

15

Data analysis*

Without the hard little bits of marble which are called ‘facts’ or ‘data’ one cannot compose a mosaic; what matters, however, are not so much the individual bits, but the successive patterns into which you arrange them, then break them up and rearrange them.

Arthur Koestler

15.1 Introduction

The analysis of data and extraction of relevant results are goals of particle physics and astroparticle physics experiments. This involves the processing of *raw detector data* to yield a variety of final-state physics objects followed by the application of selection criteria designed to extract and study a signal process of interest while rejecting (or reducing to a known and manageable level) background processes which may mimic it. Collectively, this is referred to as an *analysis*. Physics analyses are performed either to measure a known physical quantity (e.g. the lifetime of an unstable particle), or to determine if the data are compatible with a physics hypothesis (e.g. the existence of a Higgs boson). At each stage, however, a variety of ‘higher-order’ issues separate particle physics and astroparticle physics analyses from a brute-force application of signal-processing techniques.

15.2 Reconstruction of raw detector data

All physics analysis starts from the information supplied from the data-acquisition system, the raw detector data. In contemporary collider

* Steve Armstrong, CERN, now at Société Générale de Surveillance (SGS), contributed to this chapter. It is an updated version of the original data analysis chapter in the first edition of this book originally written by Armin Böhrer, Siegen.

experiments or cosmic-ray experiments, these raw detector data consist of the digitised output of detector electronic signals. Signals are induced in the detector electronics by the passage of particles, which leave ‘hits’ in active detector elements. Modern detectors are frequently highly granular, resulting in dozens or even hundreds of hits per particle per detector system.

The process referred to as *event reconstruction* aims to produce meaningful physics objects from the binary data associated with these hits while coping with electronic noise from the detectors themselves and the inherent physical processes associated with the passage of the final-state particles through the detector material. Raw detector data must also be merged with other predefined sets of data in the event-reconstruction process. A *detector description* contains detailed information on the geometry, position and orientation of active detector elements. Calibration and alignment data contain intrinsic quantities related to detector components which influence its performance (e.g. gas purity, high voltage values, temperatures, etc.); frequently, these data are assumed to be constant over a specific duration of data-taking time, referred to as a *run*.

In the previous chapters, a wide variety of detector technologies have been reviewed. A few examples of information which can be extracted from each kind of detector are presented for illustration:

- **Silicon microstrip detectors (SMDs)** As described in Sect. 7.5, particles traversing an SMD ionise the bulk silicon material liberating electron–hole pairs which drift to implanted strips generating measurable signals. Raw data include pulses recorded on strips around the signal region to permit the use of interpolation techniques. When combined with knowledge of the local strip positions as well as the global position of the silicon wafer, two- or three-dimensional coordinates may be extracted.
- **Multiwire proportional chambers (MWPCs)** As described in Sect. 7.1, particles traversing a MWPC ionise the gas. The electrons initiate avalanches that create a signal when the charge is collected at the electrodes. Raw data include the drift time, the wire and pad positions as well as the arrival time of the pulse and the charge at both ends of the wire. Combining this input with calibration constants, such as drift velocity and the moment of the intersection t_0 , the location of the electron initiating the avalanche both in the plane perpendicular to the wire and along the wire may be obtained. When the global position of the wire is included, three-dimensional coordinates may be extracted.

- **Time-projection chambers (TPCs)** As described in Sect. 7.3.3, charged particles traversing a TPC volume leave ionisation trails in the gas. Ionisation electrons drift towards an end plate containing an MWPC. Raw data contain both the drift time and the profile of signals on wires and cathode pads to permit determination of the z and ϕ coordinates. When combined with knowledge of the wire and pad position, the r position may also be extracted. The pulse heights of hits may also be included, and, when considered along a putative track trajectory, may yield dE/dx information providing valuable particle-identification discrimination.
- **Electromagnetic and hadron calorimeters** (see Chap. 8) In many high energy physics experiments, energy measurements are made using a combination of electromagnetic and hadron calorimeters. Usually, sampling calorimeters are used, where absorber material is interspersed with chambers or scintillator material providing analogue information proportional to the energy deposited. The accuracy of the location where a particle has passed through the detector is limited by the number of calorimeter cells, or granularity, of the calorimeter. The granularity is typically determined by the intrinsic nature of the calorimeter (e.g. electromagnetic, hadron, compensating hadron) as well as the number of readout channels that can be handled (e.g., the liquid-argon electromagnetic calorimeter for the ATLAS experiment at LHC will have more than $2 \cdot 10^5$ calorimeter cells). Each cell which exceeds a certain threshold is read out and becomes part of the detector raw data. Once position information of the cells is factored in, along with calibration information, cells may be clustered together to determine localised energy depositions.
- **Time-of-flight (TOF) detectors** Scintillators, resistive-plate chambers (RPCs), planar spark counters or spark chambers frequently have the ability to record with high precision the time of passage of a charged particle. When placed at sufficient distance from each other or from a known interaction point, they can provide valuable timing information, which, when used in conjunction with a measurement of the particles' momentum, may be used for particle-identification discrimination. These types of detectors are also often used as trigger counters.
- **Specialised particle identification detectors** In addition to the TPC and TOF detectors discussed above, additional particle-identification information can be obtained from Cherenkov counters in which the position and diameter of Cherenkov rings are recorded, or transition radiation detectors in which the yield of X-ray pho-

tons from charged particles traversing media of differing dielectric constants is measured.

High *granularity* and *hermeticity* are required for modern collider-detector experiments. Hence, the total number of electronic channels can easily exceed 10^8 , and, when processed by front-end and intermediate electronics, can lead to raw event sizes in the range of several Mbytes. The demand for large amounts of data in which to search for rare signatures of known or new physics requires high interaction rates; for example, the LHC interaction rate will be 40 MHz. To reduce this to the range of 100 Hz for offline storage requires sophisticated trigger systems, frequently consisting of multiple *levels* of hardware and software.

Once written to storage, events must undergo further processing. Raw event data are fed into reconstruction algorithms which process the events. The previous generation of experiments (i.e. the LEP experiments) used the FORTRAN programming language. Current and future experiments have migrated their reconstruction software to object-oriented languages such as C++ and Java. In either case, event reconstruction yields basic physics objects such as charged-particle trajectories in a tracking detector or clusterised energy depositions (*energy-flow objects*) in a calorimeter. These basic physics objects are the fundamental building blocks of analysis, and are discussed further below.

15.3 Analysis challenges

Once the reconstructed physics objects are available for analysis, *selection criteria* must be designed and applied to them. Most often, these selection criteria are designed on the basis of a *Monte Carlo simulation* of relevant underlying physics processes as well as a simulation of the response of the detector to final-state physics objects. The choice of selection criteria is often a balance between many complementary challenges.

The first challenge is the optimisation and enhancement of the statistical significance of a signal process which involves achieving a high efficiency for signal as well as a high rejection power for background. Frequently, these selection criteria consist of *cuts* on a variety of kinematic features derived from the four-momenta of the final-state physics objects. Additional criteria may be placed upon other characteristics such as particle-identification information. The higher the signal efficiency and background rejection, the fewer data are required to achieve a physics result. In the era of expensive or limited data-taking opportunities, this endeavour has recruited advanced multivariate techniques to exploit fully the information within the data.

The second challenge is the understanding of systematic uncertainties induced into the physics result by the choice of selection criteria. If the modelling of the kinematic behaviour of the underlying physics processes is flawed within the simulation upon which selection criteria are based, a bias is introduced into the final result. Uncertainties associated with this type are referred to as *theoretical* systematic uncertainties. If the modelling of the detector response is imperfect, a further bias may be introduced into the final result; this is referred to as *experimental* systematic uncertainties. Finally, in the era of complex and large detectors, the size of simulated event samples is frequently limited by available computing resources. Frequently, in simulation, only a handful of important events of a relevant physics process are retained after selection criteria are imposed. This induces a *statistical* systematic uncertainty on the final result.

In contemporary high-statistics experiments special attention should be paid to a careful estimate of possible systematic uncertainties. Let us consider as an example a recent publication of the Belle Collaboration describing the first observation of a rare decay of the τ lepton: $\tau^- \rightarrow \phi K^- \nu_\tau$ [1] (see also Chap. 13). The analysis is based on a data sample of 401 fb^{-1} corresponding to 3.58×10^8 events of the process $e^+e^- \rightarrow \tau^+\tau^-$ produced at a centre-of-mass energy of 10.58 GeV. For this study events are selected where one τ lepton decays purely leptonically (tag side) while the other one decays into the $K^+K^-K^\pm\nu_\tau$ final state (signal side). To obtain the number of decays under study, the K^+K^- invariant-mass spectrum, containing the ϕ -meson peak smeared by the detector resolution and a smooth background, was fit. Then the number of signal events (often referred to as a signal yield), $N_{\text{sig}} = 573 \pm 32$, is derived after subtracting the peaking backgrounds coming from the $\tau^- \rightarrow \phi\pi^-\nu_\tau$ decay and the $q\bar{q}$ continuum. As can be seen, the statistical uncertainty of N_{sig} is higher than just $\sqrt{N_{\text{sig}}}$. Figure 15.1 shows the result of this fitting procedure.

