij} is the angle between particle i and particle j;
- E_i is the energy of particle i;
- E_j is the energy of particle j;
- $E_{visible}$ is the sum of the energies of all the particles of the event, seen by the different detectors.

The closest particles, i.e. the particle pair with the smallest y_{ij} is replaced by a pseudo-particle with four-momentum equal to the sum of the four-momenta of the two particles i and j, if y_{ij} is smaller than a pre-chosen cut-off value y_{cut}. This is repeated until all distances y_{ij} are larger than the preset cut value.

DURHAM

The DURHAM algorithm is very similar to the JADE algorithm, but the jet resolution parameter, y_{ij}, is defined for all pairs of particles in the events as,

$$y_{ij} = \frac{2 \cdot min(E_i^2, E_j^2) \cdot (1 - \cos\theta_{ij})}{E_{visible}^2}$$

where:

- θ_{ij} is the angle between particle i and particle j;
- E_i is the energy of particle i;
- E_j is the energy of particle j;
- $E_{visible}$ is the sum of the visible energies of all particles in the event.

The pairing of the particles is identical to the *JADE* algorithm.

Chapter 7

B-tagging

There exists in principle two different possibilities to identify or to *tag* an event containing heavy B-particles, lepton tag or lifetime tag. After tagging, a jet is most likely emanating from quarks and thus "known" and it will be used in the analysis to *anti-tag* the gluon jet by identifying the two quark jets in threejet events. The lifetime tag is the only one used in this analysis and the semileptonic tagging is only presented for reasons of reference.

7.1 Lepton Tag

One possibility to separate events with heavy quarks, like b and c, from event with light quarks s, d and u, is the decay of the heavy quarks into a lepton, e.g. an electron or a muon. This *semileptonic* decay, see figure 7.1, provides a mean to tag B-events. However, only about 20% of the b-quarks decay to leptons so even if this tagging method is effective it will have limitations. By using energy and momentum conservation one can show that

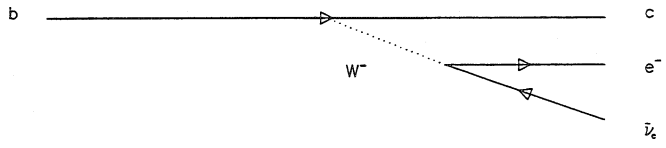

Figure 7.1: The semileptonic decay $b \rightarrow c\, e^-\, \bar{\nu_e}$ mediated by a W^-.

the transverse momentum, p_T, of the lepton, i.e. the momentum component perpendicular to the direction of the jet, must fulfil $p_T < 0.5 \cdot m_q$, see [24], where m_q is the mass of the decaying quark. Since the lepton is very difficult to detect inside a jet, it is often separated by a cut, thus making it *isolated*, by requiring the lepton to have an angle greater than an angle, θ_{min}, with respect to the jet axis. The decay of a heavy quark will in general give leptons with large transverse momenta. For example, a decaying c-quark with a mass

of $\sim$ 1.5 GeV/c^2, from a semileptonic decay would give leptons with a maximum p_T of 0.75 GeV/c while a decaying b-quark with a mass of $\sim$ 4.5 GeV/c^2, would give leptons with a maximum p_T of 2.3 GeV/c. Hence, by requiring some lower cut on the p_T, e.g. 1.0 GeV/c, one can enhance a sample of B-events. An estimation of the purity achieved by using this tagging of semileptonic decays is about 80% with an tag efficiency of about 60% for electrons and 80% for muons, for details see [25].

7.2 Lifetime Tag

The B-hadrons have a lifetime of about a picosecond. The B^0, for example, has a measured lifetime of 1.56 picoseconds. Its rest mass, m_B^0 is about 5 GeV/c^2, so if the B^0 gets a total energy, E_B of the order of 20 GeV, the particle will travel at a speed close to that of light, with a γ-factor = E_B/E_B^0, of about four. Average γ-factor distributions for different types of jets, taken from threejet events in 1994 simulated data, are shown in figure 7.2. Moving at this velocity and with the given lifetime, the B^0 will cover a distance of $c \cdot \tau$ = 468 μm in the particle's rest frame, before the particle decays, something that will be measured as a distance of a few millimetres, in the Delphi detector. The production vertex (the primary vertex), will therefore be well separated from the decay vertex (the secondary vertex) and easily separated in experiments equipped with vertex detectors. These detectors, see figure 3.9, (earlier discussed in section 3.3.1), can measure tracks in

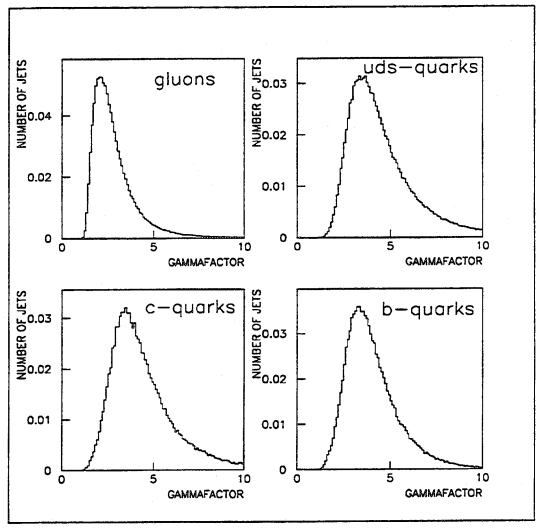

Figure 7.2: The normalized γ-factor distribution for jets emanating from gluons, light quarks, c-quarks and b-quarks, respectively.

the silicon micro strips with an accuracy of a few μm. An event with a secondary vertex is shown in figure 7.3.

The lifetime signed impact parameter, d, is defined as the shortest distance between

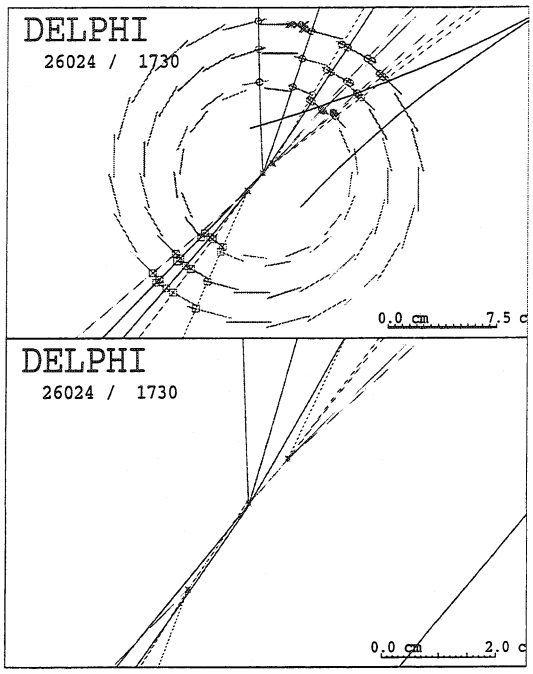

Figure 7.3: The decay of a Z^0 into B-mesons giving clearly identified secondary vertices.

the track in an event and the primary vertex, PV. This distance can be parametrised by an angle α between the direction of the impact parameter and the jetaxis and if α falls in the interval 0-90 degrees, the signed impact parameter is positive (see figure 7.4) and if α falls in the interval 90-180 degrees, the signed impact parameter is negative (see figure 7.5). There will, of course, be an uncertainty in the impact parameter, due to the finite resolution of the detectors. To create a dimensionless variable, the impact parameter is divided by its standard error, σ_d, and this new variable is called the *significance*, S. Tracks with negative significance (or negative impact parameter since the standard error is always positive) are either badly reconstructed (do not belong to the jet in question) or emerges directly from the primary vertex. These tracks have an equal chance of being assigned a negative or a positive significance, depending on which side of the origin the extrapolated track pass. Tracks with $S < 0$ will therefore describe the background distribution of the tagging technique built on a secondary vertex, since they essentially reflects the resolution of the vertex detectors, see figure 7.6. Since the position of the primary vertex is crucial for a correct measurement of the impact parameters, it is of great importance that the primary vertex is reconstructed correctly. To improve this, and the reconstruction of the secondary vertex, many different constrains are used and a lot of different approaches tried.

The tagging technique in DELPHI assumes a symmetrically distributed functional dependence on the significance, f(S), around S equals zero for non B-events. Ideally the negative significance distribution should have a Gaussian shape, but typically, f(S)

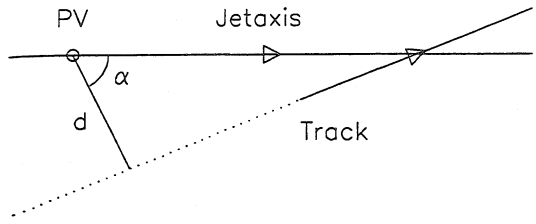

Figure 7.4: The signed impact parameter is positive, $d > 0$.

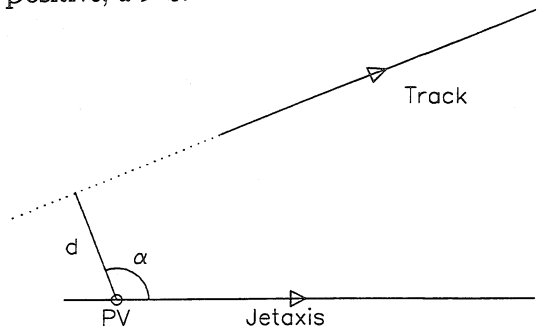

Figure 7.5: The signed impact parameter is negative, $d < 0$.

will have a maximum at S=0 and decrease rapidly for large values of S. There is a non-Gaussian tail, due to tracks measured wrongly by the tracking detectors and from particles from secondary decays and interactions. This part of the distribution can be changed by different cuts.

The negative significance distribution is used to define the *track probability function*, P(S_0). B-events with their large positive impact parameters and therefore large significance values add an excess of large positive values of S compared to the background from the negative S distribution and the difference will increase for larger S values. If f(S) is properly normalized it can be integrated from minus infinity up to a value S_0 to be used as an interpretation of probability. With the following definition of P(S_0) it is clear that P(S_0) is almost one for small S_0 values and close to zero for large S_0 values.

$$If \;\; S_0 < 0, \quad P(S_0) = \int_{S<S_0} f(S)dS$$

$$If \;\; S_0 > 0, \quad P(S_0) = P(-S_0)$$

A convenient interpretation is now that P(S_0) gives the probability that a track from the primary vertex will have $S > S_0$. Since tracks from a B-decay will have large positive significance values for S, say S_B, the probability, P(S_B), that these tracks originate from the primary vertex will be small. By choosing a small enough value of P(S_B) (i.e. make a *cut* on P(S_B)) the chances of finding tracks from a B-decay will increase. Using this track probability function and the value of the significance, the probability for each track in the

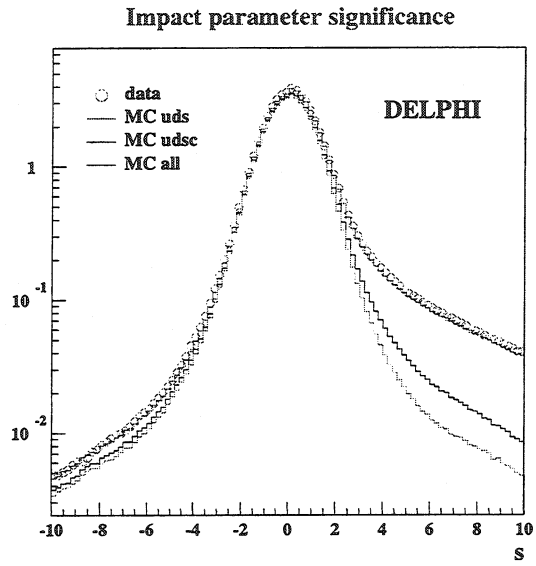

Figure 7.6: The impact parameter significance for light quarks, light+c-quarks and all quarks, showing the obvious enhancement for b-quarks.

event can be calculated. Using this track probability interpretation an even better tagging variable, the *N-track probability*, P_N, can be constructed, taking all track probabilities in the event into consideration.

$$P_N \equiv \prod \cdot \sum_{j=0}^{N-1} (-\ln \prod)^j / j!$$

$$where \quad \prod = \prod_{i=0}^{N} P(S_i)$$

The variable P_N gives the probability that a group of N tracks, with individual values $P(S_i)$ are all coming from the primary vertex, [26]. A flat distribution of P_N is expected if the N tracks are coming from the primary vertex and their significance are uncorrelated. If, however, some of the tracks are coming from a secondary vertex, the distribution will have high values for small values of P_N. Therefore the distribution will be approximately flat for the light quarks, u, d, s and have high values when P_N is close to zero for the heavier c-quark and very high values for the even heavier b-quark, see figure 7.7. There is a slight excess at low P_N even for light quarks, due to residual tracks from neutral particles like K^0 or Λ, decaying into a charged pair outside the detector or from interactions in the detector material, but P_N is still significantly lower than for the heavy quarks.

7.3 The tagging method used in the analysis

The standard program package used in Delphi for b-tagging [26] is based on the algorithm proposed by D.Brown and M.Frank [27] and it is implemented in the package AABTAG by G.Borisov [28]. Only the routines used in the analysis are briefly discussed here. They use the impact parameters of the tracks, which have large positive values in events containing b-quarks due to the long lifetime of the B-hadrons for tagging. The probability can be used to tag the whole event, but also for the tagging of individual jets.

7.3.1 The event tagging

The main routine for b-tagging of the event, BTAGAA, is reconstructing the jets and the primary vertex. Using all the tracks, it calculates the probability that all well-measured tracks belonging to the event are originating from this vertex. The output parameters are:

PROBN, the probability computed with the tracks with negative impact parameters;

PROBP, the probability computed with the tracks with positive impact parameters;

PROBS, the event probability computed for all tracks.

The variable PROBP is used for the actual tagging of B-events, using the fact that the value of PROBP is much smaller for b-quark events than for event coming from light quarks or c-quarks, see figure 7.7. The B-sample can be selected with a desired degree of purity (taking the corresponding efficiency into consideration) by a cut on the PROBP. The cut $PROBP < 0.01$, for example, selects B-events with a purity of around 90% and an efficiency of around 50%, for 94 data. A typical efficiency-purity diagram for the b-tagging in the DELPHI detector can be seen in figure 7.8.

7.3.2 The individual tagging of the jets in the event

The main routine for the b-tagging of the individual jets, AAJPRB, is computing the probabilities with signs of the impact parameters for one or several jets in the event. Input information to the algorithm is the number of jets, the direction of the jets, the number of tracks and to which jet number in the event the tracks belongs. This can be used for tagging the individual jet by calculating the probability for each jet separately, that all well-measured tracks belonging to the jet are originating from the primary vertex. The output parameters of AAJPRB are:

PROBJN, the probability computed for the given group of tracks with negative impact parameters;

PROBJP, the probability computed for the given group of tracks with positive impact parameters;

PROBJS, the probability computed for the given group of all tracks.

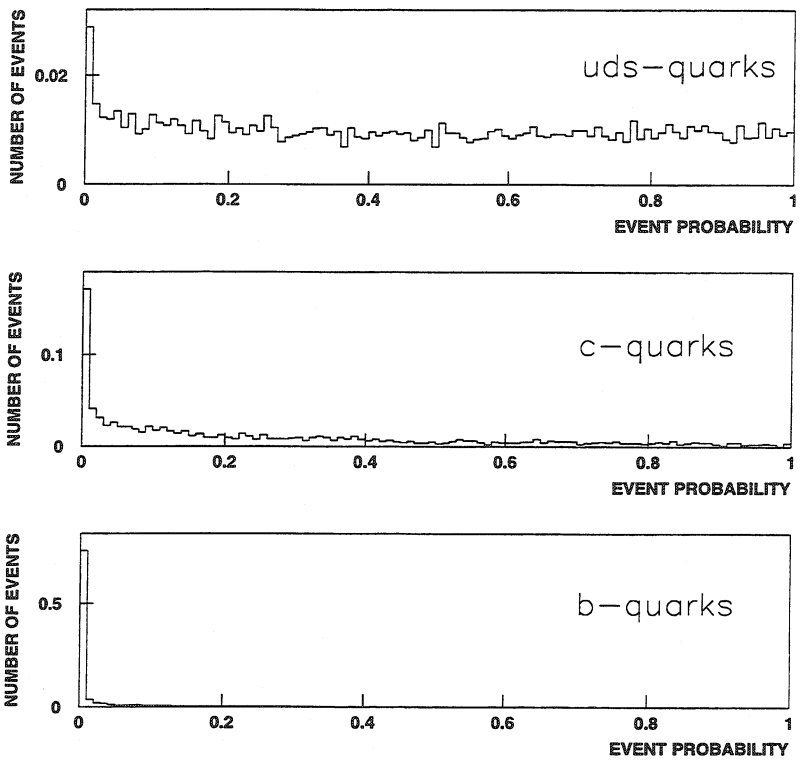

Figure 7.7: The normalized event probability for uds-, c- and b-quarks in simulated 1994 data.

The variable PROBJP is used for the actual tagging of individual jets, see figure 7.9.

The tagged gluon sample, used in the following analysis, is obtained by first using an event cut in order not to collect too many non B-events, before the jet-algorithm, searching for threejet events, is applied. Then individual jets are tagged and the idea is to select among all threejet events, the events containing exactly two jets satisfying the experimental signature for being initiated by b-quarks. The remaining jet is then assumed to be the gluon jet. This procedure is described in some detail in section 8.4 where different settings of the tagging parameters are also discussed, since this affect the purity of the tagging method.

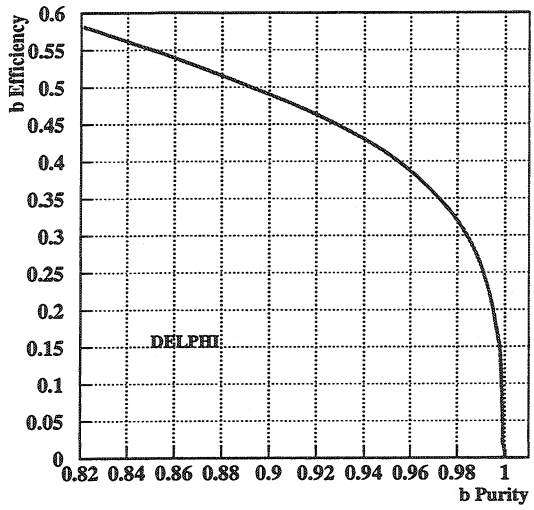

Figure 7.8: A typical efficiency-purity diagram for b-tagging in DELPHI.

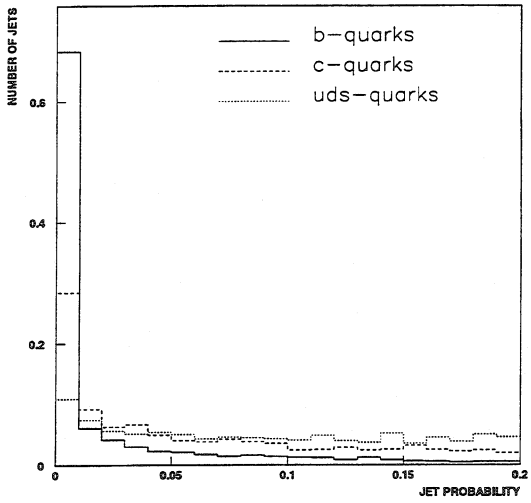

Figure 7.9: The normalized jet probability for b-, c- and uds-quarks in 1994 simulated data.

Chapter 8

Quark/Gluon jet samples

The basic idea in the method of collecting quark jet and gluon jet samples in threejet events, is that pairs of quark jets are identified, where the third jet must then come from a gluon. To obtain jet samples, quality cuts are applied in the particle selection as well as the standard DELPHI cuts in the hadronic event selection and finally the cuts to select threejet events. These criteria were all previously discussed in chapter 5.

8.1 Association of quarks/gluons to jets in $q\bar{q}g$ MC events

If Monte Carlo events are used, the flavour of the quarks in the event is in principle known. After the hadronization it is, however, not obvious which particle belongs to which jet and to which parton. If the quarks are heavy b- or c-quarks one can use the simulation history of the decaying heavy hadrons to associate a jet to a specific quark and this method is described in this section. A flow cart of the method used is seen in figure 8.1.

The simulated data was produced by using Jetset 7.4 with the CMS energy equal to the Z^0 mass in order to generate the decay chain of $Z^0 \rightarrow q\bar{q}g$ events.

Since the event description is known in detail in the simulation, the whole event can be reconstructed and all the necessary values of the jet parameters can be calculated. This makes it possible to follow the history of the original quark pair, via the breaking of the string and the decay of the unstable particles until only stable, detectable particles remain and are seen in the DELPHI detector. This is done by starting with a $q\bar{q}$-parton pair of a selected flavour, by choosing the corresponding flavour code in Jetset (for a complete description of the Jetset variables see [7]) and then follow the evolution until all hadrons are hanging on a string. Finally all particles are decayed until only stable particles remain in the jets. These are the phases of the hadronization process discussed earlier, see figure 2.6. Events starting with a $u\bar{u}$-, $d\bar{d}$- or a $s\bar{s}$-parton pair, the *light quarks*, are grouped together and separated from the heavy $c\bar{c}$-parton pair and the heaviest $b\bar{b}$-parton pair, which are forming their own groups.

Two tests with different selections of particles were done. One used all detected particles and the second only the charged particles. One of the reasons for this investigation was to see if any of these samples would introduce a bias in the performance of the jet

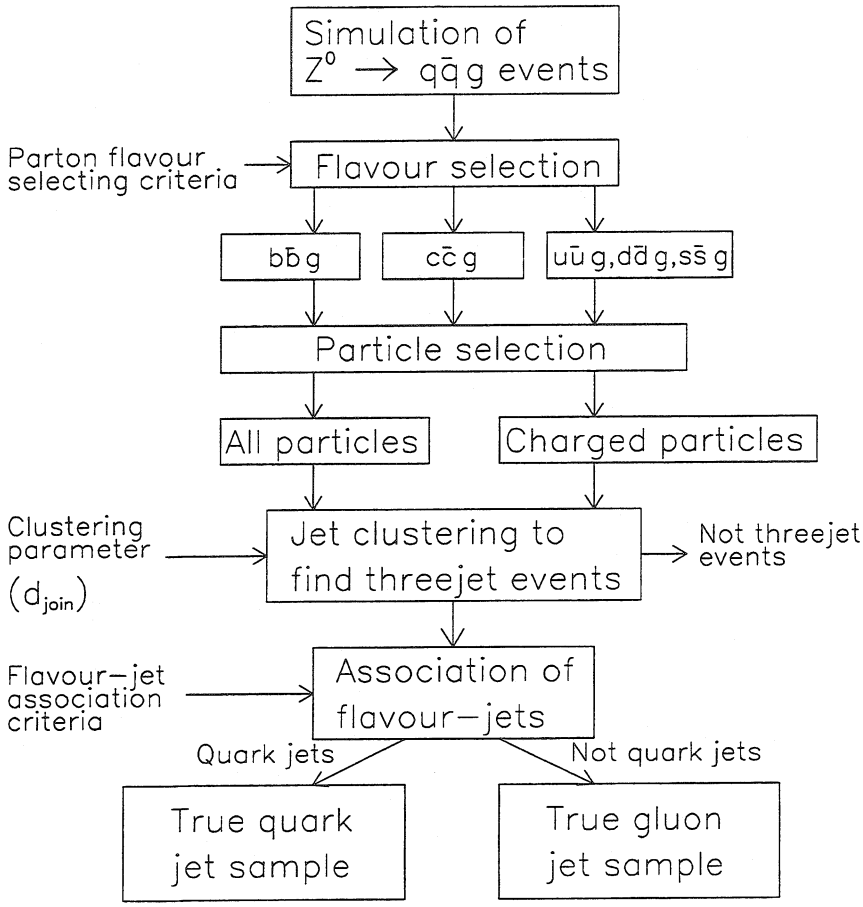

Figure 8.1: How to obtain the pure quark sample and the pure gluon sample.

reconstruction and in the association procedure.

The reconstruction of jets has been made by using the LUCLUS jet algorithm, described in section 6.6, with a d_{joint} value of 3.5 GeV/c. The LUCLUS jet algorithm was used because it was considered the best algorithm for reconstructing jets from heavy quarks.

To illustrate how the association procedure works, a typical $Z^0 \rightarrow b\bar{b}g$ Monte Carlo event has been chosen. The particles produced in this event are shown in figure 8.6. The result, after the LUCLUS algorithm was applied to the selected three jet event, is shown in figure 8.2. More precise information about the same event is given in the following boxes.

The energy of JET 1 is 40.2 GeV and its mass is 7.5 GeV/c^2.
It consists of the following 14 stable particles:
3 π^+, 3 π^- and 8 γ

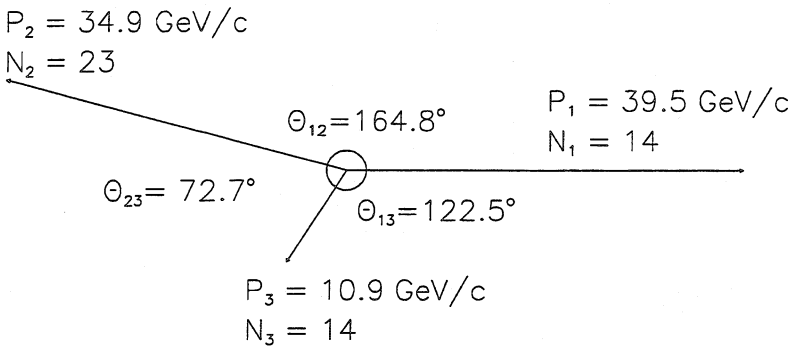

Figure 8.2: The selected threejet event with all the relevant information.

The energy of JET 2 is 36.0 GeV and its mass is 8.8 GeV/c^2. It consists of the following 23 stable particles: 5 π^+, 4 π^-, 10 γ, 1 K^+, 1 K^-, 1 e^- and $\bar{\nu}_e$ (not used in the jet finding process)
The energy of JET 3 is 13.3 GeV and its mass is 7.6 GeV/c^2. It consists of the following 14 stable particles: 2 π^+, 2 π^-, 2 n^0, 2 $\bar{n}^0$ and 6 γ

In order to be able to compare the values for this selected event, with other $Z^0 \rightarrow b\bar{b}g$ events, the distribution of the average number of particles in a jet, the average values of the angles Θ_{12}, Θ_{13} and Θ_{23}, and the average energy distribution for the three jets in threejets events for the 1994 simulated data, are shown in figures 8.3, 8.4 and 8.5, respectively.

There are several possibilities to identify the quark jets and hence the gluon jet in simulated threejet events. One obvious method to use the start with a heavy parton pair and to use the simulation history of the decaying B/C-hadrons and identify the jet that contains every decay products of the B/C-hadron as being the B/C-jet.

The event in figure 8.6, which is used as an example, contains two heavy B-hadrons (i.e. B-mesons, see table 2.6), a B_S^* and a B^-.

The decay products of the neutral B_s^*, consists of the following 17 stable particles (including the undetected $\bar{\nu}_e$) : 4 π^+, 2 π^-, 1 K^-, 1 e^-, 1 $\bar{\nu}_e$ and 8 γ
The decay products of the charged B^-, consists of the following 11 stable particles: 2 π^+, 3 π^-, 6 γ

In this event, all 17 particles (with 16 detected) from the decaying neutral B-hadron are all in JET2, so this jet can easily be identified as a B-jet and all 11 particles from the charged B-hadron are all in JET1, so this jet can also be interpreted as a B-jet.

