PUB 94-165

T

Measurement of R at the Z_0 peak using Neural Networks [1]

Francisco Ariztizábal
Universitat Autònoma de Barcelona
Institut de Física d'Altes Energies
Edifici C E-08193 Bellaterra (Barcelona) Spain

Sep 1994

[1]Memòria presentada per a optar al grau de Doctor en Ciències, Secció de Físiques, per la Universitat Autònoma de Barcelona. Septembre 1994.

Abstract

In this thesis a determination of the quotient of the hadronic and leptonic widths of the Z_0 is presented, through the measurement of the ratio of cross sections $\sigma(e^+e^- \rightarrow hadrons)/\sigma(e^+e^- \rightarrow leptons)$ where leptons are muons or taus. The classification of the different Z_0 decay modes is based on the Neural Network technique. An algorithm of "learn and grow" has been developed to train the Neural Network. The data used were collected by the ALEPH detector at CERN's LEP accelerator during the 1991 running period.

Resum

En aquesta tesi, presentem una determinació del quocient de les amplades hadròniques i leptòniques del bosó Z_0, mesurant el quocient de seccions eficaces $\sigma(e^+e^- \rightarrow hadrons)/\sigma(e^+e^- \rightarrow leptons)$ on leptons són muons i taus. La classificació dels diversos modes de desintegració del bosó Z_0 ha estat basada en la tècnica de xarxes neuronals. Per entrenar la xarxa, un algorisme de "creix i aprèn" ha estat desenvolupat. Les dades han estat recollides pel detector ALEPH situat a l'accelerador LEP al CERN durant l'any 1991.

Acknowledgments

This work was done under the economic support of the "Tito Ariztizábal" foundation.

I specially want to thank my advisor, Dr. Manuel Delfino who has given me his friendship and cleverness to solve problems.

I would like to thank the staff of the Institut de Fisica d'Altes Energies, specially Dr. Vicenc Gaitan who introduce me to Neural Nets world. I want to thank Lic. Frederic Teubert for many suggestions on the theoretical layout of this work. I've benefited of Dr. Salvador Orteu's permanent support in solving software problems and discussions in the experimental part of this work. Dr. Manel Martinez was very kind in helping the in fitting part of the work. The help of Dr. Pere Comas in the LaTex design of this thesis is also acknowledged. Finally , I have enjoyed the atmosphere of friendship and companionship of the Institut.

A Marcela y Lucía

Laberinto

No habrá nunca una puerta. Estás adentro
Y el alcázar abarca el universo
Y no tiene ni anverso ni reverso
Ni externo muro ni secreto centro.
No esperes que el rigor de tu camino
Que tercamente se bifurca en otro,
Que tercamente se bifurca en otro,
Tendrá fin. Es de hierro tu destino
Como tu juez. No aguardes la embestida
Del toro que es un hombre y cuya extraña
Forma plural da horror a la maraña
De interminable piedra entretejida.
No existe. Nada esperes. Ni siquiera
En el negro crepúsculo la fiera.

Jorge Luis Borges, 1969

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The quantity R is defined as the ratio of the hadronic to the leptonic partial widths of the Z_0,

$$R \equiv \frac{\Gamma_h}{\Gamma_l} \tag{1.1}$$

or in others words the quotient of the probabilities of the Z_0 decaying to hadrons versus a lepton-antilepton pair. The decay to hadrons is the sum of all accessible quark-antiquark final states.

The interest in measuring R lies in the fact that it is sensitive to the strong coupling constant α_s. Considering QCD radiative corrections to order $O(\alpha_s^3)$ in the $\overline{MS}$ scheme [8],

$$R = R_0[1 + 1.060(\alpha_s/\pi) + (0.9 \pm 0.1)(\alpha_s/\pi)^2 - 15(\alpha_s/\pi)^3] \tag{1.2}$$

where $R_0 = 19.943 \pm 0.03$ is the value of R when no final state strong interactions are considered. The principal advantage of determining α_s from R in e^+e^- annihilation is that there is little dependence on fragmentation models or jet algorithms, which in turn means a considerable reduction in theoretical uncertainties, while from an experimental point of view a reduction in systematic errors can be achieved. The main experimental disadvantage of this quantity is that the error is expected to be dominated by the statistical uncertainty in measuring the relatively small leptonic width. The statistical error in α_s is given by $\Delta\alpha_s \approx \pi\Delta R/R$.

There are methods to experimentally determine α_s from the event topology. The various measurements of event shapes at LEP can be combined to give a value of

$\alpha_s(M_Z)$=0.115±0.008 [9]. This is an unweighted average of the LEP results. The error is dominantly systematic, arising from hadronization and scale uncertainties.

In e^+e^- collisions, R is not a directly measurable quantity. There is, however, a closely related observable, namely the ratio of the number of hadronic and leptonic events observed at center-of-mass energies near the Z_0 pole,

$$Q_l = \frac{\sigma(e^+e^- \to hadron)}{\sigma(e^+e^- \to l^+l^-)}, \tag{1.3}$$

where l stands for muons or taus. This quantity is directly related to the ratio of the hadronic width to the leptonic width and is very sensitive to its value. In order to get the value of α_s with a precision of 10 % , one needs to measure this ratio with a precision of about .5 % .

Two other methods have been proposed to measure α_s from the hadronic width at LEP. One can extract the hadronic width from the value of the total width of the Z_0, obtained from a fit to the line shape. Assuming that one can measure the total width with an error $\delta\Gamma_Z = 50MeV$, the corresponding uncertainty in the hadronic width is

$$\frac{\delta\Gamma_h}{\Gamma_h} = \frac{50}{1789} = 2.8\%, \tag{1.4}$$

and since the size of the radiative QCD corrections to Γ_h is small, on the order of $\frac{\alpha_s}{\pi} = 0.035$, the corresponding uncertainty in the determination of α_s is large, approximately 70%.

A second method is to extract the hadronic width from a direct measurement of the hadronic cross section. In the peaking approximation we obtain

$$\int \sigma(e^+e^- \to hadrons)dE = \frac{6\pi^2}{M_Z^2}\frac{\Gamma_h\Gamma_{e^+e^-}}{\Gamma_Z}. \tag{1.5}$$

In the right hand side of eq. 1.5 the hadronic width appears both in the numerator and in the denominator, as the total width is simply the sum of all partial widths. Therefore there is a reduced sensitivity to the strong corrections. In leading order in α_s one has

$$\frac{\Gamma_h}{\Gamma_Z} = \frac{\Gamma_h^0}{\Gamma_Z}(1 + \frac{\alpha_s}{\pi}\frac{\Gamma_l}{\Gamma_Z}) \tag{1.6}$$

This method is well applicable only if the absolute luminosity is very well measured, so as to compensate for the above mentioned effect.

It can be concluded that the technique of obtaining α_s from R is the most sensitive, and therefore this work will concentrate on a high precision measurement of the related observable, namely Q_l.

In Chapter 2, a theoretical expression for Q_l as a function R and a set of four other parameters is developed within the framework of the Standard Model. This relation demonstrates that Q_l is highly sensitive to R but rather insensitive to the other parameters.

Chapter 3 is devoted to the description of the experimental apparatus, the ALEPH detector, describing in detail the subdetectors most relevant to the analysis.

The original contribution is a new method to classify the different decays of the Z_0, using the Neural Network technique, which is described in chapter 4. A new algorithm of "learn-and-grow" to train a Neural Network has been developed and is fully described in appendix A.

Finally, the results and main conclusions of this work are given in chapter 5.

Chapter 2

Theory

As introduced in the previous chapter, the quotient R of hadronic width and leptonic width of the Z_0, is not a directly measurable quantity in e^+e^- collisions. Instead, it is possible to measure quantities which are very sensitive to R, that is the quotients of the hadronic cross section and the leptonic cross sections at the Z_0 peak

$$Q_\mu = \frac{\sigma(e^+e^- \to hadrons)}{\sigma(e^+e^- \to \mu^+\mu^-)}, \tag{2.1}$$

$$Q_\tau = \frac{\sigma(e^+e^- \to hadrons)}{\sigma(e^+e^- \to \tau^+\tau^-)}, \tag{2.2}$$

where leptons are muon and tau particles. It is then necessary to compute the cross section for hadronic and leptonic events at the Z_0 peak in terms of the partial widths already mentioned. It is also necessary to include all radiative corrections, as they are larger than the experimental uncertainties. It is possible to compute these cross sections in two different ways. One of them is to adopt the Standard Model (SM) while the other is a model independent parametrization, which will be the form finally used. Nevertheless, it is instructive to follow the Standard Model formulation, and then derive from it in a natural way the model independent parametrization.

2.1 The Standard Model

The Standard Model [1] is based on the invariance of the lagrangian under transformation of the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which is the direct product

of Quantum Chromodynamics (QCD) and the Glashow-Weinberg-Salam (GWS) electroweak theory.

The GWS model is a non abelian gauge theory with spontaneous symmetry breaking, based on the gauge group $SU(2)_L \otimes U(1)_Y$ of weak isospin and hypercharge that unifies electromagnetic and weak interactions. Some of the vector gauge bosons resulting from the local gauge invariance acquire a mass through the Higgs mechanism, preserving the gauge invariance needed for renormalizability.

QCD is also a non abelian gauge theory based, in the color group $SU(3)_C$. The eight gauge bosons (gluons) obtained as a result of the local invariance are responsible for the strong interactions among quarks.

The constituents of the minimal Standard Model are:

- the fermionic fields of matter grouped into three families of leptons and quarks[1] and placed into representations of the gauge group in the following way:
  - Leptons
    - isospin doublets and color singlets

$$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}_{Y=-\frac{1}{2}} \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}_{Y=-\frac{1}{2}} \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}_{Y=-\frac{1}{2}}$$

    - isospin singlets and color singlets

$$\begin{pmatrix} e_R \end{pmatrix}_{Y=-1} \begin{pmatrix} \mu_R \end{pmatrix}_{Y=-1} \begin{pmatrix} \tau_R \end{pmatrix}_{Y=-1}$$

  - Quarks[2]
    - isospin doublets and color triplets

$$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}_{Y=\frac{1}{6}} \begin{pmatrix} c_L \\ s'_L \end{pmatrix}_{Y=\frac{1}{6}} \begin{pmatrix} t_L \\ b'_L \end{pmatrix}_{Y=\frac{1}{6}}$$

    - isospin singlets and color triplets

$$\begin{matrix} \begin{pmatrix} u_R \end{pmatrix}_{Y=\frac{2}{3}} & \begin{pmatrix} c_R \end{pmatrix}_{Y=\frac{2}{3}} & \begin{pmatrix} t_R \end{pmatrix}_{Y=\frac{2}{3}} \\ \begin{pmatrix} d'_R \end{pmatrix}_{Y=-\frac{1}{3}} & \begin{pmatrix} s'_R \end{pmatrix}_{Y=-\frac{1}{3}} & \begin{pmatrix} b'_R \end{pmatrix}_{Y=-\frac{1}{3}} \end{matrix}$$

[1] The number of families is not constrained by the theory. However the recent measurements given by the LEP experiments restricts the number of families with light neutrinos to three

where $\{d', s', b'\}$ are linear combinations of the $\{d, s, b\}$ quarks, the mass eigenstates. The unitary matrix which relates the isospin and mass eigenstates is the Cabibbo-Kobayashi-Maskawa matrix, which depends on three angles and a phase that have to be measured experimentally. In the leptonic sector these angles are not observed in the limit of massless neutrinos.

- Gauge boson fields, mediators of the interactions:
  - the photon γ responsible for the electromagnetic interaction.
  - the bosons $W^{\pm}, Z$ responsible for the charged and neutral weak interaction respectively.
  - and the eight gluons g_{ab} responsible for the strong interaction.
- Higgs fields[2], producing spontaneous symmetry breaking.

2.1.1 Interactions between gauge bosons and fermions

The gauge fields associated to the generators of the electroweak gauge group $SU(2)_L \otimes U(1)_Y$ are a triplet of vector gauge fields $\vec{W}_\mu$ for $SU(2)_L$ and a scalar gauge field B_μ for $U(1)_Y$ with coupling constants g and g' respectively. The hypercharge Y is related to the electromagnetic charge Q by the expression,

$$Q = I_3 + Y$$

where I_3 is the third isospin component.

After spontaneous symmetry breaking of the electroweak lagrangian, the physical content of the mass terms for the vector bosons becomes transparent by performing the transformation from $\vec{W}_\mu, B_\mu$ to the "physical" fields,

$$W_\mu^{\pm} = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2)$$

and,

$$Z_\mu = \cos\theta_W W_\mu^3 + \sin\theta_W B_\mu$$

$$A_\mu = -\sin\theta_W W_\mu^3 + \cos\theta_W B_\mu$$

[2]The top quark (t) and the Higgs fields are predicted only.

where θ_W is known as the weak mixing angle defined as the ratio of the g and g' weak coupling constants,

$$\tan\theta_W = \frac{g'}{g}$$

or, equivalently,

$$\cos\theta_W = \frac{M_W}{M_Z} \tag{2.3}$$

where M_W and M_Z are the masses of the gauge fields $W_\mu^{\pm}$ and Z_μ respectively.

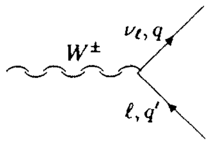
Identifying the massless field A_μ with the photon which couples to the electron via the electric charge e, the following relationship holds:

$$e = \frac{gg'}{\sqrt{g'^2 + g^2}}$$

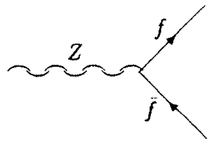
or, analogously

$$e = g\sin\theta_W \qquad e = g'\cos\theta_W.$$

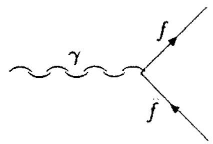
From the Standard Model lagrangian describing the coupling of the fermion fields to vector bosons, we get the Feynman rules for the gauge field-fermion interactions,



$$= ie\gamma_\mu(1-\gamma_5)\frac{1}{2\sqrt{2}\sin\theta_W}$$



$$= ie\gamma_\mu(v_f - a_f\gamma_5)$$



$$= -ieQ_f\gamma_\mu$$

where the neutral current coupling constants are given by,

$$
\begin{aligned}
v_f &= \frac{I_3^f - 2Q_f \sin^2\theta_W}{2\sin\theta_W\cos\theta_W} \equiv \frac{1}{2\sin\theta_W\cos\theta_W}\, g_{v_f} \\
a_f &= \frac{I_3^f}{2\sin\theta_W\cos\theta_W} \equiv \frac{1}{2\sin\theta_W\cos\theta_W}\, g_{a_f}
\end{aligned}
\tag{2.4}
$$

where I_3^f and Q_f denote the third isospin component and the electric charge of a given fermion species f.

2.2 Total cross section with non-photonic corrections

For the total cross section [40] we are interested in the following channels

$$
\begin{aligned}
e^+e^- &\to \mu^+\mu^- \\
e^+e^- &\to \tau^+\tau^- \\
e^+e^- &\to hadrons
\end{aligned}
$$

The diagrams contributing at tree level to the process $e^+e^- \to f\bar{f}(f \neq e)$ are those shown in figure 2.1.

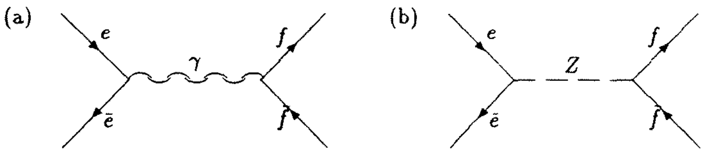


Figure 2.1: Feynman diagrams contributing at tree level to the process $e^+e^- \to f\bar{f}(f \neq e)$.

For the calculations we use the on-shell (OS) renormalization scheme. The lepton masses are known experimentally. The quark masses represent a suitable parametrization of the dispersion relation results for the hadronic vacuum polarization. We also require M_W, which is obtained from the muon decay constant G_μ and

the quantity Δr defined through the relation for effective fine structure constant at scale M_Z.

$$
\alpha(M_Z) = \frac{\alpha}{1-\Delta r}
\tag{2.5}
$$

where Δr includes the radiative corrections due to a change of the scale from low energy to the Z_0 mass. The results will depend on the unknown parameters of the Standard Model: M_Z, the mass of the Z_0 boson; m_t, the mass of the top quark; M_H, the mass of the Higgs boson; and α_s, the QCD coupling constant.