In this way the branching ratio is obtained from the formula

$$\mathcal{B} = \frac{N_{\text{sig}}}{2N_{\tau^+\tau^-}\varepsilon} , \quad (15.1)$$

where $N_{\tau^+\tau^-}$ is the number of produced $\tau^+\tau^-$ pairs and ε is the detection efficiency obtained from a Monte Carlo simulation. The result is

$$\mathcal{B} = (4.05 \pm 0.25) \cdot 10^{-5} , \quad (15.2)$$

where the error is statistical, only determined by the number of selected signal events and that of subtracted background events. Systematic uncertainties are estimated as follows. The systematic error on the signal yield

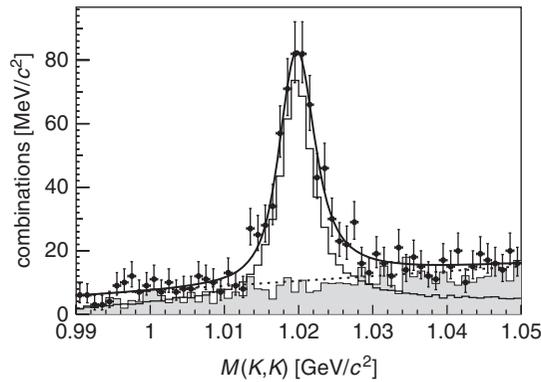


Fig. 15.1. K^+K^- invariant-mass distributions for $\tau^- \rightarrow \phi K^- \nu_\tau$. Points with error bars indicate the data. The shaded histogram shows the expectations from $\tau^+\tau^-$ and $q\bar{q}$ background MC simulations. The open histogram is the signal MC with $\mathcal{B}(\tau^- \rightarrow \phi K^- \nu_\tau) = 4 \cdot 10^{-5}$ [1].

in the numerator of Eq. (15.1) equals 0.2% and is determined by varying the value of the ϕ -meson width and the shape of background parametrisation. The systematic uncertainty of $N_{\tau^+\tau^-}$ originates from the uncertainty of the integrated luminosity (1.4%) and our inexact knowledge of the theoretical cross section of the process $e^+e^- \rightarrow \tau^+\tau^-(\gamma)$ (1.3%). The dominant uncertainty is due to the detection efficiency, which is affected by various factors: trigger efficiency (1.1%), track-finding efficiency (4%), lepton and kaon identification (3.2% and 3.1%, respectively), the branching fraction of the $\phi \rightarrow K^+K^-$ decay (1.2%), and Monte Carlo statistics (0.5%). A total systematic uncertainty of 6.5% is obtained by adding all uncertainties in quadrature, assuming that they are not correlated. The resulting branching fraction is

$$\mathcal{B} = (4.05 \pm 0.25 \pm 0.26) \cdot 10^{-5} . \quad (15.3)$$

Note that most of the uncertainties listed above are determined by using various control data samples.

15.4 Analysis building blocks

Different physics analyses can frequently be separated and identified by applying specific selection criteria imposed upon quantities related to various standard final-state physics objects provided by reconstruction of raw detector data. After the identification of these objects and the determination of relevant parameters (e.g. four-momenta, impact parameters with respect to an origin, or decay point of a long-lived particle),

relevant results-oriented quantities may be determined. This section discusses objects most commonly encountered in the context of analyses of data from high-energy general-purpose collider detector experiments. Most described techniques also apply to cosmic ray and astroparticle physics experiments.

15.4.1 Charged-particle trajectories

A variety of detector technologies exist which can aid in the reconstruction of the trajectories of charged particles (hereafter also referred to as *tracks*) through a given volume. Such detectors may be grouped together to form a *tracking system* in which tracks are reconstructed from their measured spatial coordinates. Tracking systems are usually immersed in powerful magnetic fields of known strength at each point in the fiducial volume so that the electric charge and momentum may be measured.

Events resulting from the high-energy collisions or interactions may contain anywhere from several dozen to several thousand charged particles, which leave hits in the tracking system. To illustrate the extreme complexity, an event in the Inner Detector of the future ATLAS experiment is shown in Fig. 15.2. This detector, typical of contemporary

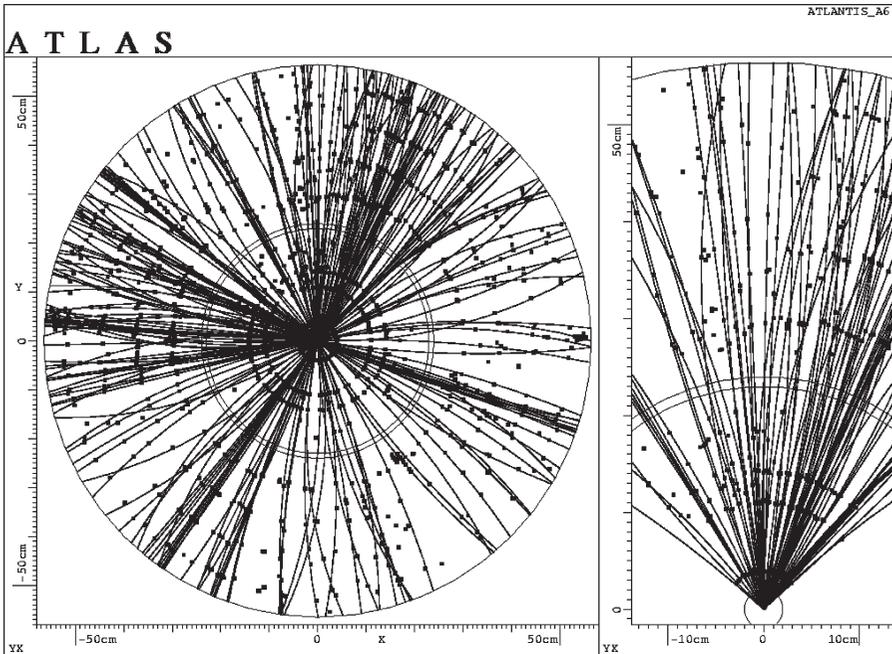


Fig. 15.2. Reconstruction of charged-particle trajectories in a typical event in the Inner Detector of the ATLAS experiment [2].

tracking chambers, consists of multiple concentric subdetectors utilising different technologies. It combines high-resolution silicon pixel and microstrip detectors at the inner radii (pixel and silicon tracker (SCT)) with a straw-tube tracker at the outer radii (transition radiation tracker (TRT)). Figure 15.2 shows a display in the transverse plane of a simulated event with typical charged-particle multiplicity. The tracking chambers are immersed in a 2 T axial magnetic field parallel to the beam line. Raw measured spatial coordinates are shown with the dots while reconstructed tracks are denoted by the curved lines intersecting relevant spatial coordinates.

Pattern-recognition algorithms attempt to group tracking-detector hits together, first forming two- or three-dimensional coordinates from which track candidates may be found. The challenge then becomes how to group these coordinates together to form tracks. There are two extreme possibilities.

The straightforward method of taking all possible combinations of hits is too time-consuming. The number of combinations for thousands of hits is immense and all possible track candidates must be validated so as not to use hits twice, i.e. by several tracks. The other extreme point of view is the global method where a classification of all tracks is done simultaneously. For points close in space, characteristic values (e.g. coordinates) are entered in an n -dimensional histogram. Hits belonging to the same track should be close in parameter space. A simple example would be the reconstruction of tracks coming from the interaction point without a magnetic field. The ratios $(y_i - y_j)/(x_i - x_j)$ calculated for all i, j pairs of points (Fig. 15.3) and plotted in a histogram would show peaks at the values of the slopes expected for straight tracks.

In practice a method lying between these two approaches is chosen. Its implementation depends heavily on the chamber layout and physics involved.

One method that is commonly used is the *road method*. It is explained most easily for the example of the muon chambers consisting of two double layers of staggered drift tubes (see Fig. 15.3). Reconstructed spatial coordinates for a charged particle outside a magnetic field lie essentially on a straight line. Possible tracks are found from the permutation list of four points lying on a road of a width which corresponds roughly to the spatial resolution (mm or cm).