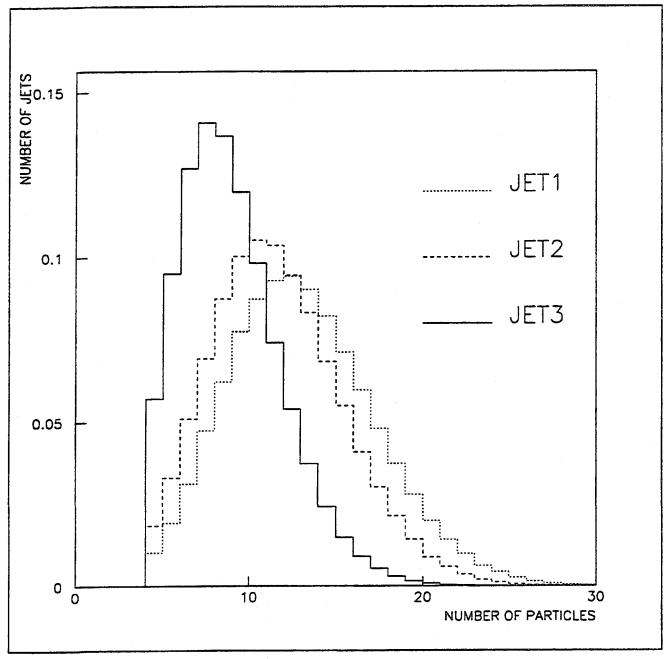

Figure 8.3: The normalized distribution of the number of particles of the jet in simulated threejet events.

A crucial question for this jet association method is if this is the case every time, i.e. are all the decay products of the B-hadrons always contained in one jet or are some particles from the decay chain sometimes found in another jet?

To investigate this, a variable, *bfrac*, was introduced in order to measure if the decaying B-hadrons were fully contained in the two B-jets or if there were some "cross-talk" between the B-jets and the non B-jet (the third jet). The bfrac variable was defined as:

$$bfrac = \frac{Number\ of\ B-hadron\ decay\ products\ in\ the\ non\ B-jet}{Number\ of\ all\ B-hadrons\ decay\ products}$$

The decay of the lighter c-quark was also investigated by the construction of a similar *cfrac* variable. Figures 8.7 and 8.8, show the variable bfrac/cfrac as a function of the minimum separating jet angle (*angmin*), to the relevant B/C-jet.

The conclusion from the study was clear. The association of the jet to a heavy B-particle works well since most decay products of the B-hadrons are fully contained in the two B-jets as long as the angle between the jets are not too small. This is even more true with the lighter C-hadrons since their decay is more "pencil like" due to their lower mass.

Another method, which also uses the decay chain, was utilized to identify the light quark jets. The natural choice is then to measure "the distance" between the momentum vectors of the original partons with the momentum vectors of the three jets and associate the "closest" of the two jet directions to the parton directions as the quarkjet directions.

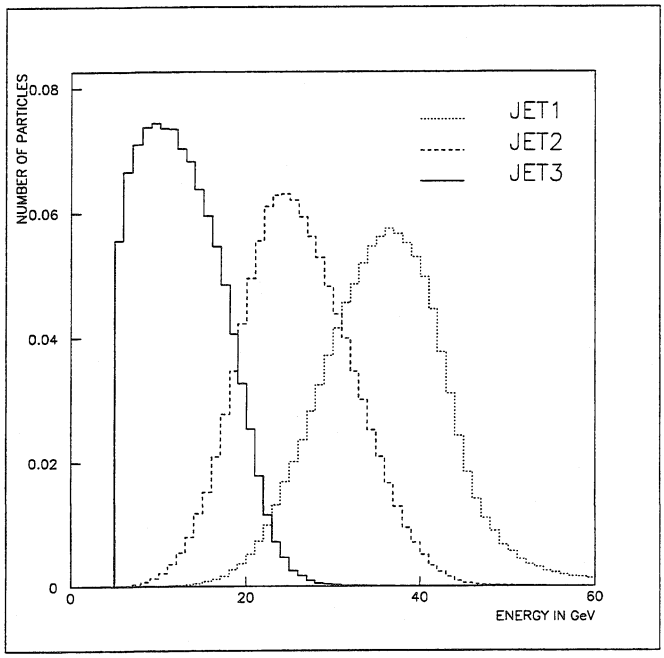

Figure 8.4: The normalized distribution of the energy of the jet in simulated threejet events.

The angle between the jets can be used as a measure of closeness.

In a previous section it was shown that most likely all particles from the decay of a heavy B-hadron are found in the same jet. Even at the smallest allowed angle between the jets, the cross-talk between the jets is completely negligible and do not in any way jeopardize the associated procedure where the jet that contains the majority of the decay products of the B-hadron is identified as the B-jet. Thus, since two jets will have quark origin, they are considered to be *quark jets.*

Using the two quark jet associations described above, it is obvious to assign the remaining jet, in a three jet event, as the *gluon jet.*

8.2 Differencies between the true quark/gluon jet samples

Differencies between the quark and gluon samples can be studied on a statistical basis. If the quarkjets are subdivided into light quarks, c-quarks and b-quarks, studies can be made concerning their energy jet distributions, their jet masses or their particle multiplicities of the jet, see figures 8.9, 8.10 and 8.11. In table, 8.1 are summarized the average values of the energy, the jet mass, the number of particles and the γ-factors of the jet. It is clear that there is a significant differences between the gluon jet and the different quark jets

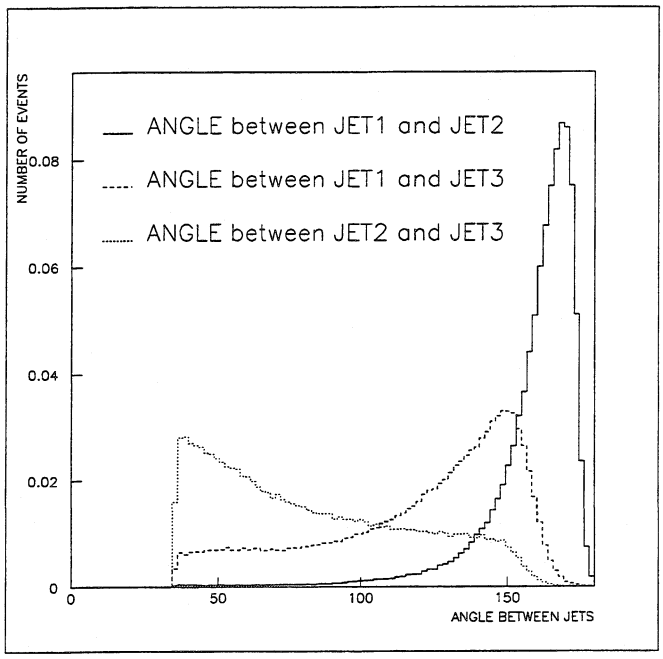

Figure 8.5: The distribution of the angles between the jets in simulated threejet events.

and it is this difference that will be used in the forthcoming analysis.

Jets from	Energy (GeV)	Jet mass (GeV/c^2)	Multiplicity	γ-factor
gluons	14.3	5.3	9.6	2.8
light quarks	30.8	7.4	11.8	4.4
c-quarks	28.8	7.0	11.8	4.3
b-quarks	28.4	7.5	12.9	4.0

Table 8.1: Average energies, jet masses, jet multiplicities and γ-factors for the different jets in 1994 simulated data.

One often measured quantity in LEP experiment, is the ratio of the number of particles in a quarkjet compared to the number of particles in a gluonjet, see for instance, [1], [2] or [3]. Gluonjets contain more particles than a quarkjet *of the same energy.* This can be seen in figure 8.12 which shows the number of particles as a function of the jetenergy.

8.3 Selection of quark jets in $q\bar{q}\gamma$ events

In this analysis not only gluon jets but also jets from light quarks are needed. Using the DELPHI data, it is easy to obtain a completely pure sample of quark jets from twojet

events like $e^+e^- \rightarrow q\bar{q}$. Such a sample is of limited use, however, because the jetenergy is restricted to E_{beam} and these two quark jets will, in general, be at a higher energy than the energy of the gluon jets. Quark jets of lower energies can instead be obtained from events where one of the quarks has radiated a hard photon. Such a $q\bar{q}\gamma$ event seen in the DELPHI detector is depicted in figure 8.14. These events are, however, quite rare, especially if a high E_γ (i.e. a low quark jet energy) is required.

The flow chart in figure 8.13, describes the quark jet selection procedure. The real data from the DELPHI detector, collected between 1992–1994, has to pass the standard quality cut presented in section6.2. The simulated data was produced by using Jetset 7.4, with the CMS energy equal to the Z^0 mass. The Monte Carlo data had to pass the same cut as the real data.

The selection of $q\bar{q}\gamma$ events has been made by using the LUCLUS jet algorithm with a d_{joint} value of 3.5 GeV/c. There exists a possibility to test the selection by removing the γ-particle from the event and then apply LUCLUS on the remaining particles. If the number of jets is still three after this procedure, this will indicate that the γ-particle was probably coming from the fragmentation process, instead of being a final state radiation (FSR) γ-particle, as desired. Events with two jets and an isolated particle (classified here as a threejet event) are quite rare (of the order of some hundred events a year), if the photon energy is large.

Isolated particle identified as a photon.

After finding a jet with one single particle in threejet events, requirements are made on this particle. It is necessary to distinguish a neutral hadron, which gives a signal in the hadron calorimeter, from a γ-particle, which gives a signal in the electromagnetic calorimeter, while a π^0 (an unresolved gamma pair, $\pi^0 \rightarrow \gamma\,\gamma$) and a γ-particle can be distinguished because of the cluster separation done in the HPC.

Therefore, Monte Carlo events with a final state radiation γ have been investigated, i.e. $Z^0 \rightarrow q\bar{q}\gamma$ events and the suitable cuts for this type of events have been investigated. These events, with a single particle in one of the jets, should of course, like all other threejet event, be lying in the same plane, due to momentum conservation. The sum of the angles between the jets should therefore add up to 360°, see figure 8.16. A cut on the sum of the three angles is obviously very useful and it was decided that the sum $\theta_{12} + \theta_{1\gamma} + \theta_{2\gamma}$ should be $> 357°$ in order to enhance the final state radiation sample. The two remaining hadronic jets have to pass the normal jet selection cuts and are therefore required to contain at least 4 particles and the jet energy must be > 5 GeV for each of the jets.

For the γ-particle, a cut on the minimum γ-energy and a cut on the angle between the γ-particle and the nearest jet was chosen by looking at figure 8.15, which shows the energy spectrum for the γ-particles in $q\bar{q}\gamma$-events from 1993 simulated data. The number of isolated photons as a sensitive function of the energy and thus the lowest allowed γ-energy has to be selected with extra care and the γ-particle energy was finally required to be > 3 GeV and it was decided the that the angle between the photon and the nearest jet direction should be $> 15°$, thus clearly separating it from the quark jet. The tagged quark jet sample was obtain by simply taking the jets without an identified single photon

giving an almost completely pure tagged quark jets sample. The purity of the quark jet sample has been estimated to be on average 92% using Monte Carlo data. In total 3211 quark jets with $E_{jet} < 35$ GeV were tagged in DELPHI data from 1992–1994.

8.4 Selection of gluon jets in $b\bar{b}g$ events

As had already been pointed out, the improved vertex detector technology together with a high statistics of heavy quark events at LEP, has made it possible to collect a gluon jet sample with a broad energy spectrum and low quark contamination. The idea is to select among all reconstructed threejet events the events which contain exactly two jets satisfying the experimental signature for being initiated by b-quarks. In each of the selected threejet event the remaining jet, i.e. the jet which has *not* been tagged (i.e. *anti-tagged*) as being initiated by a b-quark, is associated to the gluon jet. See the flow chart in figure 8.17 for a description of the procedure in detail. The lifetime tag has been described in a previous sections where jets originating from heavy quarks were recognizable through a secondary vertex, see 7.2.

First, real data from the DELPHI detector, collected between 1992–1994, have to pass the standard quality cuts, as well as the hadronic cuts, presented in sections 6.2 and 6.1.

Before applying the jet clustering algorithm, events which are not B-events are removed and the Monte Carlo events are used to test the purity of this B-event tagging. After some studies the event cut $PROBP < 0.032$ was chosen in order to achieve a high purity of B-events with a reasonable efficiency.

The reconstruction of jets in $b\bar{b}g$ events has been made by using the LUCLUS jet algorithm with a d_{joint} value of 3.5 GeV/c.

The lifetime tag has been described previously. The following requirements was considered for what will be referred to as the *hard* btagging in order to obtain a very pure sample of quark jets:

- two of the three jets in the event are required to have a probability of less than 1% not to contain decay products from long lived hadrons;
- the third jet is required to have a probability greater than 10% to contain no decay products from long lived hadrons.

This is obtained with the following cuts the variable PROBJP for the individual jet tagging:

- $PROBJP < 0.01$, in order to tag one of the jets as a quark jet;
- $PROBJP < 0.01$, in order to tag one of the two remaining jets as a quark jet;
- $PROBJP > 0.1$, in order to anti-tag the gluon jet.

The purity of the gluon jet sample with this hard b-tagging is estimated by using the Monte Carlo simulated data in the DELPHI detector and the result is plotted in figure 8.18, as a function of the jet energy. As can be seen, the gluon purity for low jet energies is above 95%, falling to around 87% for jet energies above 30 GeV. A softer tagging

procedure was also tested in order to achieve higher statistics. The following cuts on the variable PROBJP were used:

- $PROBJP < 0.1$, in order to tag one of the jets as the quark jet;
- $PROBJP < 0.1$, in order to tag one of the two remaining jets as a quark jet;
- $PROBJP > 0.1$, in order to anti-tag the gluon jet.

The purity of the gluon jet sample with this soft b-tagging is also estimated by using the Monte Carlo simulated data. The gluon purity for low jet energies is just under 90% falling to about 73% for jet energies around 30 GeV.

Finally, the hard tagging was chosen because the required purity in this analysis needed to be high over the full energy spectrum. The softer tagging can, however, be used in special topologies, e.g. Mercedes events (where all the jets have approximately the same energy).

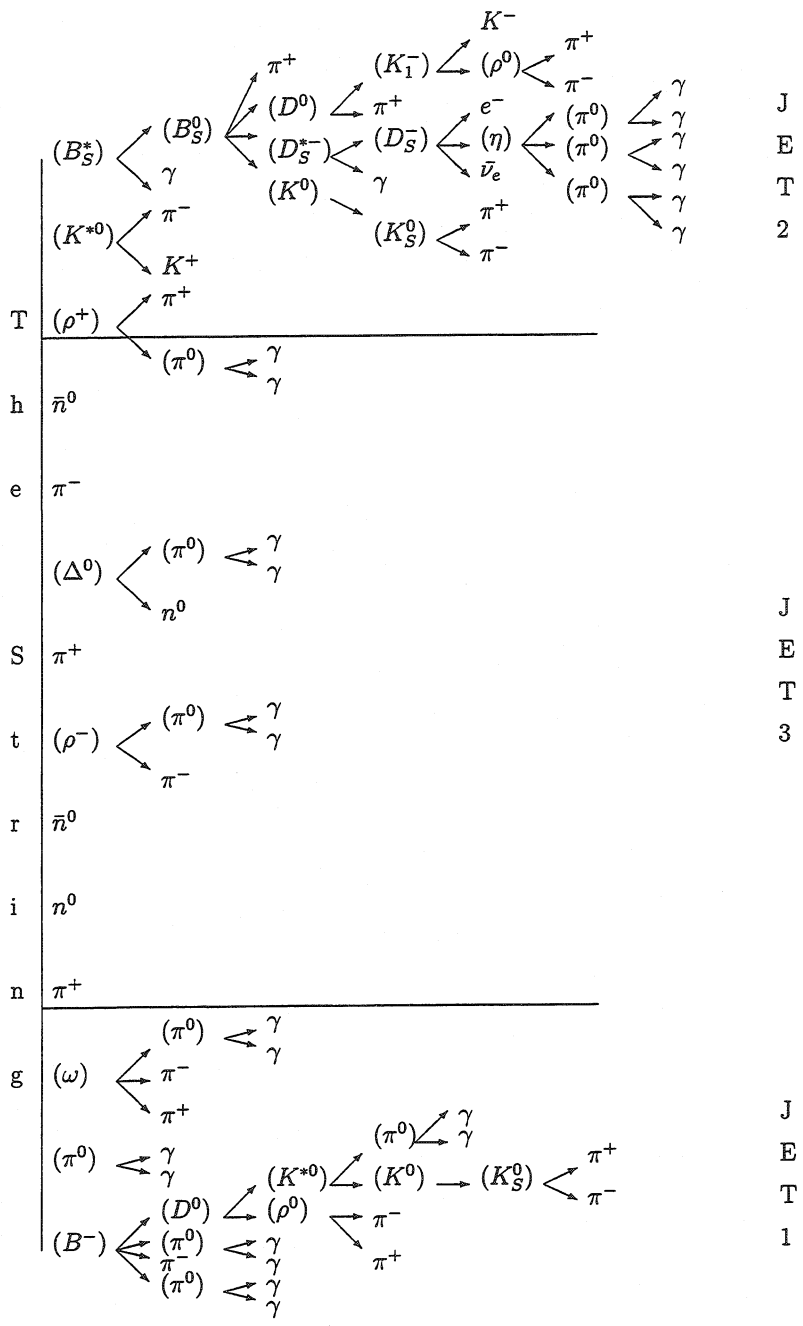

Figure 8.6: A string cascading into a threejet event.

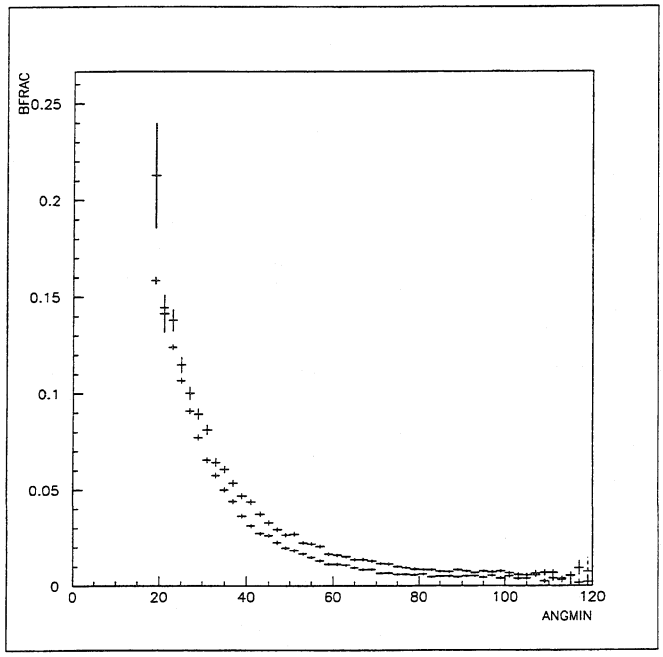

Figure 8.7: The variable bfrac as a function of the smallest angle, angmin, to a jet. The lower points are the result of using all particles, the upper the result of using only the charged particles.

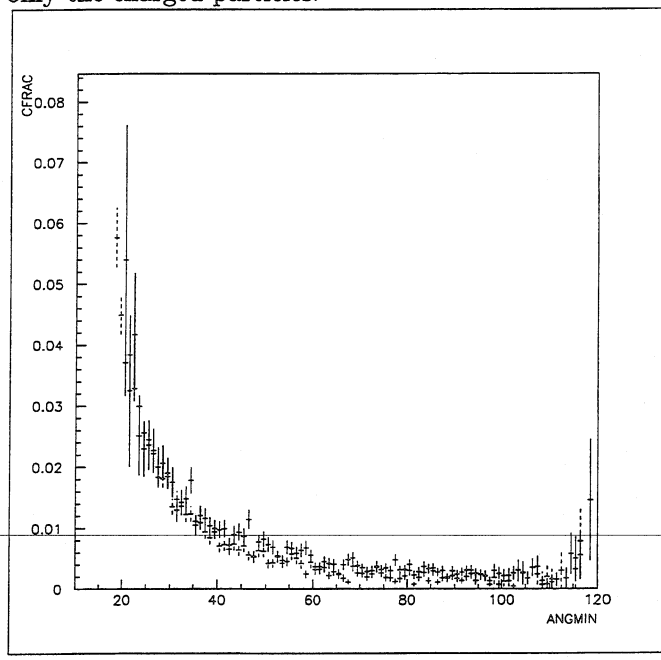

Figure 8.8: The variable cfrac as a function of the smallest angle, angmin, to a jet. The lower points are the result of using all particles, the upper the result of using only the charged particles.

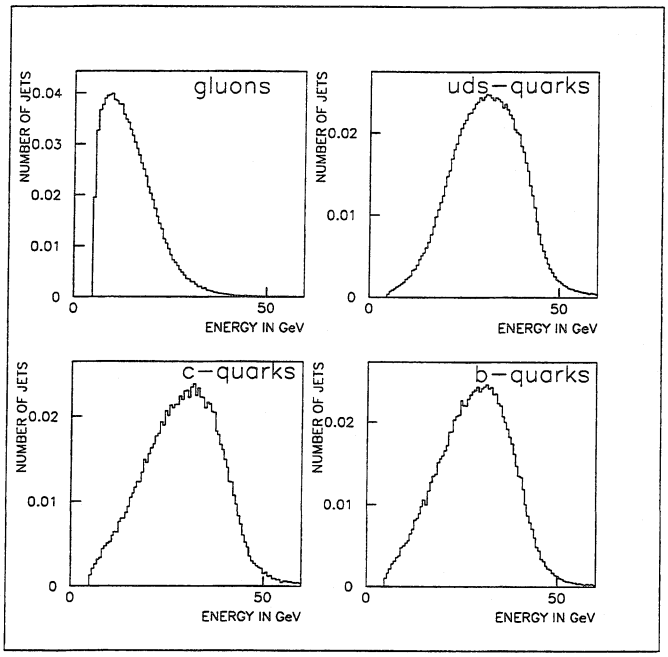

Figure 8.9: The normalized energy distribution of jets emanating from gluons, light quarks, c-quarks and b-quarks, respectively.

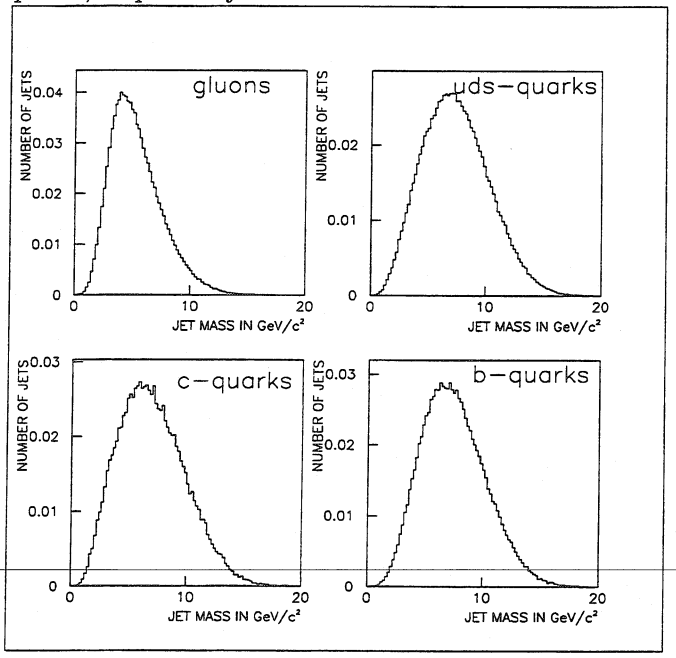

Figure 8.10: The normalized mass distribution of jets emanating from gluons, light quarks, c-quarks and b-quarks, respectively.

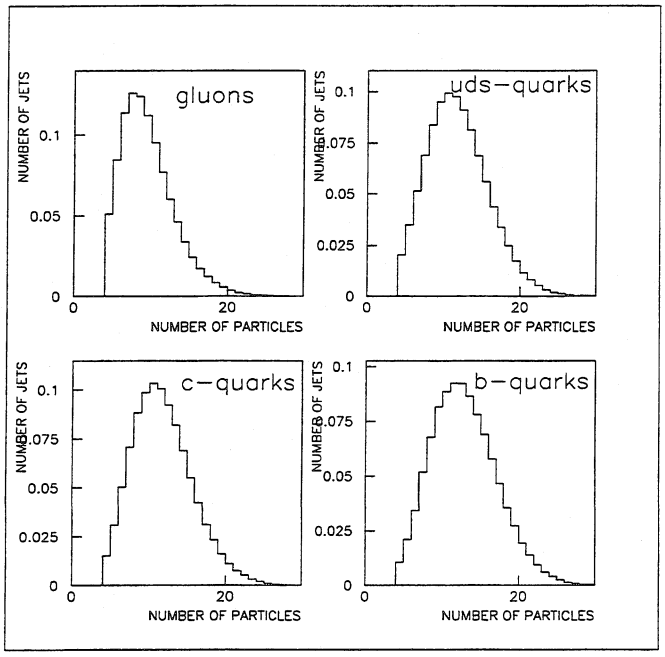

Figure 8.11: The normalized multiplicity of particles in jets emanating from gluons, light quarks, c-quarks and b-quarks, respectively.

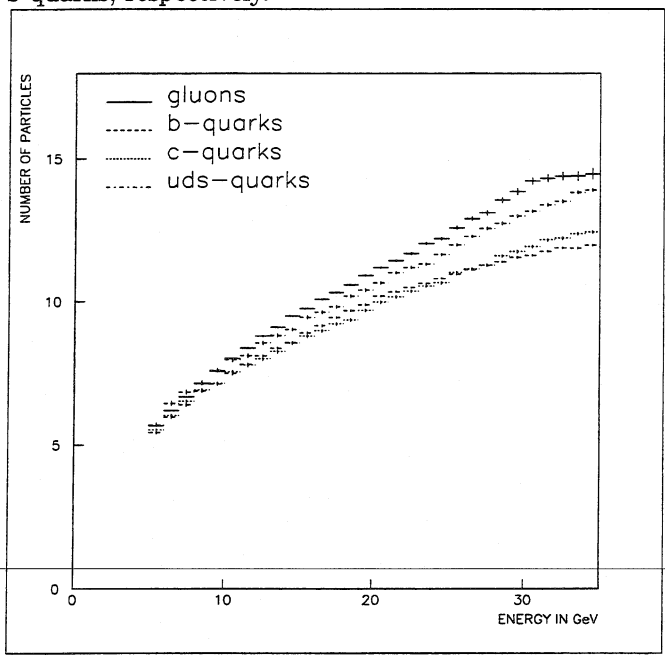

Figure 8.12: The number of particles as a function of the jetenergy in jets emanating from gluons, b-quarks, c-quarks and light quarks.

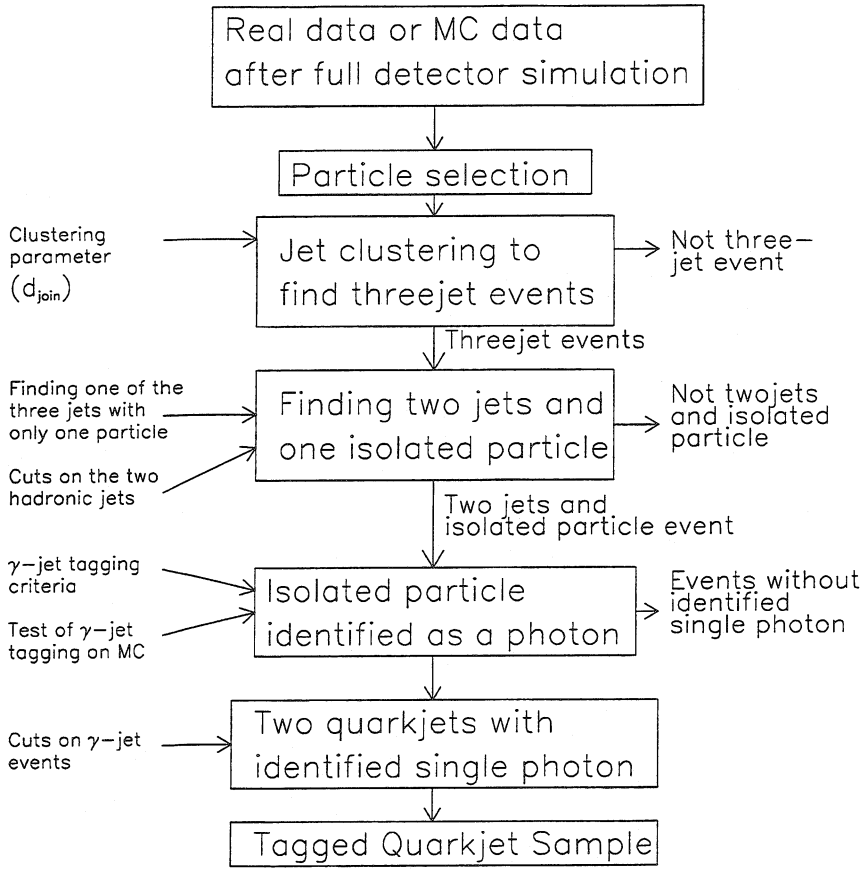

Figure 8.13: How to obtain the tagged quark sample.