2.2.1 Lowest order widths and cross sections

The relation between M_W and G_μ, when electroweak radiative corrections in muon decay are considered, reads

$$
M_W^2 sin^2\theta_W = \frac{\pi\alpha}{\sqrt{2}G_\mu}\frac{1}{1-\Delta r}
\tag{2.6}
$$

Relation (2.6) is used to determine M_W or equivalently (in OS scheme)

$$
s_W^2 \equiv sin^2\theta_W = 1 - \frac{M_W^2}{M_Z^2}.
\tag{2.7}
$$

In lowest order, the partial width of a Z_0 decaying into a fermion pair is given by [40]

(2.8)

$$
\Gamma^{(0)}_{Z\to f\bar{f}} = \frac{\alpha}{6} N_c M_Z \sqrt{1-\frac{4m_f^2}{M_Z^2}}\left((g_f^-)^2 + (g_f^+)^2 + \frac{m_f^2}{M_Z^2}\left(6g_f^- g_f^+ - (g_f^-)^2 - (g_f^+)^2\right)\right).
$$

with,

$$
\begin{aligned}
g_f^- &= (I_f^3 - Q_f sin^2\theta_W)/(sin\theta_W cos\theta_W) \\
g_f^+ &= -Q_f sin^2\theta_W/(sin\theta_W cos\theta_W)
\end{aligned}
\tag{2.9}
$$

N_c representing the number of colors and m_f the mass of the fermion.

Making use of the relation (2.6) and eq. (2.7), the lowest order width can also be written in terms of G_μ

$$
\begin{aligned}
\Gamma^{(0)}_{Z\to ff} = N_c \frac{G_\mu M_Z^3}{24\pi\sqrt{2}}\sqrt{1-\frac{4m_f^2}{M_Z^2}} \\
\left(1-\frac{4m_f^2}{M_Z^2} + \left(2I_f^3 - 4Q_f s_w^2\right)^2\left(1+\frac{2m_f^2}{M_Z^2}\right)\right)
\end{aligned}
\tag{2.10}
$$

For a fixed M_Z, m_t and M_H one can determine Δr and consequently from eqs. (2.6) and (2.7) M_W and s_w^2. The total width is given by

$$\Gamma_Z^{(0)} = \sum_f \Gamma_{Z\to f\bar{f}}^{(0)} \tag{2.11}$$

It should be noted that in the massless fermion case the partial width (2.8) reduces to:

$$\Gamma_{Z\to f\bar{f}}^{(0)} = \Gamma_- + \Gamma_+ = \frac{\alpha}{6} N_c M_Z \left(|g_f^-|^2 + |g_f^+|^2\right), \tag{2.12}$$

where the $-$, $+$ sign refers to the helicity of f, with $\bar{f}$ having the opposite. The total cross section in lowest order for $e^+e^- \to f\bar{f}$ with massless fermions reads

$$\begin{aligned}\sigma_0(s) = \frac{4\pi\alpha^2}{3s}\frac{N_c s^2}{4} & \left\{\left|\frac{g_e^- g_f^-}{s-M_Z^2+iM_Z\Gamma_Z} + \frac{Q_eQ_f}{s}\right|^2\right. \\ & + \left|\frac{g_e^- g_f^+}{s-M_Z^2+iM_Z\Gamma_Z} + \frac{Q_eQ_f}{s}\right|^2 \\ & + \left|\frac{g_e^+ g_f^-}{s-M_Z^2+iM_Z\Gamma_Z} + \frac{Q_eQ_f}{s}\right|^2 \\ & \left. + \left|\frac{g_e^+ g_f^+}{s-M_Z^2+iM_Z\Gamma_Z} + \frac{Q_eQ_f}{s}\right|^2\right\}\end{aligned} \tag{2.13}$$

The four terms in the cross section correspond to the following helicity combinations for e^+, e^-, $\bar{f}$ and f: (+ - + -), (+ - - +), (- + + -), and (- + - +).

Taking for $\Gamma_Z \approx \Gamma_Z^{(0)}$ and using the definition (2.12) for the decay widths into specific helicity states we can rewrite the total cross section in terms of widths and partial widths:

$$\sigma_0(s) = \frac{N_c}{48\pi}\sum_{\lambda_e\lambda_f}\left|\frac{C_{\lambda_e\lambda_f}}{s-M_Z^2+iM_Z\Gamma_Z} + \frac{Q_eQ_f}{s}\right|^2 \tag{2.14}$$

where

$$C_{\lambda_e\lambda_f} = \pm\lambda_e\lambda_f\frac{24\pi\Gamma_{\lambda_e}^{\frac{1}{2}}\Gamma_{\lambda_f}^{\frac{1}{2}}}{N_c^{\frac{1}{2}}M_Z} \tag{2.15}$$

where the $\pm$ sign correspond to $I_f^3 = \pm\frac{1}{2}$ of the final state and $\lambda_e\lambda_f$ refer to the helicities of the e^- and f. Carrying out the summations over those states, the total

cross section in lowest order for $e^+e^- \to f\bar{f}$ reads

$$\sigma_0(s) = \frac{sN_c}{(s-M_Z^2)^2 + M_Z^2\Gamma_Z^2}\left[\frac{12\pi\Gamma_e\Gamma_f}{M_Z^2 N_c} + \frac{I(s-M_Z^2)}{s}\right] + \frac{4\pi Q_f^2\alpha^2 N_c}{3s} \tag{2.16}$$

with

$$I = \frac{\pm 4\pi Q_e Q_f \alpha}{N_c^{\frac{1}{2}} M_Z}\left(\Gamma_+^{\frac{1}{2}}(e) - \Gamma_-^{\frac{1}{2}}(e)\right)\left(\Gamma_+^{\frac{1}{2}}(f) - \Gamma_-^{\frac{1}{2}}(f)\right) \tag{2.17}$$

The first term in eq. (2.16) is the Breit-Wigner form for a spin 1 resonance of mass M_Z and width Γ_Z, while the last term is the pure QED cross section. The interference term I is positive for the expected value of s_w^2 and smallest for lepton pair production.

2.2.2 The corrected partial and total widths

The first order corrections to $\Gamma_Z^{(0)}$ can be divided in four classes:

1. Non-photonic loop corrections
2. Photonic loop corrections and radiative decays
3. QCD corrections
4. Decay into three or more particles

The non photonic correction in OS scheme have been discussed extensively in refs. [4, 5]. The value of $\Gamma_Z^{(0)}$ is generally larger than the one obtained from (2.10). However the one loop correction is small such that the corrected widths end up almost the same. Therefore eq. (2.10) is a good representation of the width at least for $m_t < 150GeV$.

The photonic corrections give a multiplicative factor $1 + \delta_{QED}$, where

$$\delta_{QED} = \frac{3\alpha}{4\pi}Q_f^2 \tag{2.18}$$

which is smaller than 0.17%. It should be noted that although we discuss the QED corrections to the cross sections in section 2.3.1, we include here the QED correction to the width. The reason is that it will have an effect on the propagator as we will see in the next section. The QCD corrections are obtained by multiplying the quark decays by the factor $1 + \delta_{QCD}$. We will be back over this point later. Decays into three or more particles also represent corrections to the lowest order width. Two types were already treated above in the form of QED and QCD corrections. Of the other possible decays only the decay (ref.[6]) $Z \to Hf\bar{f}$ is of relevance, but only when $M_H \leq 10 GeV$. The others decays are negligible [7].

2.2.3 Total cross section with electroweak corrections

The lowest order total cross section for massless fermions given by eq. (2.13) is of order α^2 except at $s = M_Z^2$ where it is of order α^0. Since Γ_Z is related to the imaginary part of the self-energy, the one-loop corrections to the propagator (see figure 2.2) are not sufficient and two-loop corrections should be taken into account in the resonance region.

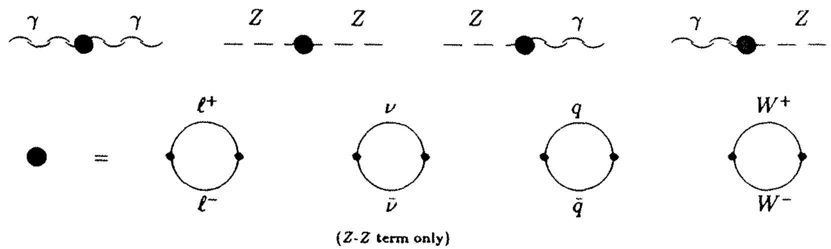


Figure 2.2: Propagator corrections to s-channel driven processes in a physical gauge.

Besides the modifications of coupling constants by vertex corrections (see figure 2.3),

the introduction of very small box diagrams (see figure 2.4), the electroweak corrections amount to the replacements

$$\frac{1}{s} \to \frac{1}{s + \Sigma_\gamma(s)}, \tag{2.19}$$

$$\frac{1}{s - M_Z^2 + iM_Z\Gamma_Z} \to \frac{1}{s - M_Z^2 + \Sigma_Z(s)}, \tag{2.20}$$

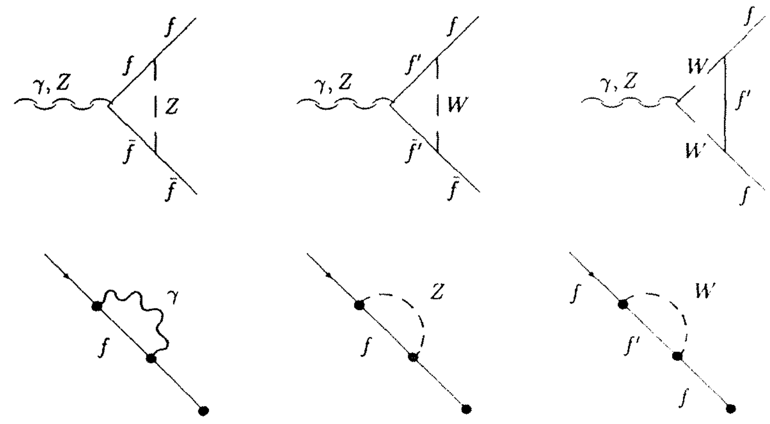


Figure 2.3: Vertex corrections and fermion self-energy insertions (neutral Higgs boson neglected).

where,

$$\Sigma_\gamma(s) = \Sigma_{\gamma\gamma}(s) - \frac{\Sigma^2_{\gamma Z}(s)}{s - M_Z^2 + \Sigma_{ZZ}(s)} \tag{2.21}$$

$$\Sigma_\gamma(s) = \Sigma_{ZZ}(s) - \frac{\Sigma^2_{\gamma Z}(s)}{s + \Sigma_{\gamma\gamma}(s)} \tag{2.22}$$

These expressions are obtained from a Dyson series summation involving the renormalized one particle irreducible self-energies $\Sigma_{\gamma\gamma}$, Σ_{ZZ} and $\Sigma_{\gamma Z}$. Besides the above propagators, one has also to include a γ-Z mixing propagator, which takes the form

$$D_{\gamma Z}(s) = \frac{-\Sigma_{\gamma Z}(s)}{s\,(s - M_Z^2 + \Sigma_Z(s))}. \tag{2.23}$$

In these expressions, the real parts of Σ_γ and Σ_Z are taken in first order. The imaginary part of Σ_Z is considered up to second order. That is also the imaginary

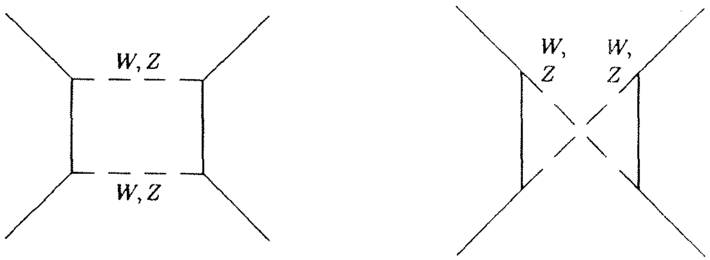


Figure 2.4: Box corrections to s-channel driven processes.

part of Σ_{ZZ} should be evaluated in second order. These is done by the following approximation

$$Im\Sigma^{(2)}_{ZZ}(s) = \frac{s}{M_Z^2} Im\Sigma^{(2)}_{ZZ}(M_Z^2), \tag{2.24}$$

where the latter expression is related to the first order correction to the width. All corrections to the width contribute to eq. (2.24) except for the wave function renormalization of the Z and γZ mixing contributions. These can be seen by expanding (2.20) in the resonance region

$$\frac{1}{s - M_Z^2 + \Sigma_Z(s)} = \frac{1}{1 + \Pi_Z(M_Z^2)} \frac{1}{s - M_Z^2 + \frac{iIm\Sigma_Z(s)}{1+\Pi_Z(M_Z^2)}}, \tag{2.25}$$

where

$$\Pi_Z(M_Z^2) = \frac{\partial Re\Sigma_Z}{\partial s}(M_Z^2) \tag{2.26}$$

and

$$M_Z\Gamma_Z = \frac{Im\Sigma_Z(M_Z^2)}{1 + \Pi_Z(M_Z^2)}. \tag{2.27}$$

The denominator in eq. (2.27) represents the wave function renormalization of the Z. It gives a first order correction to Γ_Z.

Besides the propagator effects, vertex corrections replace the couplings g_f^- and g_f^+ by s dependent form factors. The box diagrams turn out to be very small and can be neglected in the total cross section.

2.2.4 Approximate expressions

As mentioned in the previous section, the modification of the total cross section (2.13) due to electroweak corrections is mainly caused by the introduction of s-dependent form factors, which replace the coupling constants $g^{\pm}$, and the changes to the propagators of (2.19), (2.20) and (2.23). Thus, in several places s-dependent quantities replace the original constants. Also, the values for $s = M_Z^2$ are different from the lowest order quantities.

In the region $s = M_Z^2$, the s-dependence of $\Sigma_Z(s)$ in eq. (2.20) is important while the other s-dependences have a small influence. The s-dependences of $Im\Sigma_Z(s)$ near the resonance can be well approximated by the replacement

$$\frac{1}{s - M_Z^2 + iM_Z\Gamma_Z} \to \frac{1}{s - M_Z^2 + is\Gamma_Z/M_Z}. \tag{2.28}$$

The other s-dependent corrections can be evaluated at $s = M_Z^2$. Effectively, we get new coupling constants $g^{\pm}(M_Z^2)$ and

$$e^2(M_Z^2) = \frac{e^2}{1 + \Pi_\gamma(M_Z^2)}, \tag{2.29}$$

where

$$\Pi_\gamma(s) = \frac{Re\Sigma_{\gamma\gamma}(s)}{s}. \tag{2.30}$$

In the constants $g^{\pm}(M_Z^2)$ also form factors effects are in corporated. This is not done for the coupling to photons. The imaginary parts of the form factors and Σ_γ are neglected at this point. Finally, we have the following approximation for the electroweak corrected total cross section for massless fermions:

$$\sigma(s) = \left\{ \frac{s}{(s - M_Z^2)^2 + s^2\Gamma_Z^2/M_Z^2} \left[\frac{12\pi\Gamma_e\Gamma_f}{M_Z^2} + \frac{IN_c(s - M_Z^2)}{s} \right] + \frac{4}{3}\pi Q_f^2 \frac{\alpha^2(s)N_c}{s} \right\} (1 + \delta_{QCD}), \tag{2.31}$$

Where Γ_e and Γ_f are the electroweak corrected partial widths without QED and QCD corrections. The quantity I is given by eq. (2.17) , where α is $\alpha(M_Z^2)$ and the helicity partial widths contain electroweak corrections.

2.3 The QED corrections to the total cross section

The various possible QED corrections are discussed (figure 2.5). Then the most important one, initial state photon radiation, is considered.

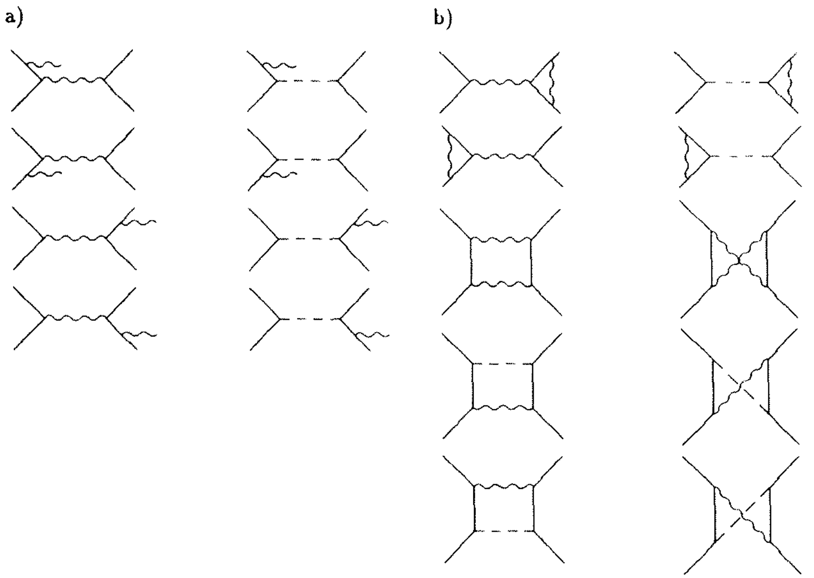


Figure 2.5: QED corrections to s-channel driven processes. a) Real bremsstrahlung. b) Virtual photon exchange.