To all four points on a road with coordinates x_1, \dots, x_4 one has measurements y_i with errors σ_i (here Gaussian errors are assumed). In a straight line fit [3, 4] the expected positions η_i with respect to the measured y_i are linear functions of the x_i :

$$\eta_i = y_i - \epsilon_i = x_i \cdot a_1 + 1 \cdot a_2, \quad (15.4)$$

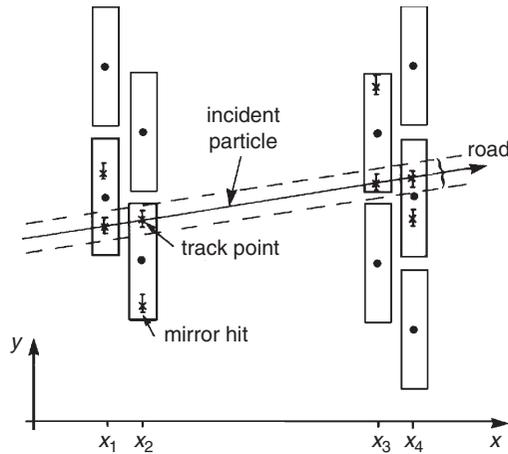


Fig. 15.3. Track finding with the road method and straight line fit. Due to the left–right ambiguity of drift chambers, two coordinates per hit are reconstructed: one being the true track point, the other a mirror hit.

or

$$\vec{\eta} = \vec{y} - \vec{\epsilon} = \mathcal{X} \cdot \vec{a} , \quad (15.5)$$

where a_1 is the slope and a_2 is the axis intercept. The matrix \mathcal{X} contains the coordinates x_i in the first and values 1 in the second column. For independent measurements the covariance matrix \mathcal{C}_y is diagonal:

$$\mathcal{C}_y = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix} =: \mathcal{G}_y^{-1} . \quad (15.6)$$

One obtains the values of \vec{a} by the least-squares method, minimising

$$\chi^2 = \vec{\epsilon}^T \mathcal{G}_y \vec{\epsilon} , \quad (15.7)$$

which follows a χ^2 distribution with $4 - 2 = 2$ degrees of freedom:

$$\vec{a} = (\mathcal{X}^T \mathcal{G}_y \mathcal{X})^{-1} \mathcal{X}^T \mathcal{G}_y \vec{y} , \quad (15.8)$$

and the covariance matrix for \vec{a} is given by

$$\mathcal{C}_a = (\mathcal{X}^T \mathcal{G}_y \mathcal{X})^{-1} =: \mathcal{G}_a^{-1} . \quad (15.9)$$

As shown in Fig. 15.3, several track candidates may be fitted to the data points, because of hit ambiguities. To resolve these, the χ^2 can be translated into a confidence limit for the hypothesis of a straight line to be true

and one can keep tracks with a confidence level, for example, of more than 99%. The more commonly used choice is to accept the candidate with the smallest χ^2 . By this method mirror hits are excluded and ambiguities are resolved.

Points which have been used are marked so that they are not considered for the next track. When all four-point tracks have been found, three-point tracks are searched for to allow for inefficiencies of the drift tube and to account for dead zones between them.

For larger chambers with many tracks, usually in a magnetic field, the following track-finding strategy is adopted. The procedure starts in those places of the drift chamber, where the hit density is lowest, i.e. farthest away from the interaction point. In a first step three consecutive wires with hits are searched for. The expected trajectory of a charged particle in a magnetic field is a helix. As an approximation to a helix, a parabola is fitted to the three hits. This is then extrapolated to the next wire layer or chamber segment. If a hit matching within the errors is found, a new parabola fit is performed. Five to ten consecutive points form a track segment or a chain. In this chain at most two neighbouring wires are allowed not to have a hit. Chain finding is ended when no further points are found or when they do not pass certain quality criteria. When the track-segment finding is complete, the segments are linked by the track-following method. Chains on an arc are joined together and a helix is fitted. Points with large residuals, i.e. points that deviate too much in χ^2 , are rejected and the helix fit is redone. The track is extrapolated to the closest approach of the interaction point. In the final fit variations in the magnetic field are included, and a more sophisticated track model is used. In the ALEPH experiment [5], for example, the closest approach to the beam line in $r\varphi$ is denoted by d_0 with the z coordinate at that point z_0 (the z coordinate is measured parallel to the magnetic field along the beam, see Fig. 15.4). The angle φ_0 of the track in the $r\varphi$ plane with respect to the x axis at closest approach, the dip angle λ_0 at that point, and the curvature ω_0 complete the helix parameters: $\vec{H} = (d_0, z_0, \varphi_0, \lambda_0, \omega_0)$. For some applications the set $(d_0, z_0, p_x, p_y, p_z)$ is used, with p_x, p_y, p_z being the components of the track's momentum at closest approach. This procedure also provides the covariance matrix C for the helix.

The knowledge of the position of the interaction vertex is of particular importance, if one is interested in determining the particles' lifetimes. For colliders the position of the incoming beams is known to be $\approx 200\ \mu\text{m}$ or better, while the length of the colliding bunches may range from a few millimetres to half a metre. The vertex is fitted using all tracks with closest approach to the beam line of less than typically $200\ \mu\text{m}$. This restriction excludes particles not coming from the primary vertex such as

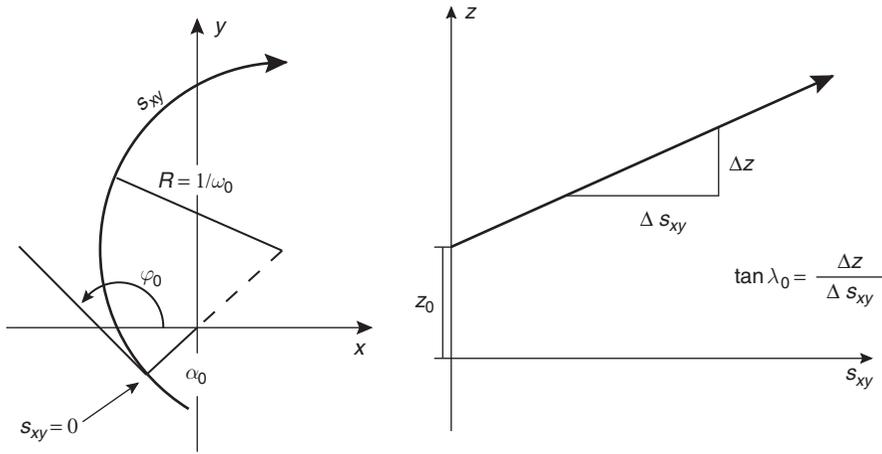


Fig. 15.4. Definition of helix parameters. On the left, the projection of the helix onto the xy plane orthogonal to the magnetic field and the beam is given. The figure on the right shows the z coordinate versus s_{xy} .

K_s^0 , A^0 , \bar{A}^0 , called V^0 , and photon conversions, which produce a pair of oppositely charged tracks.

15.4.2 Energy reconstruction

Calorimeter detector systems provide not only a measurement of energy deposition, but also position measurements where the energy was deposited. This information may be used in either a local or global sense. When used in a local sense, energy depositions within calorimeters may be grouped together to form clusters of energies to associate to tracks or to neutral particles. Furthermore, the profile of the energy deposition may be used for particle identification (see below).

Aggregate quantities such as total event energy as well as *missing energy* (which might originate from neutrino-like objects) are vital for many physics analyses. This requires the vectorial summation of all visible depositions of energy in calorimeters and the correction for signatures in outer muon systems. This yields energy imbalances with respect to known collision energies or the total event energy balance.

Large deviations from expected imbalances may indicate production of new Weakly Interacting Massive Particles such as supersymmetric neutralinos. On the other hand, detailed examination of expected imbalances helps to identify regions of the event where energetic neutrinos may have escaped detection. To find the energy of a possible neutrino one must detect all other particles in the detector. To each energy deposition in the calorimeter one assigns a vector with length proportional to the measured

energy, and its direction given by a line connecting the interaction point with the fired calorimeter cell. A non-zero sum of these vectors in a collider experiment with beams of equal energy and opposite momentum indicates the presence and direction of missing energy. If this is the case, it may be attributed to a neutrino. It must be assumed that no particle escaped, for example, through the beam pipe. Since this cannot be assured, especially for $p\bar{p}$ collider experiments, one usually restricts oneself to the analysis of the momentum transverse to the beam.

In the hard scattering of proton and antiproton only one quark and antiquark collide. The other constituents fragment as jets close to the beam line and partially escape detection. Consequently, the event has a longitudinal imbalance and only the transverse momentum of the neutrino can be used. Certainly, also other corrections have to be taken care of: muons deposit only a small fraction of their energy in the calorimeter. The missing energy must be corrected in this case using the difference between the muon momentum measured in the tracking chamber and its energy seen in the calorimeter.

15.4.3 Quark jets

Quarks produced from or participating in high-energy collisions may manifest themselves as collimated jets of hadrons at sufficiently high energies; this was first observed at centre-of-mass energies near 7 GeV [6]. Quarks may also bremsstrahl gluons thereby creating additional jets in a hadronic event. Primary quarks and any gluons which they may radiate are referred to as *initial partons*. Initial partons carry colour charges and cannot exist in isolation since Nature apparently permits only colour-neutral states to exist freely. Non-perturbative QCD processes convert the coloured initial partons into colour-singlet hadrons. This is referred to as *hadronisation*.