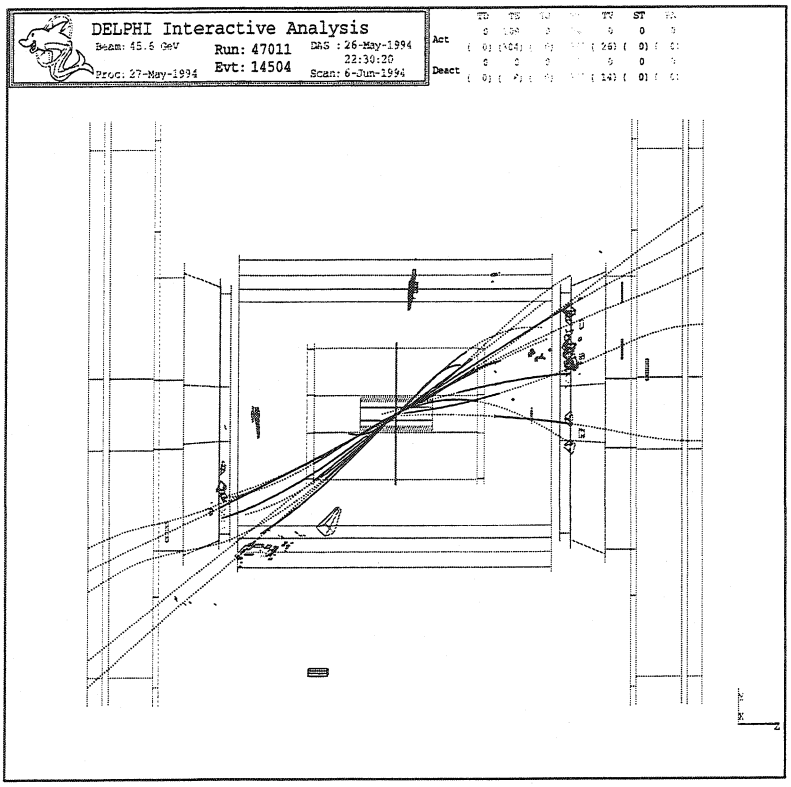

Figure 8.14: A $Z^0 \rightarrow q\bar{q}\gamma$ event, seen in the DELPHI detector. This event has two well collimated jets at a polar angle of about 45 o and a hard γ of $\sim$ 15 GeV is seen in the HPC.

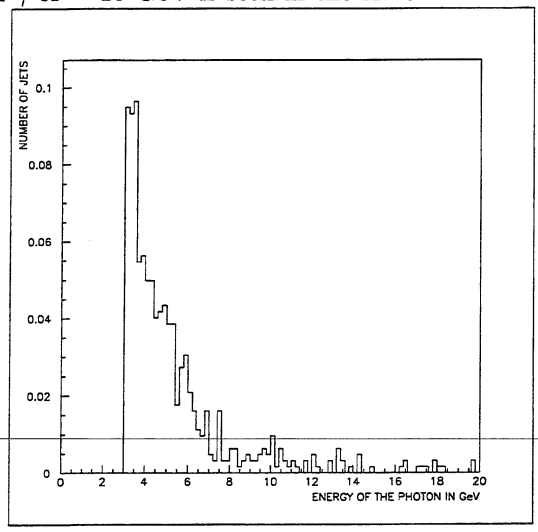

Figure 8.15: The energy spectrum of the γ in $q\bar{q}\gamma$-events from 1993 simulated data.

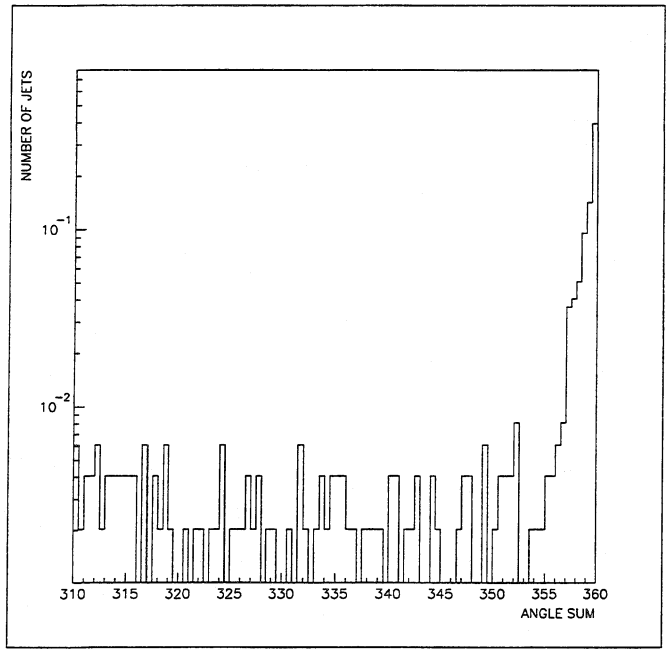

Figure 8.16: The sum of the three angles in the $q\bar{q}\gamma$-events from 1993 simulated data.

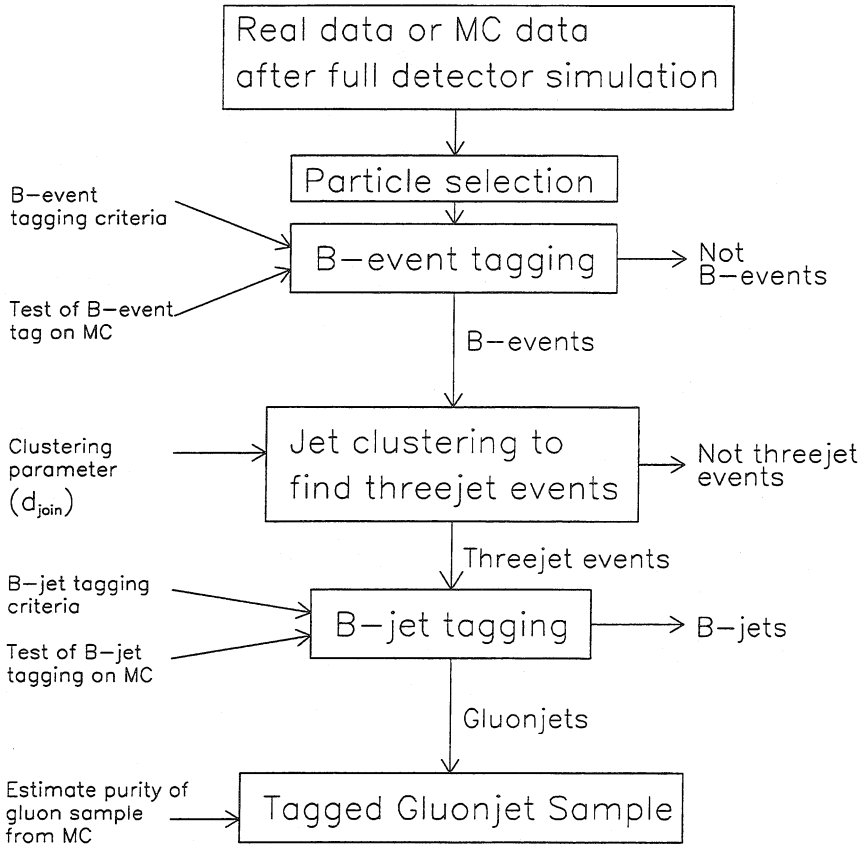

Figure 8.17: How to obtain the tagged gluon sample.

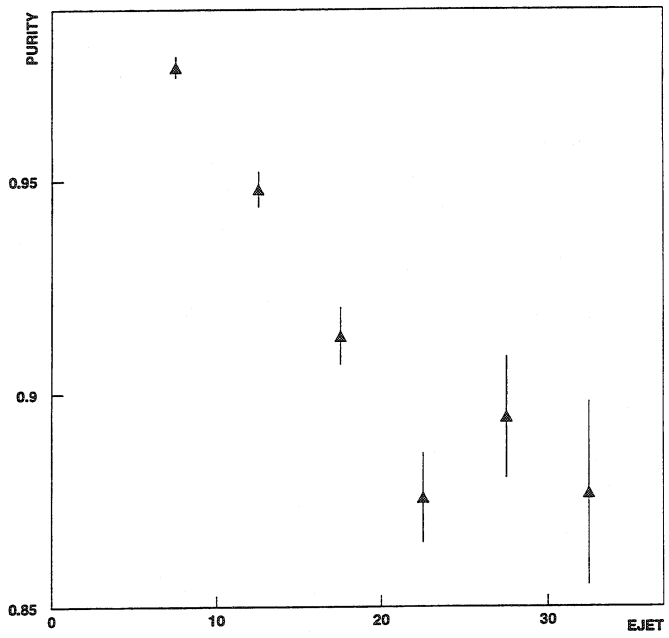

Figure 8.18: The purity of the hard tagged gluon jet sample, as a function of the measured jetenergy. The tagging corresponds to the cuts given in the text.

Chapter 9

Jet classification using the jet energy

The jet energy is by far the most discriminant variable in quark/gluon jet classification because the gluon jets are normally produced in a bremsstrahlung like process and hence more likely to have lower energy than the radiating parton (which can be a quark or another gluon). The gluon jet energy spectrum is therefore much softer than the quark jet energy spectrum. The jet energy, thus, provides a good separation between quark jets and gluon jets as can be seen from figure 9.1, which shows the energy spectrum of quark jets and gluon jets in $e^+e^- \rightarrow q\bar{q}g$ events.

Since the strategy of this analysis is to divide the quark/gluon jet identification into one part based on the jet energy and another part based on the fragmentation, these two parts will be discussed separately. This chapter is devoted to the quark/gluon jet classification based on the energy.

First, the definitions of efficiency and purity will be discussed. The study of the jet energy based classification itself is based on data simulated using the second order QCD matrix element option in the JETSET 7.4 [7] Monte Carlo without any detector simulation.

9.1 Definition of efficiency and purity

Given the jet energy spectrums $f_q(E_{jet})$ for quark jets and $f_g(E_{jet})$ for gluon jets, the efficiency (ε) for a specific jet energy cut E_{cut}, is the fraction of the jets passing the cut. Formally, it is defined as

$$\varepsilon_g(E_{cut}) = \frac{\int_0^{E_{cut}} (f_q(E) + f_g(E))\, dE}{\int_0^{\infty} (f_q(E) + f_g(E))\, dE}$$

for gluon jets and

$$\varepsilon_q(E_{cut}) = \frac{\int_{E_{cut}}^{\infty} (f_q(E) + f_g(E))\, dE}{\int_0^{\infty} (f_q(E) + f_g(E))\, dE}$$

for quark jets.

The purity (π), i.e. the relative number of correctly identified jets (either as a gluon jet or a quark jet) of all the jets accepted by the cut E_{cut}, is formally defined as

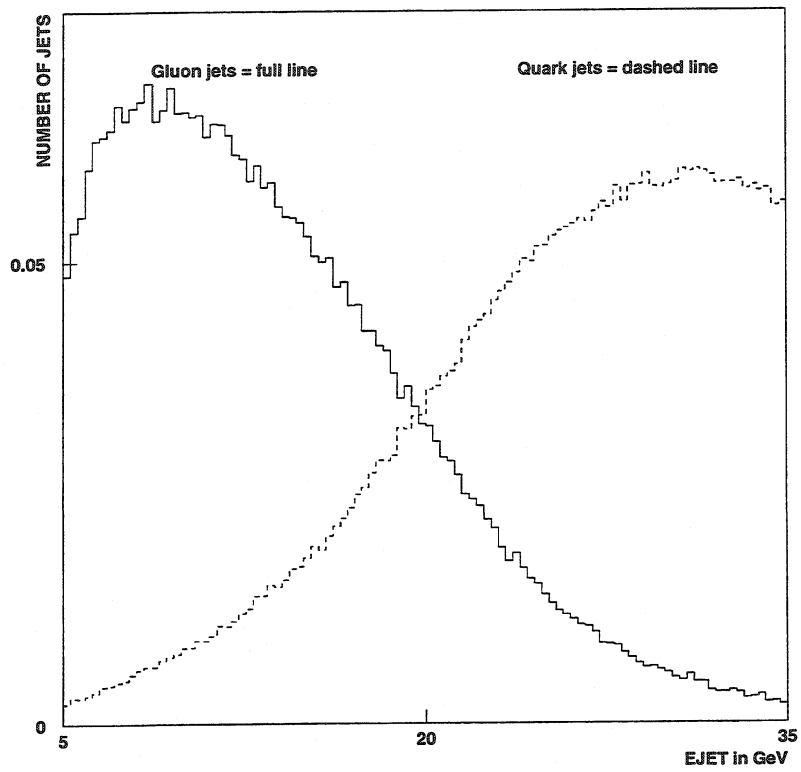

Figure 9.1: The quark jet and gluon jet energy spectrums. (The integral is normalized to one).

$$\pi_g(E_{cut}) = \frac{\int_0^{E_{cut}} f_g(E)\,dE}{\int_0^{E_{cut}} (f_q(E) + f_g(E))\,dE}$$

for gluon jets and

$$\pi_q(E_{cut}) = \frac{\int_{E_{cut}}^{\infty} f_q(E)\,dE}{\int_{E_{cut}}^{\infty} (f_q(E) + f_g(E))\,dE}$$

for quark jets.

In order to account for the actual number of quark jets and gluon jets in the test sample, it is assumed that $\int_0^\infty f_q(E)\,dE = N_q$ (the number of quark jets) and $\int_0^\infty f_g(E)\,dE = N_g$ (the number of gluon jets) in these formulas.

The definitions above may sometimes be a bit confusing. The E_{cut} does not have to be the same for quark jets and gluon jets but it should of course be chosen so that $E^g_{cut} \leq E^q_{cut}$.

When the two cuts are equal, $\varepsilon_g + \varepsilon_q = 1$, which means that a jet is always identified,

either as a quark jet or a gluon jet and consequently the classifier works with a 100% efficiency. However, the individual efficiencies ε_g and ε_q are always less than 100%.

The lowest limit of identification, the *random limit*, is equal to the total fraction of jets in each class, i.e. $N_g/(N_g + N_q)$ for gluon jets and $N_q/(N_g + N_q)$ for quark jets, and it will be indicated for all result in the following. With an equal number of gluon jets and quark jets, the random limit is 50% and a method giving a purity of 50% has not provided any identification at all.

9.2 Efficiency and purity on a jet to jet basis

With the f_g and f_q functions normalised to one, the calculated efficiency/purity refer of course to a sample of equal number of quark jets and gluon jets. The calculated efficiency and purity therefore depend on whether the functions have been normalised or not. Table 9.1 summarises the corresponding efficiency-purity diagrams before and after the actual proportion of quark jets and gluon jets is taken into account.

Event type	π_g (%)			π^{rl} (%)
	$\varepsilon_g = 10\%$	$\varepsilon_g = 30\%$	$\varepsilon_g = 50\%$	
3-jet	91.6 ± 0.3	80.6 ± 0.2	63.1 ± 0.2	33
3-jet Q:G=1:1	95.9 ± 0.3	92.4 ± 0.2	84.4 ± 0.2	50

Table 9.1: Quark/gluon separation using the jet energy. Gluon purities (π_g) are calculated for the natural proportion of quarks and gluons is shown on the first line. The second line shows the purities calculated from normalised f_q and f_g which corresponds to an equal proportion (1:1) of quark jets and gluon jets. The corresponding *random limits* (π^{rl}) are indicated in the rightmost column. The errors are statistical.

9.3 Efficiency/purity on an event basis

If the number of quark and gluon jets are known in an event it is reasonable to use this information in the jet classification. A typical example is the $e^+e^- \rightarrow q\bar{q}g$ threejet event which always contains exactly one gluon jet and two quark jets. Assigning the jet with the lowest energy in a threejet event as the gluon jet gives an overall purity of 78%. The efficiency is obviously 100% since there is always a lowest energetic jet in an event. Figure 9.2 shows how the probability of correctly assigning the lowest energetic jet as the gluon jet in threejet events as a function of the jet energy of the lowest energetic jet.

Evidently the probability for correctly assigning the lowest energetic jet as the gluon decreases as the jet energy approaches the energy of the other jets in the event. In highly non symmetric $e^+e^- \rightarrow q\bar{q}g$ threejet events the probability exceeds 90% for correctly assigning the lowest energetic jet as gluon whereas in twofold symmetric "Y"-events (see

figure 9.3) or threefold symmetric "Mercedes"-events (see figure 9.4) where the corresponding probability is close to the limit value π^{rl}, the assignment is completely random. The purity may be increased at the expense of a lower efficiency if a cut on the lowest jet energy is introduced. There are then two choices, either an event is accepted whenever there is *at least* one jet which fulfil the jet energy cut or an event is accepted whenever there is *exactly* one jet which fulfil the cut. In the first case a 100% efficiency is achieved when the cut is set to $E_{tot}/3$ whereas in the second case a 100% efficiency is not reachable. For instance, with $E_{cut} = E_{tot}/3$ all three jets fulfil $E_{jet} \leq E_{cut}$. Note that a hard cut on the lowest jet energy implies a bias towards asymmetric threejet topologies.

The overall conclusion is that jet energy is effective in identifying gluon jets in asymetric three jet events but to improve the gluon identification for symmetric threejet events one clearly needs to make use of the fragmentation differences between quarks and gluons, e.g. use the variables defined in a following chapter.

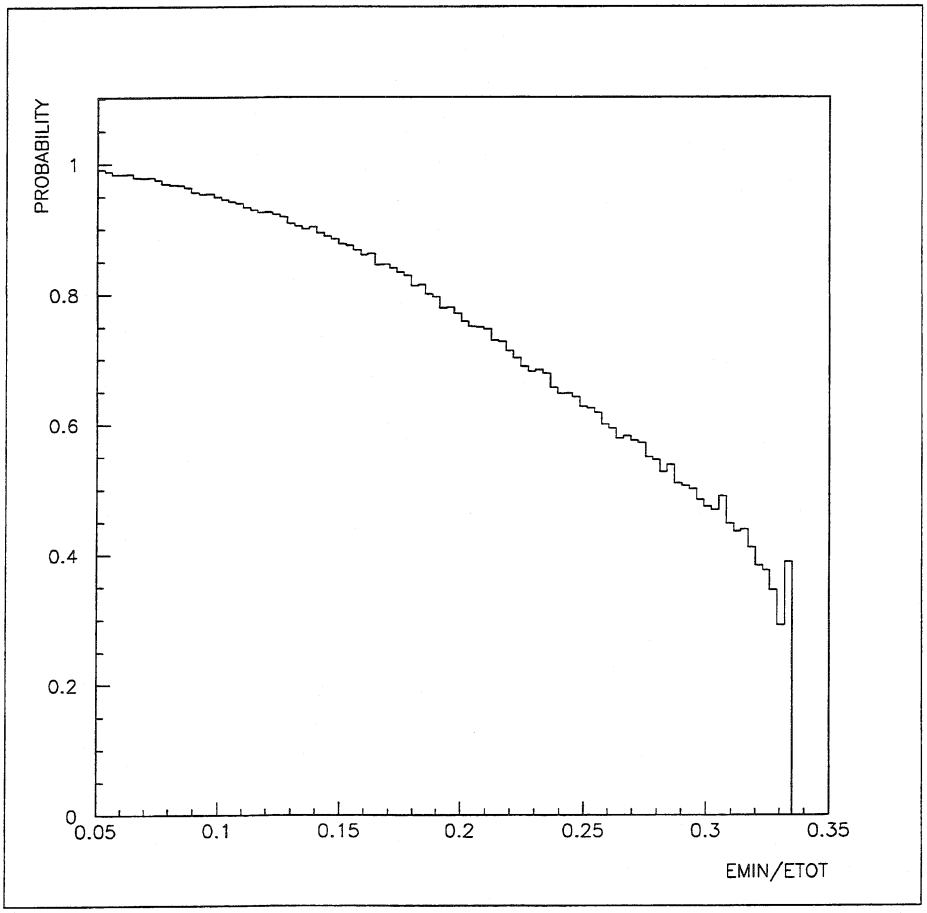

Figure 9.2: The probability of correctly assigning the lowest energetic jet as a gluon jet in $e^+e^- \rightarrow q\bar{q}g$ threejet events, versus the scaled energy of the lowest energetic jet. At $E_{MIN}/E_{TOT} = 0.33$, the probability is 1/3 and the assignment is completely random.

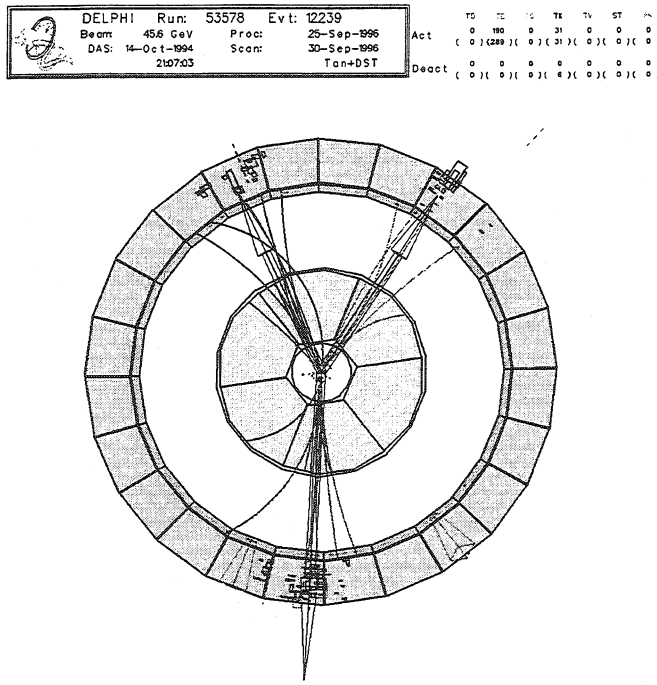

Figure 9.3: An Y-event, where JET2 and JET3 have about the same energy.

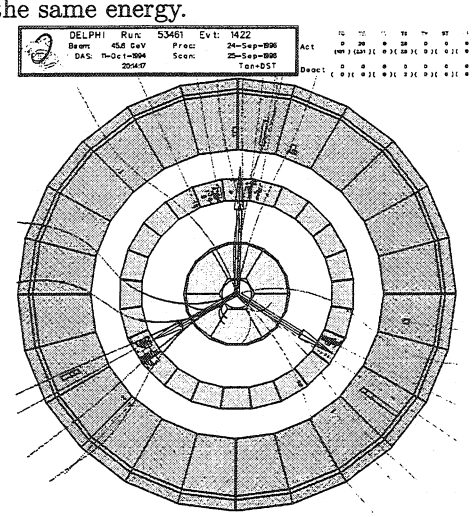

Figure 9.4: A "Mercedes-event" where all three jets have about the same energy.

Chapter 10

Fragmentation variables

Many of the expected differences between gluon jets and quark jets are due to their different colour charge, which leads to differencies in the way that they hadronize. The gluons are expected to give rise to broader jets with a larger number of particles with a softer momentum spectrum than the quarks of the same energy. While the jet energy spectrum can be calculated with pertubative QCD only phenomenological models like the Lund string model are available for the fragmentation studies. For this reason a comparison between Monte Carlo and real data is of vital importance. Described earlier is the selection of samples of jets. In this chapter different variables sensitive to the fragmentation will be described.

10.1 Fragmentation sensitive variables for gluon and quark jets

Fragmentation differencies may be seen already from the four-momentum of the most energetic particle, the *leading* particle of the jet, see [29]. A rather natural choice is therefore to chose a set of variables based on the four-momentum of the most energetic particles of the jet and study their energy dependence.

A set of more complex variables can improve, at least in theory, the identification by combining the four-momentum of several particles in different groups. These groups are discussed in section 10.1.2.

Finally, the two sets of fragmentation sensitive variables have been tested for jet energies from 5 GeV (which is the low cut on the jet energy) up to 35 GeV. (This energy interval is hereafter referred to as the full energy range).

10.1.1 Variables based on leading particles

The first set of fragmentation sensitive variables that was tried consisted of:

Q_M^2, the square of the invariant jetmass;

p_{Li}, the longitudinal momentum of the i:th particle of the jet relative to the jet axis;

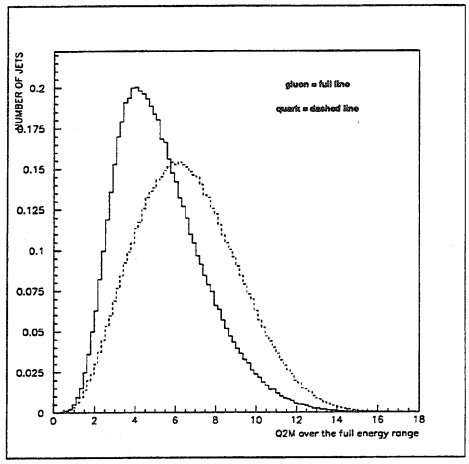

Figure 10.1: The distributions of the fragmentation variable Q_M^2, for gluon jets and quark jets over the full energy range.

p_{Ti}, the transverse momentum of the i:th particle of the jet relative to the jet axis.

The same set of variables has been used by Fodor et al. [30]. The invariant mass of the jet, Q_M^2, which is a collective (since it is composed of all particles of the jet) scalar variable, is defined as,

$$Q_M^2 = E_{jet}^2/c^4 - p_{jet}^2/c^2$$

where:

- E_{jet}, is the total energy of all the particles of the jet;
- p_{jet}, is the total momentum of all the particles of the jet.

Figure 10.1 shows the Q_M^2 distribution for gluon jets and quark jets over the full energy range. The difference between quarks and gluons is significant but the distributions are to a large extend overlapping.

The longitudinal momentum of the particle in the jet has not the same identification power as the transverse momentum when classifying B-events, see section 7.1. Figure 10.2 shows the normalized longitudinal momentum of the i:th particle, p_{Li}/p_{jet}, for the six leading particles of the jet, i.e. i=1, 2, 3, 4, 5 and 6. Not only the number of particles used were restricted but also the minimum momentum, since a cut was applied such that only particles of the jet with a total momentum, p_i, fulfilling

$$p_i \geq 0.05 \cdot p_{jet}$$

were considered.

The identification power of p_{Li}/p_{jet} as a quark/gluon jet separation variable is illustrated in figure 10.2. There seem to be little difference between quarks and gluons for the

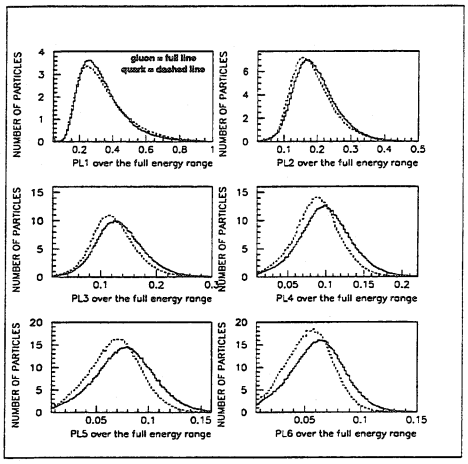

Figure 10.2: The distributions of the normalized longitudinal momentum of the i:th particle, p_{Li}/p_{jet}, for gluon jets and quark jets over the full energy range for the six leading particles of the jet.

most leading particles, but the difference seems to increase for the less energetic particles.

The transverse momentum is an important quantity forming the base for one of the B-tagging methods described earlier in chapter 7. The transverse momentum of the leading particles, p_{Ti} is expected to be of use also in quark/gluon separation. Figure 10.3 shows the normalized transverse momentum of the i:th particle, p_{Ti}/p_{jet}, for gluon jets and quark jets over the full energy range, for the six leading particles of the jet. The same momentum cut, as was used for the longitudinal momentum, was applied. The sensitivity of the transverse momentum to the fragmentation differencies (compared to the longitudinal momentum) seems to be much more higher, something that was also found in [31].