2.3.1 The importance of various QED contributions

When one considers the full first order QED correction to the cross section, one finds a contribution from initial state radiation, final state radiation, and from the interference between initial and final state radiation. When no cuts on the outgoing fermions are imposed, the final state radiative correction is just eq. (2.18) and it is small. When a stringent cut on the fermion pair invariant mass is applied the correction can become negative and large. The size of the interference contribution to the total cross section is negligible. Thus, only the initial state radiative correction

remains (see figure 2.6). It is sizeable due to the occurrence of large logarithms of the type:

$$L = ln\frac{s}{m_e^2}. \tag{2.32}$$

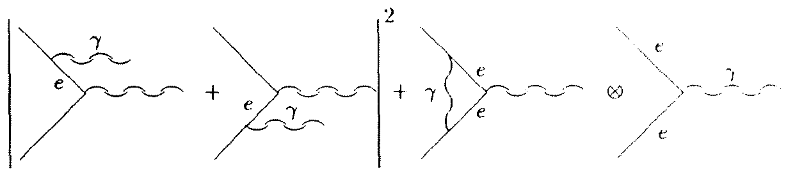


Figure 2.6: Initial state bremsstrahlung. The propagator of the internal electron lines gives a contribution proportional to L.

2.3.2 Initial state photon radiation

A standard second order calculation is needed in order to consider initial state photon radiation. One has to calculate double bremsstrahlung, the one-loop corrections to single bremsstrahlung and the two-loop vertex corrections. The result can be written in the form

$$\sigma(s) = \int_{z_0}^{1} dz\sigma_w(sz)G(z), \tag{2.33}$$

where, the cross section including weak corrections is denoted by $\sigma_w(sz)$, i.e. eq. (2.31). The invariant mass of the produced fermion pair is given by

$$s' = sz \tag{2.34}$$

where,

$$4\frac{m_f^2}{s} \leq z_0 \leq z \leq 1. \tag{2.35}$$

Unless specified differently, the cut-off invariant mass z_0 will be $4m_f^2$. In general the function describing the photonic radiation $G(z)$ has the following expansion

$$G(z) = \delta(1-z) + \frac{\alpha}{\pi}(a_{11}L + a_{10}) + \left(\frac{\alpha}{\pi}\right)^2(a_{22}L^2 + a_{21}L + a_{20}) + \ldots + \left(\frac{\alpha}{\pi}\right)^n \sum_{i=0}^{n} a_{ni}L^i + \ldots \tag{2.36}$$

In the full second order calculation the coefficients a_{1i}, $i = 0,1,2$ are obtained. Instead of performing the explicit second order QED calculation, one may apply the QCD structure approach to QED problems. Usually this method is used to obtain the leading logarithms, i.e. the terms

$$\left(\frac{\alpha}{\pi}\right)^n a_{nn} L^n \tag{2.37}$$

in eq. (2.36) up to a certain n. When a number of terms of the form (2.37) has been obtained, certain parts of a_{nn} generalize to higher n values. Then it is often possible to carry out the summation over n in eq. (2.36) for those parts of the terms a_{nn}, which are related to soft photons. The latter represents a specific exponentiation of some terms in the first order result.
We adopt the exponentiation as in refs. [41] and [42] where the complete subleading logarithms have been obtained in the structure function approach. The radiator function reads

$$G(z) = \beta(1-z)^{\beta-1}\delta^{V+S} + \delta^H, \tag{2.38}$$

with

$$\beta = \frac{2\alpha}{\pi}(L-1), \tag{2.39}$$

$$\delta^{V+S} = 1 + \delta_1^{V+S} + \delta_2^{V+S}, \tag{2.40}$$

$$\delta^H = \delta_1^H + \delta_2^H, \tag{2.41}$$

$$\delta_1^{V+S} = \frac{\alpha}{\pi}\left(\frac{3}{2}L + 2\zeta(2) - 2\right), \tag{2.42}$$

$$\delta^{V+S} = \left(\frac{\alpha}{\pi}\right)^2 \left[\left(\frac{9}{8} - 2\zeta(2)\right)L^2 + \left(-\frac{45}{16} + \frac{11}{2}\zeta(2) + 3\zeta(3)\right)L \right. \\ \left. -\frac{6}{5}\zeta(2)^2 - \frac{9}{2}\zeta(3) - 6\zeta(2)ln2 + \frac{3}{8}\zeta(2) + \frac{19}{4}\right], \tag{2.43}$$

$$\delta_1^H = -\frac{\alpha}{\pi}(1+z)(L-1), \tag{2.44}$$

$$\delta_2^H = \left(\frac{\alpha}{\pi}\right)^2 \frac{1}{2}(L-1)^2 \left[(1+z)(3lnz - 4ln(1-z)) - \frac{4}{1-z}lnz - 5 - z\right]. \tag{2.45}$$

In these definitions the Riemann zeta function has been introduced, $\zeta(2) = \pi^2/6$ and $\zeta(3) \approx 1.202$.

The terms δ_1^{V+S}, δ_2^{V+S} originate from first and second order virtual and soft photon corrections. Similarly δ_1^H and δ_2^H originate from single and double hard bremsstrahlung.

The convolution (2.33) of the cross section with the radiator function (2.38) results in an important deformation of the z_0 line shape, lowering the observable cross-section by about 30% at the peak.

2.3.3 Results for σ_l and σ_h cross sections

The radiative cross section for fermion pair production (2.31), can be modified to obtain a model independent formula as in ref [48]. The idea is to have an expression for the cross sections that depends only on the four parameters Γ_f, Γ_Z, Γ_e, and M_Z.

Firstly, the QCD corrections enter through a multiplicative factor $(1 + \delta_{QCD})$ but, when this factor multiplies the hadron width it is absorbed to have the fully corrected physical width, with the replacement

$$\Gamma_h(1 + \delta_{QCD}) \rightarrow \Gamma_h \tag{2.46}$$

The pure photon exchange term is considered as a known background and computed from pure QED. There is no way to write the interference term I in terms of the four above mentioned parameters. A small model dependency is introduced by computing this term in the standard model (i.e. (2.12), (2.17)). The actual size of I is small for values of $sin^2\theta_w$ in a reasonable range.

The QED coupling constant α is taken as a function of s by applying the QED one-loop corrections to the photon propagator,

$$\alpha(s) = \frac{\alpha}{1-\Delta\alpha} \tag{2.47}$$

with,

$$\Delta\alpha = \frac{\alpha}{3\pi}\sum_l Q_l^2\left(ln\frac{s}{m_l^2}\right) + \Delta\alpha_h \tag{2.48}$$

where, the index l run over charged leptons and α is taken in the Thompson limit (i.e. $\alpha = 1/137.036$), while the contribution from quarks (ref. [3]) is

$$\Delta\alpha_h = a + b\left[ln\frac{s}{s_0} + c\left(\frac{s}{s_0} - 1\right)\right] \tag{2.49}$$

where s_0 is taken to be $92^2 Gev^2$.

To apply the initial-state radiation correction, the cross section should be convoluted with the QED initial state radiator. However, QED corrections to Γ_e would be included twice. In order to avoid this double correction the following substitution has been applied

$$\Gamma_e \rightarrow \frac{\Gamma_e}{(1+\delta_{QED})} \tag{2.50}$$

where, $\delta_{QED} \simeq 0.17\%$ as defined by eq. (2.18).

In order to include final state corrections up to $O(\alpha)$ to the terms that are not Z exchange, the pure QED term is multiplied by the factor $(1+\delta_{QED})$ for lepton pair production and $(11/3 + (35/27)\delta_{QED})$ for hadron production. The interference term for lepton pair production in turn reads

$$I_l = \frac{4\pi Q_l \alpha(s)\alpha(M_Z^2)}{6}\cdot(g_e^+ + g_e^-)\cdot(g_l^+ + g_l^-)\cdot(1+\delta_{QED}) \tag{2.51}$$

while for hadrons

$$I_h = \frac{-12\pi\alpha(s)\alpha(M_Z^2)}{6} \cdot (g_e^+ + g_e^-)\sum_{q=u,d,s,c,b}(g_q^+ + g_q^-)Q_q(1+Q_q^2\delta_{QED}) \tag{2.52}$$

where g_f^+ and g_f^- are given by eqs. (2.9).

QCD corrections enter as a multiplicative factor $1+\delta_{QCD}$ where

$$\delta_{QCD} = a_1\frac{\alpha_s}{\pi} + a_2(\frac{\alpha_s}{\pi})^2 + a_3(\frac{\alpha_s}{\pi})^3 + \ldots \tag{2.53}$$

The coefficients a_i are taken from ref. [8] $a_1 = 1.060, a_2 = 0.90, a_3 = -15$.

After including all the modifications mentioned above, and following the prescriptions of ref. [48], the explicit expression for the non-radiative cross section for lepton pair production becomes

$$\begin{aligned}\sigma_l(s) = &\frac{s}{(s-M_Z^2)^2 + s^2\Gamma_Z^2/M_Z^2}\\ &\left[\frac{12\pi\Gamma_e\Gamma_l}{M_Z^2}\cdot\frac{1}{1+\delta_{QED}} + \frac{I_l(s-M_Z^2)}{s}\right]\\ &+ \frac{4\pi Q_l^2\alpha^2(s)}{3s}.(1+\delta_{QED}),\end{aligned} \tag{2.54}$$

whereas for hadron production it is

$$\begin{aligned}\sigma_h(s) = &\frac{s}{(s-M_Z^2)^2 + s^2\Gamma_Z^2/M_Z^2}\\ &\left[\frac{12\pi\Gamma_e\Gamma_h}{M_Z^2}\cdot\frac{1}{1+\delta_{QED}} + \frac{I_h(s-M_Z^2)}{s}(1+\delta_{QCD})\right]\\ &+ \frac{4\pi\alpha^2(s)}{3s}\cdot\left(\frac{11}{3} + \frac{35}{27}\delta_{QED}\right)(1+\delta_{QCD}).\end{aligned} \tag{2.55}$$

If we evaluate the expressions above at the Z_0 peak (i.e. $s = M_Z^2$) we have for the di-leptonic final state

$$\sigma_l = \sigma_l^0\frac{1}{1+\delta_{QED}} + \sigma_l^{QED} \tag{2.56}$$

where,

$$\sigma_l^0 = \frac{12\pi\Gamma_e\Gamma_l}{M_Z^2\Gamma_Z^2} \tag{2.57}$$

is the non-radiative peak cross section from Z exchange,

and

$$\sigma_l^{QED} = \frac{4\pi\alpha^2(M_Z^2)}{3M_Z^2}(1+\delta_{QED}) \tag{2.58}$$

is the pure QED cross section with the radiative corrections corresponding to dileptonic final states.

For the hadronic final states we have

(2.59)
$$\sigma_h = \sigma_h^0 \frac{1}{1+\delta_{QED}} + \sigma_h^{QED}$$

where,

(2.60)
$$\sigma_h^0 = \frac{12\pi\Gamma_e\Gamma_h}{M_Z^2\Gamma_Z^2}.$$

is the non-radiative peak cross section from Z exchange,

and

(2.61)
$$\sigma_h^{QED} = \frac{4\pi\alpha^2(M_Z^2)}{3M_Z^2}\cdot\left(\frac{11}{3}+\frac{35}{27}\delta_{QED}\right)(1+\delta_{QCD})$$

is the pure QED cross section with the radiative corrections corresponding to hadronic final states. For massless leptons $\Gamma_l = \Gamma_e = \Gamma_\mu = \Gamma_\tau$, and their cross-sections are all equal. Taking the mass of the lepton different from zero but small, $m_l \ll M_Z$, in eq. (2.10), the following substitution should be made

(2.62)
$$\Gamma_l \rightarrow \Gamma_l(1-\frac{6m_l^2}{M_Z^2}).$$

The difference $1-\frac{6m_l^2}{M_Z^2}$ is considerably different from 1 only for the tau lepton, taking the value 0.9977 . The model independent cross sections (2.54) and (2.55) have to be convoluted according to (2.33) in order to take initial state radiation into account.

2.3.4 Sensitivity of Q_μ and Q_τ to R

To have an idea of the dependence of Q_μ and Q_τ on the four physical parameters M_Z, Γ_Z, σ_h^0 and R, let's compute the quotients Q_μ and Q_τ using the expressions for the cross section at the Z_0 peak without initial state radiation (2.56) and (2.59). Neglecting the mass of the leptons, Q_l reads

(2.63)
$$Q_l = \frac{R+\frac{\sigma_h^{QED}}{\sigma_l^0}}{1+\frac{\sigma_l^{QED}}{\sigma_l^0}}.$$

Considering $\sigma_l^{QED}/\sigma_l^0 \ll 1$, neglecting the term proportional to $(\sigma_l^{QED}\sigma_h^{QED})/(\sigma_l^0)^2$ and making use of the relation

(2.64)
$$\frac{\sigma_h^0}{\sigma_l^0} = \frac{\Gamma_h}{\Gamma_l} \equiv R$$

we have

(2.65)
$$Q_l = R\left[1+\left(\sigma_h^{QED} - R\sigma_l^{QED}\right)\frac{1+\delta_{QED}}{\sigma_h^0}\right]$$

Although expression (2.65) ignores initial state radiation, it serves to illustrate the functional dependence of Q_l on the parameters. The main dependence is linear with respect to R. The dependence on Γ_h^0 is small of order 10^{-2}, as it in this approximation, is through the ratio $\sigma_h^{QED}/\sigma_h^0$. In this approximation, Q_l is independent of M_z and Γ_z.

Then evaluating the QED cross sections and taking $\sigma_h^0 = 41.6$nb we have the parabola

(2.66)
$$Q_l = 1.00108\ R - 2.84 \times 10^{-4}\ R^2$$

That parabola can be approximated to a straight line in the interval [20, 21], with a slope of 0.989 and an intercept of 0.12.

Finally, the cross sections (2.54) and (2.54) have been convoluted as indicated in (2.33) by numerical methods to get the final fitting formula. The sensitivity of Q_μ and Q_τ to R has been studied by varying separately each of the four parameters M_Z, Γ_Z, σ_h^0 and R (figure 2.7). As we see in the figure, Q_l once again practically depends only on R and its dependence is linear. The dependence of Q_l on R represents a sensitivity to α_s (eq. (1.2)) as indicated in figure 2.7 d.

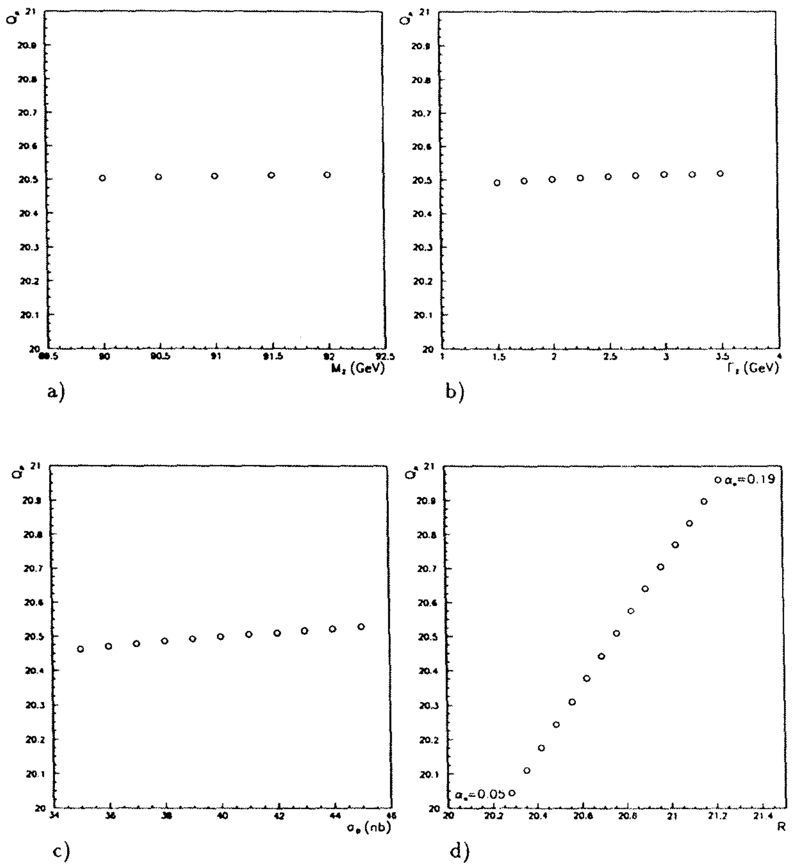


Figure 2.7: Sensitivity of $Q\mu$ respect to M_Z, Γ_Z, σ_h^0 and R at the Z_0 peak;
a)$\Gamma_Z = 2.501GeV$, $\sigma_h^0 = 41.6nb$ and $R = 20.69$;
b)$M_Z = 91.187Gev$, $\sigma_h^0 = 41.6nb$ and $R = 20.69$;
c)$M_Z = 91.187Gev$, $\Gamma_Z = 2.501GeV$, and $R = 20.69$;
d)$M_Z = 91.187Gev$, $\Gamma_Z = 2.501GeV$ and $\sigma_h^0 = 41.6nb$

Chapter 3

The ALEPH experiment

3.1 The LEP collider

LEP, the Large Electron Positron storage ring [10], is a circular accelerator sited at the European Laboratory for Particle Physics (CERN, Geneva) and spanning the French and Swiss territories (see figure 3.1). It is located inside a tunnel of 26.7

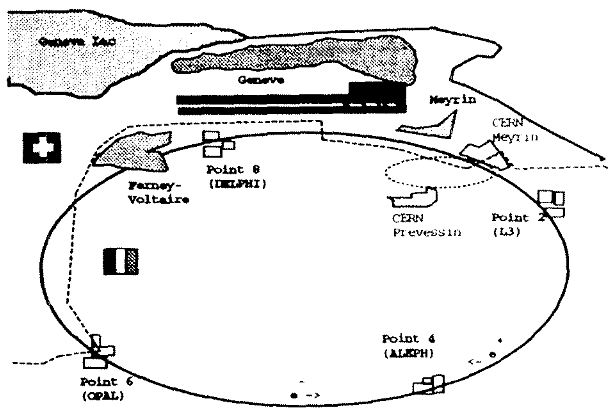


Figure 3.1: The LEP ring.

km of circumference whose depth below ground ranges from 50 to 170 m due to the rise and fall of the terrain. LEP is, in fact, a big (nearly circular) octagon which consists of 8 arcs and 8 straight sections. Electrons and positrons are accelerated in four bunches each (changed to eight since September 1992), in opposite directions. These bunches are steered to collide at each of the four experimental areas where the detectors L3 [11], ALEPH, OPAL [12] and DELPHI [13] (see figure 3.1) are located.