Although the hadronisation process is not well understood, phenomenological models exist. Examples of these models are the string model [7] (implemented in the JETSET Monte Carlo programme [8]) and the cluster model [9] (implemented in the HERWIG Monte Carlo programme [10]). In the string model, for example, the confining nature of the strong interaction dictates that the colour potential of the initial partons becomes proportional to their separation at large distances. As the initial partons fly apart from the interaction point, it becomes energetically favourable for additional quark pairs to be produced from the vacuum. Ultimately, the initial coloured partons are transformed into bound colour-singlet hadronic states.

Jets originating from quarks are qualitatively different than jets from gluons. In Quantum Chromodynamics (QCD), the gluon self-interaction

coupling is proportional to the colour factor C_A while the quark–gluon coupling is proportional to the colour factor C_F . The values of the colour factors are determined by the structure of the colour gauge group $SU(3)$. The ratio C_A/C_F is predicted to be $9/4 = 2.25$; this is in good agreement with experimental measurements [11]. Hence gluons are more likely to radiate softer gluons in the hadronisation process, and a gluon jet is consequently broader with a higher particle multiplicity than a light-quark jet of the same energy. These features of gluon jets have been observed in experiment [12].

Although events exhibit jet structure which may have properties qualitatively indicative of the initial parton, jets are intrinsically ill-defined objects. It is impossible to assign all of final-state particles rigorously to a single initial parton. Algorithms exist which cluster charged and neutral particles in an event together to form jets from which an overall four-momentum and other characteristics (e.g. *track multiplicity, jet shapes, etc.*) may be determined. These *jet-clustering algorithms* form the basis of most analyses dealing with hadronic events which rely upon the clustered jets to approximate the direction and energies of the initial partons in an event.

Many commonly used jet-clustering schemes are based upon the JADE algorithm [13]. This recursive algorithm begins by considering each instance of energy deposition (e.g. charged particle associated to a calorimeter cluster or candidate neutral particle cluster) in an event to be a pseudo-jet. Pairs of pseudo-jets are then combined according to a metric defined as

$$y_{ij} = \frac{2E_i E_j (1 - \cos \theta_{ij})}{E_{\text{vis}}^2}, \quad (15.10)$$

where i and j are two pseudo-jets, and E_{vis} is the visible energy in the event (i.e. the sum of the energy of all energy-flow objects). The numerator is essentially the invariant mass squared of the two pseudo-jets. The energy and three-momentum of the new pseudo-jets are determined according to a combination scheme from the energy and three-momenta of a previous pseudo-jet and an energy-flow object, yielding a new set of pseudo-jets. In the E scheme the simple sum of three-momenta and energy is used. The combination procedure is iterated until all y_{ij} are larger than a specified threshold which is referred to as y_{cut} .

Several variants of the JADE scheme exist and have been extensively studied in the context of QCD-related measurements and predictions [14]. DURHAM, one of the JADE variants, has several advantages (e.g. reduced sensitivity to soft-gluon radiation) [15]. In this scheme, the JADE clustering metric is replaced by

$$y_{ij} = \frac{2 \min(E_i, E_j)^2 (1 - \cos \theta_{ij})}{E_{\text{vis}}^2} . \quad (15.11)$$

The numerator is essentially the square of the lower-energy particle's transverse momentum k_{Tij}^2 with respect to the higher-energy particle.

15.4.4 Stable-particle identification

Another important input for the analysis is the identification of particles. Various methods were described in Chap. 9, such as energy-loss measurements dE/dx , use of Cherenkov counters and transition-radiation detectors. The different longitudinal and lateral structure of energy deposition in calorimeters is used to separate electrons from hadrons. The simplest method is to introduce cuts on the corresponding shape parameters. More sophisticated procedures compare the lateral and longitudinal shower shape with a reference using a χ^2 test or *neural networks*. In this case (and in the physics analysis, see below), in contrast to track finding, multilayer feed-forward networks are used. (For pattern recognition feed-back networks are applied.) The input neurons – each neuron represents an energy deposit in a calorimeter cell – are connected with weights to all neurons in a next layer and so forth until one obtains in the last layer one or a few output neurons. The result, which can vary between zero and one, indicates whether the input originated from a pion or an electron. The weights from the neuron connections are adjustable and are obtained by minimising a cost function. This is done by an iterative learning algorithm called *backpropagation* [16–18].

A comparison of these procedures to separate electrons from pions in a calorimeter can be found in [19].

15.4.5 Displaced vertices and unstable-particle reconstruction

The advent of precision tracking detectors with the ability to provide track-impact-parameter resolutions under 50 microns has permitted the use of displaced vertices in a wide spectrum of analysis contexts, most notably in heavy flavour physics. In this method, information from the tracking detectors is extracted not only about the track momentum but also on its precise location. Using an ensemble of tracks, one can fit a hypothesised common origin or *vertex* for these tracks, and compare it to a known collision position or *interaction point*. Vertices with significant displacement from the interaction point result from the decay of beauty and/or charm hadrons. An example of this is shown in Fig. 15.5.

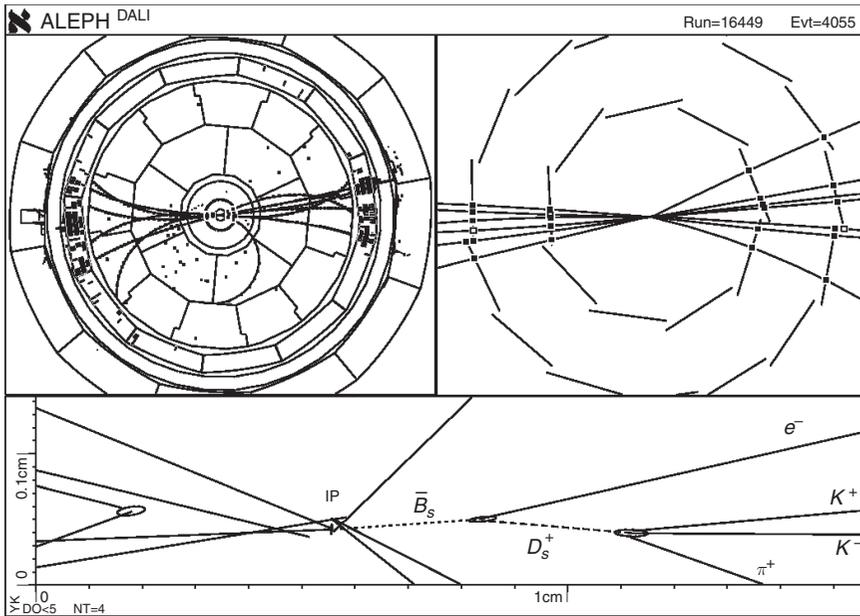


Fig. 15.5. An event display from the ALEPH detector showing the displaced vertex topology of heavy-quark hadron decay [5]. The scale of the upper left event is ≈ 10 m. The upper right-hand display shows tracks in the silicon vertex detector ($\varnothing \approx 20$ cm), while the lower event reconstruction shows the decays of B_s and D_s mesons with typical lengths of ≈ 200 μm .

A common subclass of displaced vertices are the extremely displaced vertices characterised by only two oppositely charged tracks, referred to as V^0 s, which is indicative of, for example, a $\Lambda \rightarrow p\pi^-$ decay; in this context, photon conversions to electron–positron pairs may also be thought of as V^0 s. A search for V^0 s within an event and subsequent calculation of its invariant mass with the charged-track pair is a form of particle identification.

The V^0 decay point, measured in a tracking chamber, is well separated from the primary interaction point. Their decay products are recorded with high precision and allow the reconstruction of the particle’s properties. Typical candidates are weakly decaying particles such as B , D , and V^0 (K_s^0 , Λ^0) mesons and baryons. Converting photons produce a similar pattern: a photon may convert to an e^+e^- pair in the wall of a tracking chamber, beam pipe, etc. The conversion probability in typical detectors is on the order of a few per cent. Neglecting the masses of the positron and electron and the recoil of the nucleus, the e^+e^- tracks are parallel. This can be seen from the reconstructed photon mass squared: $m_\gamma^2 = 2p_{e^+}p_{e^-}(1 - \cos\theta)$, where θ is the opening angle between electron and positron. Figure 15.6 shows a sketch of a photon conversion in

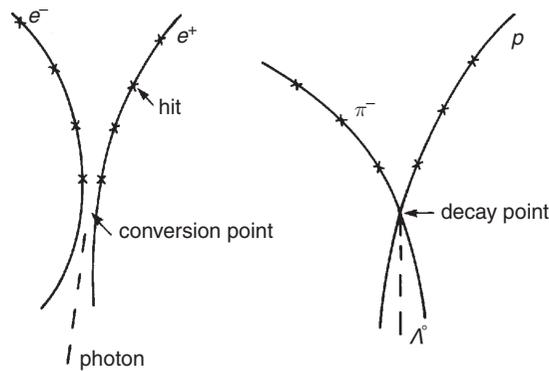


Fig. 15.6. Sketch of a photon conversion and a Λ^0 decay.

comparison to a Λ^0 decay. The two reconstructed tracks from a photon conversion can either intersect or may not have a common vertex because of measurement errors. The conversion point is found as the point where the two tracks are parallel ($m_\gamma^2 = 0$) in the plane orthogonal to the magnetic field. The photon momentum is the vector sum of the e^+ and e^- track at or closest to the conversion point.