10.1.2 The Fodor moments

The second set of fragmentation variables consists of a type of moments, proposed by Fodor et al. [32]. These moments, *Fodor moments*, M_{nm}, are based on the same jet variables as above but are assumed to be more sensitive to the fragmentation differences. They are defined as,

$$M_{nm} = \sum_{i=1}^{N_{jet}} (\frac{p_{Ti}}{E_{jet}})^n \cdot \eta_i^m$$

where the sum is taken over all particles, N_{jet}, of the jet and where:

- p_{Ti} is the transverse momentum of the i:th particle of the jet;

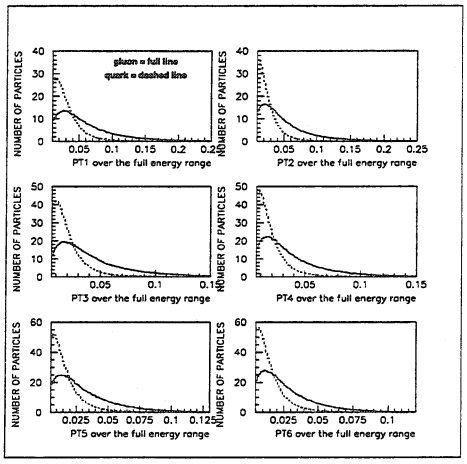

Figure 10.3: The distributions of the normalized transverse momentum of the i:th particle, p_{Ti}/p_{jet}, for gluon jets and quark jets over the full energy range for the six leading particles of the jet.

- η_i the pseudo rapidity of the i:th particle of the jet;
- E_{jet} is the total jet energy.

The pseudo rapidity, η_i, of a particle is defined as,

$$\eta_i = -\frac{1}{2} \cdot \ln \frac{(E_i + p_{Li})}{(E_i - p_{Li})}$$

where:

- E_i is the energy of the i:th particle;
- p_{Li} is the longitudinal momentum of the i:th particle of the jet.

After a study of the 10×10 first M_{mn} the following eight, M_{10}, M_{11}, M_{14}, M_{15}, M_{16}, M_{25}, M_{26} and M_{27}, were seen to show the largest difference between quark jets and gluon jets, [33] and were therefore chosen for the analysis.

Negative signed p_L

Using LUCLUS (and for that matter any jet algorithm) does not imply that all p_{Li} for the particles of the jet are positive and the Fodor moments can occasionally be negative if $p_{Li} < 0$. One way of avoiding this is to reassign all particles with negative p_{Li} to the jetaxis to which it has the greatest p_L. The jetaxis itself can not be changed, because, if the problem is solved for one particle, this can cause similar problems for another particle when changing the axis, so this is unfortunately not a recipe that will work in every situation. Especially in events with two hadronic jets and a final state photon, the

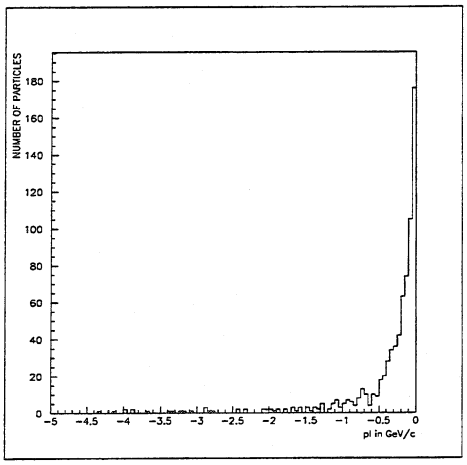

Figure 10.4: The distribution of the 748 particles in the 1994 simulated data with a $p_L < 0$.

probability of having particles with $p_L < 0$ is not zero, although these particles often have small p_L and thus contribute little to the Fodor moments.

In 295 721 threejet events, the number of particles with a $p_L < 0$ was 748 (in the 1994 simulated data) and in addition the p_L-values were rather low, see figure 10.4, so in the end, it was decided to simply exclude these particles in the calculation of the Fodor moments.

Normalization of the Fodor moments

The Fodor moments generally depend on the jet energy and it is convenient to have the Fodor moments scaled to a similar size (~ 1), that is to normalize them to the average of all jets of the same energy, $< M_{nm}(E_{jet}) >$. This quantity will also permit an easier use in the multivariable discrimination methods that will be applied later in the analysis. These *normalized* Fodor moments, F_{nm}, are defined as

$$F_{nm}(E_{jet}) = \frac{M_{nm}(E_{jet})}{< M_{nm}(E_{jet}) >}$$

Here the normalization is calculated from the jet sample itself and one could therefore expect that some of the model dependence on the absolute value of the Fodor moments is divided out. The residual energy dependence in $F_{mn}(E_{jet})$ which is apparent in the upper plot in figure 10.5 could in principle have been avoided if the scale factor ($< M_{mn}(E_{jet}) >$) had been calculated with an equal amount of gluon jets and quark jets, rather than taking all jets in each energy bin. The advantage with the rescaling used, is that it can be applied to both real data as well as to Monte Carlo generated data.

The following eight normalized Fodor moments, F_{10}, F_{11}, F_{14}, F_{15}, F_{16}, F_{25}, F_{26} and F_{27} were chosen as input for the future analysis. Figures 10.6 and 10.7 show the difference

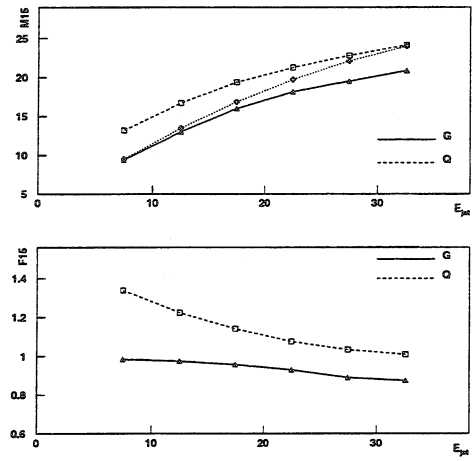

Figure 10.5: The upper plot shows the average value of M_{15} versus the jet energy for detector simulated data. Gluon jets = full line, quark jets = dashed line. The dotted line represents the scale factor ($< M_{15} >$) which is calculated as an average over all jets in each of the energy bins. The lower plot shows the rescaled average Fodor moment $F_{15} = M_{15} / < M_{15} >$ for quark and gluon jets. Gluon jets = full line, quark jets = dashed line.

in the fragmentation variables of these eight selected normalized Fodor moments for gluon jets and quark jets over the full energy range. It is obvious from these figures, that the identification power is very different for different moments and F_{15} is for example much more powerful than F_{25}. What is not clear from these figures is the correlation between the variables. As a cross check the multivariable analysis described later was therefore compared to an analysis using only one variable (F_{14}).

10.2 The energy dependence of the fragmentation variables

The dominating role played by the jet energy in the classification of gluon jets and quark jets has already been demonstrated. Therefore the energy spectrum was divided up in a number of bins. Six bins were used, each with a width of 5 GeV, i.e. 5-10 GeV, 10-15 GeV, 15-20 GeV, 20-25 GeV, 25-30 GeV and 30-35 GeV were the selected energy intervals. In the following, they will be referred to as the six energy bins. The obvious next thing to study is the distribution of the fragmentation variables for gluon jets and quark jets in these six energy bins.

Figure 10.8 shows the fragmentation variable Q_M^2, (defined in section 10.2) for gluon jets and quark jets in the six energy bins. The identification power of Q_M^2 seems to increase with increasing jet energy.

Figure 10.9 shows the difference between the gluon jet and the quark jet distributions

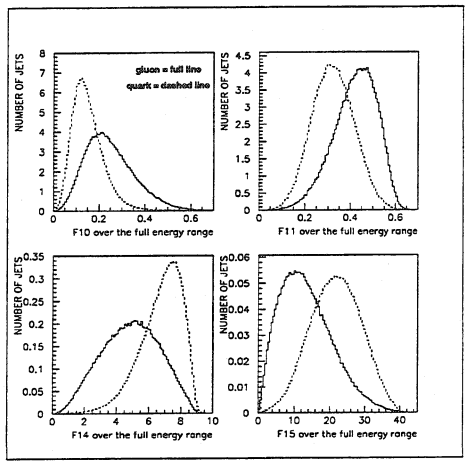

Figure 10.6: The distributions of the fragmentation variables, the normalized Fodor moments, F10, F11, F14 and F15 for gluon jets and quark jets, respectively, over the full energy range.

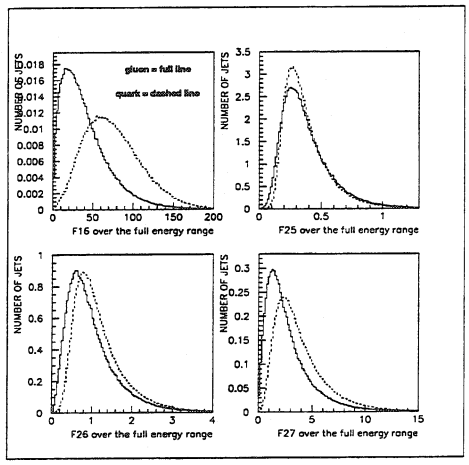

Figure 10.7: The distributions of the fragmentation variables, the normalized Fodor moments, F16, F25, F26 and F27 for gluon jets and quark jets, respectively, over the full energy range.

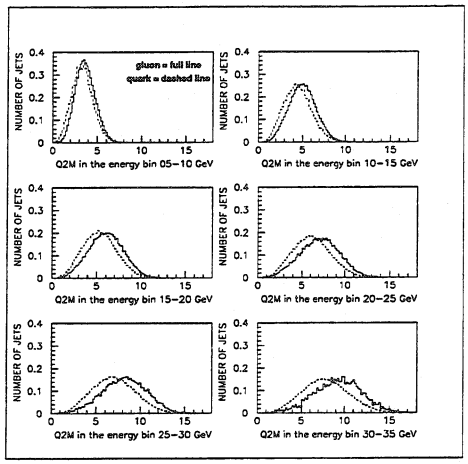

Figure 10.8: The distributions of the invariant jetmass, Q_M^2, for gluon jets and quark jets in the six energy bins.

of the normalized longitudinal momentum (p_{L_1}/p_{jet}) for the leading particle of the jet in the six energy bins. The differencies between gluon jets and quark jets, that were small already over the full energy range, are still there but it is clear that the most energetic particles are the most important, since the differencies for the second to the sixth leading particle of the jet are steadily decreasing and are hardly significant, see appendix B.

The normalized transverse momentum distributions (p_{T_1}/p_{jet}) for the leading particle of the jet are given in figure 10.10. The differencies between gluon jets and quark jets, that were quite large over the full energy range are clearly observable in all bins and the same is true for the other leading particles of the jet, see appendix C.

Figures 10.11 and 10.12 show the normalised Fodor moments F15 and F25 for gluon jets and quark jets in the six energy bins. It is noticeable that the identification power seems to be much higher for F15 than F25 and indeed all moments with the p_T/E_{jet}-factor squared, results in moments with a smaller difference between quarks and gluons. See appendix D for all the other Fodor moments.

The Fodor moments seem to have a similar power as quark/gluon jet classification variables in all energy bins.

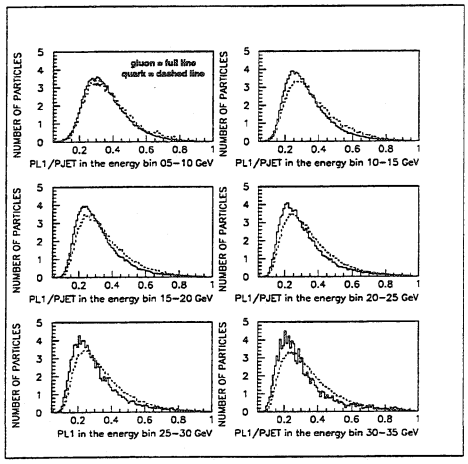

Figure 10.9: The distributions of the normalized longitudinal momentum, p_{L1}/p_{jet}, for the leading particle of the jet for gluon jets and quark jets, in the six energy bins.

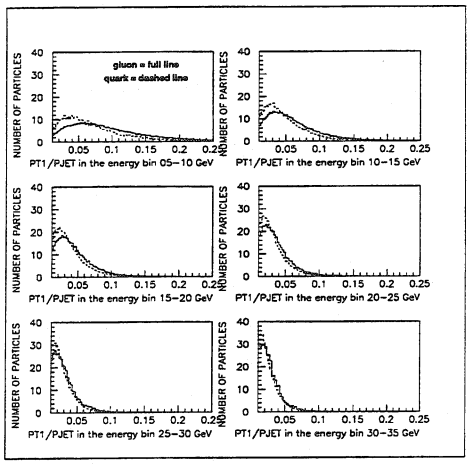

Figure 10.10: The distributions of the normalized transverse momentum, p_{T1}/p_{jet} for the leading particle of the jet for gluon jets and quark jets, in the six energy bins.

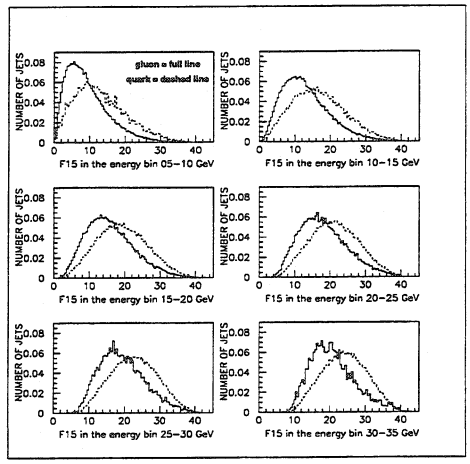

Figure 10.11: The distributions of the normalised Fodor moment, F15, for gluon jets and quark jets in the six energy bins.

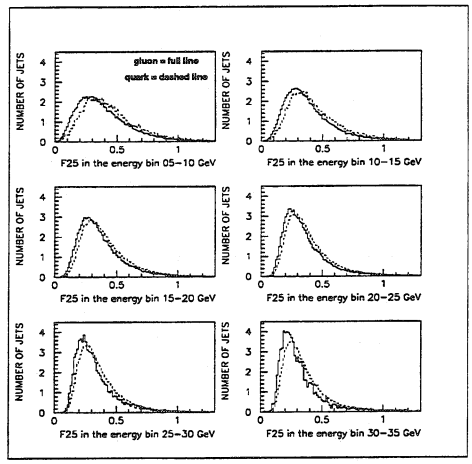

Figure 10.12: The distributions of the normalised Fodor moment, F25, for gluon jets and quark jets in the six energy bins.

Chapter 11

Jet classification using fragmentation

As can be seen in the overview flow-chart of the analysis in figure 5.1, it is at this step of the analysis that the major part of the study, i.e. the test of the fragmentation as a quark/gluon jet separator, will take place. Both linear and nonlinear multivariable analysis (see chapter 4) has now been described and the selection of jets and events has been made (see chapter 6). The quark/gluon jet separation using jet energy has been studied (see chapter 9) and the fragmentation variables has furthermore been presented (see chapter 10). What remains to be done is to demonstrate how the quark/gluon jet classifier performs on real data without using the jet energy. In contrast to the jet energy, it is not expected that all the fragmentation properties can be contained in a single variable, although M. A. Graham et al. [34, 35] did some attempts to do so. The quark/gluon jet classification using the fragmentation is a multidimensional statistical problem which may be solved by using methods such as Fisher statistical discrimination or Artificial Neural Networks. The effect of adding fragmentation variables to a classification based only on energy is often marginal in most experimental situations since the jet energy provides such a high degree of separation between quark and gluon jets. It is therefore better to use fragmentation variables at colliders (e.g. LEP) for symmetric threejet events where the quark jets and gluon jets have almost the same energy. Fodor et al. [30] showed that what had been achieved in [29] by using an ANN, could be a obtained with a single cut on the jet energy. Naturally, a combination of a cut on the jet energy together with an ANN, trained to see fragmentation differences has been exploited, [30], as well as an ANN, trained on both jet energy and fragmentation at the same time [29].

In an attempt to use only fragmentation differencies, the jet energy has been hidden during the training of the ANNs. This was made possible by using training samples containing equal numbers of quark jets and gluon jets at each jet energy, which is also the approach in [30, 33]. Threejet events are mostly *asymmetric* since the gluon jet often has a lower energy than the gluon jets and a large training sample and a long learning time is needed for the ANN to perform well on *symmetric* events. It is also difficult to hide the E_{jet} information in order to avoid that the ANN "discovers" correlations between the jets in the event and consequently the jets are always presented one at a time to the ANN. In the training of the ANN to identify jets from the fragmentation differencies, a special E_{jet}-*balanced* training sample was produced. This sample consisted of an equal amount (i.e. 500) of quark jets and gluon jets at each jet energy (in bins of 5 GeV) and in this

way the jet energy was hidden from the ANN.

11.1 Using the Fisher discrimination method

The linear statistical discrimination method i.e. the Fisher discrimination analysis, earlier described in section, 4.1, have been used to classify jets. Since the purpose of applying the method is to separate gluon jets from quark jets the following arbitrary choice was made:

- the quark jet is the signal;
- the gluon jet is the background.

Using the results presented earlier (see section 10.1.2), the following eight fragmentation sensitive discrimination variables, M10, M11, M14, M15, M16, M25, M26 and M27 were used in the Fisher analysis.
The quark jet variables x_i^Q are associated with the Fodor moment (using only quark jets) as input variables:

- $x_1^Q = M10^Q$;
- $x_2^Q = M11^Q$;
- $x_3^Q = M14^Q$;
- $x_4^Q = M15^Q$;
- $x_5^Q = M16^Q$;
- $x_6^Q = M25^Q$;
- $x_7^Q = M26^Q$;
- $x_8^Q = M27^Q$.

In table 11.1 typical values and their rms values are shown for the Fodor moments for quark jets, in the six energy bins.
In the same way the gluon jet variables x_i^G are associated with the Fodor moment (using only gluon jets) as input variables:

- $x_1^G = M10^G$;
- $x_2^G = M11^G$;
- $x_3^G = M14^G$;
- $x_4^G = M15^G$;
- $x_5^G = M16^G$;

$< M^Q >$	5-10 GeV	10-15 GeV	15-20 GeV	20-25 GeV	25-30 GeV	30-35 GeV
M10	.246 (.113)	.202 (.093)	.172 (.076)	.153 (.065)	.140 (.058)	.129 (.052)
M11	.413 (.100)	.380 (.099)	.351 (.094)	.330 (.089)	.312 (.085)	.296 (.080)
M14	4.77 (1.75)	5.66 (1.61)	6.29 (1.39)	6.68 (1.22)	6.70 (1.10)	7.21 (.976)
M15	13.2 (7.22)	16.7 (7.39)	19.4 (7.07)	21.3 (6.76)	22.8 (6.50)	24.2 (6.18)
M16	39.4 (29.8)	52.8 (32.8)	63.7 (33.6)	72.0 (33.9)	79.2 (34.1)	85.7 (33.8)
M25	.448 (.232)	.433 (.216)	.405 (.194)	.382 (.181)	.362 (.174)	.350 (.171)
M26	1.17 (.732)	1.23 (.718)	1.21 (.662)	1.18 (.634)	1.16 (.619)	1.15 (.615)
M27	3.26 (2.45)	3.67 (2.58)	3.79 (2.46)	3.84 (2.42)	3.89 (2.41)	3.97 (2.41)

Table 11.1: The average values and the rms value (in parentheses) of the Fodor moments for quark jets in the six energy bins, 5-10 GeV, 10-15 GeV, 15-20 GeV, 20-25 GeV, 25-30 GeV and 30-35 GeV.

$< M^G >$	5-10 GeV	10-15 GeV	15-20 GeV	20-25 GeV	25-30 GeV	30-35 GeV
M10	.306 (.115)	.251 (.099)	.211 (.084)	.187 (.073)	.174 (.064)	.162 (.057)
M11	.461 (.085)	.427 (.090)	.395 (.091)	.370 (.089)	.355 (.085)	.339 (.082)
M14	3.76 (1.63)	4.80 (1.59)	5.56 (1.47)	6.05 (1.34)	6.34 (1.19)	6.58 (1.06)
M15	9.39 (6.00)	13.0 (6.58)	16.0 (6.75)	18.1 (6.72)	19.5 (6.41)	20.8 (6.22)
M16	25.5 (22.6)	38.1 (27.1)	49.3 (29.9)	58.3 (31.7)	64.0 (31.6)	70.5 (32.8)
M25	.393 (.219)	.385 (.204)	.372 (.191)	.356 (.187)	.334 (.181)	.321 (.184)
M26	.928 (.653)	.999 (.651)	1.03 (.636)	1.04 (.649)	1.01 (.637)	1.01 (.663)
M27	2.36 (2.08)	2.77 (2.22)	3.05 (2.28)	3.23 (2.41)	3.22 (2.40)	3.32 (2.36)

Table 11.2: The average values and the rms value (in parenthesis) of the Fodor moments for gluon jets in the six energy bins, 5-10 GeV, 10-15 GeV, 15-20 GeV, 20-25 GeV, 25-30 GeV and 30-35 GeV.

- $x_6^G = M25^G$;
- $x_7^G = M26^G$;
- $x_8^G = M27^G$.

In table 11.2 typical values and the rms value are shown for the Fodor moments for quark jets, in the six energy bins.

Using the technique described in section 4.1, the Fisher coefficients are calculated and they are presented in table 11.3 for the six energy bins.

The Fisher function, F, using the Fisher coefficients above, is finally constructed from

$$F = f0 + f1 \cdot M10 + f2 \cdot M11 + f3 \cdot M14 + f4 \cdot M15 +$$

$$f5 \cdot M16 + f6 \cdot M25 + f7 \cdot M26 + f8 \cdot M27$$

Typical outputs of the normalised distributions of the Fisher function, F, for simulated gluon jets and quark jets in the six energy bins, are shown in figure 11.1. The overall

$F_{coefficient}$	5-10 GeV	10-15 GeV	15-20 GeV	20-25 GeV	25-30 GeV	30-35 GeV
f1	-4.987	-5.867	-6.211	-6.057	-7.863	-8.774
f2	-5.868	-5.568	-5.013	-4.371	-5.285	-5.550
f3	0.2917	0.3151	0.3151	0.2970	0.3821	0.4243
f4	0.07527	0.07239	0.06516	0.05613	0.06732	0.06939
f5	0.02048	0.01816	0.01526	0.01244	0.01433	0.01417
f6	2.078	2.136	1.970	1.640	1.848	1.727
f7	0.5849	0.5664	0.5032	0.4093	0.4572	0.4235
f8	0.1708	0.1556	0.1327	0.1048	0.1153	0.1052
f0	-0.4796	-0.5005	-0.4536	-0.370	-0.4102	-0.3832

Table 11.3: The normalised Fisher coefficients in the six energy bins.

impression from these plots is that although the shape of the Fisher function distribution varies with energy, the separation between the quark and gluon distributions remains more or less the same in all bins. The quark and gluon distributions are also to a large extent overlapping and one can only expect a modest performance of the method. The actual purity and efficiencies that can be obtained with this method will be presented later in this chapter together with the result from ANNs and an analysis using only one variable.

11.2 Using only one fragmentation variable

In the Fisher discrimination method, eight variables was used in order to separate gluon jets from quark jets. Obviously one tries to avoid multi dimensional analysis in favour of simple one dimensional cuts when possible and indeed, in [35] it was found that a cut on an appropriately chosen single variable gave almost the same separation as an ANN, using several input variables and that conclusion also will be tested in this analysis. One-dimensional cuts on highly overlapping distributions can, however, become very inefficient.

It has already been pointed out that some Fodor moments appear to have a better discrimination power than others. A test with the best Fodor moment was made and F14 was chosen for this purpose. The normalised distributions for the single variable F14, for gluon jets and quark jets, in the six energy bins are shown in figure 11.2. The separation between the gluon and quark distributions does not seem be much worse than what was obtained after constructing the Fisher function distributions shown in figure 11.1.

11.3 Using artificial neural networks

In order to hide the jet energy, two different methods were used. In the first method, the jet energy interval was divided into bins and an individual ANN was trained in each bin as proposed in [30]. This method was applied to both sets of input variables, described in section 10.1. The jet energy interval (5-35 GeV) was divided into six bins of 5 GeV. For each E_{jet} bin an ANN was trained on a sample containing an equal number of quark

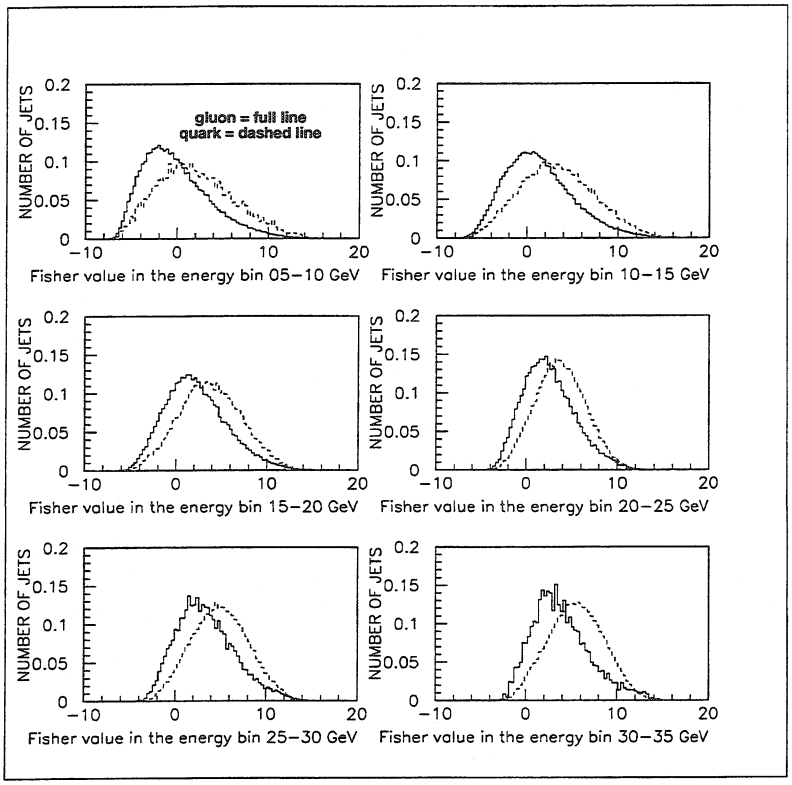

Figure 11.1: The normalised Fisher output distributions for simulated gluon jets and quark jets in the six energy bins.

jets and gluon jets (the sizes of the training samples varied between 2000 and 5000 jet of each kind).

From an experimental point of view there is a disadvantage with having individual ANNs in each energy bin since most likely the output from two or more independent ANNs must be compared with each other in a multi-jet event and there is no guarantee that the gluon output distribution for one ANN does not entirely overlap the quark output distribution from another ANN. Thus, it is in general not possible to apply the principle of ordering the jets according to their distance to the gluon and quark target values and assign the most "gluon-like" as gluons etc. Instead each one of the jets must be classified using some predetermined cut in the output characteristics for the specific ANN used. The probability of correctly identifying all the jets in the event will therefore be equal to the product of the individual jet identification probabilities. In practice such an identification method is not very efficient for events containing more than two jets. A typical output from the ANN can be seen in figure 11.3. After the training when the back propagating algorithm has given the variation of the weights of the network, which are then fixed, the network is used for a test on both real and simulated data.