The accelerator program is comprised of two phases. In the first (current) phase, it accelerates, stores and collides electrons and positrons up to a beam energy of 55 GeV. At a center-of-mass energy around 90 GeV, LEP produces Z bosons with a luminosity of the order of $10^{31}cm^{-2}s^{-1}$. In a second phase, an increase of a center-of-mass energy up to 200 GeV, above W-pair production threshold, and a luminosity of approximately $10^{32}cm^{-2}s^{-1}$ are planned. Polarized beams are also foreseen.

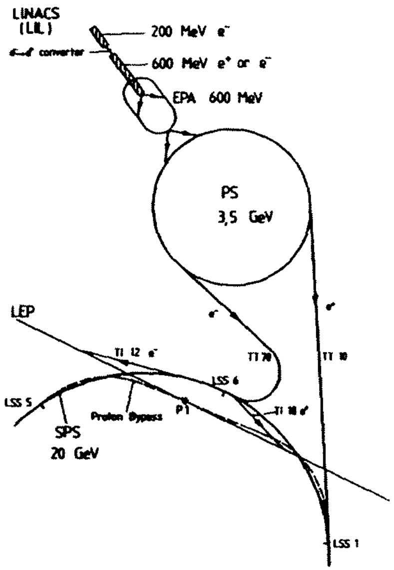


Figure 3.2: Scheme of the LEP injectors and accelerators.

As shown in figure 3.2, the LEP injection chain starts in the LINear ACcelerator (LINAC) which accelerates electrons and positrons in two stages. The electrons are first accelerated up to 200 MeV. Part of the electrons are used to produce positrons and the rest, together with the positrons, are accelerated up to 600 MeV in a second stage. After the LINAC, the particles are injected in a small circular e^+e^- accelerator, the Electron Positron Accumulator (EPA), where they are accumulated. From there, they are inserted to the Proton Synchroton (PS) accelerator, where the energy is taken up to 3.5 GeV. The particles are injected into the Super Proton Synchroton (SPS) storage ring, reaching an energy of 20 GeV. Finally, they are injected into the LEP main ring and accelerated to a maximum of $\simeq$ 55 GeV with a current up to 2.9 mA per beam.

Table 3.1 gives the main parameters of LEP.

Parameter	Value
Circumference	26658.883 m
Average radius	4242.893 m
Bending radius in the dipoles	3096.175 m
Depth	50–170 m
Number of interaction points	8
Number of experimental areas	4
Number of bunches per beam	4–8
RMS Bunch length	11.67 mm
Horizontal bunch sigma	200 μm
Vertical bunch sigma	12 μm
Injection Energy	20 GeV
Maximum beam energy (phase I)	55 GeV
RF Frequency	353 MHz
Total current per beam	0.029 A
Luminosity	$10^{31}cm^{-2}s^{-1}$
Vertical β_V^*	5 cm
Horizontal β_H^*	25 × β_V^* cm

Table 3.1: Main LEP parameters.

3.2 The ALEPH detector: general description

ALEPH (ALEPH: "A detector for LEp PHysics" [14], [15]) is one of the four large detectors using the LEP accelerator. It has been designed as a general purpose detector for e^+e^- interactions: to study in detail the parameters of the Standard Model, to test QCD at large Q^2 and to search for new phenomena (such as the top quark, the Higgs boson or supersymmetric particles). Therefore, the detector has been conceived to have good track momentum resolution, fine calorimetric granularity, almost 4π angular coverage and hermeticity.

The ALEPH detector is shown in figure 3.3. It has a $12 \times 12 \times 12$ m^3 cylindrical structure around the beam pipe with the interaction point in the center. In the ALEPH reference system (ARS) the z direction is along the beam line, positive in the direction followed by the e^-. The positive x direction points to the center of LEP, and is horizontal by definition. The positive y direction is orthogonal to z and x and is very close to vertical up.

It consists of the following subdetectors (in the order in which a particle leaving the interaction point would encounter them):

- The Mini Vertex DETector (VDET), fully operational since 1991, is a double sided silicon strip device with two layers of $r\phi$ and z strips around the beam pipe, providing a very accurate vertex tagging of tracks coming from the interaction point ($\sigma_{r\phi} = 10\mu$m, $\sigma_z = 13\mu$m). It is not shown in figure 3.3 due to its small size.
- The Inner Tracking Chamber (ITC) is a cylindrical multiwire drift chamber. It is used to provide up to eight precise $r\phi$ coordinates per track, with an accuracy of 100 μm per coordinate. It contributes to the global ALEPH tracking and is also used for triggering of charged particles coming from the interaction region.
- The Time Projection Chamber (TPC), the central track detector of ALEPH, is a very large three-dimensional imaging drift chamber. It provides a three dimensional measurement (up to 21 coordinate points) of each track (single-coordinate resolutions of 173 μm in the azimuthal direction and 740 μm in the longitudinal direction are achieved) and, from its curvature in the magnetic field, gives a measurement of transverse particle momenta p_T with an

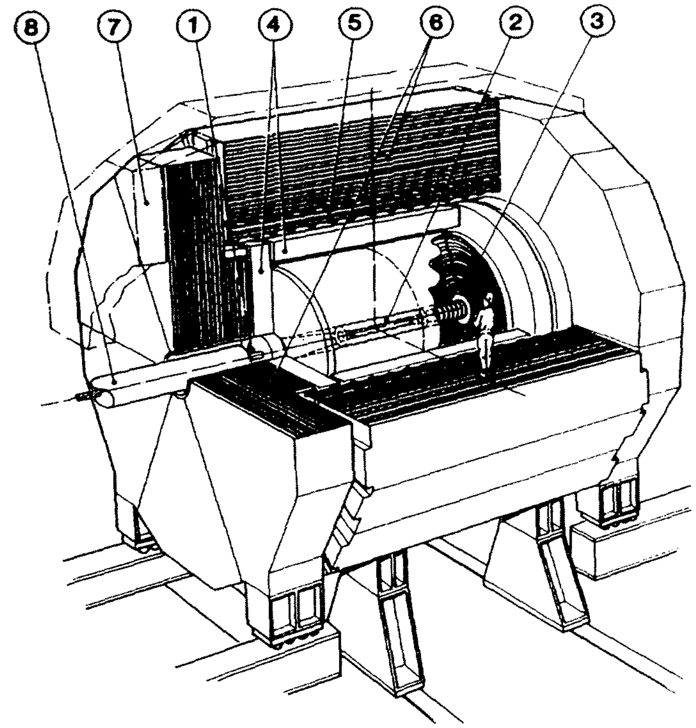


Figure 3.3: Schematic view of the ALEPH detector. (1) Luminosity Monitor. (2) Inner Tracking Chamber. (3) Time Proportional Chamber. (4) Electromagnetic Calorimeter. (5) Superconducting Coil. (6) Hadronic Calorimeter. (7) Muon Chambers. (8) Beam Pipe.

accuracy of $\Delta p_T/p_T^2 = 0.8 \times 10^{-3}$ (GeV/c)$^{-1}$ at 45 GeV if it is used together with the ITC (1.2×10^{-3} if used alone). The chamber also contributes to particle identification through measurements of energy loss (dE/dx) derived from about 340 samples of the ionization for a track traversing the full radial range.

- The Electromagnetic CALorimeter (ECAL) is a sampling calorimeter consisting of alternating lead sheets and proportional wire chambers read out in projective towers to obtain a very high granularity (about $1° \times 1°$). It measures the energy and position of electromagnetic showers. The high position and energy resolutions achieved ($\Delta E/E \simeq 0.18\ GeV^{1/2}/\sqrt{E}$) leads to good electron identification and allows to measure photon energy even in the vicinity of hadrons.

- The superconducting coil is a liquid-Helium cooled superconducting solenoid creating, together with the iron yoke, a 1.5 T magnetic field in the central detector.

- The Hadronic CALorimeter (HCAL) is a sampling calorimeter made of layers of iron and streamer tubes. It provides the main support of ALEPH, the large iron structure serving both as hadron absorber and as return yoke of the magnet. It measures energy and position for hadronic showers ($\Delta E/E \simeq 0.84\ GeV^{1/2}/\sqrt{E}$) and, complemented with the muon chambers, acts as a muon filter. The readout is performed twice: using cathode pads forming projective towers and using digital readout of the streamer tubes for muon tracking and also for triggering.

- The MUON chambers (MUON), outside HCAL, are two double layers of limited streamer tubes which identify muons and measure their positions and angles.

Precise measurements of the electroweak parameters require accurate knowledge of the beam luminosity which is provided by four detectors for small angle Bhabha scattering installed around the beam pipe:

- The Luminosity CALorimeter (LCAL), the main luminosity calorimeter before SICAL installation, is a lead/wire calorimeter similar to the ECAL in its operation. It consists of two semi-circular modules placed around the beam pipe at each end of the detector. As SICAL, it is used to measure the integrated luminosity accumulated by the detector.

- The Small Angle Monitor of the BAckground (SAMBA) is positioned in front of the LCAL at either end of the detector and replaced the Small Angle TRacking device (SATR) in 1992. It consists of two multi-wire proportional chambers at each end, read out in two rings of 8 pads per ring. It is used as a background monitor.

- The SIlicon Luminosity CALorimeter (SICAL) is a new luminosity monitor installed in September 1992 on each side of the interaction region. It uses 12 silicon/tungsten layers to sample the showers produced by small angle Bhabhas. It improves the statistical precision of the luminosity measurement by sampling at smaller angles than the LCAL. The systematic error of the luminosity is also reduced thanks mainly to the greater internal precision of its components.

- The very small Bhabha CALorimeter (BCAL) located after the final focus quadrupoles, is used to give a measurement of the instantaneous and specific luminosity and also as a background monitor. It is a sampling calorimeter made of tungsten converter sheets sandwiched with sampling layers of plastic scintillator. A single plane of vertical silicon strips is used to locate the shower position.

Finally, the Beam Orbit Monitors (BOMs), located around the circumference of LEP, measure the mean position and angle of the beam orbits which are used by LEP to optimize the beam conditions, and by ALEPH to determine the (x, y) position of the beam spot as a starting point for offline reconstruction of the primary vertex.

Trigger System

The purpose of the trigger system [14] is to produce a signal that starts the readout of the events. It is desirable to keep all the electron-positron collisions and to reduce as much as possible the rate of background events. For these reasons the trigger system has been organized in a three-level scheme:

- **Level one** decides whether or not to read out all the detector elements. Its purpose is to operate the TPC at a suitable rate. The decision is taken approximately 5 μs after the beam crossing from pad and wire information from the ECAL and HCAL and hit patterns from the ITC. The level one rate must not exceed a few hundred Hz.

- **Level two** refines the level one charged track trigger using the TPC tracking information. If level one decision cannot be confirmed with better precision, the readout process is stopped and cleared. The decision is taken approximately 50 μs after the beam crossing (the time at which the TPC tracking information is available). The maximum trigger rate allowed for level two is about 10 Hz.

- **Level three** is performed by software. It has access to the information from all detector components and is used to reject background, mainly from beam-gas interactions and off-momentum beam particles. It ensures a reduction of the trigger rate to 1–2 Hz, which is acceptable for data storage.

This trigger scheme has to be rather flexible since it has to be able to reject the background and keep signals from possible new physics events. Therefore the available electronic signals from different ALEPH detector components allow for a variety of triggers which together cover all possible types of events.

Data Acquisition System and Event Reconstruction

The ALEPH modular structure is also used in the data acquisition system, allowing each subdetector to take data independently. The DAQ [16] architecture is highly hierarchical. Following the data and/or control flow from the bunch crossing of the accelerator down to storage device, the components found and its tasks are briefly described below:

- Timing, Trigger and Main Trigger Supervisor: synchronize the readout electronics to the accelerator and inform the ReadOut Controllers (ROCs) about the availability of the data.

- ROCs: initialize the front-end modules, read them out and format the data.

- Event Builders (EBs): build a subevent at the level of each subdetector and provide a "spy event" to a subdetector computer.

- Main Event Builder (MEB): collects the pieces of an event from the various EBs and ensures resynchronization and completeness.

- Level three trigger or Event Processor: as seen, performs a refined data reduction.

- Main host and subdetector computers: The main machine (a VAXCluster) initializes the complete system, collects all data for storage and provides the common services. The subdetector computers get the "spy events" and perform the monitoring of the large subdetectors (TPC, ECAL, HCAL).

The event reconstruction is performed in a quasi-online way. Due to the event rate (1 Hz) and the large size of the events (50 Kbytes) a large computing facility is needed. A system based on a Local Area VaxCluster initially configured with 12 satellite VaxStations 3100/M30 ($\simeq$ 6 CERN units[1]) and upgraded in 1992 to 6 VaxStation 4000/M60 and 4 VaxStation 3100/M76 ($\simeq$ 20 CERN units) for parallel processing is implemented (FALCON, Facility for ALeph COmputing and Networking [17], [18]) as seen in figure 3.4.

The two boot nodes (MicroVAX 3600 and MicroVAX 3800) control four dual-ported RA90 1.2 Gbytes disks, where the second port is connected to the DAQ VAXCluster. The disks are alternatively mounted on both systems under software control, and the raw data files written to one of these disks by the DAQ are made available to FALCON shortly after the end of the run. After a preliminary scan of the data to produce an event directory, contiguous subsets of events are then assigned for reconstruction in parallel in each of the FALCON processors, exploiting the fact that events are independent from one another. Each of the processors runs the full ALEPH reconstruction program JULIA (Job to Understand Lep Interactions in ALEPH) [19] which, for each event of the raw data, processes all the information from the different subdectectors and, basically, associates observed coordinates in the tracking chambers with charged particle tracks and energies deposited in the calorimeters.

After their reconstruction, the events are written in POT (Production Output Tape) data files and transmitted to the CERN computer center where they are converted into different data types more suitable for physics analysis and stored on tape by ALPROD (ALeph data PRODuction job). The events are ready to be analyzed only a few hours after having been taken.

[1]A CERN unit is equivalent to an IBM 168 CPU unit, approximately 1/6 of an IBM 3090 processor or about 1.2 Mflops.

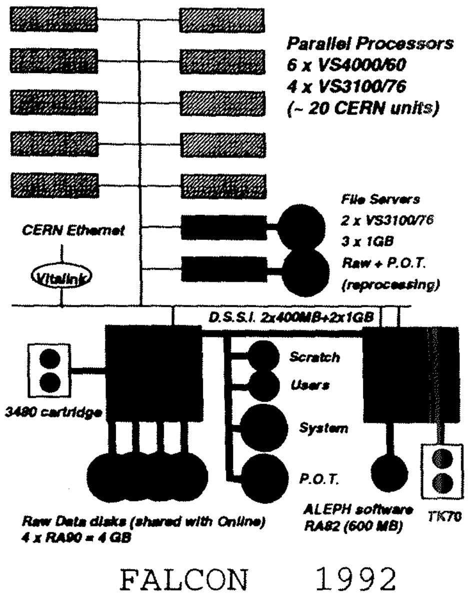


Figure 3.4: Schematic representation of the FALCON cluster.

The Event Simulation

In addition to these detector reconstructed events, the different physics analyses use Monte Carlo simulated events. Simulated data are used to evaluate background contaminations, compute acceptances and efficiencies and, in general, compare theoretical models to the experimental results. GALEPH [22] is the ALEPH generation program.

The steps needed to get simulated data are the following:

- Generation of the event kinematics. The different particle four-momenta are generated according to the different physics processes (in parentheses the

names of the computer programs used).