For massive particles (e.g. Λ^0 with $m = 1.116 \text{ GeV}/c^2$) the opening angle is finite and the trajectories of proton and pion intersect. The closest approach of the two trajectories in space is a good approximation for the decay point. A more precise procedure, however, is to perform a *geometrical fit* using the parameters of the two tracks as obtained from the track fit (\vec{H}_i) and their error matrix (\mathcal{C}_i) and to perform a χ^2 fit. With two tracks, including, for example, ten measurements $\mathcal{H} = (\vec{H}_1, \vec{H}_2)$ (see Fig. 15.4) and nine parameters $\mathcal{Q} = (\vec{D}, \vec{p}_1, \vec{p}_2)$ to be determined (decay point \vec{D} , and two momenta \vec{p}_1, \vec{p}_2), one has a fit with one degree of freedom. The calculation [3, 4, 20] is similar to the straight line fit discussed above. The covariance matrix, however, is non-diagonal as the five track variables are correlated. It is a 10×10 matrix consisting of two submatrices of dimension 5×5 . A very important difference is that the expectation values of the nine parameters \mathcal{Q} are not linear functions of the measurements \mathcal{H} . Therefore one must obtain the parameters by Taylor series expansion and approximate \mathcal{X} , see Eq. (15.5), from the first derivative $\delta\mathcal{H}/\delta\mathcal{Q}$. This matrix is evaluated at an assumed starting value of \mathcal{Q}_0 , which is derived from an educated guess. Improved parameters \mathcal{Q}_1 are obtained using the least-squares method and the procedure is iterated.

With the Λ^0 mass known one can include the mass as a further constraint in the fit. In addition, the origin of the Λ^0 may be known; it is usually the primary vertex. Therefore a kinematical fit can use the fact

that the direction of flight of the A^0 coincides with the momentum sum of the decay products $\vec{p}_1 + \vec{p}_2$. This procedure allows one to obtain samples of V^0 with both high purity and high efficiency.

15.5 Analysis components

Particle physics data analysis consists of several distinct components, excluding the actual design, construction and operation of detector experiments. These include the Monte Carlo generation of events corresponding to physical processes of interest, the simulation of the detector response to these events, the reconstruction of raw and simulated data, the design and application of selection criteria frequently using multivariate techniques, and the statistical interpretation of results.

15.5.1 Monte Carlo event generators

The generation and study of the four-vectors of final-state particles associated with a physics process of interest is required for developing a particle physics analysis. A wide variety of packages exist which generate a list of particles and their four-vectors associated with well-known, putative or purely hypothetical particle physics processes. These packages build upon decades of theoretical and phenomenological research and constant revision based upon new experimental observations and measurements. At the heart of these packages is the numerical technique developed by Stanislaw Ulam and referred to as *Monte Carlo techniques* [21, 22]. An excellent overview of modern Monte Carlo techniques in particle physics is provided in [23]. A brief summary of some commonly used packages is given in Appendix 4; in each case, the packages have extensive development histories and are continuously subject to revisions and updates.

Beyond the accelerator domain – in the field of astroparticles – particle interactions have also been modelled to describe the propagation of energetic cosmic-ray particles with energies in excess of PeV through the Earth's atmosphere. The measured primary cosmic-ray spectrum extends to $\approx 10^{21}$ eV, corresponding to centre-of-mass energies around 1000 TeV, energies that will not be in reach of earthbound accelerators in the near future. In these models the hadronic interactions are described by a set of sub-interactions between the participating particles (mainly proton–air or heavy-nucleus–air interactions). These processes are dominated by soft interactions plus occasional semi-hard or even hard interactions, where only the latter can be described using perturbative QCD. These processes are modelled in terms of the formation of a set of colour strings. For the soft interactions semi-empirical phenomenological models must be used. The implementation of such approaches based on extrapolations

from accelerator data has been used for simulations of extensive air showers (QGSJET, SIBYLL, DPMJET, VENUS, NEXUS, FLUKA [24–26], see also Appendix 4). Commonly, such models have been integrated into simulation packages such as CORSIKA [27].

15.5.2 Simulation of detector response

The list of particle four-vectors provided by Monte Carlo event-generator packages form a phenomenological basis for a proposed analysis. They may also be used in the context of a *fast simulation* where rough parameterised detector responses are used to smear particle parameters. Finally, they may be fed into full simulations of the response of a specific detector. This last step involves the precise modelling of not only the nature and response of a detector experiment but also the passage of final-state particles through the matter the detector is composed of; this modelling is generally done with Monte Carlo techniques as well with packages such as GEANT [28] and FLUKA [26].

Both Monte Carlo event generation and detector simulation are computationally intensive. Hence, it is common for large detector experiments to employ hundreds or thousands of computers from member institutes to produce Monte Carlo samples. The size of these samples directly impacts the uncertainties associated with signal efficiencies and background-rejection rates in data analysis.

15.5.3 Beyond the detector

Limitations of detector apparatus may often be overcome or at least ameliorated through insight into the nature of the physical processes under study. Such insights are hard to quantify, but in this section some examples of such techniques are presented.

Mass constraint refits

When tracks or jets are known to be from a known particle, their four-vectors may be refit with a *mass constraint*. This technique is common when dealing with decay products of W^\pm or Z bosons.

At the Large Electron–Positron Collider (LEP) and, in particular, at the Large Hadron Collider and the Tevatron the event topologies can be quite complicated. In the search for the Higgs boson or supersymmetric particles one has to work out the invariant mass of some anticipated new object very carefully. The better the mass resolution the higher the probability to find a signal associated with low statistics. If it is known for some reason that in the final state a W or a Z has been produced, one can reconstruct these particles from their decay products. For leptonic W

decays this can be rather difficult, since one first has to reconstruct the energy and momentum of the missing neutrino. But also for Z decays into jets the association of low-energy particles to jets may not be unambiguous. Also the jet-energy determination using electromagnetic and hadronic calorimetry may not be very accurate. Therefore the reconstruction efficiency and, in particular, the overall mass resolution would benefit from a kinematic refit to the event assuming the known exact masses for those particles that are known to have been produced.

Extracting efficiency from data

Data can sometimes be used to extract the efficiencies. In many cases one would prefer to obtain the efficiencies and, in general, the characteristics of the various detector components from measurements in a test beam, where the particle type and its momentum are well known. In large experiments with hundreds of detector modules this is difficult to achieve. On top of that, the properties of detector modules may change during the experimental runs, depending on the ambient temperature, radiation levels, pressure and other parameters. Certainly, these parameters are monitored by slow control but, still, one wants to have an on-line calibration to analyse the data with the best set of calibration constants. This can be achieved by using known particles or particle decays. For example, at LEP running at the Z resonance, the decay of the Z into muon pairs presented a very interesting sample of penetrating tracks of known momentum and interaction properties. The *track efficiency* of the tracking device (e.g. a time-projection chamber) could easily be determined from the muon tracks. The efficiency of the muon chambers mounted behind the hadron calorimeter, which also serves as flux return for the magnetic field, could clearly be worked out, since the penetration through the iron for muons of about 46 GeV each (for a Z decay at rest) is guaranteed. Also, detailed properties of the time-projection chamber like magnetic-field inhomogeneities or edge effects due to possible problems of the field cage could be investigated and refitted.

In very much the same way the decay of neutral kaons into charged pions can be used to check on the *track-reconstruction efficiency* or, equivalently, the decay of neutral pions into two photons can be used to investigate the properties of electromagnetic calorimeters in the detector. A particularly clean sample of data is obtained if the photons convert into electron-positron pairs in the gas of a time-projection chamber where the density of the target is well known, and the electron and positron momenta are determined from the track curvature and their energies are then measured

in the electromagnetic calorimeter. Such a redundancy gives confidence in the reliability of the calibration parameter.

Finally, it should be mentioned that energetic cosmic-ray muons, which are also available when the accelerator is not running, can be used to check out the detector for efficiency and uniformity of response.