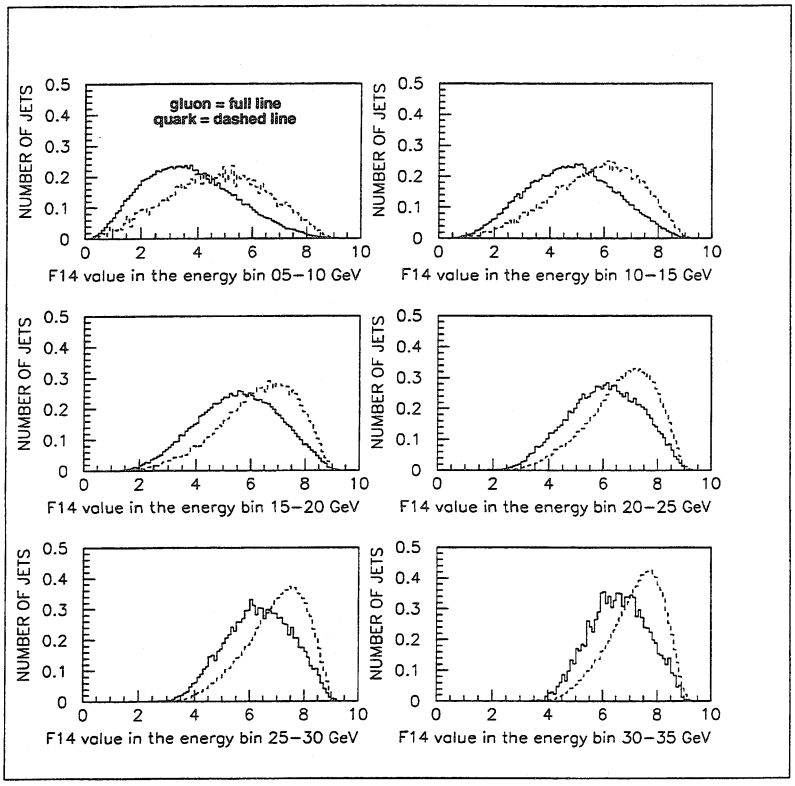

Figure 11.2: The normalised distributions for the single variable F14, for gluon jets and quark jets, in the six energy bins.

The second training method has been described in [33]. It has the advantage that the classifier consists of only one ANN. At the expense of slightly decreased performance on individual jets, a single ANN was trained over the whole energy range but with the training sample still containing an equal number (500) of jets at each jet energy (within 5 GeV bins). In this way the ANN was presented to quark jets and gluon jets with a "flat" energy spectrum.

This method was applied only to the set of input variables which consisted of the variables F_{mn}. In order to help the ANN to correct for the the residual energy dependence (see figure 10.5) the jet energy was added as an extra input variable. The jet energy itself could not be used by the ANN since the training sample contained an equal amount of jets in each energy bin and hence the fact that gluon jets have lower energy, on average, than quark jets was hidden from the ANN. With other words, the jet energy difference between quarks and gluons was not seen by the ANN but the energy dependence of the F_{mn} variables could be used by the ANN.

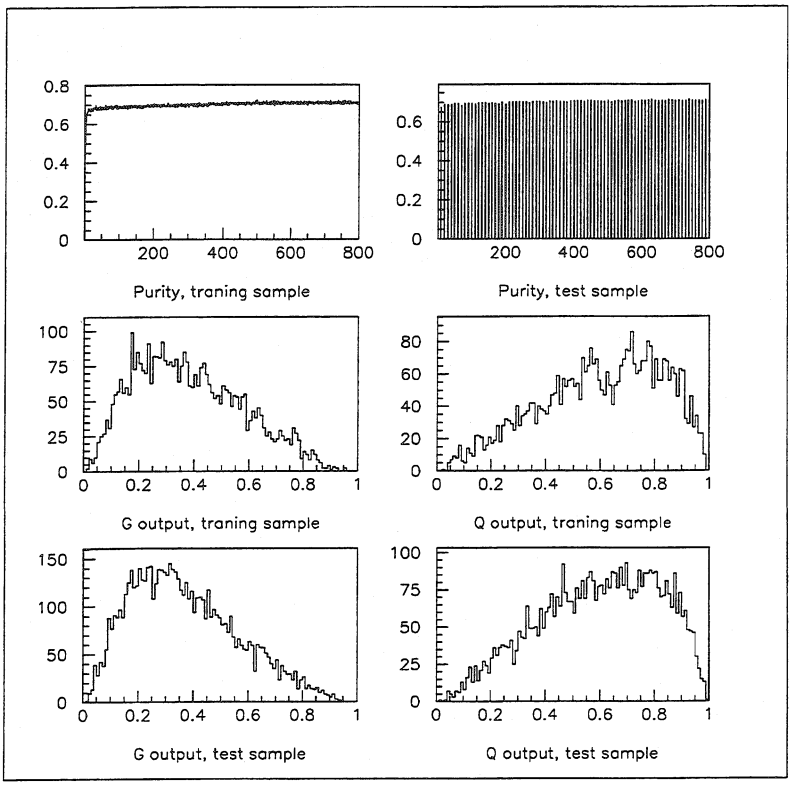

Figure 11.3: The output of the PLPTBIN(10-15) network during the training and testing period.

11.3.1 The training of an ANN

Two types of ANNs have mainly been used in earlier quark/gluon jet identification studies:

- feed forward multi-layered perceptron (MLP) with gradient descent back-propagation learning rule were used in [29, 36, 35, 33];
- learning vector quantization (LVQ) was used by Fodor et al. [30].

There are no strong arguments for the preference of one type over the other. It seems that the MLP type has become the most popular among experimental physicists [37, 38, 39, 40] and it will also be used in this study.

To train the ability of the ANN to distinguish gluon jets from quark jets a Monte Carlo program was used to generate a sample of both jets. The output of the Monte Carlo was a set of variables which were used for the discrimination. These variables were previously defined in section 10.1. The separation of gluon jets and quark jets in any of these variables is modest even though the differencies in the distributions are quite clear. These variables were the input to a multi-layer perceptron with one hidden layer and

one single output. The network was then trained with a back propagating network using some hundreds of gluon jets and about the same amount of quark jets. In order to avoid that the ANN "discovers" correlations between the jets in the event, the jets are always presented one at a time.

In a *naturally* distributed training sample, threejet events are mostly **asymmetric**, i.e. the gluon jets are low energetic and a large training sample and long learning is needed for the ANN to perform well on **symmetric** events. It is furthermore difficult to *hide* the E_{jet} information. As mentioned earlier, a special E_{jet}-**balanced** training sample had therefore to be produced when trying to train an ANN to identify jets only from their fragmentation. The artificial E_{jet}-balanced training sample was composed of an equal amount (500) of quark and gluon jets in each jet energy (in bins of 5 GeV).

It is possible that the ANN introduces a bias in the subsequent analysis due to the classification done by the ANN since the quark/gluon jet separation is normally only one of the steps in an analysis. Consider, for example, an analysis where the difference in particle multiplicity between quark and gluon jets is to be measured. If the quark jets and gluon jets are identified with an ANN and one of the ANN input variables is correlated with the multiplicity of particles in the jet there is a risk for a bias in the measurement towards a higher value. The reason for this bias is of course that the ANN during the training has learned to identify jets with a high/low multiplicity of particles, as gluon/quark jets.

11.3.2 JETNET

The artificial neural network program, the Lund JETNET program (version 3.0), has been used throughout the analysis. The architecture is a feed-forward ANN with a Fortran implementation of the standard gradient descent back propagating algorithm for the weights. The program package has been used with the following settings:

- initial learning rate, $\eta = 0.01$, (this parameter was dynamically decreased during the learning phase);
- momentum term for stabilisation of the learning, $\alpha = 0.5$;
- the weights were randomly initialized in $[-0.5, 0.5]$;
- number of patterns per update = 10;
- normal gradient descent back-propagation with summed square error measure;
- the length of the training phase was 400 epochs.

More details about JETNET and the meaning of the above parameters can be found in [21].

Background (gluons) and signal (quarks) were presented in equal proportion. In order to obtain an universal quark gluon jet classifier, the ANNs are trained on individual jets. This was the approach also in the previous studies [29, 36, 30, 35, 33]. However, when using the ANN in a specific analysis it is clear that as much as possible of the event

information should be used in order to improve the performance. An example of such information could be that there is exactly one gluon jet in a $e^+e^- \rightarrow q\bar{q}g$ threejet event and hence one should normally assign the gluon jet as the one giving the ANN output closest to the gluon target value. Another example is the jet energy. In e^+e^- experiments, where often the total jet energy (E_{tot}) adds up close to the center of mass energy (E_{CMS}), it is normally sufficient to give the jet energy as an input variable to the ANN, whereas in experiments where this is not the case (e.g. hadron colliders), the scaled jet energy (E_{jet}/E_{tot}) is normally better.

11.3.3 Description of the different ANNs used

To distinguish between the three different types of ANNs in the subsequent sections the following notation will be used. For all ANNs, the output target values were chosen to be "1" for the signal (quark jets) and "0" for the background (gluon jets).

Definition of the PLPTBIN ANN

PLPTBIN consists of six different ANNs trained in each E_{jet} bin on the input variable set

$$Q^2_M,\ p_{L1}/p_{jet},\ p_{T1}/p_{jet},\ p_{L2}/p_{jet},\ p_{T2}/p_{jet},\ p_{L3}/p_{jet},\ p_{T3}/p_{jet},$$
$$p_{L4}/p_{jet}, p_{T4}/p_{jet},\ p_{L5}/p_{jet},\ p_{T5}/p_{jet},\ p_{L6}/p_{jet}, p_{T6}/p_{jet}.$$

The layout of the PLPTBIN ANN was 13 input, 26 hidden and one single output node, see figure 11.4. Figure 11.5 shows typical output distributions from the PLPTBIN ANN for gluon jets and quark jets in the six energy bins.

Definition of the FMBIN ANN

FMBIN consists of six different ANNs trained in each E_{jet} bin on the input variable set

$$E_{jet},\ F_{10},\ F_{11},\ F_{14},\ F_{15},\ F_{16},\ F_{25},\ F_{26},\ F_{27}.$$

The layout of the FMBIN ANN had 9 input, 18 hidden and one single output node, see figure 11.6. Figure 11.7 shows typical output distributions from the FMBIN ANN for gluon jets and quark jets in the six energy bins.

Definition of the FMONE ANN

FMONE consists of one single ANN trained over the entire E_{jet} interval, with the same number of quark jets and gluon jets (in each energy bin) on the input variable set

$$E_{jet},\ F_{10},\ F_{11},\ F_{14},\ F_{15},\ F_{16},\ F_{25},\ F_{26},\ F_{27}.$$

The layout of the FMONE ANN had 9 input, 18 hidden and one single output node and looks like the FMBIN ANN, see figure 11.6. Figure 11.8, shows typical output distributions from the FMONE ANN for gluon jets and quark jets in the six energy bins.

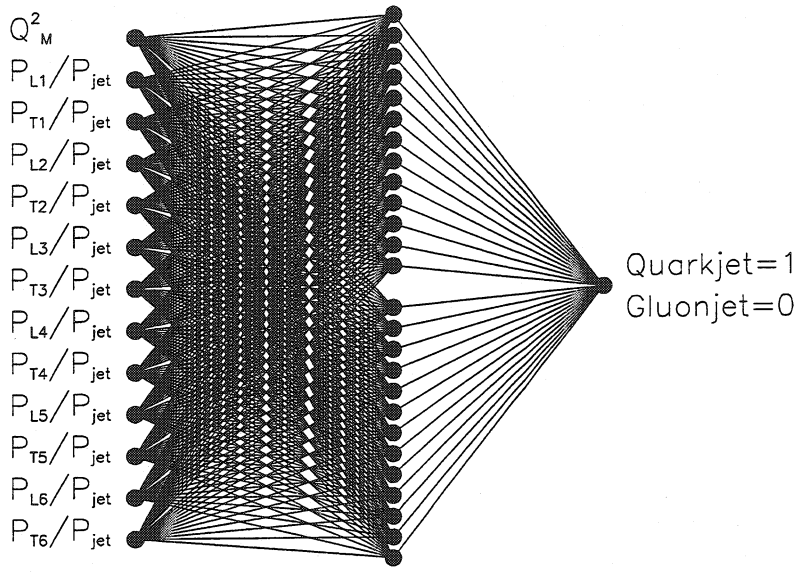

Figure 11.4: The layout of the PLPTBIN ANN with 13 input, 26 hidden and one single output node, with the target values, quark jet=1 and gluon jet=0.

11.4 Comparison of Monte Carlo data with real data

Despite the significant theoretical efforts to prove the functionality of quark/gluon jet separation using ANNs, very few experimental paper on the methods have been published. One reason for this low experimental interest is the lack of clean samples of gluon and quark jets samples to test the method, i.e. the method relies heavily on the assumption that the Monte Carlo simulation used to train the ANNs describes the data perfectly. A test on clean quark jets and gluon jets is therefore essential in order to verify if this assumption is correct. This is important because the calculations performed by the ANN are rather complex since it explores any non linear correlations and anti-correlations found in the input training sample (normally consisting of simulated data).

Therefore, before the presentation of the efficiencies and purities achieved with these fragmentation based quark/gluon jet classifiers, it will be checked that the ANNs behave similarly on real and Monte Carlo data. Samples of gluon jets and quark jets from DELPHI have been collected using the methods described in chapter 8. In total, 28625 lifetime tagged gluon jets were collected and used as input to the ANNs. The resulting ANN output distributions for the lifetime tagged gluon jets in the data together with the tagged and "true" (from the simulation history) gluon jets from PLPTBIN, FMBIN and FMONE in the Monte Carlo are shown in figures 11.9, 11.10 and 11.11, respectively.

The ANN outputs for the lifetime tagged gluon jets in real data follow closely the simulated gluon jets for all three ANNs. The quark jet contamination, of the order of 10% in the high energy bins of the lifetime tagged gluon jets, is manifested as a slight deviation towards the quark jet target value (=1). Figures 11.12, 11.13 and 11.14 shows

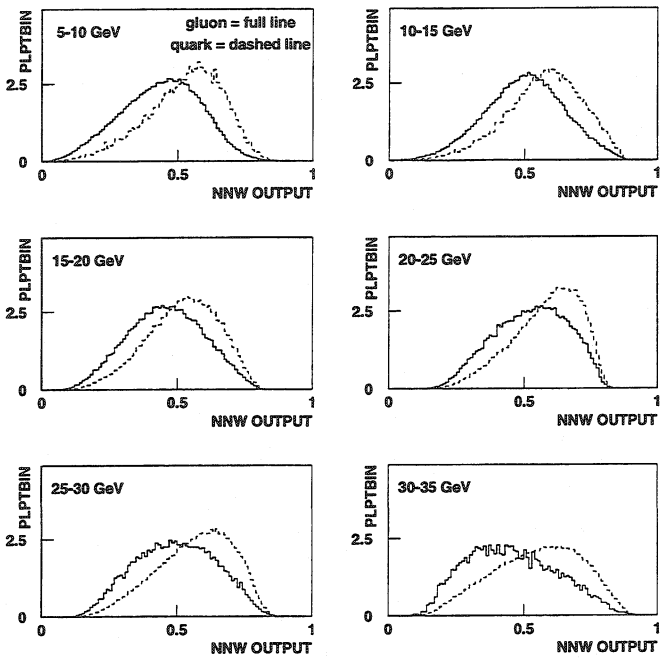

Figure 11.5: Typical output distributions of the PLPTBIN ANN, for an equal number of detector simulated quark jets and gluon jets, in the six energy bins.

the output distributions for the ANN output distributions from the PLPTBIN, FMBIN and FMONE classifiers, respectively, when the tagged quark jets from experimental data and the Monte Carlo were used as input (in the six energy bins). The overall conclusion from this comparison of data with Monte Carlo samples is that the simulated events give a very good description of the data even after the ANNs have been applied. This is of fundamental importance for the analysis.

11.5 Performance of the classifiers used on Monte Carlo data

Since the output distributions from the ANNs tested on data and Monte Carlo are very similar it is meaningful to continue the analysis with an estimation of the efficiency and the purity that the methods can achieve. Given the samples of gluon jets and quark jets tagged according to the method explained earlier in section 8.1, the measurement of the purity was firstly done on Monte Carlo generated jets. This was done for five classifiers, i.e., Fisher, cuts on F14, ANNs trained in each bin with both Fodor moments as input (FMBIN) and the momentum of the leading particles as inputs (PLPTBIN) and an ANN trained with a "flat" energy spectrum with the Fodor moments as input (FMONE).

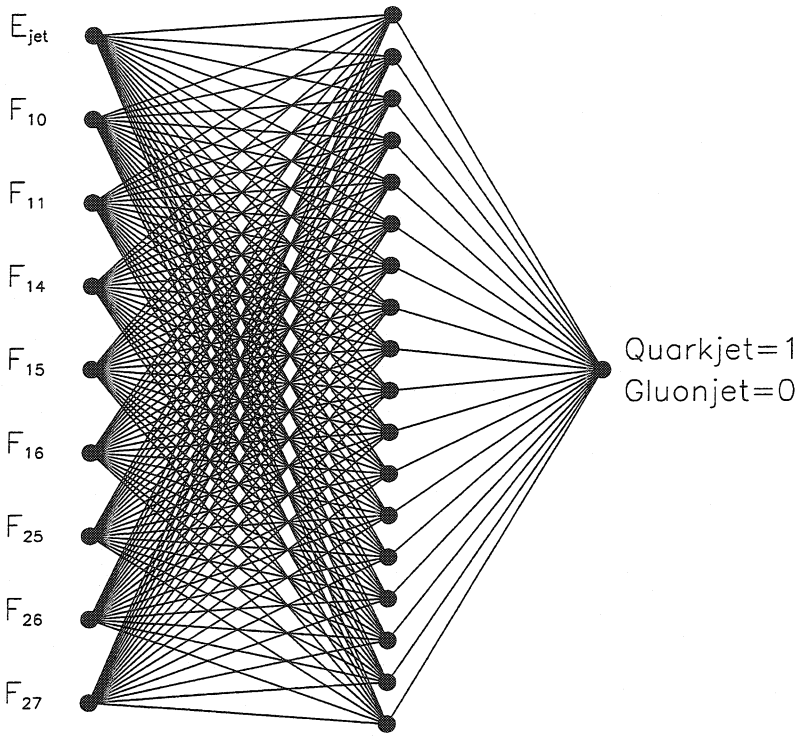

Figure 11.6: The layout of the FMBIN (and FMONE) ANN with 9 input, 18 hidden and one single output node, with the target values, quark jet=1 and gluon jet=0.

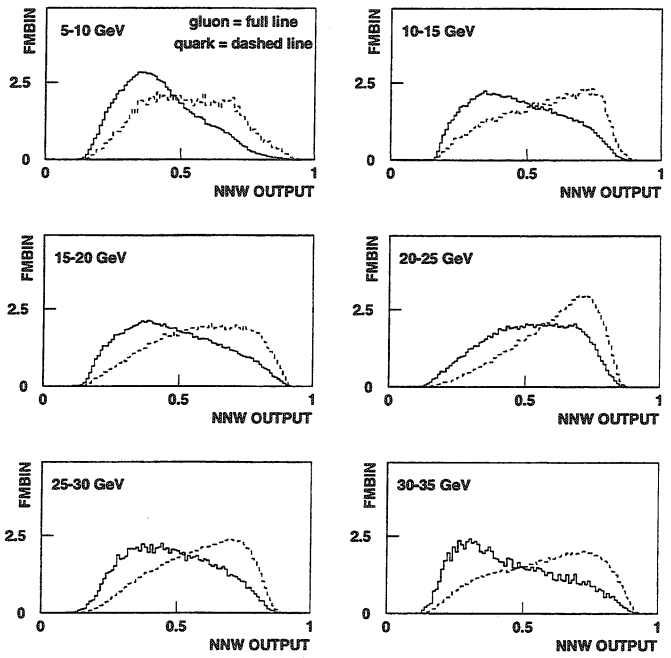

Figure 11.7: Typical output distributions of the FMBIN ANN for an equal number of detector simulated quark jets and gluon jets, in the six energy bins.

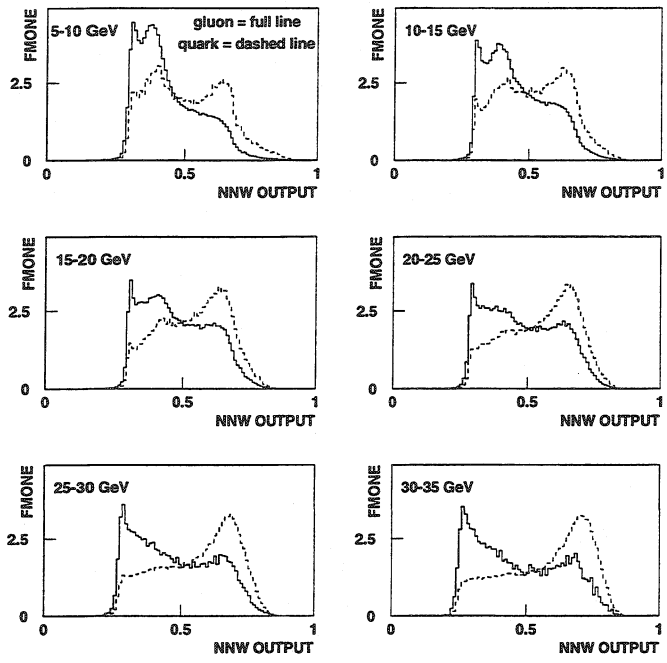

Figure 11.8: Typical output distributions of the FMONE ANN for an equal number of detector simulated quark jets and gluon jets, in the six energy bins.

11.5.1 Purity on a jet to jet basis

The discriminative power of a quark/gluon jet classifier based on the fragmentation is best measured on a sample of jets with similar energy. The jet energy interval 5-35 GeV was divided into six 5 GeV-bins (same binning as in the training). In each jet energy bin the normalised output distributions (f_q and f_g) for quark jets and gluon jets were used to calculate the purity (the limit for random assignment is 50%). The purities presented here refer to 100% jet identification efficiency, i.e. each jet was always classified either as a quark jet or a gluon jet ($\varepsilon_q + \varepsilon_g = 1$). The cut on the discrimination variable (e.g. the ANN output) was therefore the same for quark jets and gluon jets. In addition, it was chosen such as to give about equal efficiency for quark jets and gluon jets in the Monte Carlo data ($\varepsilon_g \approx \varepsilon_q \approx 0.5$). The same cut will later be used on the jets from the experimental data (in the following section 11.6). Figure 11.15 show the results for the three different ANNs for the detector simulated Monte Carlo data. The purity is plotted versus the jet energy. It can be seen in this figure that the FMBIN classifier gave the best separation between quark jets and gluon jets. This may not be surprising since this classifier had access to most information. The PLPTBIN classifier was provided with information from only the six leading particles in the jet and can thus achieve the same separation as FMBIN only for low-multiplicity (≤ 6 particles in the jet) jets. The FMONE classifier was provided with fewer jets in the individual energy bins during the

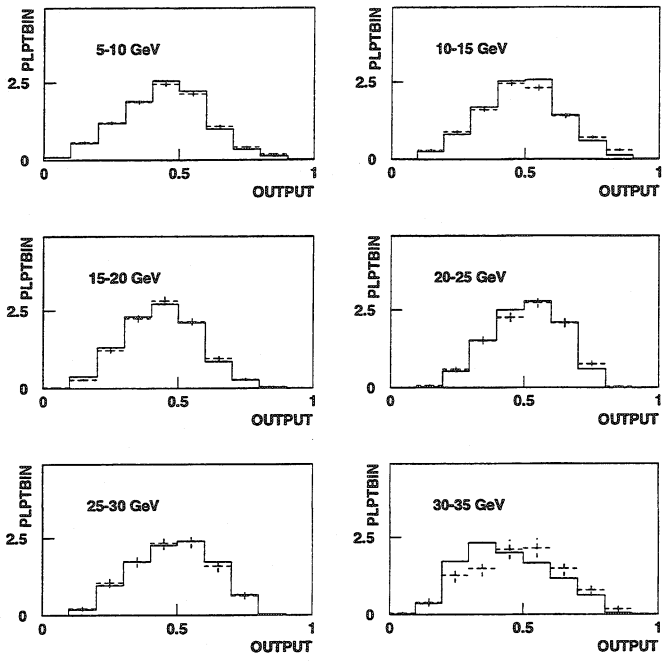

Figure 11.9: ANN output distributions from the PLPTBIN classifier for the tagged gluon jets (solid line) and data (point), in the six energy bins.

training in order to give a reasonable training time. The FMONE ANN was also of the same size (same number of nodes and weights) as each one of the six FMBIN ANNs. It is probable that FMONE could achieve the same separation as FMBIN if the size of the FMONE network (e.g. more nodes in the hidden layer) as well as the size of the training sample were extended.

In figure 11.16 the FMBIN network is compared with two other classifiers, the Fisher statistical discrimination, (FISHER), and the best fragmentation variable, (F14). The FISHER classifier is described earlier in section 4.1. The F14 classifier consisted of six different cuts in the $F14(E_{jet})$ distributions, one for each jet energy bin.

Clearly the FMBIN and FISHER classifiers did not perform much better than the best of their input variables (F14). It seems that there are almost no advantages in using several variables for quark jet and gluon jet separation on a jet to jet basis. This was also observed in [35] on a different type of input variables than the ones used in this analysis. It is surprising that FISHER has a somewhat poorer performance than F14 since this was one of its input variables. A possible explanation can be that the F_{mn} variables are strongly correlated (in some cases to more than 90%) and some of them deviate significantly from Gaussian distributions. The conclusion from this comparison is that the performance of the classifiers depends only weakly on the input variables chosen and the type of statistical discrimination used.

The importance of the jet clustering algorithm and the separation cut used was also

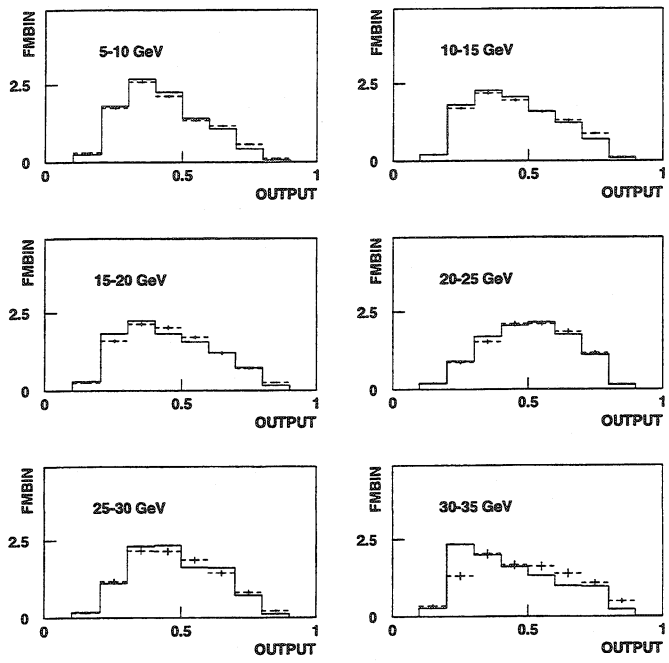

Figure 11.10: ANN output distributions from the FMBIN classifier for the tagged gluon jets (solid line) and data (point) in the six energy bins.

investigated. The five classifiers described above were all trained on jets clustered with the LUCLUS [7] algorithm using the separation parameter $d_{join} = 3.5 GeV/c$. Figure 11.17 shows the test of the FMBIN classifier on test samples of jets which were clustered differently from the jets in the training sample.

The following jet algorithms, which were discussed in section 6.5, were tested:

- DURHAM with $y_{cut} = 0.015$;
- JADE with $y_{cut} = 0.015$;
- LUCLUS with $d_{join} = 5.0\ GeV/c$.

The classifier seems to have a better performance on jets clustered with these three different algorithms than with the one that was used for the training sample. This is explained by the comparably low $d_{join} = 3.5\ GeV/c$ which was used for the clustering of the jets in the training sample. In each of the three new test samples the jet separation was higher which resulted in fewer but more separated jets (from the other jets in the event), compared to the jets in the training sample. The JADE clustering gave the best performance of all the algorithms.