  - $e^+e^- \to \mu^+\mu^-$ (KORALZ [23]).
  - $e^+e^- \to \tau^+\tau^-$ (KORALZ).
  - $e^+e^- \to e^+e^-$ (BABAMC [24]).
  - $e^+e^- \to q\,\bar{q}$ (LUND [25]).
  - $e^+e^- \to \ell^+\ell^-(\ell^+\ell^-)$ (PHOPHO [26], [27]).

  In ALEPH, all these programs have been unified through the common interface KINGAL [28].

- Simulation of the detector. This is done using GEANT [29]. Through subroutine calls, the program is informed about the geometry and materials involved in the experimental setup.
- Tracking. The primary long-lived particles are followed through the detector. Secondary particles are also produced by interaction with the detector materials. Bremsstrahlung, Compton and ionization are some of the processes simulated. GEANT and GHEISHA [30] are used to simulate the electromagnetic and nuclear interactions respectively.
- Digitization or simulation of the detector behaviour. The energy depositions are converted to measurable electrical signals. The complexity of the TPC required the development of a special package (TPCSIM) for its tracking and digitization.
- Simulation of trigger. The same conditions of the real trigger are checked.

The final output of the program has the same format as the real data so that the reconstruction program is also used on simulated data.

3.3 Subdetectors relevant to the analysis

3.3.1 The Mini Vertex Detector

The VDET [31] was the first double sided silicon microstrip detector to be installed in a colliding beam experiment.

The two concentric layers of silicon microstrips wafers are located at radii of 6.3 cm and 10.8 cm, as can be seen in figure 3.5. Particles passing through a wafer

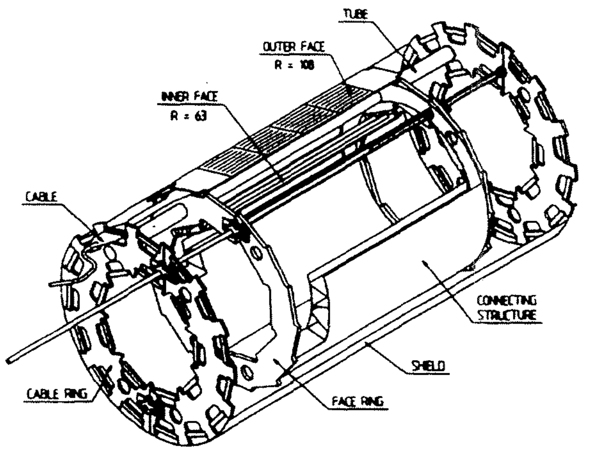


Figure 3.5: Cut-away view of the VDET.

deposit ionization energy, which is collected on each side of the wafer. On one side, the wafer is read out in the z direction, while in the other, it is read out in the orthogonal $r - \phi$ direction. Hits on the two sides are not associated by the hardware; they are added to tracks during reconstruction.

The advantage of the VDET is that it pinpoints a track's location in space quite near to the beam pipe. VDET hits are used by extrapolating a track found by the ITC and/or the TPC to the VDET and then refitting the track more precisely using VDET hits which are consistent with it. The addition of VDET to the tracking improved the momentum resolution to $\Delta p_T/p_T^2 = 0.6 \times 10^{-3}$ (GeV/c)$^{-1}$ at 45 GeV [32].

Using VDET, together with the other tracking detectors, the spatial coordinates of the origin of a charged track's helix can be found to within about 30 μm. This allows tracks produced by decay of short-lived particles to be separated from those at the primary interaction point with good efficiency.

3.3.2 The Inner Tracking Chamber

The Inner Tracking Chamber (ITC) [33] using axial wires provides up to eight $r\phi$ points for tracking in the radial region between 16 and 26 cm. It also provides the only tracking information for the level one trigger system. It is able to identify roughly the number and geometry of tracks, due to its fast response time (the trigger is available within 2–3 μs of a beam crossing) and allows non-interesting events to be quickly rejected.

The ITC is operated with a gas mixture of argon (50%) and ethane (50%) at atmospheric pressure. It is also possible to add a small fraction of alcohol or water to prolong the lifetime of the chamber by inhibiting the formation of polymers and their deposition on the wires.

The ITC is composed of 8 layers of sensing wires (operated at a positive potential in the range 1.8–2.5 kV) running parallel to the beam direction, which detect the ionization of particles passing close by. By measuring the drift time, the $r\phi$ coordinate can be measured to within about 100 μm . The z coordinate is found by measuring the difference in arrival times of pulses at the two ends of each sense wire, but with an accuracy of only about 3 cm.

The drift cells of the ITC are hexagonal, with a central sense wire surrounded by six field wires held at earth potential. Four of these field wires are shared by neighbouring cells in the same layer (see figure 3.6). The cells in contiguous layers are offset by half a cell width, which helps to resolve the left-right ambiguity in the track fitting. One field wire per cell is insulated from the end-plate and can be used to inject a calibration pulse into the chamber. The other field wires are in electrical contact with the end-plates. The cell size was kept small to reduce the drift, thereby allowing a rapid trigger decision. Finally, pairs of drift-cell layers are separated by a layer of guard wires.

The sense wires are connected to preamplifiers. The preamplifier outcoming signals are then taken to the central boards which contain the main amplifiers, discriminators, latches and time-expansion circuits. The incoming signals are amplified by a factor of about 36 and then discriminated at a constant-fraction threshold. The discriminator outputs from one end of the chamber are routed to the drift time CAMAC time-to-digital converters (TDCs).

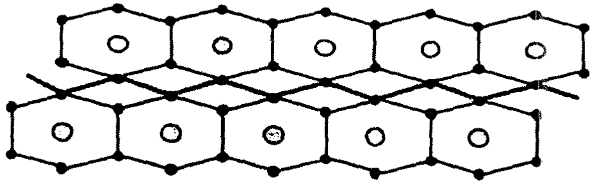


Figure 3.6: The ITC drift cells.

Four digital-to-analog converters (DACs) control the operation of time-expansion circuit for every channel. The time-expanded pulses are routed to FASTBUS TDCs for offline space-point reconstruction.

3.3.3 The Time Projection Chamber

The Time Projection Chamber (TPC) [34] is in many ways the heart of ALEPH. It was designed to obtain high precision measurements of the track coordinates, to get good momentum resolution and to measure the dE/dx depositions of charged particles coming from e^+e^- interactions.

The TPC works as follows. The electrons produced by the ionization of traversing charged particles drift towards one end-plate, where they induce ionization avalanches. These are detected and yield the impact point ($r\phi$ coordinate). The time needed for the electrons to reach the end-plate gives the z coordinate. Due to the presence of the 1.5 T magnetic field parallel to the TPC symmetry axis, the trajectory of a charged particle inside the TPC is a helix and its projection onto the end-plate is an arc of a circle. Measurement of the sagitta of this arc yields the curvature radius which is proportional to the modulus of the component of the momentum perpendicular to $\vec{B}$.

As shown in figure 3.7, the TPC has a cylindrical structure; its volume is delim-

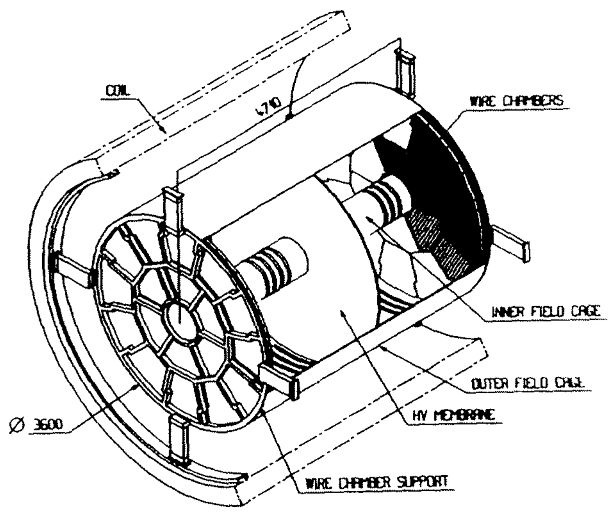


Figure 3.7: View of the Time Projection Chamber.

ited by two coaxial cylinders which hold the end-plates. The inner cylinder has a radius of 31 cm, the outer one of 1.8 m; this large radial difference allows to reach 10% resolution in transverse momentum for the highest possible momenta (muon pairs produced at a center-of-mass energy of 200 GeV). Both cylinders are 4.4 m long.

The device is divided into two half-detectors by a membrane which is situated in the plane perpendicular to the axis and midway between the end-plates. This central membrane is held at a negative high voltage (−27 kV) and the end-plates are at a potential near ground. The curved cylindrical surfaces are covered with electrodes held at potentials such that the electric field (115 V/cm) in the chamber volume is uniform and parallel to the cylinder axis.

The TPC volume is filled with a nonflammable gas mixture of argon (91%) and methane (9%) at atmospheric pressure. This mixture allows to reach high $\omega\tau$ values (ω = cyclotron frequency; τ = mean collision time of the drifting electrons). This causes the electrons to drift mainly along the magnetic field lines and thereby reduce the systematic displacements due to the electric field inhomogeneities.

The electrons produced by the ionization are amplified in the proportional wire chambers placed in the end-plates. There are 18 wire chambers ("sectors") on each end-plate. In order to get a minimum loss of tracks at boundaries, the sectors are arranged in the "zig-zag" geometry shown in figure 3.8. The gaps between the

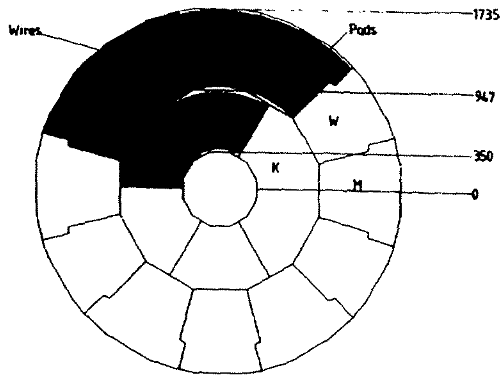


Figure 3.8: View of a TPC end-plate.

sectors are kept as small as possible. In each end-plate, there are six sectors labelled K (Kind) inside and a ring of twelve alternating sectors labelled M (Mann) and W (Weib) outside. All sectors are composed of wire chambers and cathode pads. The wire chambers consist of three layers of wires: gating wires, cathode wires and sense/field wires. Figure 3.9 shows a perspective of them.

The gating grid [35] has the purpose of preventing positive ions produced in the avalanches near the sense wires from entering the main volume of the TPC, and thereby distorting the electric field. Potentials of $V_g \pm \Delta V_g$ ($V_g \simeq -67$ V) are placed

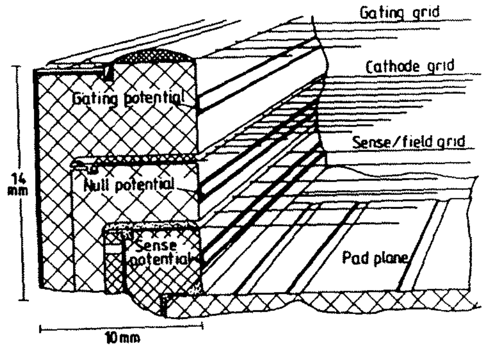


Figure 3.9: View of a TPC wire chamber.

on alternating wires of the grid. A $\Delta V_g \simeq 40$ V suffices to block the passage of the positive ions while, because of the magnetic field, a much bigger ΔV is required to block also the incoming electrons. In the open state, the grid is transparent to the drifting charged particles. When closed, positive ions are kept off of the drift volume. The gate is opened 3 μs before every beam crossing. If a positive trigger signal arrives, the gate is kept open for the maximum 45 μs drift time of the electrons in the TPC, otherwise the gate is closed.

The cathode wires keep the end-plates at null potential and, together with the central membrane, create the electric drift field.

The sense wires are kept at a positive potential to provide avalanche multiplication. They are read out to give the energy deposition (dE/dx) for particle identification and the z measurement of the tracks. For the estimation of the dE/dx a truncated mean algorithm is used, taking the mean of the 60% smaller pulses associated with a track. The estimator will be normally distributed and will be sensitive to the particle velocity.

The field wires are kept at null potential to create equipotential surfaces around the sense wires.

The ionization avalanches created around the sense wires are read out by the signal induced on cathode pads at a distance of 4 mm from the sense wires.

The cloud of charge projected onto the TPC end-plates is, therefore, measured twice: the sense wires measure track positions in an approximate rz projection, while the cathode pads measure the three-dimensional coordinates. The pulse measured in the pads contains information on charge and time. In a first step a pattern analysis is performed to find possible subpulses (in $r\phi$ and in z). Each one of the subpulses, if some conditions are fulfilled, will be a coordinate. The second step is to determine exactly the $r\phi$ and z coordinates using the charge information. Also, the errors on the coordinates are estimated.

The efficiency of the coordinate finding has been estimated with Monte Carlo events and is 92% for particles above 500 MeV and 75% for those with momentum between 100 and 200 MeV.

Once the coordinates have been calculated, the process of track finding can begin. The first step is the association of coordinates which are consistent with a helix hypothesis in what is called a chain. The second step is chain linking: chains compatible with the hypothesis of coming from the same particle form a track candidate. Finally, the track fitting tries to find the best parameters for a track candidate. This procedure may split a track candidate in two or remove points that disturb the fit.

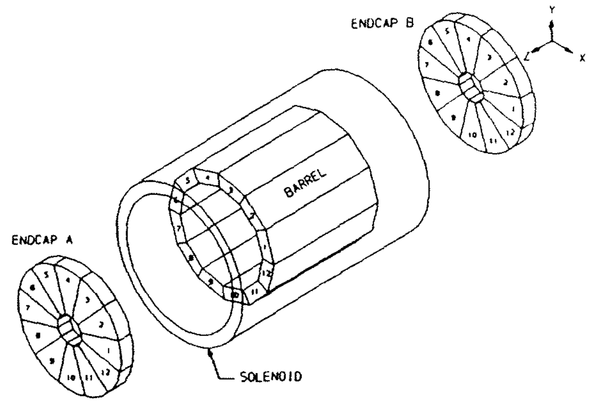
The TPC tracks found are used in the first phase of the ITC tracking: the TPC track trajectories are projected back into the ITC and a search is made for ITC coordinates around each trajectory. If no hits are found in the outer two layers of the ITC for a specific trajectory the search is abandoned. If more than three hits are found a fit is performed and the ITC track is accepted if the fit satisfies a χ^2 cut.

The readout of the TPC is based on FASTBUS [36]. The pad and wire preamplifiers on the sectors are mounted in groups of 16. The electrical signals are handled (shaping, digitization, zero-suppression and gain control) in FASTBUS modules called TPD (Time Projection Digitizer). The readout of the TPDs is controlled by an "intelligent" processor, the TPP (Time Projection Processor). One TPP serves all the TPDs corresponding to the pads, and another one, all TPDs corresponding to the wires of a sector. The tasks of the TPP are: data formatting and readout, reduction of wire data, monitoring of selected events and calibration. The data-flow

generated by the TPPs is transferred to another processor, the Event Builder. The 36 TPPs of each end-cap are read out by one EB, whilst a third one controls the common activities.

3.3.4 The Electromagnetic Calorimeter

The Electromagnetic CALorimeter (ECAL) [37] is located around the TPC and inside the coil. It is divided into a central barrel region closed at both ends with end-caps, as shown in figure 3.10. Both barrel and end-caps are divided into modules



A

Figure 3.10: Electromagnetic CALorimeter, overall view.

of 30° in azimuthal angle ϕ with the end-cap modules rotated 15° with respect to the barrel modules. The entire calorimeter is rotated by $-1.875°$ with respect to the HCAL in order to avoid the overlap of crack regions. The barrel is a 4.8 m long cylinder with an inner radius of 1.85 m and an outer radius of 2.25 m. The dimensions of the end-caps are: 0.56 m length, 0.54 m inner radius and 2.35 m outer radius.

Each module consists of 45 layers of lead and wire chambers. The wire chambers are made of open-sided aluminium extrusions and filled with a gas mixture of xenon (80%) and carbon dioxide (20%). Ionization from an electromagnetic shower developed in the lead sheets is amplified in avalanches around 25 μm diameter gold-plated tungsten wires. The signals are read out via the extrusions' open faces with cathode pads covered by a graphited mylar sheet. The structure of a typical single layer of the calorimeter is shown in figure 3.11.

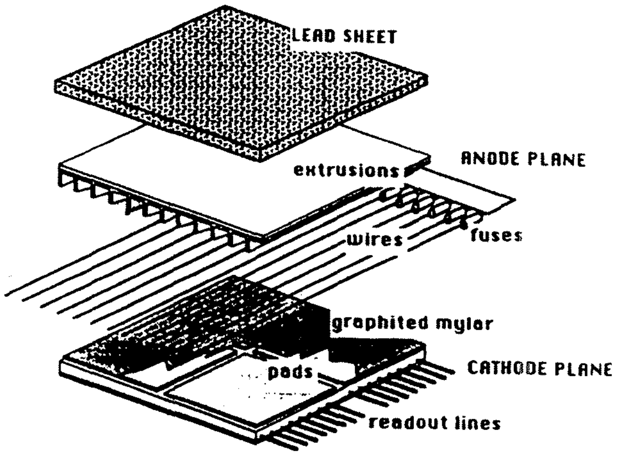


Figure 3.11: View of an ECAL stack layer.