Reconstruction of missing particles

Information from undetected particles may often be recovered when considering a specific physical process. If, for example, an event has been exclusively reconstructed and if, say, a W has decayed leptonically, the total observed event energy does not match the centre-of-mass energy: some energy, i.e. that of the neutrino, is missing. From the knowledge of the centre-of-mass energy and the four-momenta of all detected particles, the energy and momentum of the missing particle can be inferred. This also allows to work out the mass of the parent particle. This *missing-energy* or *missing-momentum technique* works in a clean environment and can be applied to many circumstances, like also in the recovery of neutrino momentum in semileptonic decays of B hadrons or in the search for supersymmetric particles, where the lightest supersymmetric particle is supposed to be stable and normally will escape detection due to its low interaction cross section.

15.5.4 *Multivariate techniques*

Frequently, a particle physics analysis makes use of a variety of discriminating variables some of which may be partially correlated with each other. Unless clearly motivated by straightforward kinematics or a striking separation between signal and background, the choice of which ones to use and what selection criterion to place on each becomes difficult and often arbitrary.

Multivariate techniques allow selection criteria to be chosen by a prescribed method which frequently reduces many variables to a single discriminant. A wide variety of multivariate techniques exists and these have been used in particle physics analysis. We present a brief summary of some of the most commonly used techniques and then discuss two techniques in more detail below:

- Maximum Likelihood;
- Artificial Neural Networks;
- Genetic Programming [29];
- Genetic Algorithms;

- Support Vector Machines [30, 31] are one of the most innovative recent developments in multivariate data analysis;
- Kernel Probability Density Estimation [32];
- Linear Discriminant Analysis;
- Principal Component Analysis [33].

Maximum-likelihood techniques

In the technique of *maximum likelihood* a likelihood function is introduced that is supposed to characterise the data. The likelihood of a data sample is described by the probability to obtain such a sample under the assumption that the assumed probability distribution describes the data well. The chosen probability distribution normally has a set of parameters that can be adjusted. It is the aim of the maximum-likelihood technique to adjust these parameters in such a way that the likelihood of the sample takes on a maximum value. The actual values of the best fit parameters are called *maximum-likelihood estimates*.

Since this procedure starts with an assumed probability distribution, this method is based on an analytic expression describing the maximisation. It can be applied to any set of data where a smooth function is anticipated to be the best description of the experimental values.

Since the best model assumption is a priori not known, various likelihood functions can be used to test various hypotheses. Within these model assumptions one has the freedom of adjusting the free parameters. The maximum-likelihood estimates are frequently normal distributed so that approximate sample variances and confidence levels can be calculated.

As with many statistical methods the maximum-likelihood technique has to be treated with care for small event samples. The technical problems of required computer time for optimising the model distributions and adjustment parameters which presented problems in the early days of data analysis have been overcome with fast computers [34].

Neural networks

An elegant and effective way to deal with multivariate problems is the use of an *artificial neural network* (NN). NNs are inspired by, and are very crude approximations of, biological cortical neural systems. They can be trained to utilise information available from multiple variables. They take into account correlations between variables and learn to rely upon the given information alone when other variables are not available. Depending on the application, NNs can be trained to identify events of a given topology while reducing the number of background events. At the

same time NNs provide additional information like efficiency and purity of an event sample for a given final-state hypothesis.

A discussion of the general NN theory and principles is given elsewhere [17, 35]; here we present a brief overview of the most commonly used NNs in particle physics analyses: *simply connected feedforward back-propagation* NNs. A variety of packages exist for the development and training of NNs for use in physics analyses. These include JETNET [36], SNNS [37], MLPfit [38] and others.

An NN is composed of neurons or *nodes* arranged in *layers*. Two given nodes i and j , which are usually in adjacent layers, are connected to each other via *links* which are assigned a *weight* w_{ij} . Each node is the site of the evaluation of an *activation function* Y which is, dependent upon the values of the activation functions of any neurons it is connected to, multiplied by their weights. For node j , this is given as

$$Y_j = g \left[\left(\sum_i w_{ij} x_i \right) - \theta_j \right] , \quad (15.12)$$

where x_i is the value of the activation function of node i and θ_j is referred to as the neuron bias. The choice of the activation function is usually a sigmoid function such as $g(x) = \tanh x$ [30].

The first layer is referred to as the *input layer*; each node in acquired values based upon a discriminating variable is derived from final-state physics objects in data or Monte Carlo simulation; one input node is assigned to each discriminating variable. These discriminating variables must be scaled such that the bulk of the distribution lies in the range $[0, 1]$ (or $[-1, 1]$) for effective processing by the NN. There may be multiple *hidden layers*, the nodes of which are fully connected to all of the nodes in previous layers. The final layer is referred to as the *output layer* and provides the discriminating variable which may be used in the selection criteria of the analysis. The choice of overall NN architecture follows no formal rule; instead, a trial-and-error approach or architecture based on previous NN experience is usually chosen.

The NN architecture nomenclature is of the form X - Y_1 -...- Y_N - Z where X denotes the number of nodes in the input layer, Y_i denotes the number of nodes in the i th hidden layer, and Z denotes the number of nodes in the output layer. Figure 15.7 illustrates the architecture of a 6-10-10-1 NN.

Once the NN architecture is specified, the NN is trained, usually using events from Monte Carlo simulation. A *pattern* of input variables selected from a *training sample* is presented to the NN along with information related to the desired output for each pattern (e.g. an output value of 1 for signal patterns and 0 for background patterns). A *learning scheme* adjusts the weights of the links connecting neurons in order to minimise

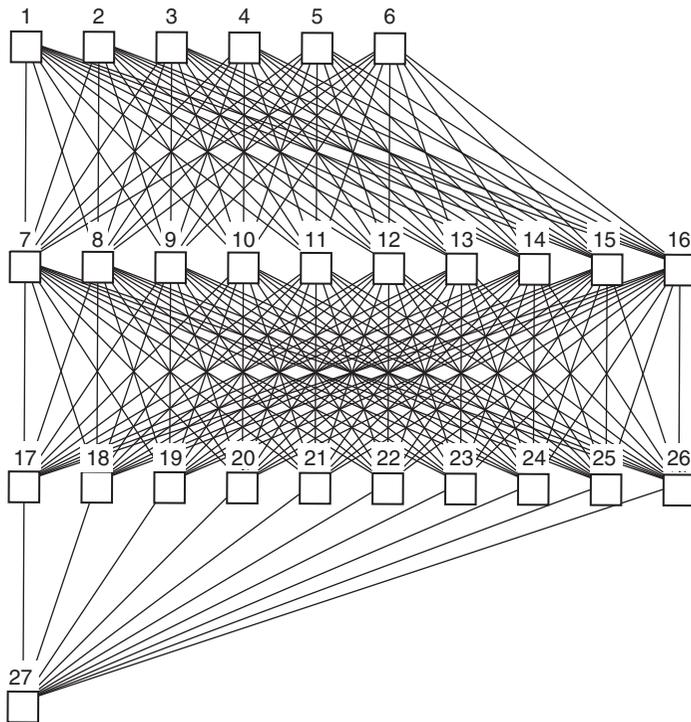


Fig. 15.7. A diagram illustrating the architecture of a 6-10-10-1 neural network. Neurons are depicted as squares while links are the lines joining the neurons. Neurons 1–6 are the six input nodes. Neurons 7–16 comprise the first 10-node hidden layer. Neurons 17–26 comprise the second 10-node hidden layer. Neuron 27 is the output node.

an overall metric denoting how well the NN produces the desired output. As mentioned above, the most commonly used learning scheme is the backpropagation scheme [16–18]. This process of presenting patterns and adjusting weights is repeated many times for a specified number of *training cycles*. Once NN training is finished, NN performance must be evaluated using an independent *testing sample* to avoid bias.

There are many common pitfalls associated with the NN training. These include the correlated issues of NN *overtraining*, testing sample bias and training sample size. Overtraining refers to the use of too many training cycles. In this case, the NN learns the fine-structure details specific to the training sample used rather than providing a more general character of the training sample. This will degrade the performance when evaluated on the testing sample. However, if the number of training cycles is varied and repeatedly evaluated on the testing sample, the testing sample itself becomes biased, and there is a need for another independent *validation*

sample. Furthermore, if a large number of input variables is used, a sufficiently large training sample must be used to populate fully the resultant high-dimension domain. If the training sample is too small, the NN will quickly become overtrained or fail to converge in the learning cycle.

15.6 Analysis in action

Perhaps the best way to illustrate some of the methods described above is to consider the case of a sophisticated analysis of collider data. Here the example of an analysis is chosen designed to search for evidence of the production of the *Standard Model Higgs Boson* in electron–positron collisions at centre-of-mass energies near 200 GeV recorded with the LEP detectors in the year 2000 [39].

The Higgs mechanism plays a central rôle for the unification of weak and electromagnetic interactions. Among others it generates the masses for the intermediate vector bosons W and Z . In the electroweak theory the symmetry is broken by the Higgs mechanism. Within this scheme the existence of a single neutral scalar particle, the Higgs boson, is required. The theory, however, gives no clue for the mass of this object.