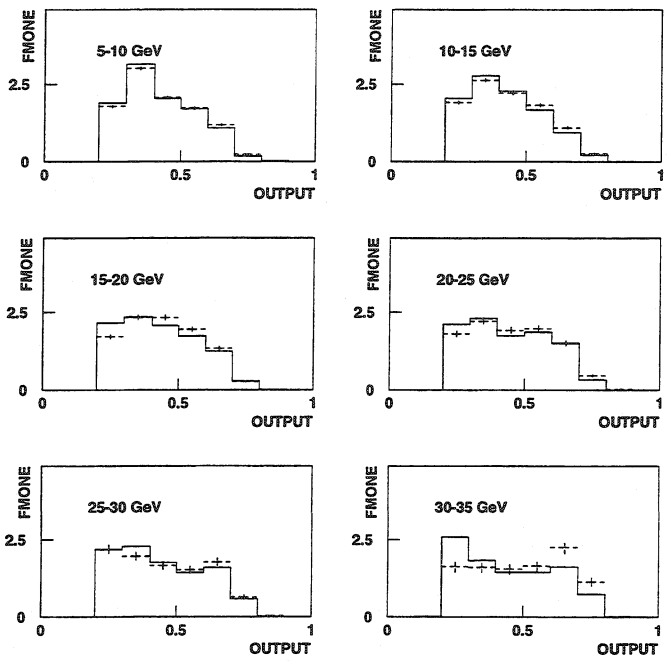

Figure 11.11: ANN output distributions from the FMONE classifier for the tagged gluon jets (solid line) and data (point), in the six energy bins.

11.5.2 Purity on an event basis

The event based identification presented in section 9.3 can also be applied to fragmentation based classifiers. For each of the five classifiers the jets in a $e^+e^- \rightarrow q\bar{q}g$ threejet event were ordered in the output variable and the jet which was closest to the gluon target value was assigned as the gluon jet. The efficiency is then 100% and the random limit for the purity is 33%. Figure 11.18 shows the purity for each of the five classifiers versus the energy of the lowest energetic jet (E_{min}) in the event. In this type of plot the fragmentation based classifiers can be directly compared with the jet energy based classifier (denoted *JET3* in the figure), bearing in mind that a common feature of the jet algorithms are that they order the jets according to their energy, see section 6.5. The following observations can be made:

- the fragmentation based classifiers can only compete with *JET3* in events where $E_{min} \geq 25$ GeV;
- for the fragmentation based classifiers the purity is nearly the same in highly asymmetric events (low E_{min}) as in "Mercedes" events and this shows that the jet energy was not used in the classification;
- the "bin"-trained classifiers (*FMBIN*, *PLPTBIN*, *FISHER* and *F14*) show a peculiar

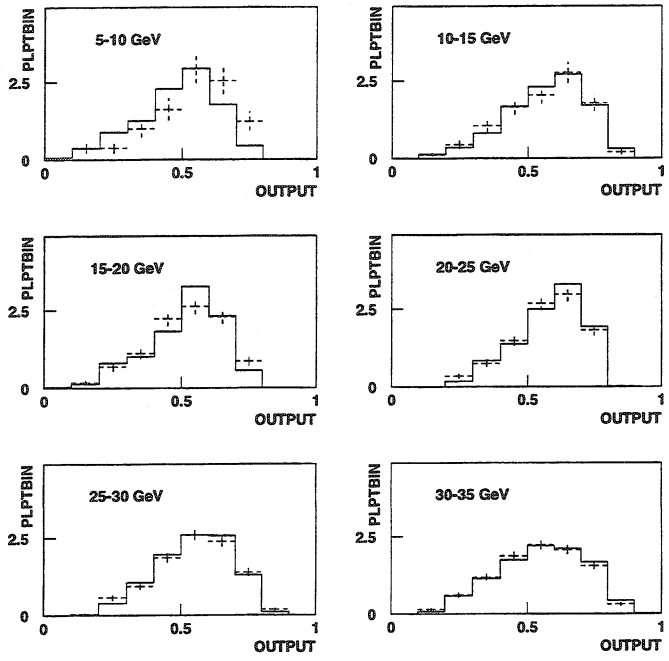

Figure 11.12: ANN output distributions from the PLPTBIN classifier for the tagged quark jets (solid line) and data (point) in the six energy bins.

jagged behaviour which becomes flatter with increasing E_{min} and only the *FMONE* classifier seems to be more or less independent of E_{min}.

The jagged behaviour of the bin-trained classifiers come about because the output from different ANNs (or Fisher statistical discriminators) were compared when ordering the jets. Consider for instance a threejet event where the three jet energies were 8, 24 and 35 GeV[1]. The first jet was then fed into the ANN (or Fisher discriminator) for the jet energy bin 5-10 GeV, the second jet into the 25-30 GeV ANN and the third jet into the 30-35 GeV ANN. Thus, the output from three different ANNs were used in the ordering. Numerical differences between the ANN output distributions are inevitable and this may spoil the gluon jet identification based on ANN output ordering. Because the jet energy bins must be of finite size (5 GeV in this analysis) the distribution of quark jet and gluon jet energies within each training sample were not totally uniform. This can be seen in figure 11.19 which shows the jet energy distribution in each bin. The variation of the energy within a bin explains why the purity is always peaked at the low energy edge of the bin. The effect is most pronounced in the first bin (5 - 10 GeV).

The jagged behaviour disappeared only in symmetric threejet events (high E_{jet}) where all three jets are in the same energy bin, i.e. were fed into the same ANN. These are

[1]The jet energies seldom add up to the center-of-mass energy (91.2 GeV) because of the limited acceptance and energy-momentum resolution of the detector.

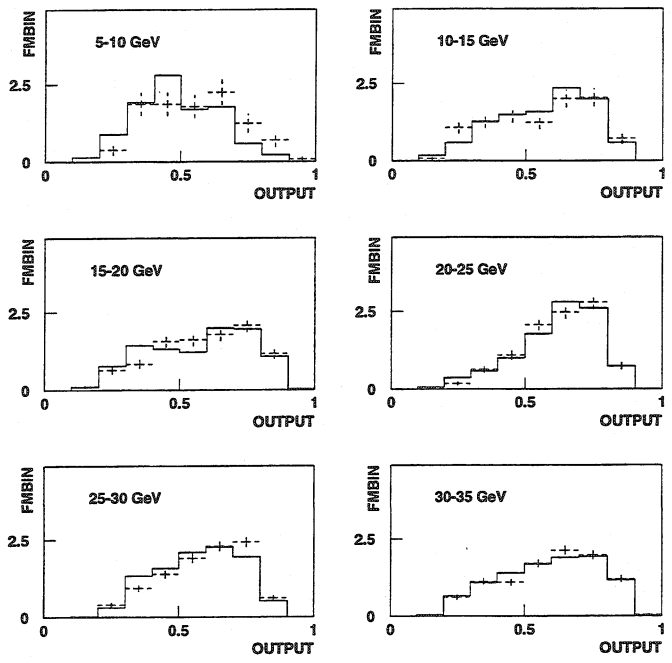

Figure 11.13: ANN output distributions from the FMBIN classifier for the tagged quark jets (solid line) and data (point), in the six energy bins.

normally the most interesting topologies in which to use fragmentation based quark/gluon jet separation since in other topologies the jet energy (*JET3*) provides a much stronger discrimination.

11.6 The classifiers used on experimental data

Given the samples of gluon and quark jets tagged in the DELPHI data according to the method explained in section 8.1, the same tests of the purity on a jet to jet basis for each of the five classifiers which were done for Monte Carlo generated jets, can also be done with real data. This may reveal differences between experimental data and the Monte Carlo data which was used to train the classifiers. It is also a test of the stability of the classifiers with respect to the input data.

Although the output distributions in figures 11.9, 11.10 and 11.11 showed that there is a good agreement between data and the simulated Monte Carlo sample of gluons, as well as the figures 11.12, 11.13 and 11.14 indicated that there is a good agreement between data and the simulated Monte Carlo sample of quarks, it is interesting to get it quantified in terms of efficiency and purity. Deviations which are larger than the statistical errors in any of these two quantities must be accounted for as systematic errors for the corresponding classifier. The results for the three ANN classifiers (*FMBIN*, *PLPTBIN* and *FMONE*) on

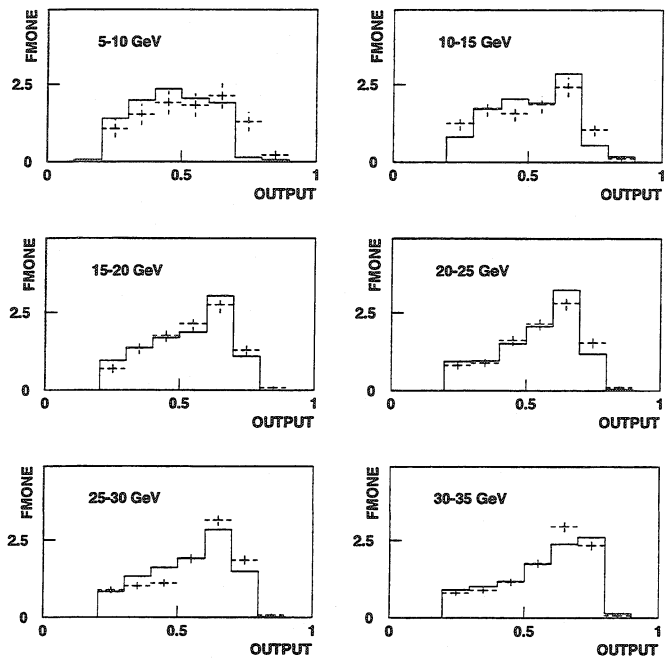

Figure 11.14: ANN output distributions from the FMONE classifier for the tagged quark jets (solid line) and data (point), in the six energy bins.

jets from the real data (showing the resulting purity versus energy for the whole energy range) are presented in figure 11.20 where the purities were calculated using the same cuts in the output distributions as in figure 11.15. In a similar way the results for the other three classifiers (*FMBIN*, *FISHER* and *F14*) are presented in figure 11.21.

From the tests on real data the systematic errors on the efficiency and purity for the different classifiers were estimated. The same tagging techniques for collecting the real gluon jet and quark jet samples were applied to the Monte Carlo data and the maximum deviation of the central values from the corresponding values in figure 11.15 and figure 11.16 was taken as the systematic error.

The imperfection of the Monte Carlo model for the fragmentation was also estimated from the difference in the results with the tagged gluon jets and quark jets in the real data and the Monte Carlo data and again the maximum deviation for each classifier is taken as the systematic error.

Table 11.4 summarises the results for each of the five classifiers tested on data and Monte Carlo for quark jets and gluon jets, in the energy range 25-30 GeV. The cuts in the classifier output distributions (cf. E_{cut} in chapter 9 were chosen to give about 50% efficiency ($\varepsilon_q \approx \varepsilon_g \approx 0.5$) on the Monte Carlo data. Since the same cuts were also used in the tests on the experimental data any systematic global displacement of the output distributions compared to the tests on Monte Carlo data will result in efficiencies which

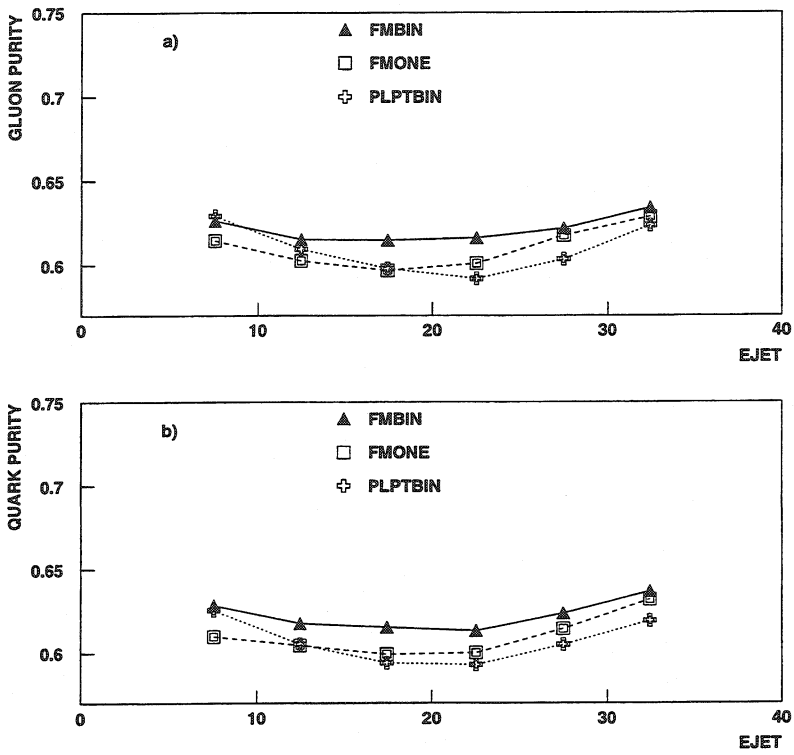

Figure 11.15: Performance (on Monte Carlo data) of the fragmentation based quark/gluon jet classifiers presented in this paper. The purity is measured with a 100% efficiency (a jet is always identified, either as a quark jet or a gluon jet). The random limit is at 50%.

deviate from 50%. An example of such global displacement could be a badly tuned Monte Carlo which in average gives higher average multiplicity for both quark jets and gluon jets. Comparing the two rightmost columns in table 11.4 gives an idea of how well the Monte Carlo model describes data. *PLPTBIN* is clearly the best classifier in this respect, *FMBIN* and *FMONE* are the worst and *FISHER* and *F14* falls in between. Simple input variables like p_L and p_T of the leading particles in the jet gives, as expected, less model dependence than the rather complex F_{mn} variables. A further observation, though less significant, is that non-linear discriminators like artificial neural networks (*FMBIN* and *FMONE*) are more model dependent than linear discriminators like Fisher statistical discrimination (*FISHER*) or a cut on a single variable (*F14*).

For completeness the results on Monte Carlo data with the JADE clustering algorithm ($y_{cut} = 0.015$) are given in table 11.4. Among the four different clustering methods and cuts in the jet distance measures (see section 6.5) which were tested, JADE with $y_{cut} = 0.015$ was found to give the best performance compared with the default LUCLUS with $d_{join} = 3.5$ *GeV/c*. Note that the JADE results are better (higher purity) than the LUCLUS ones. This was also the case for the other two clustering choices. The

Classifier	$25 \leq E_{jet} \leq 30$ GeV			
	Simulated	Simulated+JADE	Simulated+tagging	Data+tagging
FMBIN	$\varepsilon_g = 50.4 \pm 0.2$	49.5 ± 0.3	55 ± 1	49 ± 1
	$\pi_g = 62.2 \pm 0.1$	66.3 ± 0.2	64 ± 1	66 ± 1
	$\varepsilon_q = 49.6 \pm 0.2$	50.5 ± 0.3	45 ± 1	51 ± 1
	$\pi_q = 62.4 \pm 0.2$	66.0 ± 0.4	67 ± 1	65 ± 1
FMONE	$\varepsilon_g = 49.3 \pm 0.2$	49.7 ± 0.3	54 ± 1	48 ± 1
	$\pi_g = 61.8 \pm 0.1$	65.6 ± 0.2	64 ± 1	67 ± 1
	$\varepsilon_q = 50.7 \pm 0.2$	50.3 ± 0.3	46 ± 1	52 ± 1
	$\pi_q = 61.4 \pm 0.2$	65.4 ± 0.4	66 ± 1	65 ± 1
PLPTBIN	$\varepsilon_g = 50.3 \pm 0.2$	50.6 ± 0.3	54 ± 1	55 ± 1
	$\pi_g = 60.4 \pm 0.1$	64.1 ± 0.2	61 ± 1	61 ± 1
	$\varepsilon_q = 49.7 \pm 0.2$	49.4 ± 0.3	46 ± 1	45 ± 1
	$\pi_q = 60.5 \pm 0.2$	64.5 ± 0.4	63 ± 1	63 ± 1
FISHER	$\varepsilon_g = 49.8 \pm 0.1$	49.0 ± 0.3	51 ± 1	48 ± 1
	$\pi_g = 62.1 \pm 0.1$	66.4 ± 0.2	64 ± 1	66 ± 1
	$\varepsilon_q = 50.2 \pm 0.1$	51.0 ± 0.3	49 ± 1	52 ± 1
	$\pi_q = 62.0 \pm 0.2$	65.7 ± 0.4	64 ± 1	65 ± 1
F14	$\varepsilon_g = 48.6 \pm 0.1$	49.8 ± 0.3	51 ± 1	47 ± 1
	$\pi_g = 62.1 \pm 0.1$	65.8 ± 0.2	63 ± 1	66 ± 1
	$\varepsilon_q = 51.4 \pm 0.1$	50.2 ± 0.3	49 ± 1	53 ± 1
	$\pi_q = 61.5 \pm 0.2$	65.7 ± 0.4	64 ± 1	64 ± 1

Table 11.4: Jet to jet classification with the five classifiers in the jet energy bin $25 \leq E_{jet} \leq 30$ GeV. The purities and efficiencies are given in percent. The first column shows the results on simulated quark jets and gluon jets. The second column shows the results on simulated quark jets and gluon jets which have been clustered using the JADE algorithm with $y_{cut} = 0.015$. The third and fourth column shows the results on gluon jets and quark jets tagged according to the method described in the section 8.1 for Monte Carlo and experimental data.

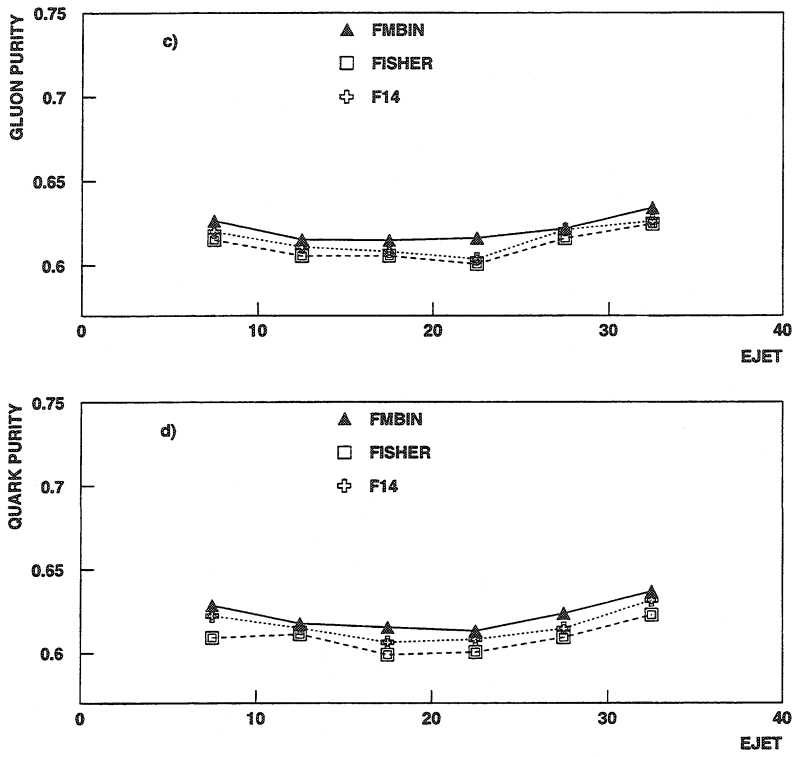

Figure 11.16: Performance (on Monte Carlo data) of the fragmentation based quark/gluon jet classifiers presented in this analysis. The purity is measured with a 100% efficiency since a jet is always identified, either as a quark jet or a gluon jet. The random limit is at 50%.

efficiency-purity results obtained with default choice of LUCLUS 3.5 represents the worst case.

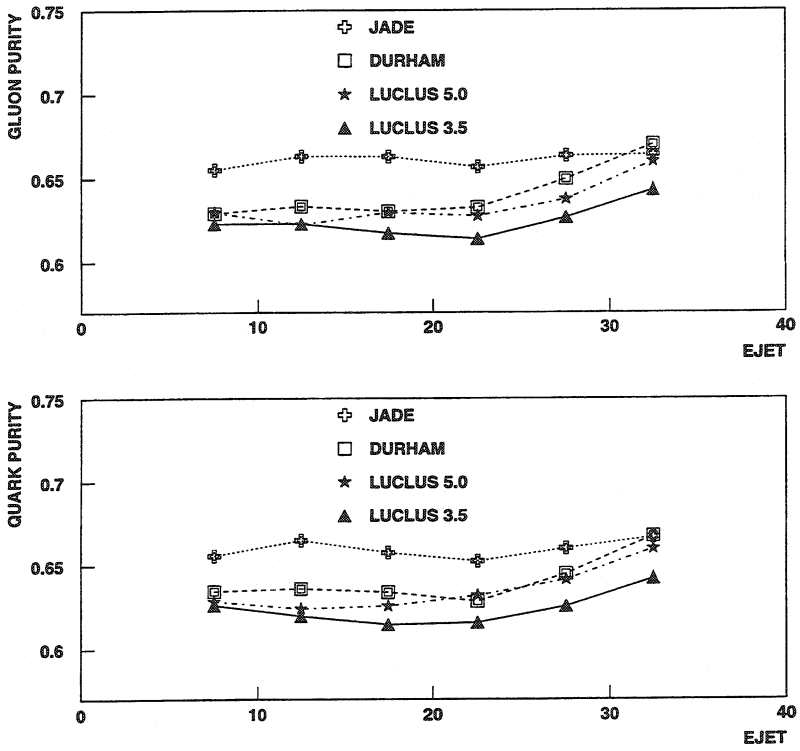

Figure 11.17: The dependence of the FMBIN performance on the jet clustering. The classifier was trained using jets clustered with the LUCLUS algorithm and $d_{join} = 3.5\ GeV/c$. After the training it was tested on samples of jets with similar clustering (LUCLUS 3.5) as well as three different ones: LUCLUS with $d_{join} = 5.0\ GeV/c$ (LUCLUS 5.0), JADE with $y_{cut} = 0.015$ (JADE) and DURHAM with $y_{cut} = 0.015$ (DURHAM). The upper plot shows the gluon purity and the lower plot the quark purity. The random limit is 50%.

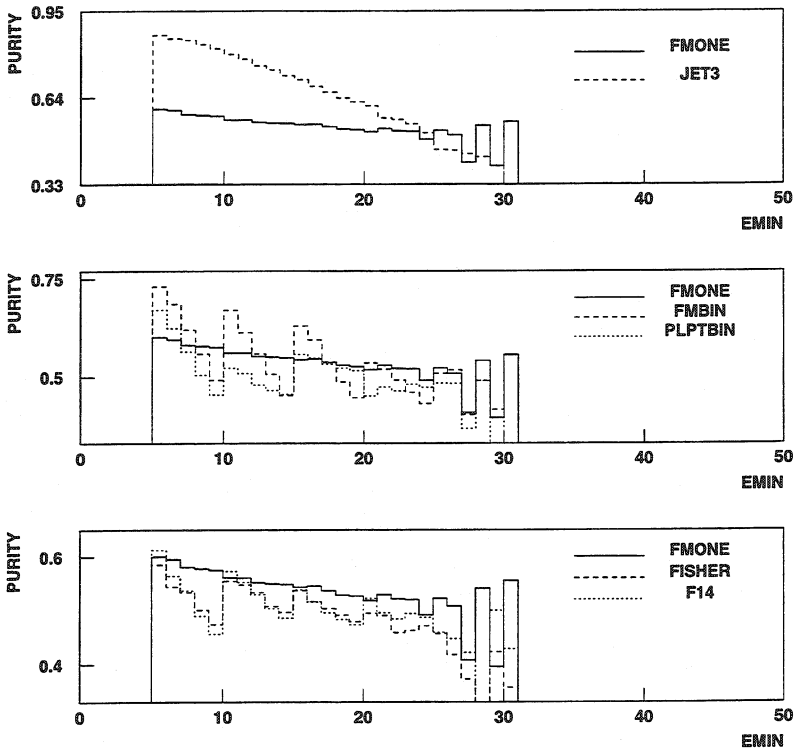

Figure 11.18: The identification of the gluon jet in $e^+e^- \rightarrow q\bar{q}g$ threejet events. In the upper plot, the jet energy based identification *JET3* is compared with the *FMONE* classifier. The middle plot shows the three ANN based classifiers. In the lower plot, *FMONE* is compared with the single variable classifier, *F14* and the Fisher statistical discrimination classifier, *FISHER*. The lower limit of the y-axis in all plots is set to the random limit, i.e. 0.33.

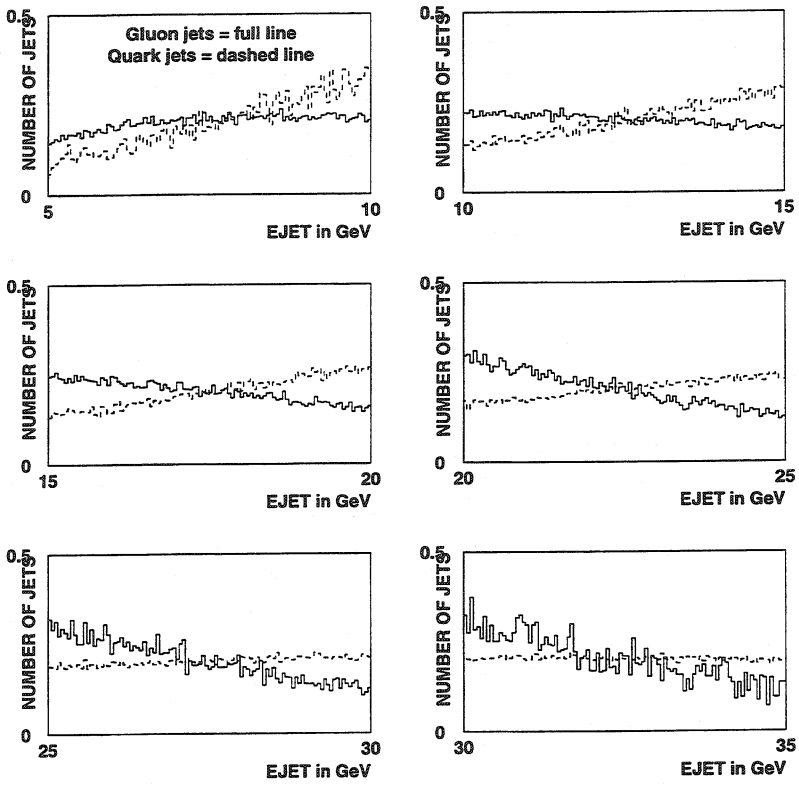

Figure 11.19: The quark jet and gluon jet energy spectrums with the integral normalised to one in the six energy bins.

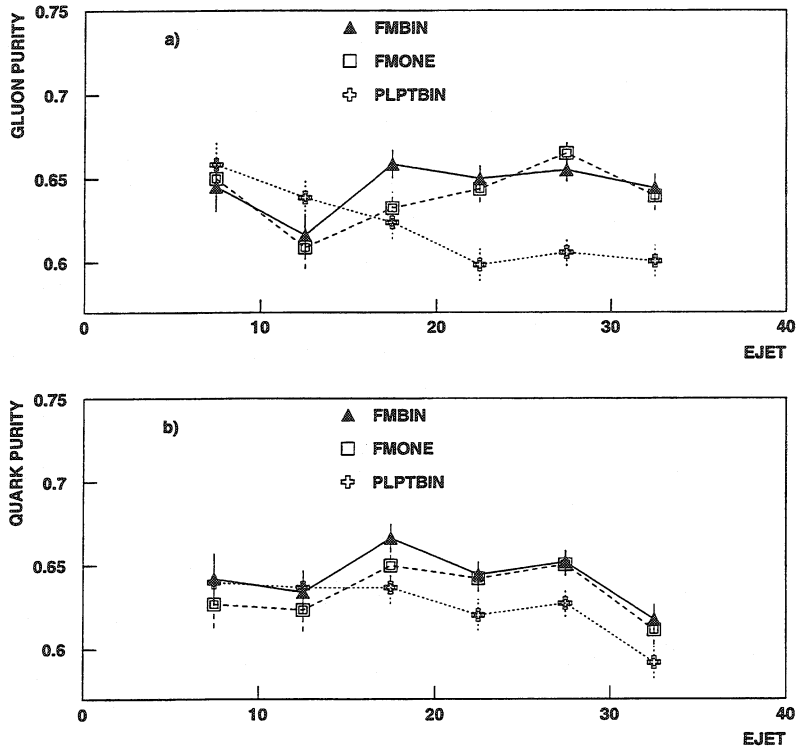

Figure 11.20: Performance of the different fragmentation based classifiers on quark jets and gluon jets from experimental data. The purities were calculated with the same cuts on the output distributions as were used in figure 11.15. The upper plot shows the gluon purity and the lower plot the quark purity. The random limit is 0.5.