The cathode pads are connected internally to form "towers" which point to the interaction point. Each tower is read out in three sections in depth ("storeys") corresponding to the first four, the middle nine and the last nine of the 22 radiation lengths (X_0) nominal thickness. The size of the pads is approximately 30×30 mm^2 leading to a high granularity (73 728 towers). In addition to the signal of the pads, an analog signal is also available from each anode wire plane (these signals are used for testing and calibrating the modules and also for triggering).

The achieved energy resolution is

$$\frac{\sigma_E}{E} \simeq 1.0\% + \frac{18\%\ \mathrm{GeV}^{1/2}}{\sqrt{E}} \qquad (3.1)$$

and the position resolution for charged tracks with $|\cos\theta_{track}| < 0.98$ is [38]

$$\sigma_\phi = \sigma_\theta / \sin\theta = 0.32\ \mathrm{mrad} + \frac{2.7\ \mathrm{mrad\ GeV}^{1/2}}{\sqrt{E}}.$$

At the reconstruction level, the signals of the triggered towers are combined in clusters: a cluster is defined as the set of towers which are geometrically connected by at least one corner.

To associate clusters with charged tracks, the track is extrapolated step-by-step to the ECAL region. At each step, the ECAL geometry package is used to determine which storeys are intercepted by the track. Then the clustering algorithm is used to determine if the storey, or its neighbours, are hit and to which cluster they belong. A track and a cluster are associated if one point of this track is in one storey of the cluster or in a storey which has at least one corner in common with the cluster.

The particle identification relies on the fact that the structure of the clusters is quite different for electromagnetic and hadronic showers. A complete description of the algorithms used for particle identification can be found in [39]. The efficiency for identifying electrons is close to 100%, with very little contamination.

3.3.5 The hadronic calorimeter and the Muon Detector

The magnet iron (see figure 3.12) is instrumentedd with 23 layers of plastic (limited-)streamers tubes, separated by iron sheets 5 cm thick. The tubes layers are equipped with pad readout, summed in towers for a localized measurement of the total deposited energy. Each tube is also coupled capacitatively to strips paralell to the wires. It is their hit pattern that is read out. Wire information is read out plane by plane (end-caps) and two planes by two planes (barrel) and used for the trigger. The muon detector is composed of two double-layers streamer tubes, wich are outside the magnet, behind the last layer (10 cm) of the HCAL. Each single layer reads out two orthogonal coordinates using strips. There are 4788 towers with an iron depth near $\theta = 90°$ and $\phi = 0°$ of 120 cm. The size of each tower is $3.7° \times 3.7°$. The tipical accuracy of energy measurement $\sigma_E/E = 0.84 GeV^{-1/2}/\sqrt{E}$ while the tipical spatial accuracy for strip coordinate measurement (perpendicular to the strip direction) is $\Delta = 0.35 cm$. There are 94 double-layer muon chamber with a distance between the two layers of 50 cm for the barrel and middle angles and 40 cm for end-caps. The tipical accuracy for a muon exit angle measurement is 10 mrad while the tipical misidentification probability to take a π for μ is of 0.7 % and the probability to take a K for a μ is 1.6 %. The fraction of solid angle covered by the sensitive part of the muon chamber above 15° is for the inner layer of 92 % while for the outer layer of 85 %.

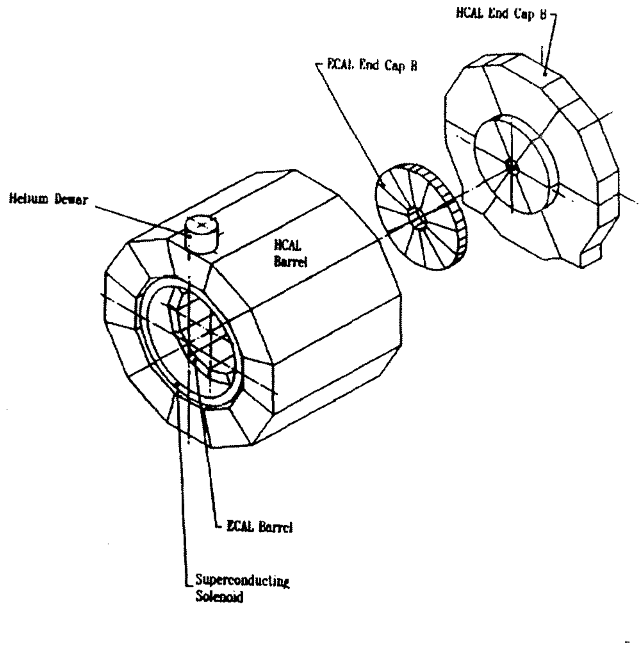


Figure 3.12: Hadronic CALorimeter, overall view.

Chapter 4

Experimental Analysis

In order to measure R, a selection is needed which separates the five kinds of visible final states that we have in e^+e^- collisions at the Z_0 pole, namely:

$$e^+e^- \to e^+e^-,$$
$$e^+e^- \to q\bar{q},$$
$$e^+e^- \to \mu^+\mu^-,$$
$$e^+e^- \to \tau^+\tau^-,$$

as well as the so-called two-photon events

$$e^+e^- \to e^+e^-f^+f^-.$$

The Neural Network (NN) method provides a way to deal with a large number of variables, exploiting combinations between them which are unknown a priori, and which are more suitable to separate the different kinds than simple binary cuts. After training a NN, we have a transformation between the space of measured variables and a space with dimension equal to the number of kinds to separate. In this space, the distribution of events is less overlapped than in the original variable space, which has the advantage that the systematic errors due to miss-simulations are reduced.

4.1 Event classification

The event classification is based on a set of variables that are the outcome of the ALEPH energy flow algorithm, combined with methods to identify electrons, photons and muons (ref. [52]). As a result of these algorithms, for each event we have a set of tetra-momenta, charges and labels identifying the particle. Particles are identified as belonging to one of the following categories:

- electron,
- muon,
- non-identified charged track,

for charged tracks;

- photon,
- non-identified neutral cluster,

for neutral clusters; and

- luminosity cluster,

for energy depositions in the luminosity calorimeter, outside the acceptance of the charged tracking devices.

4.1.1 Energy flow algorithm

For each event, the energy flow is reconstructed using charged particle tracks and calorimeter clusters in the following way:

- charged particle tracks, with at least four space coordinates reconstructed in the TPC and originating from the beam-crossing point within 7 cm along the beam direction and 2.5 cm in the transverse direction, are counted as charged energy;
- V^0's (long-lived neutral particles decaying into two oppositely-charged particles) are kept if they point to the interaction vertex within the same tolerance as those defined for charged particles tracks;

- photons, identified in the electromagnetic calorimeter through their characteristic longitudinal and transverse shower profiles, are counted as neutral electromagnetic energy;
- The remaining neutral hadronic energy is finally determined from calorimeter clusters, defined as sets of calorimeter cells which are topologically connected. The typical size of a cell is smaller than $1° \times 1°$ in the electromagnetic calorimeter, and $3° \times 3°$ in the hadron calorimeter. In a given cluster, let E_{ecal} be the energy in the electromagnetic calorimeter not attributed to photons, and E_{hcal} the energy in the hadron calorimeter, and let $E_{charged}$ be the energy of the charged tracks, if any, topologically associated to the cluster. The difference

$$E_{neutral} = E_{hcal} + rE_{ecal} - E_{charged}$$

is counted as neutral hadronic energy if $E_{neutral} > \xi\sqrt{E_{charged}}$, with $E_{neutral}$ and $E_{charged}$ in GeV . Here r is the ratio of the responses for electrons and pions in the electromagnetic calorimeter ($r \approx 1.3$), and ξ is related to the energy resolution of the calorimeters for hadronic showers: $\xi = 0.5$ for a shower fully contained in the electromagnetic calorimeter, $\xi = 1.0$ for a shower fully contained in the hadron calorimeter, and $\xi = (0.5rE_{ecal} + E_{hcal})/(rE_{ecal} + E_{hcal})$ in the general case.

4.1.2 Selection of Variables

Three different types of variables can be obtained from the outcome of the energy flow algorithm and the particle identification methods [53]. One type of variable is obtained by counting the number of particles belonging to each category, and will be called a "counting variable". If we sum over quantities related to each tetra-momentum in order to have a quantity for the whole event, we have an "adding variable". Finally, it is possible to build variables that reflect the event topology, or "shape variables". A variable is selected if it contributes to the separation of at least two of the five kinds enumerated above.

Given the set of tetra-momenta, the event is divided in two hemispheres by the plane normal to the thrust axis. The thrust T is obtained by finding the direction of the unitary vector $\hat{\mathbf{t}}$ in 3-momentum space which maximizes the sum in absolute value of projections of 3-momentum vectors,

$$T = MAX\left(\frac{\sum_i |\mathbf{p}_i.\hat{\mathbf{t}}|}{\sum_i |\mathbf{p}_i|}\right) \tag{4.1}$$

where the direction of $\hat{\mathbf{t}}$ is the thrust axis.
If we apply again the thrust algorithm, but only to the projections of the 3-momentum vectors over that plane, we have the major value

$$M = MAX\left(\frac{\sum_i |[\mathbf{p}_i - (\mathbf{p}_i.\hat{\mathbf{t}})\hat{\mathbf{t}}].\hat{\mathbf{m}}|}{\sum_i |\mathbf{p}_i|}\right), \tag{4.2}$$

where the direction of the unitary vector $\hat{\mathbf{m}}$ is the major axis.
Finally we have a third axis orthogonal to both the thrust and major axes which is the minor axis. If we take the unitary vector along this axis $\hat{\mathbf{n}}$, the minor value reads

$$N = \frac{\sum_i |\mathbf{p}_i.\hat{\mathbf{n}}|}{\sum_i |\mathbf{p}_i|}. \tag{4.3}$$

Once we have the event divided in two hemispheres, it is possible to compute variables related to each one. In the following, the subscript hi with $i = 1, 2$ will indicate that the variable is taken in hemisphere hi.

The counting variables are:

- $n_{e,hi}$, the number of identified electrons, which contributes to the separation of bhabha events from muon events, and to a lesser extent to the separation of bhabha events from tau events. (figure 4.1 a)

- $n_{\mu,hi}$, the number of identified muons, which contributes to the separation of bhabha events and muon events, as well as to the separation of muon events and tau events. Combined with $n_{e,hi}$ it separates the tau events which has an electron in one hemisphere and a muon in the other one. (figure 4.1 b)

- $n_{ch,hi}$, the number of non-identified charged tracks, gives the charged track multiplicity when added to $n_{e,hi}$ and $n_{\mu,hi}$, being the main variable to separate hadronic events from the other kinds. (figures 4.1 c and d)

- $n_{\gamma,hi}$, the number of identified photons, contributes to separate hadron events, tau events, bhabha events, and muon events. (figure 4.2 a)

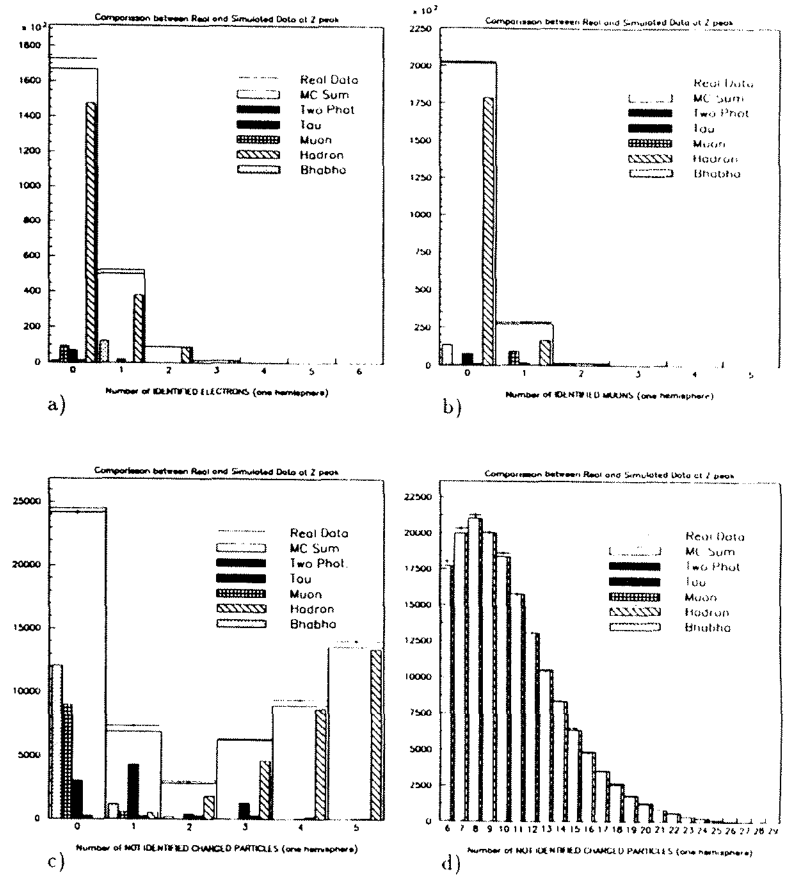


Figure 4.1: Counting variables
a) Number of identified electrons;
b) Number of identified muons;
c) Number of non-identified charged tracks, low multiplicity ($n_{ch,hi} \leq 5$);
d) Number of non-identified charged tracks, high multiplicity ($n_{ch,hi} > 5$).

- $n_{ne,hi}$, the number of non-identified neutral tracks, is similar to $n_{\gamma,hi}$, with the sum of the two being equal to the neutral multiplicity. (figure 4.2 b)
- $n_{lu,hi}$, the number of luminosity clusters, used to monitor energy depositions at small angles.

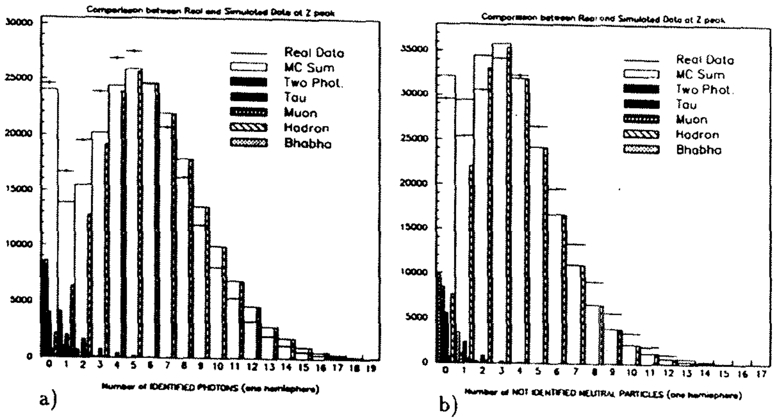


Figure 4.2: Counting variables
a) Number of identified photons;
b) Number of non-identified neutral particles.

The adding variables are:

- $E_{ch,hi}$, the energy of charged tracks, contributes to separate Bhabha events, muon events, tau events and two-photon events. (figure 4.3 a)
- $E_{ne,hi}$, the energy of neutral tracks, contributes to separate hadron events, muon events and to a lesser extent to the separation of Bhabha events and two-photon events. (figure 4.3 b)
- $E_{lu,hi}$, the energy of luminosity clusters, takes into account the energy deposited at small angles.
- W^2_{hi}, the invariant mass squared, where $W^2_{hi} = (E_{ch,hi} + E_{ne,hi} + E_{lu,hi})^2 - (P_{hi})^2$, contributes to the separation of Bhabha events, hadron events and muon events. (figure 4.3 c)
- $P_{t,hi}$, the transverse component of the hemisphere momentum, contributes to the separation of two-photon events from the other kinds. (figure 4.3 d)
- $cos\theta_{hi}$, the cosine of the polar angle of the total momentum vector, mostly used for monitoring purposes.