Measurements at LEP energies below 200 GeV gave no evidence for the production of the Standard Model Higgs boson. In the last year of data taking at LEP large data samples were collected by the four LEP experiments at centre-of-mass energies beyond 200 GeV. The main production process for the Higgs at these energies is supposed to be *Higgsstrahlung* $e^+e^- \rightarrow HZ$. Small additional contributions are expected from W and Z boson fusion. The signal processes were simulated extensively by Monte Carlo techniques. For the energy regime of LEP the Higgs boson is expected to decay predominantly into a pair of b quarks, but decays into tau pairs, charm-quark pairs, gluons or W pairs (with one virtual W) are also possible. The main channel of investigation was the decay of the Higgs into b -quark pairs and the decay of the Z into two jets. In addition, the decay channels $H \rightarrow b\bar{b}$ and $Z \rightarrow \nu\bar{\nu}$ characterised by missing energy, and the channels where the Z decays to lepton pairs, or the Higgs to tau pairs, were investigated.

The search for the Higgs is plagued by background processes which could easily mimic the Higgs production. For example, the WW or ZZ production, which is kinematically possible at centre-of-mass energies exceeding 200 GeV, also produces four-jet final states. Two-photon processes and radiative returns to the Z might also produce signatures that look like the signal.

The main purpose of the analysis is to reduce the background without cutting too much into a possible signal. The identification of b quarks played an important rôle in reducing the background. Due to the high spatial resolution of the vertex detectors displaced vertices from heavy-quark decays could be used for this discrimination. Knowing the signature of the Higgs from Monte Carlo studies, very selective cuts could be applied. For example, the masses of the W and Z could be very close to a possible Higgs mass. Therefore, the reconstruction of their masses was essential to separate a possible signal from more mundane processes. In addition to the *classical cut stream analysis*, also multivariate techniques such as likelihood analysis and, in particular, *neural networks* were extensively used.

It is very important that the different strategies to search for the Higgs using cuts and neural networks have to be established and frozen – after estimating background and optimising selection criteria using Monte Carlo events – *before* the analysis of real events. Modifying the cuts or training the neural networks after the data have been taken and the analysis has been started might introduce a psychological bias, because if one wants to find something, one might be tempted to introduce unconsciously cuts tailored for a signal. Also a blind analysis and a study of real events not in the signal region would help to establish the confidence in a possible discovery.

After extensive and independent analyses of different groups within the same collaboration, ALEPH observed three candidate events consistent with the production of a Standard Model Higgs boson of a mass at around 115 GeV, while OPAL and L3 could explain their candidates with the assumption of background, even though a signal plus background hypothesis was slightly favoured, in contrast to DELPHI which recorded less events compared to the background expectation. A Higgs candidate from ALEPH is shown in Fig. 15.8.

The overall ALEPH evidence for a Higgs at around 115 GeV can be taken from Fig. 15.9, where the results of the *neural-net analysis* and that of a *cut analysis* are compared. The small excess over background at invariant masses of about 115 GeV could be an indication of the Higgs production, even though the evidence from these diagrams is not too convincing.

A discovery of the Higgs at a mass of around 115 GeV by the Tevatron or the Large Hadron Collider at CERN will for sure make the ALEPH collaboration very happy. However, taking the evidence of the four LEP experiments together, a claim for a discovery certainly cannot be made. Instead, combining the results of the four experiments, only a lower bound on the Higgs mass of 114.4 GeV can be set at 95% confidence level [39].

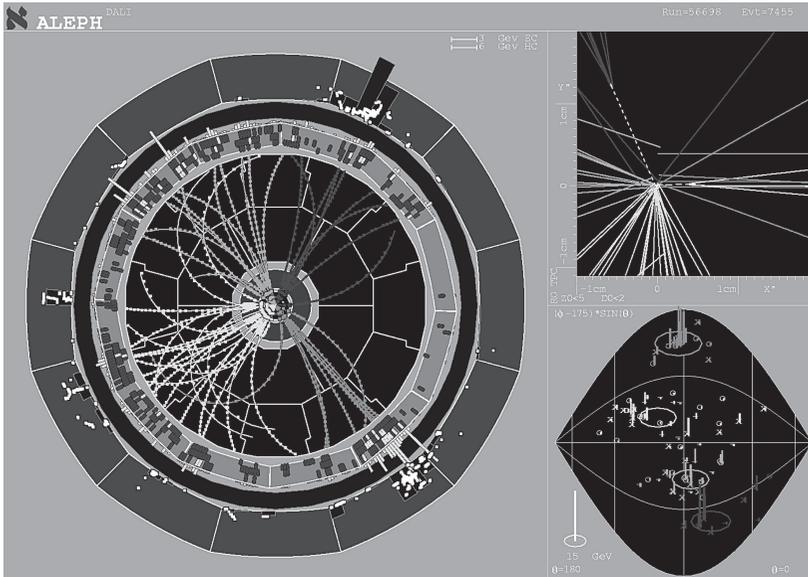


Fig. 15.8. A Higgs candidate assumed to be produced by Higgsstrahlung as observed in the ALEPH experiment at a centre-of-mass energy of 206.7 GeV. Both the Higgs and the Z appear to decay into pairs of b quarks [40].

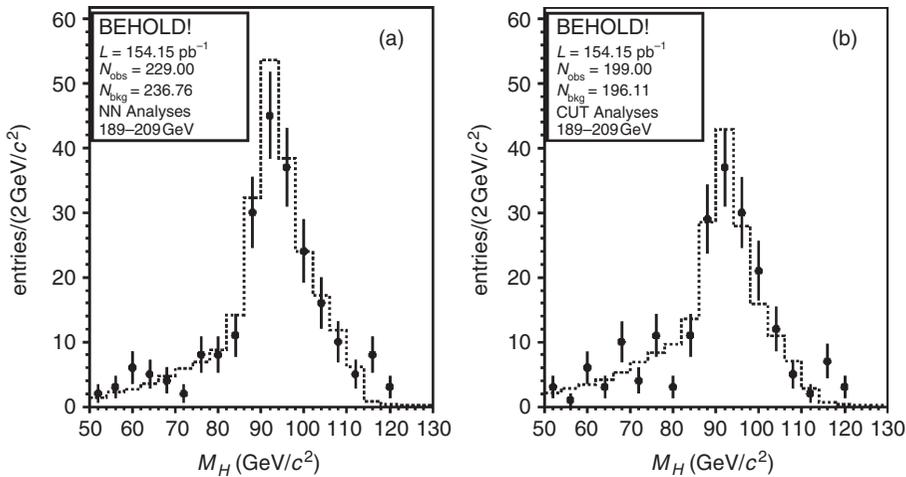


Fig. 15.9. Results of (a) the neural-network analysis and (b) the cut stream analysis in the search for the Higgs boson in ALEPH [40]. Plotted is the distribution of the invariant mass of two jets in a four-jet final state of e^+e^- interactions, see Eqs. (15.10) and (15.11). The dominant background is represented by ZZ , W^+W^- and QCD events (dotted histogram). Reconstructed Z s are the most abundant type of background. The excess events at larger masses ('Higgs candidates') originate from $b\bar{b}$ jets, which are expected to be the dominant decay mode of a Higgs particle in that mass range [41].

15.7 Problems

- 15.1** In a mixed electron–pion beam a set of electron candidates is selected by requiring a cut (e.g. on the Cherenkov yield) leading to N_{acc} accepted events out of a total of N_{tot} particles. Work out the total number of electron events if no cut were made, given that the efficiencies for electrons and pions to pass the cut were ε_e and ε_π . What would happen if $\varepsilon_e = \varepsilon_\pi$?
- 15.2** The exponential probability density of the variable t ($0 \leq t < \infty$) is given by

$$f(t, \tau) = \frac{1}{\tau} e^{-t/\tau},$$

which is characterised by the mean lifetime τ . Work out the expectation value and the variance of the exponential distribution.

- 15.3** Let the number of events in an experiment with neutrinos from a reactor be N_1 for a measurement time t_1 . This number includes background from cosmic rays of rate n_μ . With reactor off a number N_2 is obtained during a time t_2 . How can the measurement times t_1 and t_2 be optimised such that the error on the signal rate is minimal if only a total time $T = t_1 + t_2$ is available and the expected signal-to-background ratio is 3?
- 15.4** An electromagnetic calorimeter is calibrated with electrons of known energy. The following responses are obtained (in arbitrary units)

Energy [GeV]	0	1	2	3	4	5
Response	0.2	1.0	1.8	2.7	3.0	4.2

The experimental values have to be corrected for the common offset which is assumed to be 0.2. Determine the slope of the calibration and its error assuming a linear dependence through the origin. The standard deviation of all response measurements is $\sigma = 0.3$.