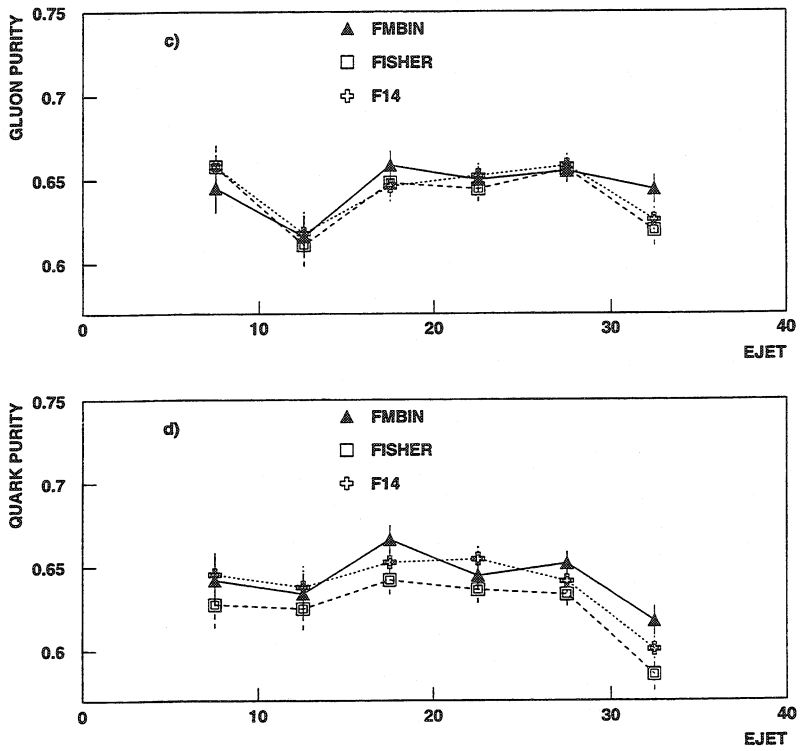

Figure 11.21: Performance of the different fragmentation based classifiers on quark jets and gluon jets from experimental data. The purities were calculated with the same cuts on the output distributions as were used in figure 11.16. The upper plot shows the gluon purity and the lower plot the quark purity. The random limit is 0.5.

Chapter 12

Applications of jet identification

The motivation for the study of jet identification, presented in the previous chapters, was to find new and unbiased methods that could be applied to various DELPHI analysis. In this chapter three physics studies in which jet identification is necessary are presented. All three analyses use heavy quark lifetime tagging as described previously.

12.1 Investigation of the splitting of quark and gluon jets

The aim of this analysis, [43], is to investigate if the initial splittings into quark and gluon jets are in agreement with QCD expectation and to study the jet evolution in order to explain the small measured value of the ratio of the number of particles in gluon and quark jets, $\frac{n_g}{n_q}$. The relative strength of quark and gluon jet splittings are determined by the colour factors (or the Casimir factors). The colour factors C_F, C_A and T_F determine the coupling strength of a gluon radiating from a quark, two gluons radiating from a gluon and a gluon splitting into a quark/antiquark pair, see figures 12.1 – 12.3. These

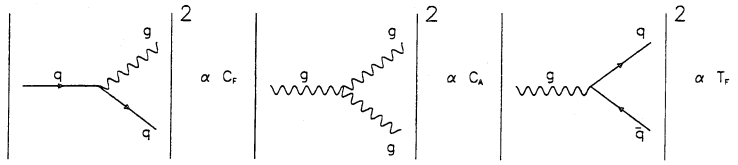

Figure 12.1: C_F. Figure 12.2: C_A. Figure 12.3: T_F.

coupling constants are determined within the SU(3) group and their predicted values are $C_F = 4/3$, $C_A = 3$ and $T_F = 1/2$ which has to be weighted with the number of active quark flavours n_F. According to QCD, more splittings are therefore expected in gluon jets than in quark jets, since $\frac{C_A}{C_F} = \frac{9}{4}$, and gluon initiated jets should consequently be different from quark jets. In particular, one expect the ratio $\frac{n_g}{n_q}$ to be approximately $\frac{C_A}{C_F}$ and this is confirmed by even more detailed calculations.

The experimental situation is, however, completely different. From threejet events in e^+e^- annihilation at Z^0 energies, recorded by DELPHI, this fraction is measured to be only about $\frac{5}{4}$, but an increase of this value with the energy has been observed. The analysis has two aims, firstly to establish that the initial splitting in quark and gluon jets are consistent with QCD and secondly to follow the jet evolution in order to understand the observed differencies in particle multiplicities. To achieve this, two types of observables are used, the scale dependence of the scaled energy distributions of stable hadrons which is related to the longitudinal part of the jet evolution and the subjet splittings, which is proportional to the transverse part of the jet evolution.

A detailed comparison of quark and gluon jet properties needs samples of quark and gluon jets with nearly the same kinematics for a direct match of their properties under similar circumstances. To fulfil this condition two different event topologies have been used. Threejet events from the decay of the Z^0, collected in the DELPHI detector during 1991–1994, are selected. After a standard hadronic event selection, similar to the one used in this thesis, the Durham jet algorithm, see section 6.6, is used to find the threejet events. With this algorithm, the minimum transverse momentum between two jets which have been resolved at a scale defined by y_{cut}, is approximately given by $k_\perp^{min} \simeq E_{vis}\sqrt{y_{cut}}$. From these events only the events with Y and Mercedes topologies, see figures 9.3 and 9.4, are used. The jet axes is projected into the event plane and the energy of each jet is calculated using the measured angles between the jets, see equation 6.6. In order to enhance events with three well defined jets, the following condition, a planar cut, is required:

$$\theta_{ij} + \theta_{ik} + \theta_{jk} \geq 355°.$$

A gluon jets sample is obtained from threejet events which originate from the Z^0 decaying to $b\bar{b}g$. The event is identified by using an event lifetime tag and the gluon jet is tagged indirectly by identifying the two other jets as b-quark jets using a jet lifetime tag. Events with a b-quark signature were selected by requiring that the event probability, see section 7.3.1, is $P_N \leq 10^{-1.5} = 0.032$, which is the same requirement as used in the thesis. The tracks corresponding to each of the reconstructed jets are then used to determine the jet probability, P_J, see section 7.3.2.

For the Y and the Mercedes events two different selection strategies are used. For the Y events, the most energetic jet is always chosen to be a quark jet. For this sample, this is correct in about 98 % of all cases. For the two remaining jets, P_J is compared with the cut 0.1 and if for one of them we have $P_J \leq 0.1$ and for the second one $P_J > 0.1$ we consider them as quarks respectively gluon jets. For the Mercedes events, two of the jets are demanded to fulfil the condition $P_J \leq 0.1$ to be consider quark jets assuming that the third jet fulfils $P_J > 0.1$.

The efficiency and purity calculations have been made using events, generated by the JETSET Monte Carlo, tuned to DELPHI data, which have passed through the full simulation program of the DELPHI detector, DELSIM and the standard DELPHI data reconstruction chain. The assignment of parton flavours to the jets is not unique, not even in simulated events. Therefore two independent ways of identifying the the gluon jets are used. The jet that makes the largest angle to the direction of the hadrons containing the heavy quarks or contains the fewest decay particles from the heavy quarks, is considered to be the gluon jet. These two ways are called the angle assignment and the history

assignment. The methods give very similar results and the purities are therefore estimated with small systematic uncertainties. Gluon jet purities of $(79.6 \pm 2.0)\%$ for Y events and of $(74.5 \pm 2.5)\%$ for Mercedes events are obtained with the tagging procedure described.

The splitting of jets can be studied by investigating the distribution of energy in the process. The study is related to the analysis of the scale dependence (scaling violation) of the fragmentation function $D_p^H(x_E, s)$ of a parton p into a hadron H described by theory, where $x_E = E_{hadron}/E_{jet}$ (and $s = E_{cms}$). For values $x_E \geq 0.5$ it can be shown that the splitting process $g \rightarrow q\bar{q}$ is disfavoured with respect to $g \rightarrow gg$. The fragmentation functions for quarks and gluons can be used to compare the relative size of the observed scaling violation. The following ratio, $r_S(x_E)$, which can be predicted by solving the evolution equations, measures the relative changes in the fragmentation functions

$$r_S(x_E) = \frac{S_g}{S_q} = \frac{\frac{d \ln D_g^H(x_E,s)}{d \ln s}}{\frac{d \ln D_q^H(x_E,s)}{d \ln s}}$$

The scaling violations in the quark and gluon jets can thus be measured by computing the relative change, r_S, in the fragmentation function and it is found to be:

$$r_S(0.23 \leq x_E \leq 0.75) = 2.8 \pm 0.8 \pm 0.6.$$

This can be compared to the QCD expectation

$$r_S(x_E \rightarrow 1) = C_A/C_F = 9/4$$

indicating that the ratio of quark and gluon splitting can indeed be explained by the colour factors.

Conclusion

The splitting processes are investigated in quark and gluon jets that are identified by the use of longitudinal and transverse observables from symmetric threejet events in DELPHI. The longitudinal part of the jet evolution is the main contributor to the scale dependency of the energy distribution of stable hadrons. The *scaling violation* in gluon jets are identified for the first time in this study. The scale energy dependence of the fragmentation function for gluons is found to be twice larger than of the fragmentation function for quark jets, something that is consistent with the QCD expectation of C_A/C_F.

12.2 Measurement of the triple gluon vertex

In this study, [44], fourjet events collected during 1992–1994 by the DELPHI experiment are analysed to determine the contribution from the triple-gluon vertex process. Diagrams showing the different four-parton final states contributions are shown in figures 12.2 - 12.6. Two of the four jets are tagged as heavy quarks (b or c), using a lifetime tag and lepton transverse momentum information.

The jets were reconstructed by the LUCLUS jet algorithm, see section 6.6, with the jet resolution parameter d_{join} set to the value of 4 GeV/c. To check the influence of the

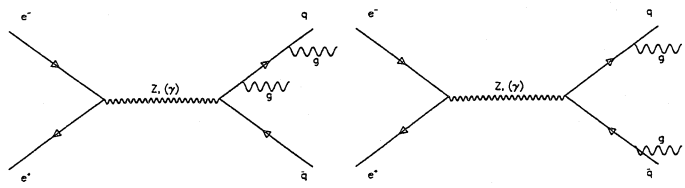

Figure 12.4: Double gluon bremsstrahlung.

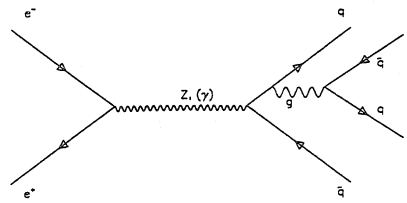

Figure 12.5: Secondary $q\bar{q}$ production.

cluster algorithm, the JADE jet algorithm was also used, with the jet resolution parameter y set to the value of 0.02, but no systematic effects were observed.

The generalized Nachmann-Reiter angle, Θ^*_{NR} is the angle between the vectors $\vec{p_1} - \vec{p_2}$ and $\vec{p_3} - \vec{p_4}$ and it distinguishes between the triple-gluon vertex contribution and secondary $q\bar{q}$-production and the angle α_{34} is the angle between the two jets of lowest energy in a fourjet event, and it distinguishes between the triple-gluon vertex contribution and the double-bremsstrahlung, see figure 12.7.

The triple-gluon vertex contribution can therefore be determined from the from the two-dimensional distribution in these two variables. These distribution in fourjet events from double gluon bremsstrahlung, secondary quark/antiquark pair and triple gluon vertex are significantly different and they are used to fit C_A/C_F, the ratio of the coupling strength of the triple gluon vertex to that of the gluon bremsstrahlung, and the ratio of the number of quark colours to the number of gluons, N_C/N_A.

In a previous analysis of fourjet events, energy ordering of the jets was used to distinguish between jets coming from primary and secondary partons. However, in the new

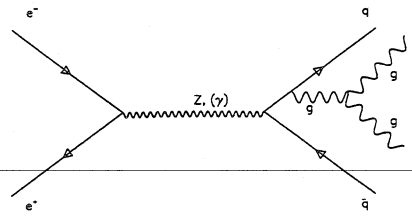

Figure 12.6: Triple gluon vertex.

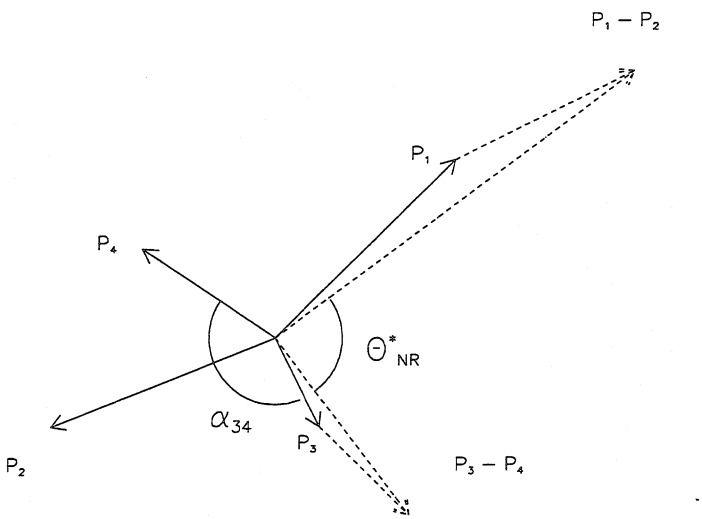

Figure 12.7: Figure showing how the generalized Nachmann-Reiter angle, Θ^*_{NR}, and the angle α_{34} are defined.

analysis, the primary jets were tagged as b-quark jets using a lifetime tag as well as in the semileptonic decay, see 7.1, in order to make the tagging as effective as possible these methods were combined using a neural network. The ANN has 13 input variables, the event lifetime probability, the lifetime probability of each jet, the transverse momentum of the electron in each jet and the transverse momentum of the muon in each jet. The network had one hidden layer and four output nodes, one for each jet. Since node one corresponds to the highest energy and so on, energy ordering was also included. The outputs were normalized to be in the interval zero to one and a jet is tagged as a primary quark jet if the associated output value is below 0.5 and the event was only accepted if exactly two of the four jets were tagged as primary quark jets. This happened in $15.6 \pm 0.2\%$ of the MC event sample and $15.6 \pm 0.2\%$ of the DELPHI data. Figure 12.8 shows the outputs from the four nodes and the good agreement between data and simulated events. This improved tagging led to larger differencies between the shapes of the distributions of the observables for gluon bremsstrahlung, the triple gluon vertex and secondary quark production and hence increasing the method's sensitivity. The efficiency of the tagging was increased by 50 % compared to the standard b-tagging which is less efficient in the case of many jets and the purity was 70%.

Conclusion

The results of the measurements were

$$C_A/C_F = 2.51 \pm 0.25(stat.) \pm 0.11(frag.) \pm 0.25(b-tagging)$$

$$T_R/C_F = 1.91 \pm 0.44(stat.) \pm 0.25(frag.) \pm 0.03(b-tagging)$$

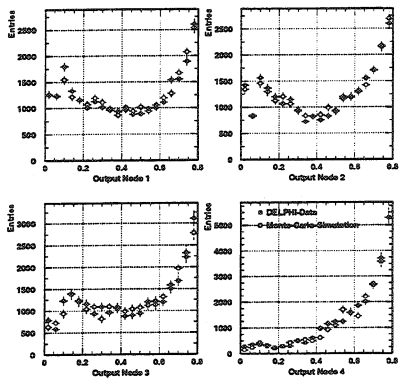

Figure 12.8: Output of the four different nodes of the neural network using DELPHI data (filled boxes) and simulation (open boxes). Nodes 1–4 correspond to jets 1–4 numbered in order of decreasing energy. A jet with output below 0.5 was classified as a quark jet. Only events with exactly two jets classed as quark jets were used in the analysis.

where $T_R = n_f T_F$. Since $T_F/C_F = N_C/N_A$ and $n_f = 5$ one obtains $C_A/C_F = 2.51 \pm 0.30$ and $N_C/N_A = 0.38 \pm 0.10$ in agreement with the QCD predictions.

12.3 Scale dependence of the multiplicity

The third analysis [45] uses three jet events selected similarly as in first study [43], where the scaling violation in jets were identified for the first time, to determine the relative coupling of the participating elementary fields and thereby making a direct test of QCD.

The data set used here was collected with the DELPHI detector at Z^0 energies during the years 1992–1995. The cuts were applied in order to select hadronic Z^0-decays of the $q\bar{q}g$ and $q\bar{q}\gamma$ type and they are very similar to the ones used in the thesis. The gluon purities vary from 95% for low energy gluons to 46% for the highest gluon jet energies. The jets in these threejet events were clustered using the Durham jet algorithm. In addition it was required that the angle θ_{12} was in the interval $100° \leq \theta_{12} \leq 170°$ and the same conditions were imposed on the angle θ_{13}, in order to obtain the requested type of events. The leading jet was not used in the analysis and events were considered to be *symmetric* if θ_{12} was equal to θ_{13}, within some tolerance (chosen to be 6°).

The jets were identified by the anti-tagging process of the the gluon jet using the heavy quark jets from $q\bar{q}g$-events enhanced by b-quark events by the use of the impact parameter technique. In the simulation, quark and gluon jets were of course also clustered with the Durham algorithm and the jets were identified at "parton level". For each parton jet the number of quarks (weight +1) and antiquarks (weight -1) were summed and the jets with the sum equal to +1 (-1) were assigned to be quark (antiquark)jets. The jets with the sum equal to 0 were assigned to be gluon jets. The small amount of events where the expected pattern (+1, -1, 0) did not appear were rejected. It was also demanded that the angle between the original parton direction and the hadron jet did not exceed a maximum

angle.

As the successive bremsstrahlung processes are directly proportional to the ratio of the radiated gluon multiplicity from a gluon and a quark source, it will asymptotically approach the ratio C_A/C_F of the colour factors. The increased radiation from gluons should manifest itself in a higher multiplicity and a stronger scaling violation of the gluon fragmentation function. Already in the first study, it was shown that this prediction is influenced by the fact that the primary parton is a quark or a gluon, due to the differencies in their fragmentation functions. These differencies must be present because quarks are valence particles of the hadron whereas gluons are not. This is evident from the behaviour of the gluon fragmentation function to charged hadrons at large scaled momentum, where it is suppressed by about one order of magnitude compared to the quark fragmentation function. At low momentum, soft radiation and the production of low energy hadrons is further suppressed which is due to a reduction of the primary splitting of gluons compared to expectation and thus responsible for the observed small hadron multiplicity ratio as was shown in the first analysis. Due to the high number of particles from the decay of a primary heavy particle this multiplicity ratio is further reduced, if heavy quarks are included. Furthermore the jet algorithms that combine hadrons to jets are likely to assign low energy particles at large angles with respect to the original parton direction to a different jet, which for gluon jets, initially broader than quark jets, will lead to a reduction of the multiplicity for gluon jets hence a corresponding enhancement of the multiplicity of the quark jets.

These effects, which will lower the ratio of the charged hadron multiplicities for gluon and quark jets, have earlier been neglected, but this study shows that they are important and a good test of the prediction that the radiated gluon to quark multiplicity ratio is equal to the colour factor ratio.

The more rapid increase of the gluon jet multiplicity with a relevant energy scale, compared to that of the quark jets have been demonstrated. The intuitive energy scale, the jet energy, has been confirmed to be valid and recently, the scale has been extended into a transverse momentum scale. This dependence of the scale on the mean hadron multiplicities in gluon and quark jets can be used for a precision determination of the colour factor ratio C_A/C_F, by an estimation of the differencies of the contribution to the multiplicities of gluon and quark jets.

The relevant scales can be determined from the jet properties and the event topology and studies of the hadron production show that the characteristics of the parton cascade depend mainly on the hardness, κ, of the process producing the jet,

$$\kappa = E_{jet} \sin \theta/2$$

where E_{jet} is the jet energy and θ is the angle to the closest jet.

This energy is related to the transverse momentum of the jet as well as to the $\sqrt{y}$ as it is used in the jet algorithm. κ is used as the relevant scale measure, see figures 12.9. This choice of scale parameter is used to determine the charged hadron multiplicity in gluon and quark jets as a function of this scale and the colour factor ratio C_A/C_F is directly observed to depend on the increase of multiplicities with this scale. The discrepancy in the multiplicity ratio between gluon jets and quark jets is explained by the different hadronization schemes.

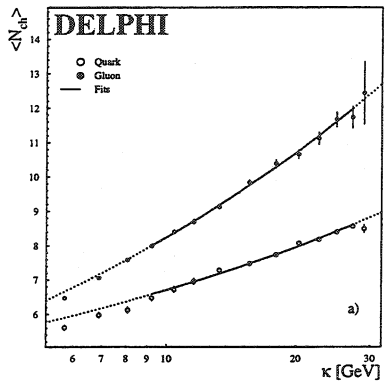

Figure 12.9: a) Average charged particle multiplicity for light quark and gluon jets as a fitted function of κ. b) Ratio of the gluon to quark jet multiplicity.

Conclusion

After taking the dependence of the charged multiplicity on the opening angle between the two jets with the smallest energies in symmetric threejet events into account the colour factor ratio, C_A/C_F, is measured to be $= 2.12 \pm 0.10$, which is in good agreement with the QCD expectation.

Chapter 13

Conclusions

Different methods for separating quark jets and gluon jets using their fragmentation differences have been tested on both Monte Carlo data and data from the DELPHI experiment. Both linear and non-linear multivariable analysis methods have been tried.

Two different types of variables were used to measure the fragmentation properties. The conclusions from this study are:

- The purity that can be obtained by using the fragmentation differences between quark jets and gluon jets is limited.
- Different input variables gives similar results.
- Artificial neural networks and Fisher statistical discrimination gives similar results.
- A multivariable discrimination is not necessary. A cut on the best single variable gives almost the same discrimination as a multivariable analysis.
- Comparisons between data and Monte Carlo showed that the output distributions of the different classifiers were in good agreement. A detailed study of the efficiency and purity obtained showed:
  - negligible model dependence for the classifier with the simplest input variables (p_L and p_T of the six leading particles).
  - significant model dependence for classifiers with complicated input variables which were calculated using all particles in the jet.
  - slightly more model dependence for non-linear discriminators (artificial neural networks) compared to linear ones (Fisher statistical discrimination and cuts on a single variable).

The purities obtained in this analysis were all around 60% which appears very low compared to most other studies where the purities often are claimed to be $\sim 95\%$. However, it must be stressed that high purities can easily be achieved with any method using the jet energy as an input. One of the major goals of this study was to remove the jet energy and all dependencies on it from the input and to find out how much the fragmentation information by itself could contribute to the identification.

Acknowledgements

Firstly I would like to thank all members of the DELPHI and the LEP collaboration. It has been a privilege to work in such a successful experimental enterprise and to have access to a supplier delivering high quality data.

My deepest gratitude goes to Professor Derek Imrie for having invited me to Brunel University and there arousing my interest in participating in a LEP experiment, by inviting me to be a research fellow at Cern. This contact was possible thanks to Professor Per Carlson, who also was a benevolent landlord during my initial period at Cern.

In the autumn 1993 I was awarded a grant from the Royal Institute of Technology with 60% work reduction as lecturer for two years 1993/95 and also a grant from the International Foundation, for a stay at CERN 1995/96, for which I am deeply grateful.

It was, also in August 1993, during an examination for Professor Cecilia Jarlskog, that the PhD plans really took off. Thanks to her kindness and personal interest even in her peripheral students, I shortly afterwards found myself being a student at Lund University and supplied with a very clever and competent supervisor on the Cern site, Olof Bärring, as well as a very understanding main supervisor in Lund, Professor Göran Jarlskog, who has supported me during all these years, whatever the difficulties. My sincere thanks goes to them both.

Vincent Hedberg has been another very important person with the same qualities. Without his help and untiring and self-sacrificing work, I would certainly not have reached the finishing line. My feelings of gratitude towards him are very profound.

Being a member of the TPC-group at DELPHI, it is impossible not to thank some colleagues of this group: Jeanne Mas, Marilisa Turluer-Faccini, Pierre Antilogus, Gerard Cosme, Patrick Roudeau, Patrick Jarry and Patrice Siegrist, who have all helped me and shared laughs and problems during shifts.

A very warm thanks to "lunchgruppen" at Cern: Hans Danielsson, Elisabeth Falk-Harris, Kristina Gunne, Magnus Hansen and Per Jonsson, with whom I have experienced many good moments and from whom I have also received a lot of help and information.

Another group of very personal friends at Cern, which has played an equally important role during my stay, consists of Per Grafström, Mats Lindroos, Thomas Pettersson, Per Werner and Mats Wilhelmsson. Bengt Sagnell, a Cern guide colleague, is a also a member of this group. I especially want to thank him for the LEP information he supplied me with, which has been a great help when being a Cern guide and when writing a part of chapter 3.

Furthermore, I have the privilege of having a friend whom I personally consider the Guru of Physics, Torbjörn Sjöstrand, who, for a very long time has answered my questions with a lot of patience and also in a way to make me understand the answers.

Persons that I consider belonging to my "family" have played a decisive role especially in difficult times. When feeling down, I could always count on the positive attitude and support from Gunn Carlsdotter and Folke Johansson. The same soothing influence also came from an old friend from CTH, Margareta Lagerkvist-Briggs. Encouraging emails and wise words from Gunnar Overgaard have also played a major part in my well-being. From my brother Tomas' poem "Elégie", attached on the wall in my office, I obtained a daily inspiration. I interpret his memorable words "Amis! Vous avez bu de l'ombre pour vous rendre visibles." as a way for particles to be seen in a detector.

Last but not least, more than thanks to my wonderful wife Hélène, who has been the centre of my universe during a fantastic time with many difficulties but with far more happy moments. She has always been standing by my side and without her positive attitude and willingness for a change of milieu, I would not even have reached first base.

Bibliography

[1] DELPHI Collaboration, Zeit. Phys. **C 70**, 179 (1996).

[2] ALEPH Collaboration, Phys. Lett. **B346**, 163 (1995).

[3] OPAL Collaboration, Zeit. Phys. **C68**, 179 (1995).

[4] DELPHI Collaboration, Eur. Phys. J **C4**, 1 (1998).

[5] DELPHI Collaboration, Phys. Lett. **B414**, 401 (1997).

[6] DELPHI Collaboration, Phys. Lett. **B449**, 383 (1999).

[7] T.Sjöstrand, Phys. Comm. **39**, 346 (1986).

[8] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, Phys. Rep. **97**, 31 (1983).

[9] T. Sjöstrand, Nucl. Phys. **B248**, 469 (1984).

[10] T. Sjöstrand, Int. J. Mod. Phys. **A3**, 751 (1988).

[11] T. Sjöstrand et al. eds. G. Altarelli, R. Kleiss and C. Verzegnassi, Z physics at LEP 1, CERN 89-08 3, 143 (1989).

[12] C. Peterson, D. Schlatter, I. Schmitt and P. Zerwas, Phys. Rev. **D27**, 105 (1983).

[13] DELPHI Collaboration Nucl. Instr. Meth. **A378**, 57 (1996).

[14] Ts. Spassov, N. Smirnov, **Skelana Users Guide**.

[15] CERN Program Library Long Writeup Y250 **HBOOK Reference Manual**.

[16] CERN Program Library Long Writeup Q121 **PAW The Complete Reference**.

[17] R. A. Fisher, The Use of Multiple Measurements in Taxonomic Problems, Contributions to Mathematical Statistics (reprinted) John Wiley&Sons, New York, (1950).

[18] F. James, Introduction to neural networks, Cern lectures, (1994).

[19] A. Cherubini and R. Odorico, in *Neural Networks: from biology to high energy physics*, EIPC, Elba (Italy), edited by O. Benhar, C. Bosio, P. Del Giudice and E. Tabet (ETS EDITRICE, piazza Torricelli 4, Pisa, Italy, 1991), pp. 515–549.