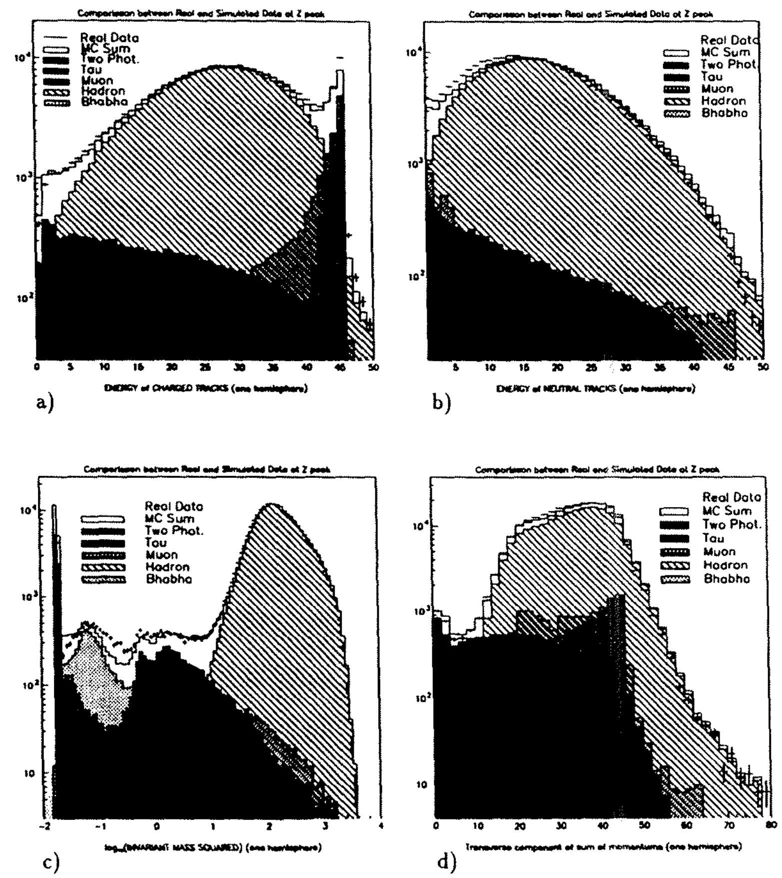


Figure 4.3: Adding variables per hemisphere
a) Energy of charged tracks;
b) Energy of neutral tracks;
c) Invariant Mass;
d) Transverse momentum.

The shape variables are:

- $Y = 2T - 1$, where T is the thrust value, mainly distinguishes hadron events from tau events. (figure 4.4 a)
- M, the major value, separates hadron events from tau events. (figure 4.4 b)
- N, the minor value, contributes to separate hadron events from tau events, like the two previous shape variables. (figure 4.4 c)
- $\cos\theta_{mopa,hi}$, the cosine of the maximum opening angle between any of the 3-momenta of the hemisphere and the total momentum in that hemisphere, separates hadron events from the others. (figure 4.5)
- S_{hi}, the sphericity of boosted particles. All 4-momenta of the hemisphere are boosted to the frame where the total hemisphere momentum is zero. Then, the momentum tensor

$$M_{ij} = \frac{1}{\sum P_1^2 + \sum P_2^2 + \sum P_3^2} \sum P_i P_j$$

is computed, where the sum runs over the particles in the hemisphere. Taking the larger eigen-value of the momentum tensor e_1, the sphericity S is defined as $S = \frac{3}{2}(1-e_1)$. This variable is introduced to have a shape variable belonging to each hemisphere, and it helps to distinguish hadron events from tau events at low multiplicity.

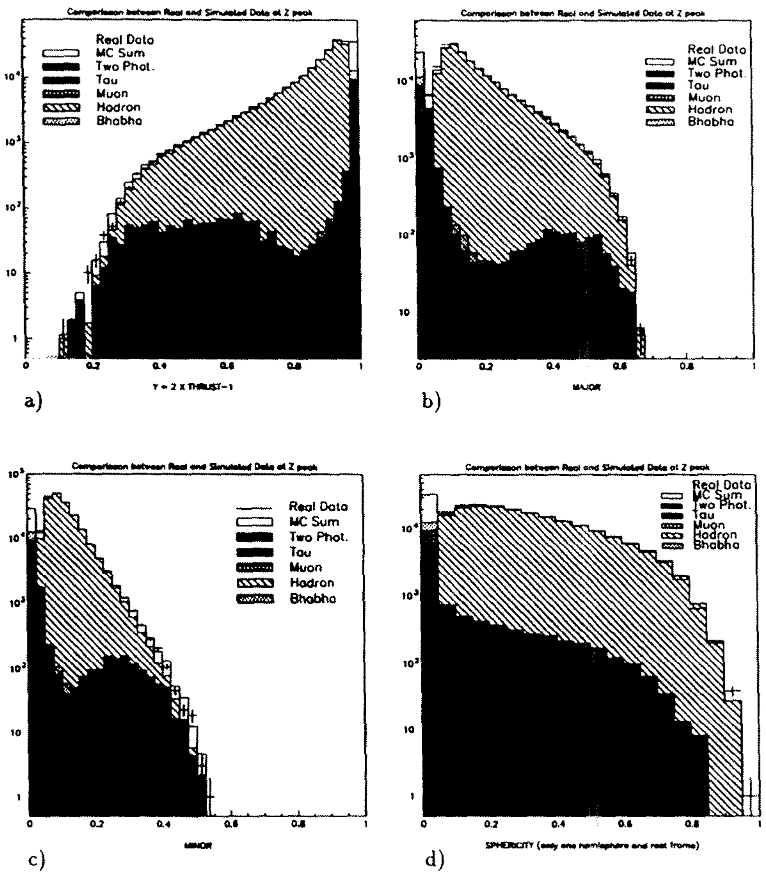


Figure 4.4: Shape variables:
a) 2(thrust-1);
b) Major value;
c) Minor value;
d) Sphericity per hemisphere.

Finally, a pair of variables combining both hemisphere momenta are defined. As pointed out in [50], in the approximation where the initial state radiation is colinear with the electron or positron and the final state radiation is colinear with the final state fermions, the longitudinal rapidity can be defined as

$$L = \frac{1}{2}\ln\frac{x_+}{x_-} = \frac{1}{2}\ln\frac{E+P_z}{E-P_z}$$

where x_+ and x_- are respectively the fractional energies left to the positron and the electron after radiation. This can be expressed as a combination of the final state fermion angles directly measured in the laboratory.

- The longitudinal rapidity reads

$$L = \frac{1}{2}\ln\left(\frac{P_{t,h1}(P_{h2}+P_{z,h2}) + P_{t,h2}(P_{h1}+P_{z,h1})}{P_{t,h1}(P_{h2}-P_{z,h2}) + P_{t,h2}(P_{h1}-P_{z,h1})}\right)$$

where

$$P_t^2 = P_x^2 + P_y^2$$

and it has been introduced as the most important variable to distinguish two-photon events from the other kinds.

In addition, the variable

- $\Delta\phi$, the difference of the azimuth angle of the hemisphere momenta, gives a measure of the acoplanarity of the two hemispheres, differentiating tau and two photon events.

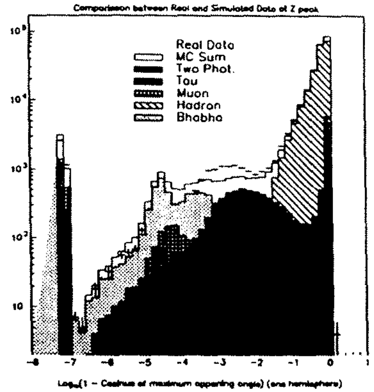
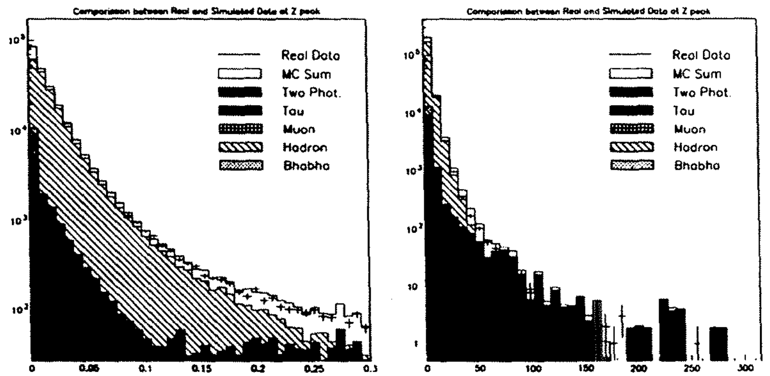


Figure 4.5: Shape variable: $cos\theta_{mopa,hi}$ amplified around $cos\theta_{mopa,hi} = 1$ by plotting $Log_{10}(1 - cos\theta_{mopa,hi})$



a) b)

Figure 4.6: Global variables:
a) Rapidity;
b) Azimuth difference between hemisphere momenta.

$e^+e^- \rightarrow$	kind number	d1	d2	d3	d4	d5
e^+e^-	1	1	0	0	0	0
$q\bar{q}$	2	0	1	0	0	0
$\mu^+\mu^-$	3	0	0	1	0	0
$\tau^+\tau^-$	4	0	0	0	1	0
$e^+e^-f\bar{f}$	5	0	0	0	0	1

Table 4.1: Assignment of desired output values

4.1.3 Neural Net (NN) training

To obtain the transformation between the space of variables and the space whose dimension is equal to the number of kinds to classify, and to have the distribution of events of each kind in this space with the least possible overlap, a NN has been trained.

A Neural Network is a set of formally identical functions (neurons) whose arguments are linear combinations of the values of the other ones (axons). The coefficients of these linear combinations are called weights (activation level of the axon), and the value of each function or activation function is called the output while its argument is called the input (see appendix A for details).

The training process, consists of varying the Net parameters to minimize the energy of the sample, which is the sum of the individual event energies defined by

$$E = \frac{1}{2} \sum_{sample} \sum_{j} (F_j - d_j)^2 \tag{4.4}$$

where j runs over neurons belonging to the output layer, F_j is the value of the output neuron j, while d_j is the desired output for that neuron. In our case, the last layer has five neurons, and the output takes the values indicated in table 4.1

Simulated sample and needed cuts

As a sample of simulated data, the standard ALEPH experiment "Monte Carlo" data was taken. Although the two-photon events have all been assigned to kind number 5, it should be remarked that six different two-photon processes have been simulated, and merged according to their cross section ratios. As these ratios are not well known, an extra systematic error is introduced. As can be seen in figure 4.6, two-photon events have a characteristically flat distribution in rapidity, quite distinct from the other kinds. Therefore an a priori cut of

- $|L| < 0.3$ is made to reduce the two-photon background.

In addition, since the two-photon process was simulated only above a certain invariant mass, the sample is similarly restricted by the cut in invariant mass

- $W > 5$ GeV.

A small background from cosmic rays which has not been simulated, is largely reduced by imposing a time window by requiring a signal coming from one of the fastest subdetectors involved in the analysis, the ITC, which remains enabled the least amount of time possible, leading to the simple cut:

- At least three hits on ITC per event.

The unphysical background, mainly the noise generated by calorimeters has been suppressed by applying the cut:

- At least one charged track per hemisphere.

Finally, a cut was applied in order to avoid the detector zones that are close to the geometric acceptance limit, to prevent miss-simulations and allow a more accurate estimation of the acceptance. The cut applied was:

- Cosine of polar angle in the Z_0 rest frame $|\cos\theta^*| < 0.9$.

The final sample of simulated events is given in table 4.2. The $e^+e^- \to e^+e^-$ process was generated in angular range defined by $|\cos\theta^*| < 0.97$, whereas the other processes were generated over the full solid angle. As can be seen, the cuts described above have little effect on the channels of interest, namely $e^+e^- \to q\bar{q}$, $e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to \tau^+\tau^-$, while eliminating a large fraction of the two-photon events.

event	Generated	after cuts
$e^+e^- \to e^+e^-$	66600	37152
$e^+e^- \to q\bar{q}$	395868	341833
$e^+e^- \to \mu^+\mu^-$	57555	48946
$e^+e^- \to \tau^+\tau^-$	50000	41675
$e^+e^- \to e^+e^- f\bar{f}$	172310	3463

Table 4.2: Number of events generated and after cuts

The training

The NN training was done to reach two goals: the obvious one, to have in the output space the least overlapped distributions for the different kinds; and to have the best relation between efficiency and background for each kind. This requires some definitions or a proper explanation.
Dividing the output space in equal hyper-volume zones, having each kind its own zone, then the efficiency is the number of events of a kind that fall in its own zone relative to the total of events of that kind. The background is defined as the fraction of events that fall in a zone other than the correct one for that kind. This fraction is computed with the number of events weighted by their actual cross sections. This question seems something that bites its own tail, which in fact it does. In an ideal case, it would be desireable to be able to measure on-line (i.e. during the minimization process) the relative amount of each kind on a real sample. In the actual case, there exist previous measurements of these quantities accurate enough for the training process.
If the ratio of efficiency over background is maximized for one kind, it would not necessarily correspond to a maximum for any other. Therefore, the differences of these ratios are computed for all pairs of kinds and these differences are minimized during the training process.

The algorithm of training is fully described in appendix A. The output variables after the training process are plotted in figure 4.7, where the different kinds of simulated data have been identified. The distributions have been weighted by luminosities and added to compare it with real data. As can be seen in the figure, the value of the output assigned to each final state has a distribution peaked at high values for simulated events of the correct final state, and peaked at low values for the other final states.

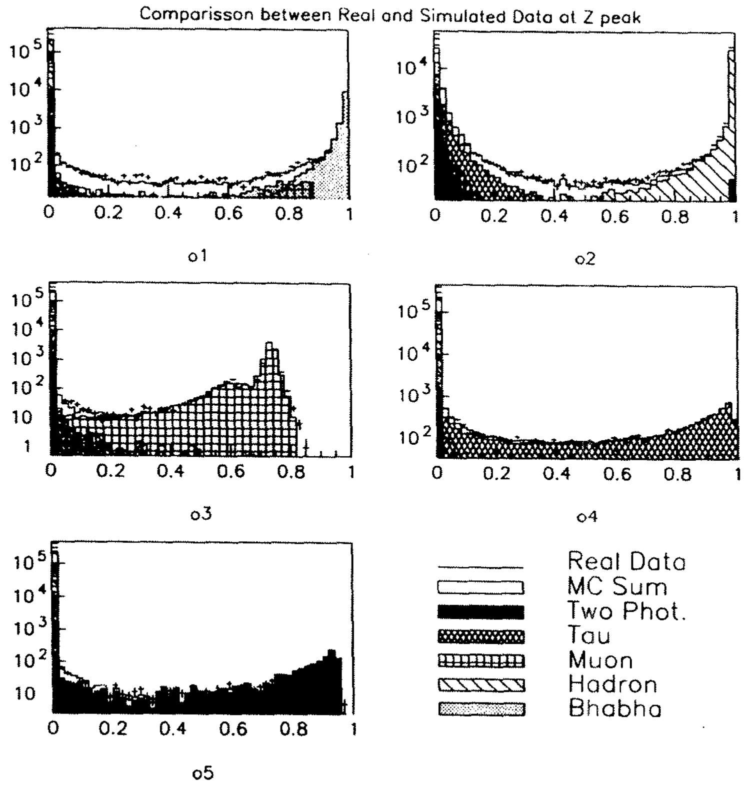


Figure 4.7: Net Output, with the different sets of simulated data identified. The samples are weighted by their luminosities and added to compare with real data

A graphical representation of the NN is shown in figure 4.8, where neurons are represented by circles and the combinations $\beta(W \cdot \mu)$, the links between neurons, are represented by coloured lines. The line colour varies with the value of the link. Only links above the indicated threshold have been plotted. The links reveal the input variables which are weighted by the NN to separate the five kinds of events. The correlation coefficients between each input variable and each output have also been computed. The sample to compute the correlations is composed of the five sets of simulated data with equal relative frequencies. The correlations can be seen in figure 4.9. In table 4.3, the correlations have been normalized to the maximum for each kind separately and they are expressed in percent with respect to this maximum.