References

- [1] K. Inami *et al.*, First Observation of the Decay $\tau^- \rightarrow \phi K^- \nu_\tau$, *Phys. Lett.* **B643** (2006) 5–10
- [2] ATLAS Collaboration: <http://atlas.web.cern.ch/atlas/>
- [3] S. Brandt, *Datenanalyse*, 4. Auflage; Spektrum Akademischer Verlag, Heidelberg/Berlin (1999); *Data Analysis: Statistical and Computational*

- Methods for Scientists and Engineers*, 3rd edition, Springer, New York (1998)
- [4] W.T. Eadie, D. Drijard, F. James, M. Roos & B. Sadoulet, *Statistical Methods in Experimental Physics*, Elsevier–North Holland, Amsterdam/London (1971)
- [5] ALEPH Collaboration, <http://aleph.web.cern.ch/aleph/>
- [6] G. Hanson *et al.*, Evidence for Jet Structure in Hadron Production by e^+e^- Annihilation, *Phys. Rev. Lett.* **35** (1975) 1609–12
- [7] X. Artru & G. Mennessier, String Model and Multiproduction, *Nucl. Phys.* **B70** (1974) 93–115; A. Casher, H. Neuberger & S. Nussinov, Chromoelectric-Flux-Tube Model of Particle Production, *Phys. Rev.* **D20** (1979) 179–88; B. Anderson, G. Gustavson, G. Ingelman & T. Sjöstrand, Parton Fragmentation and String Dynamics, *Phys. Rep.* **97** (1983) 31–145
- [8] T. Sjöstrand, High-Energy-Physics Event Generation with PYTHIA 5.7 and JETSET 7.4, *Comp. Phys. Comm.* **82** (1994) 74–89
- [9] B.R. Webber, A QCD Model for Jet Fragmentation Including Soft Gluon Interference, *Nucl. Phys.* **B238** (1984) 492–528; G.C. Fox & S. Wolfram, A Model for Parton Showers in QCD, *Nucl. Phys.* **B168** (1980) 285–95; R.D. Field & S. Wolfram, A QCD Model for e^+e^- Annihilation, *Nucl. Phys.* **B213** (1983) 65–84; T.D. Gottschalk, A Simple Phenomenological Model for Hadron Production from Low-Mass Clusters, *Nucl. Phys.* **B239** (1984) 325–48
- [10] G. Marchesini *et al.*, HERWIG 5.1 – A Monte Carlo Event Generator for Simulating Hadron Emission Reactions with Interfering Gluons, *Comp. Phys. Comm.* **67** (1992) 465–508
- [11] R. Barate *et al.* (ALEPH Collaboration), A Measurement of the QCD Colour Factors and a Limit on the Light Gluino, *Z. Phys.* **C76** (1997) 1–14
- [12] D. Buskulic *et al.* (ALEPH Collaboration), Quark and Gluon Jet Properties in Symmetric Three-Jet Events, *Phys. Lett.* **B384** (1996) 353–64; R. Barate *et al.* (ALEPH Collaboration), The Topology Dependence of Charged Particle Multiplicities in Three-Jet Events, *Z. Phys.* **C76** (1997) 191–9
- [13] W. Bartel *et al.*, Experimental Study of Jets in Electron–Positron Annihilation, *Phys. Lett.* **B101** (1981) 129–34
- [14] S. Bethke, Z. Kunszt, D.E. Soper & W.J. Stirling, New Jet Cluster Algorithms: Next-to-leading Order QCD and Hadronization Corrections, *Nucl. Phys.* **B370** (1992) 310–34
- [15] S. Catani *et al.*, New Clustering Algorithm for Multijet Cross Sections in e^+e^- Annihilation, *Phys. Lett.* **B269** (1991) 432–8; W.J. Stirling, Hard QCD Working Group – Theory Summary, *J. Phys. G: Nucl. Part. Phys.* **17** (1991) 1567–74
- [16] J. Hertz, A. Krogh & R.G. Palmer, *Introduction to the Theory of Neural Computation*, Santa Fe Institute, Addison-Wesley, Redwood City CA (1991)

- [17] R. Rojas, *Theorie der neuronalen Netze*, Springer, Berlin (1993); *Neuronal Networks. A Systematic Introduction*, Springer, New York (1996); *Theorie neuronaler Netze: eine systematische Einführung*, Springer, Berlin (1996)
- [18] D. McAuley, *The BackPropagation Network: Learning by Example* (1997) on-line: www2.psy.uq.edu.au/~brainwav/Manual/BackProp.html
- [19] H.F. Teykal, *Elektron- und Pionidentifikation in einem kombinierten Uran-TMP- und Eisen-Szintillator-Kalorimeter*, RWTH Aachen, PITHA 92/28 (1992)
- [20] B. Rensch, *Produktion der neutralen seltsamen Teilchen K_s und Λ^0 in hadronischen Z-Zerfällen am LEP Speicherring*, Univ. Heidelberg, HD-IHEP 92-09 (1992)
- [21] R. Eckart, *Stan Ulam, John von Neumann, and the Monte Carlo Method*, Los Alamos Science, Special Issue (15) (1987) 131–7
- [22] M. Metropolis & S. Ulam, The Monte Carlo Method, *J. Am. Stat. Ass.* **44** (1949) 335–41
- [23] S. Jadach, *Practical Guide to Monte Carlo*, hep-th/9906207
- [24] J.L. Pinfoldi, Links between Astroparticle Physics and the LHC, *J. Phys. G: Nucl. Part. Phys.* **31** (2005) R1–R74
- [25] H.J. Drescher, M. Bleicher, S. Soff & H. Stöcker, Model Dependence of Lateral Distribution Functions of High Energy Cosmic Ray Air Showers, *Astroparticle Physics* **21** (2004) 87–94
- [26] www.fluka.org/ (2005)
- [27] D. Heck *et al.*, Forschungszentrum Karlsruhe, Report FZKA 6019 (1998); D. Heck *et al.*, Comparison of Hadronic Interaction Models at Auger Energies, *Nucl. Phys. B Proc. Suppl.* **122** (2002) 364–7
- [28] R. Brun, F. Bruyant, M. Maire, A.C. McPherson & P. Zancarini, GEANT3 CERN-DD/EE/84-1 (1987); wwwasdoc.web.cern.ch/wwwasdoc/geant_html3/geantall.html
- [29] K. Cranmer & R. Sean Bowman, *PhysicsGP: A Genetic Programming Approach to Event Selection*, physics/0402030
- [30] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York (1995)
- [31] P. Vannerem *et al.*, *Classifying LEP Data with Support Vector Algorithms*, Proceedings of AIHENP99, Crete, April 1999, hep-ex/9905027
- [32] L. Holmström, S.R. Sain & H.E. Miettinen, A New Multivariate Technique for Top Quark Search, *Comp. Phys. Comm.* **88** (1995) 195–210
- [33] H. Wind, Principal Component Analysis and its Applications to Track Finding, in *Formulae and Methods in Experimental Data Evaluation*, Vol. 3, European Physical Society, CERN/Geneva (1984) pp. d1–k16
- [34] Engineering Statistics, NIST/SEMATECH e-Handbook of Statistical Methods, www.itl.nist.gov/div898/handbook/ (2005)
- [35] C. Peterson & T. Rönngvaldsson, An Introduction to Artificial Neural Networks, in C. Verkerk (ed.), *14th CERN School of Computing – CSC '91*, LU TP 91-23, CERN-92-02 (1991) 113–70; L. Lönnblad, C. Peterson & T. Rönngvaldsson, Pattern Recognition in High Energy Physics with Artificial Neural Networks – JETNET 2.0, *Comp. Phys. Comm.* **70** (1992)

- 167–82; S.R. Amendolia, Neural Networks, in C.E. Vandoni & C. Verkerk (eds.), *1993 CERN School of Computing* (1993) 1–35
- [36] L. Lönnblad, C. Peterson & T. Rönngvaldsson, *JETNET 3.0: A Versatile Artificial Neural Network Package*, CERN-TH-7135-94 (1994); *Comp. Phys. Comm.* **81** (1994) 185–220
- [37] A. Zell *et al.*, *SNNS – Stuttgart Neural Network Simulator User Manual, Version 4.0*, University of Stuttgart, Institute for Parallel and Distributed High Performance Computing (IPVR), Department of Computer Science, Report 6/95 (1995)
- [38] J. Schwindling & B. Mansoulié, *MLPfit: A Tool for Designing and Using Multi-Layer Perceptrons*, on-line: <http://schwind.home.cern.ch/schwind/MLPfit.html>
- [39] G. Abbiendi, The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, and the OPAL Collaboration, Search for the Standard Model Higgs Boson at LEP, *Phys. Lett.* **B565** (2003) 61–75
- [40] ALEPH Collaboration, <http://aleph.web.cern.ch/aleph/alpub/seminar/wds/Welcome.html>
- [41] R. Barate *et al.* (ALEPH Collaboration), Observation of an Excess in the Search for the Standard Model Higgs Boson at ALEPH, *Phys. Lett.* **B495** (2000) 1–17