[20] D. E. Rummelhart, G. E. Hilton and R. J. Williams, Learning Internal Representation by Error Propagation, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol.1, MIT Press, (1986).

[21] L. Lönnblad and C. Peterson and T. Rögnvaldsson, Comput. Phys. Comm. **81**, 185 (1994).

[22] DELPHI Collaboration, PROG 130, CERN DELPHI 89-15.

[23] DELPHI Collaboration, Zeit. Phys. **C70**, 179 (1996).

[24] V. Barger, R. Phillips, Collider Physics, Addison-Wesley, (1997).

[25] DELPHI Collaboration, Zeit.Phys **C 66**, 323 (1995).

[26] DELPHI Collaboration, Zeit. Phys. **C65**, 555 (1995).

[27] ALEPH Collaboration, Phys. Lett. **B313**, 535 (1993).

[28] G.V. Borisov, Lifetime Tag of events $Z^0 \rightarrow b\bar{b}$ with the DELPHI detector. AABTAG program, DELPHI 94-125.

[29] L. Lönnblad and C. Peterson and T. Rögnvaldsson, Phys. Rev. Lett. **65**, 1321 (1990).

[30] I. Csabai and F. Czako and Z. Fodor, Nucl.Phys. **B374**, 288 (1992).

[31] J. Pumplin, How to tell Quark Jets from Gluon Jets, MSUTH 91/3, (1991).

[32] I. Csabai and F. Czako and Z. Fodor, Phys.Rev. **D44**, 1905 (1991).

[33] T. Åkesson, O. Bärring and V. Hedberg, in *Application and Science of Artificial Neural Networks*, SPIE, The International Society for Optical Engineering, edited by Steven K.Rogers, Dennis W. Ruck (SPIE, P.O.Box 10, Bellingham, Washington 98227–0010 USA, 1995), Vol. 2492, pp. 1121–1131.

[34] M. A. Graham and L. M. Jones and S. Herbin, Phys. Rev. **D51**, 4789 (1995).

[35] M. A. Graham, L. M. Jones and P. R. Daumerie, Phys. Rev. **D46**, 222 (1992).

[36] L. Lönnblad and C. Peterson and T. Rögnvaldsson, Nucl.Phys. **B349**, 675 (1991).

[37] DELPHI Collaboration, Phys. Lett. **B295**, 383 (1992).

[38] ALEPH Collaboration, Phys.Lett. **B313**, 549 (1993).

[39] D. Bertrand, V. Lefébure, in *Application and Science of Artificial Neural Networks*, SPIE, The International Society for Optical Engineering, edited by Steven K.Rogers, Dennis W. Ruck (SPIE, P.O.Box 10, Bellingham, Washington 98227–0010 USA, 1995), Vol. 2492, pp. 1098–1108.

[40] DELPHI Collaboration, Nucl.Phys. **B421**, 3 (1994).

[41] C. Caso et al, The European Physical Journal **C3**, 1 (1998).

[42] R. J. Scalkoff, Pattern Recognition: Statistical, Structural and Neural Approaches, John Wiley&Sons, New York (1992).

[43] DELPHI Collaboration, Eur. Phys. **C4**, 1 (1998).

[44] DELPHI Collaboration, Phys. Lett. **B414**, 401 (1997).

[45] DELPHI Collaboration, Phys. Lett. **B449**, 383 (1999).

Appendix A

Geometry for Classification Purposes

For a complete treatment of pattern recognition, see [42]. In 3-D or $(\vec{x_1}, \vec{x_2}, \vec{x_3})$-space a plane through the origin is determined by three parameters $(\vec{\omega_1}, \vec{\omega_2}, \vec{\omega_3})$ via

$$(\vec{\omega_1}, \vec{\omega_2}, \vec{\omega_3}) \cdot \begin{pmatrix} \vec{x_1} \\ \vec{x_2} \\ \vec{x_3} \end{pmatrix} = 0$$

or simply

$$\vec{\omega}^T \cdot \vec{x} = 0$$

This equation indicates that $\vec{\omega}$ and $\vec{x}$ are orthogonal. $\vec{\omega}$ is therefore the normal to the plane. A plane through any other point X_d, represented by position vector $\vec{x_d}$, may be written

$$\vec{\omega}^T \cdot (\vec{x} - \vec{x_d}) = 0$$

or

$$\vec{\omega}^T \cdot \vec{x} - d = 0$$

where $d = \vec{\omega}^T \cdot \vec{x_d}$
This reformation is equivalent to a coordinate system transformation, where the origin is shifted to $\vec{x_d}$ and therefore measure vectors
$\vec{x'} = \vec{x} - \vec{x_d}$
Notice that $\vec{\omega}$ is orthogonal to any vector, lying in the plane.

$$\vec{\omega}^T \cdot \vec{x} - d = 0$$

can be used to determine the "distance" of H from the origin and for a given vector $\vec{x}$, ~~which "side" of H in R^d that $\vec{x}$ is on and the distance of $\vec{x}$ from H.~~
If $\vec{\omega}$ is normalised by dividing it with its norm, giving the vector $\vec{\omega'}$,

$$\vec{\omega'} = \frac{\vec{\omega}}{\|\vec{\omega}\|}$$

The vector $\vec{x_d}$ can then be written

$$\vec{x_d} = \alpha \cdot \vec{\omega}'$$

where

$$\|\vec{x_d}\| = \alpha$$

Since $\vec{x_d}$ represents a point on H

$$\vec{\omega}^T \cdot \vec{x_d} = 0$$

so

$$\alpha \cdot \frac{\vec{\omega}^T \cdot \vec{\omega}}{\|\vec{\omega}\|} - d = 0$$

This means that α can be written as,

$$\alpha = \frac{d}{\|\vec{\omega}\|} = \|\vec{x_d}\| = \frac{\|\vec{\omega}^T \cdot \vec{x_d}\|}{\|\vec{\omega}\|}$$

So any point x represented by a vector $\vec{x}$ may be written,

$$\vec{x} = \vec{x_d} + \vec{x_p} + \vec{x_n}$$

and if $\|\vec{x_n}\| = k$, the vector $\vec{x_n}$ can be written

$$\vec{x_n} = \frac{\vec{\omega}}{\|\vec{\omega}\|} \cdot k$$

and therefore

$$\vec{\omega}^T \cdot \vec{x} = \vec{\omega}^T \cdot (\vec{x_d} + \vec{x_p} + \vec{x_n}) = \vec{\omega}^T \cdot \vec{x_n} + \vec{\omega}^T \cdot \vec{x_d} = k \cdot \|\vec{\omega}\| + \alpha \cdot \|\vec{\omega}\|$$

Let, for simplicity $\alpha = 0$ (i.e. a plane through the origin). Since $\vec{x_n} = k \cdot \vec{\omega}'$ the quantity

$$\vec{\omega}^T \cdot \vec{x} = k \cdot \|\vec{\omega}\|$$

tells which side of the line H, $\vec{x}$ is on.

$$If\ \vec{\omega}^T \cdot \vec{x} > 0 \Rightarrow k > 0 \Rightarrow \vec{x}\ is\ on\ the\ "positive"\ side$$

$$If\ \vec{\omega}^T \cdot \vec{x} < 0 \Rightarrow k < 0 \Rightarrow \vec{x}\ is\ on\ the\ "negative"\ side$$

$$If\ \vec{\omega}^T \cdot \vec{x} = 0 \Rightarrow k = 0 \Rightarrow \vec{x}\ is\ in\ H$$

For an figure illustrating this discussion, see figure A.1. Notice that the plane H (given by) partitions R^2 (in general R^d) into two mutually exclusive regions, denoted R^p and R^n.

Following this mathematical outline and by looking at the figure A.1, the aim is to make the intuitive description of the linear discrimination method, outlined in section 4.1 a bit more trustworthy.

The assignment of a vector $\vec{x}$ to the "positive" or "negative" side, or along H can be implemented by using the expression $\vec{\omega}^T \cdot \vec{x} - d$ in the following way,

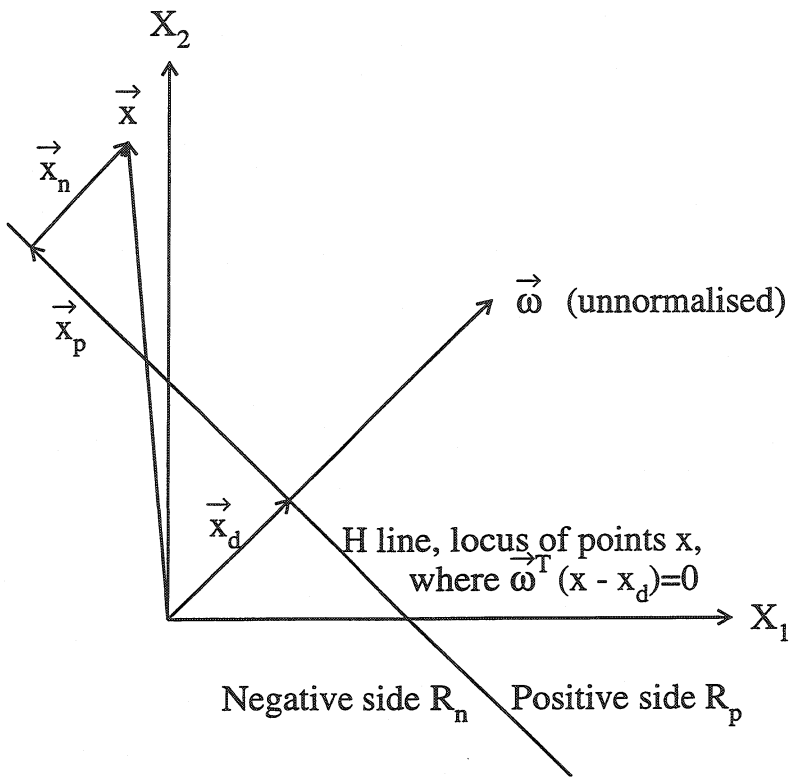

Figure A.1: The assignment of a vector $\vec{x}$ to the positive or negative side of (or along) the line H.

- $\vec{\omega}^T \cdot \vec{x} - d > 0 \; if \vec{x} \in R_p$
- $\vec{\omega}^T \cdot \vec{x} - d = 0 \; if \vec{x} \in H$
- $\vec{\omega}^T \cdot \vec{x} - d < 0 \; if \vec{x} \in R_n$

This suggests a *linear discriminant function* $F(\vec{x})$ to implement the classification of the assignment above, as

$$F(\vec{x}) = \vec{\omega}^T \cdot \vec{x} - d \tag{A.1}$$

Examination of equation A.1 and figure A.1 indicates that the linear discriminants function that yields H and is used to partition R^2 into two regions has the following characteristics

- The orientation of H is determined by $\vec{\omega}$
- The location of H is determined by d

Extension to R^d is straightforward for the linear classification of d-dimensional feature vectors $\vec{x}$. The surface H is, in this context, referred to as a *hyperplane.*

Appendix B

The energy dependence of the longitudinal momenta

Figures B.1-B.5 show the differencies in the input variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the second to the sixth particle of the jet.

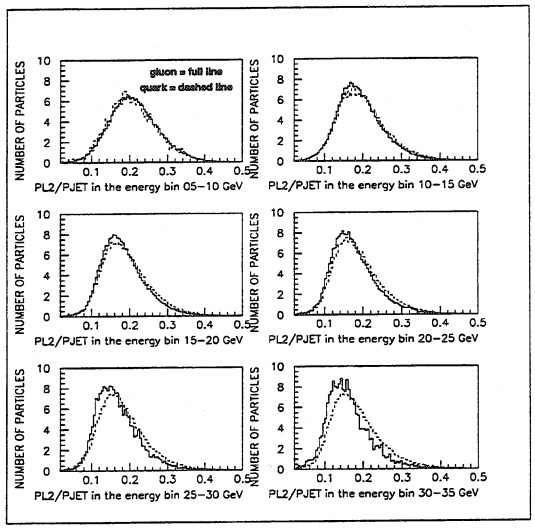

Figure B.1: The distributions of the fragmentation variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the second particle of the jet for gluonjets and quarkjets, in the six energy bins.

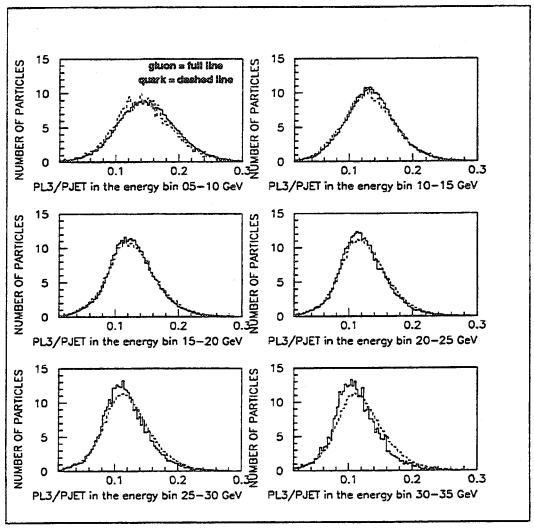

Figure B.2: The distributions of the fragmentation variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the third particle of the jet for gluonjets and quarkjets, in the six energy bins.

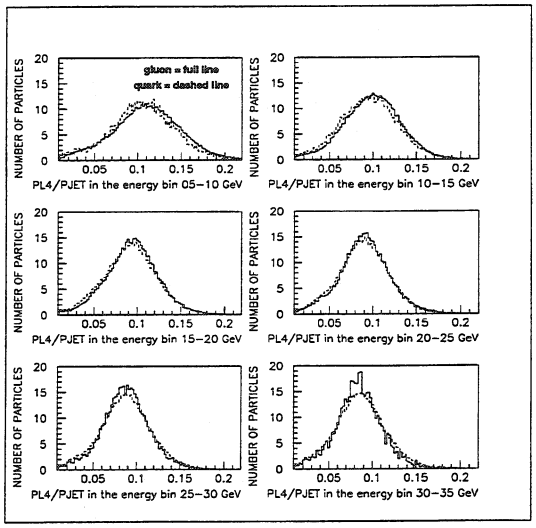

Figure B.3: The distributions of the fragmentation variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the fourth particle of the jet for gluonjets and quarkjets, in the six energy bins.

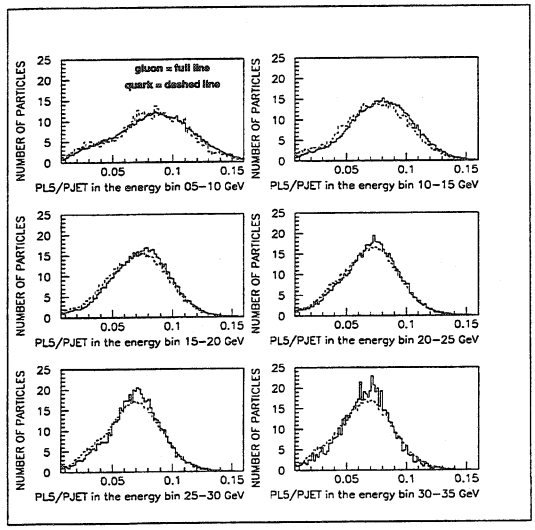

Figure B.4: The distributions of the fragmentation variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the fifth particle of the jet for gluonjets and quarkjets, in the six energy bins.

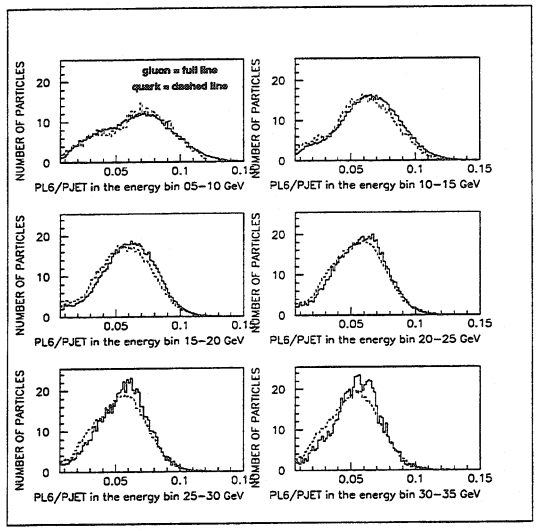

Figure B.5: The distributions of the fragmentation variable, the normalized longitudinal momentum, relative to the jet axis, p_L/p_{jet} for the sixth particle of the jet for gluonjets and quarkjets, in the six energy bins.

Appendix C

The energy dependence of the transverse momenta

Figures C.1-C.5 show the differencies in the input variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the second to the sixth particle of the jet.

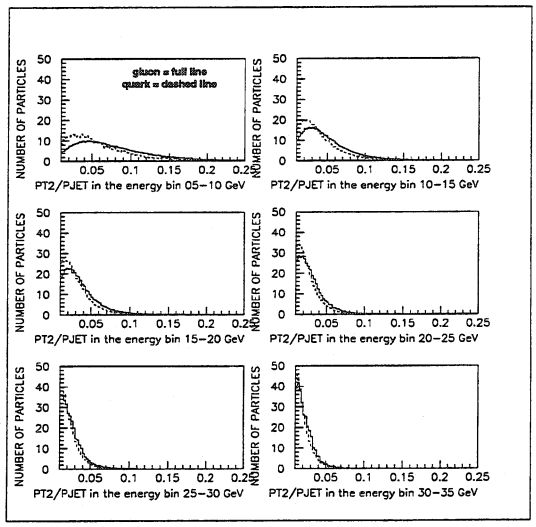

Figure C.1: The distributions of the fragmentation variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the second particle of the jet for gluonjets and quarkjets, in the six energy bins.

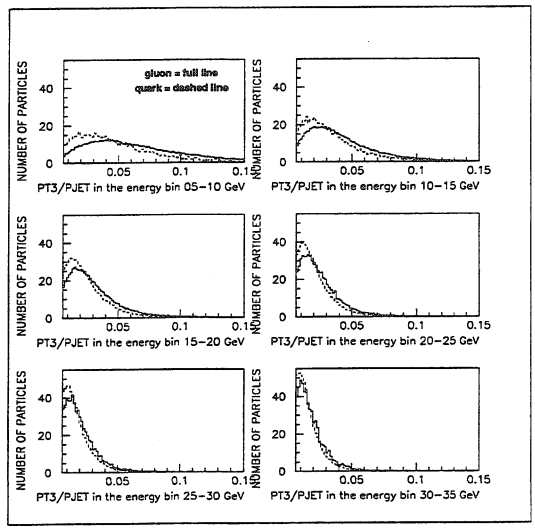

Figure C.2: The distributions of the fragmentation variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the third particle of the jet for gluonjets and quarkjets, in the six energy bins.

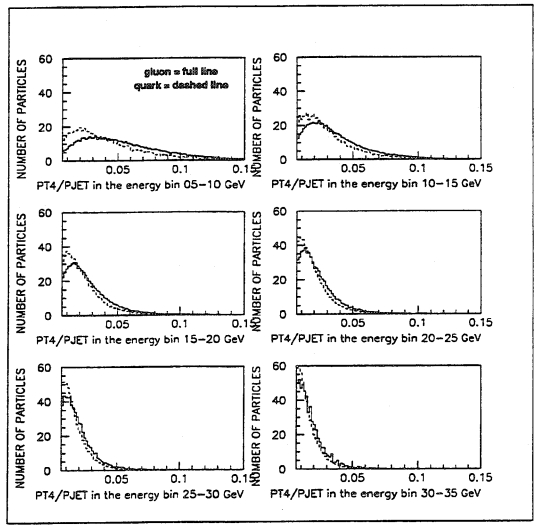

Figure C.3: The distributions of the fragmentation variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the fourth particle of the jet for gluonjets and quarkjets, in the six energy bins.

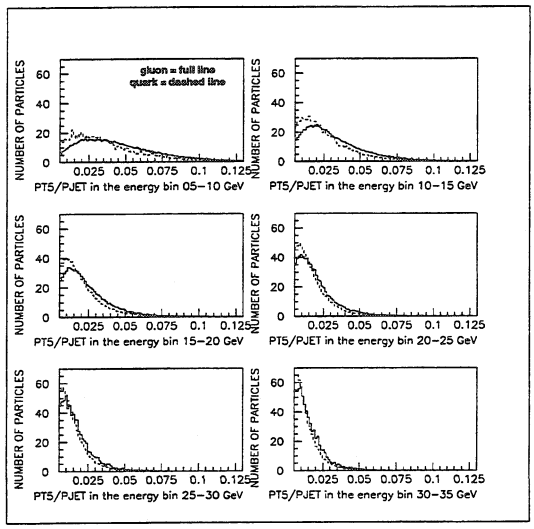

Figure C.4: The distributions of the fragmentation variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the fifth particle of the jet for gluonjets and quarkjets, in the six energy bins.

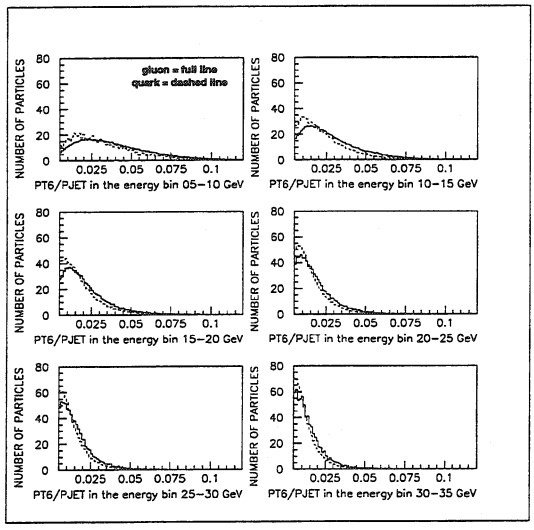

Figure C.5: The distributions of the fragmentation variable, the normalized transverse momentum, relative to the jet axis, p_T/p_{jet} for the sixth particle of the jet, for gluonjets and quarkjets in the six energy bins.

Appendix D

The energy dependence of the Fodor momenta

Figures D.1-D.6 show the differencies in the input variables, the normalized Fodor momenta, for gluon jets and quark jets in the full energy range.

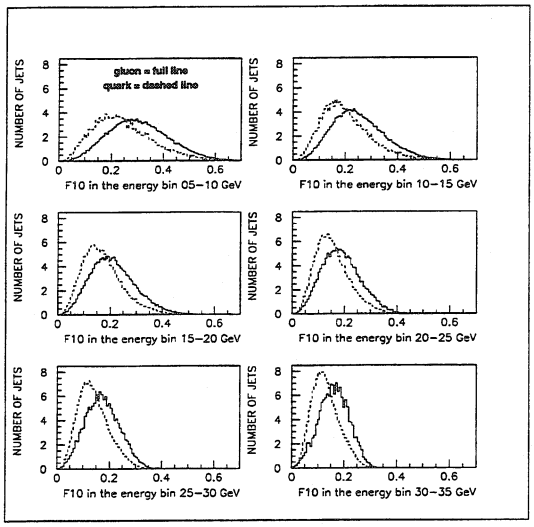

Figure D.1: The distributions of the normalised Fodor moment F10, for gluon jets and quark jets in the six energy bins.

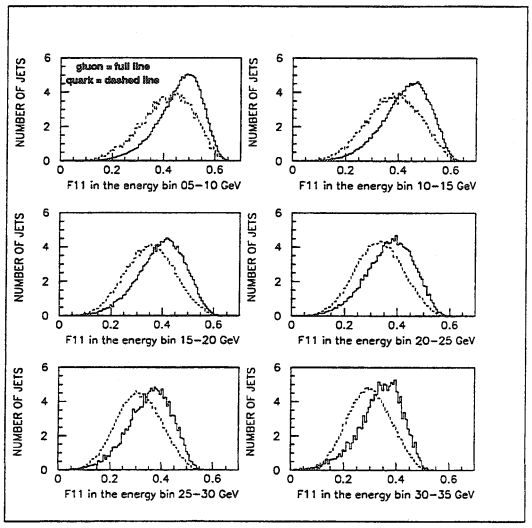

Figure D.2: The distributions of the normalised Fodor moment, F11, for gluon jets and quark jets in the six energy bins.

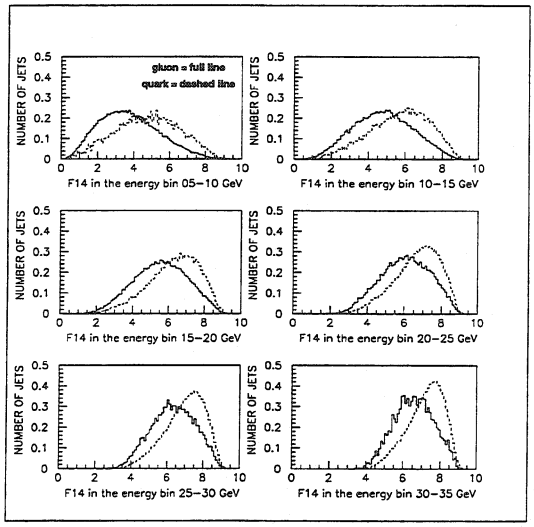

Figure D.3: The distributions of the normalised Fodor moment, F14, for gluon jets and quark jets in the six energy bins.

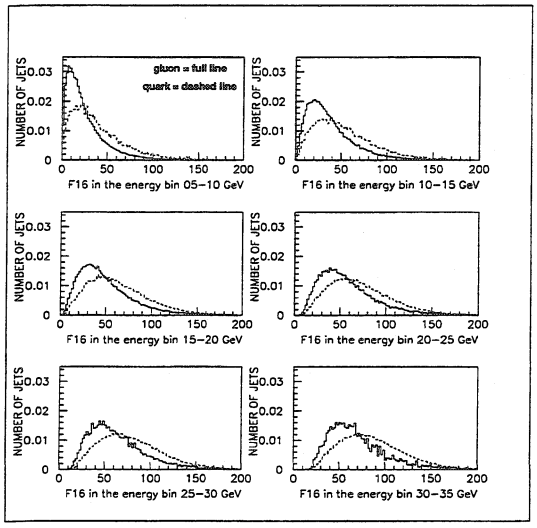

Figure D.4: The distributions of the normalised Fodor moment, F16, for gluon jets and quark jets in the six energy bins.

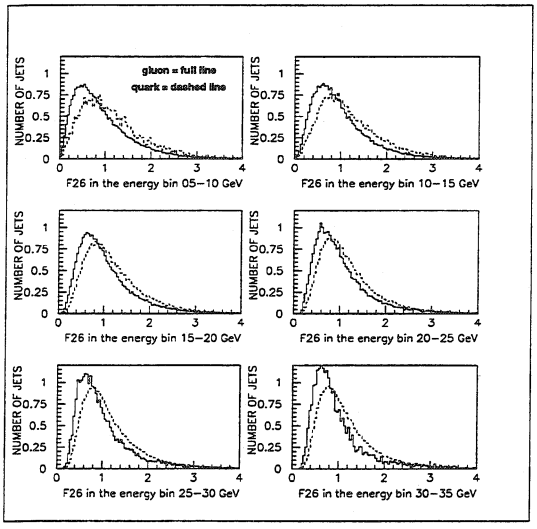

Figure D.5: The distributions of the normalised Fodor moment, F26, for gluon jets and quark jets, in the six energy bins.

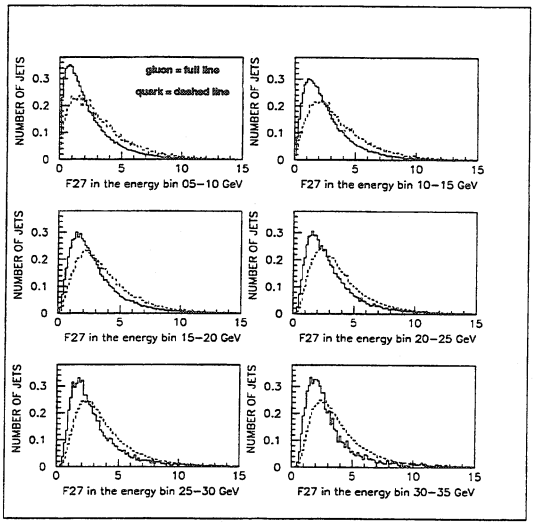

Figure D.6: The distributions of the normalised Fodor moment, F27, for gluon jets and quark jets in the six energy bins.