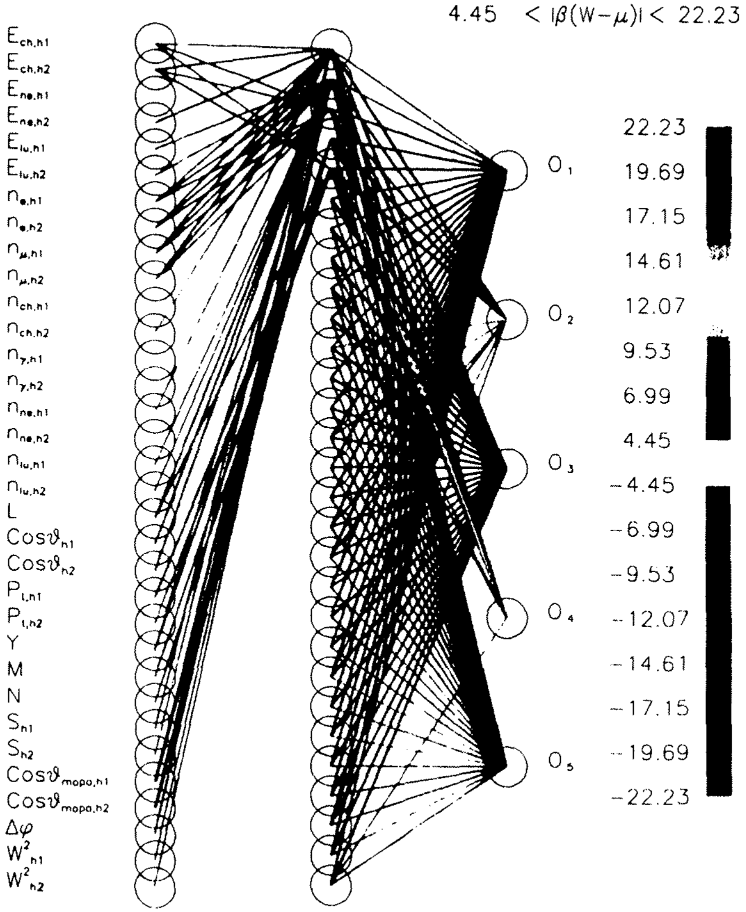


Figure 4.8: Neural Network final state after training

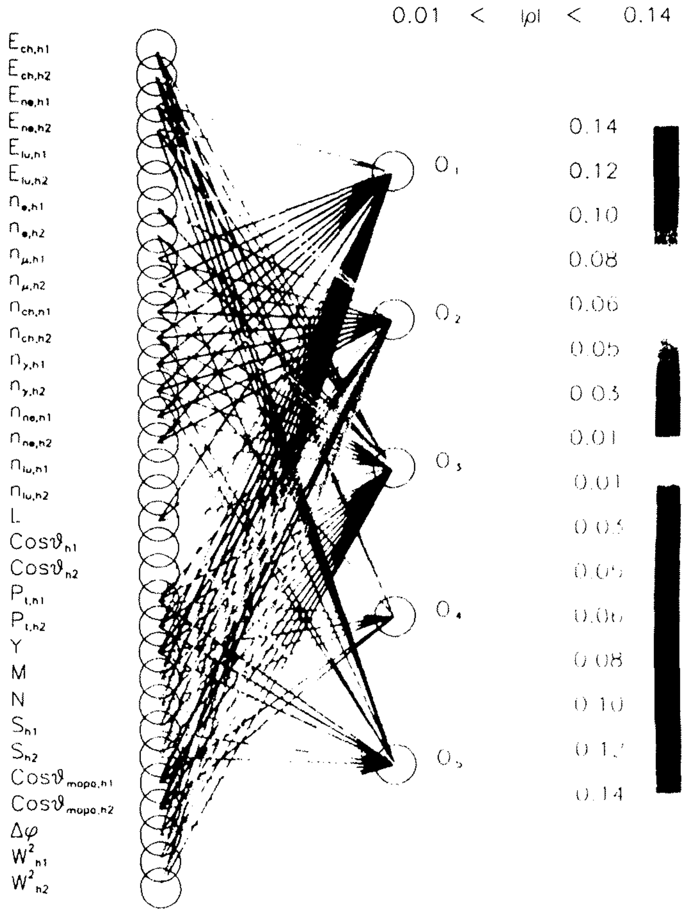


Figure 4.9: Correlation coefficients |ρ| between input and output variables

	correlation/MAX(correlation) (%)				
input variable	bhabha	hadron	muon	tau	two photons
$E_{ch,h1}$	76.0	-0.7	72.8	-90.2	-69.4
$E_{ch,h2}$	74.6	0.1	72.8	-100.0	-68.6
$E_{ne,h1}$	-18.0	64.0	-41.0	34.2	-28.0
$E_{ne,h2}$	-13.9	63.7	-41.0	20.7	-28.0
$E_{lu,h1}$	0.6	1.7	-3.9	-7.3	3.9
$E_{lu,h2}$	-0.4	1.4	-3.4	-7.1	3.5
$n_{e,h1}$	99.8	-8.0	-43.1	-38.3	-14.5
$n_{e,h2}$	100.0	-7.7	-43.0	-37.0	-16.3
$n_{\mu,h1}$	-56.2	-25.5	99.9	-27.7	-12.5
$n_{\mu,h2}$	-51.9	-25.5	100.0	-28.0	-11.9
$n_{ch,h1}$	-53.0	100.0	-39.1	-44.4	-8.5
$n_{ch,h2}$	-53.4	99.6	-39.2	-46.5	-7.6
$n_{\gamma,h1}$	-44.6	92.6	-38.9	-18.8	-17.9
$n_{\gamma,h2}$	-44.1	92.2	-38.8	-22.3	-17.8
$n_{ne,h1}$	-42.0	86.0	-35.3	-35.9	-9.8
$n_{ne,h2}$	-41.7	85.8	-34.9	-39.5	-9.8
$n_{lu,h1}$	-1.5	2.1	-5.5	-10.9	10.2
$n_{lu,h2}$	-3.7	4.7	-5.1	-10.5	7.7
L	-45.5	-14.3	-37.1	-33.6	100.0
$cos\theta_{h1}$	-1.2	0.6	-0.9	0.1	0.0
$cos\theta_{h2}$	1.4	-0.5	1.0	0.2	-0.6
$P_{t,h1}$	52.6	29.4	48.6	-52.4	-76.2
$P_{t,h2}$	53.4	29.9	48.8	-68.4	-75.6
Y	45.4	-36.4	32.9	59.7	-65.5
M	-60.1	42.7	-46.1	-60.9	75.6
N	-53.8	44.0	-41.2	-54.8	66.4
S_{h1}	-54.1	64.5	-42.0	4.6	16.3
S_{h2}	-53.3	65.0	-41.7	-6.1	16.4
$cos\theta_{mopa,h1}$	52.0	-87.5	37.9	67.7	-17.1
$cos\theta_{mopa,h2}$	52.7	-88.6	37.7	70.5	-18.0
$\Delta\phi$	-34.0	-6.7	-24.8	-5.1	58.4
W^2_{h1}	-26.7	66.3	-19.8	-38.7	-14.9
W^2_{h2}	-26.0	66.1	-19.8	-39.3	-14.7

Table 4.3: Correlation coefficients $|\rho|$ between input and output variables

4.1.4 NN output analysis

Once the NN is trained, it is necessary to divide the output space in zones, corresponding to the different kinds of final states. First of all, the distribution in output space has to be analyzed by doing projection over subspaces. Taking the output variables in pairs there are ten possible combinations since we have five output variables. Let's examine the projection of the simulated sample distribution over those two-dimensional sub-spaces, but only for the two kinds which correspond to the given pair of outputs. In the ideal case, the best cut to separate the two kinds of events in question would be a straight line with unit slope and zero intercept, or in other words one event belongs to a zone if the ratio of outputs is greater than one, to the other in the opposite case. In the actual situation, it is necessary to do a scan over that sub-space by varying the slope of the straight line, searching for the point where the ratio efficiency over background for both kinds is equal. The slopes can be parameterized by the ten variables θ_{ij}, related to the pairs of outputs o_i and o_j

$$\theta_{ij} = tan^{-1}(o_i/o_j) \tag{4.5}$$

The distributions of simulated events in this new parametrization are shown in figure 4.10 and 4.11. In the figure, the distribution of the variable θ_{ij} is only plotted for kinds i and j. As can be seen in the figure the distribution for kind i is peaked around $\theta_{ij} = 90$ while for kind j the distribution is peaked around $\theta_{ij} = 0$.

Considering only the distribution for the kind i and the kind j in the variable θ_{ij}, the efficiencies and backgrounds are computed. The distributions are divided in two zones by a cut χ_{ij} in θ_{ij}, the zone i is defined by $\chi_{ij} < \theta_{ij} < 90$ and the zone j is defined by $0 < \theta_{ij} < \chi_{ij}$. Then we have the four numbers n_{ii}, n_{ij}, n_{ji} and n_{jj}, which are respectively the number of events of kind i that are in zone i, the number of events of kind i that are in zone j, the number of events of kind j that are in zone i and the number of events of kind j that are in zone j.
The background of kind j in kind i is

$$b_{ji} = \frac{n_{ji}}{n_{ii} + n_{ji}}, \tag{4.6}$$

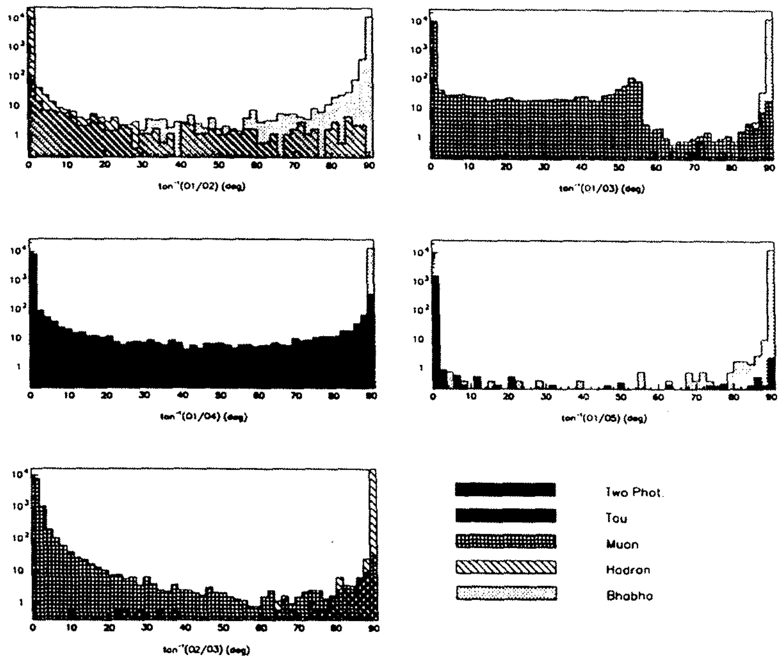


Figure 4.10: $Tan^{-1}(O_i/O_j)$ for $(i,j) = (1,2),(1,3),(1,4),(1,5),(2,3)$; in each case only the corresponding kinds i and j are plotted.

and the efficiency for kind i is

$$e_i = \frac{n_{ii}}{n_{ii} + n_{ij}}. \tag{4.7}$$

A running cut in the interval defined by $0 < \chi_{ij} < 90$ is applied to obtain the efficiencies e_i and e_j, and the backgrounds b_{ij} and b_{ji} as a function of the cut χ_{ij} (see figures 4.12, 4.13 and 4.14).

In general, the ratio of the efficiency over background e_i/b_{ji} is a monotonically decreasing function of the cut χ_{ij} and conversely the ratio of the efficiency over background e_j/b_{ij} is a monotonically increasing function of the cut χ_{ij}. Then, in the general case, there exists a value of the cut variable χ_{ij} where both ratios take

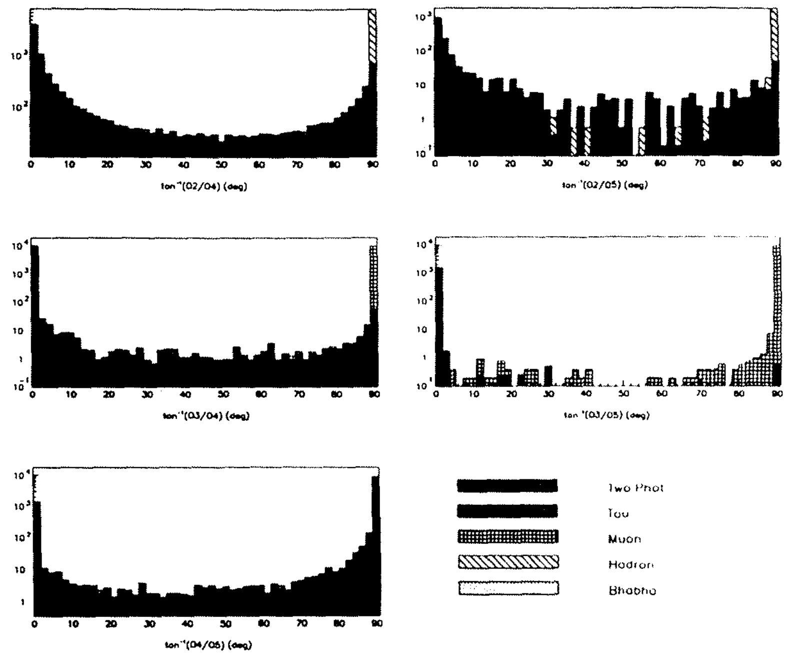


Figure 4.11: $Tan^{-1}(O_i/O_j)$ for $(i,j) = (2,4),(2,5),(3,4),(3,5),(4,5)$; in each case only the corresponding kinds i and j are plotted.

the same value. Therefore, to keep these ratios equated, the scan in cut variable should be performed in a interval around this point defined by the difference of those ratios. The relative difference of those ratios is

$$\Delta qeb_{ij} = \frac{|e_i/b_{ji} - e_j/b_{ij}|}{e_i/b_{ji} + e_j/b_{ij}} \tag{4.8}$$

The ten quantities Δqeb_{ij} are computed as the scan in cut variable is performed and they are shown in figures 4.12, 4.13 and 4.14. To keep all $\Delta qeb_{ij} < 0.6$ the ten variables χ_{ij} should be keep in the intervals defined in table 4.4. This is equivalent in a conventional analysis to avoiding cuts in distributions at places where the result will have a poor signal to noise ratio.

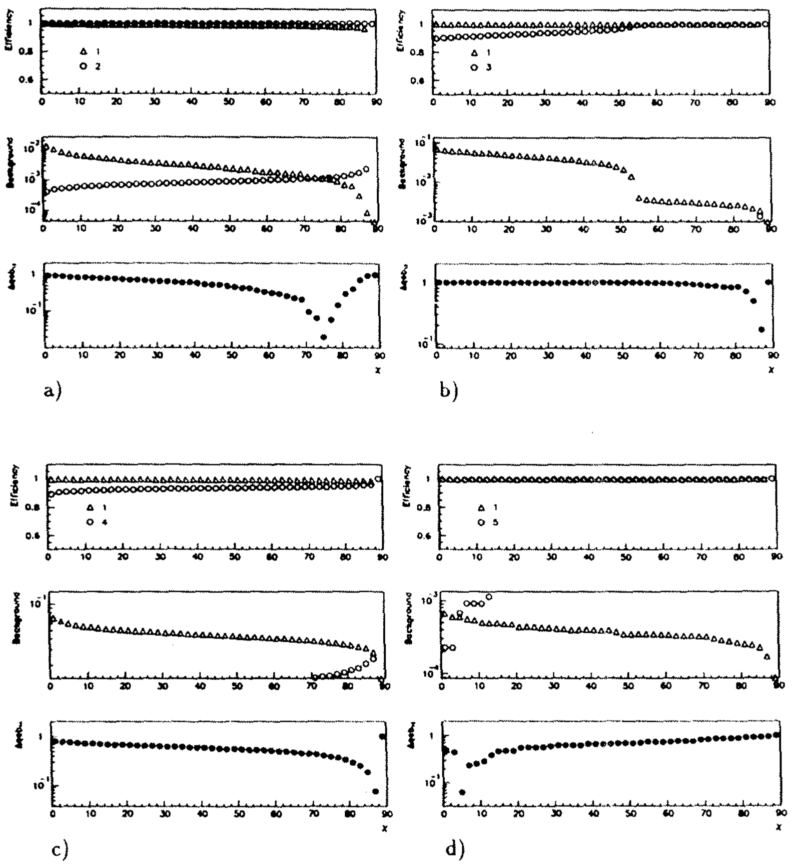


Figure 4.12: Binary Efficiency, Background and Δqeb_{ij} for
a) $i = 1$ $j = 2$;
b) $i = 1$ $j = 3$;
c) $i = 1$ $j = 4$;
d) $i = 1$ $j = 5$.

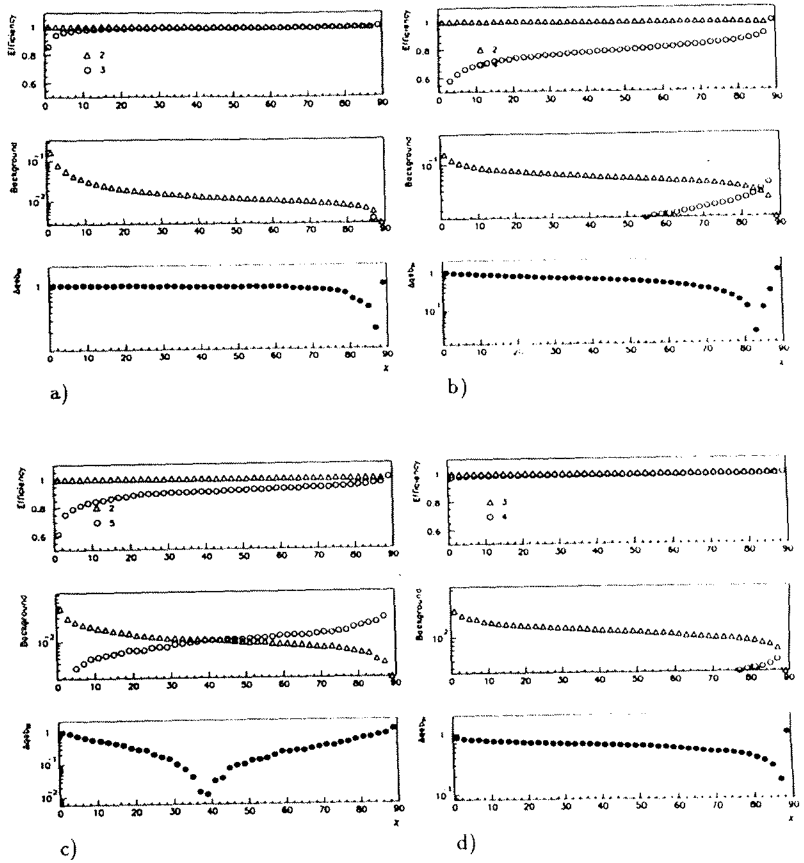
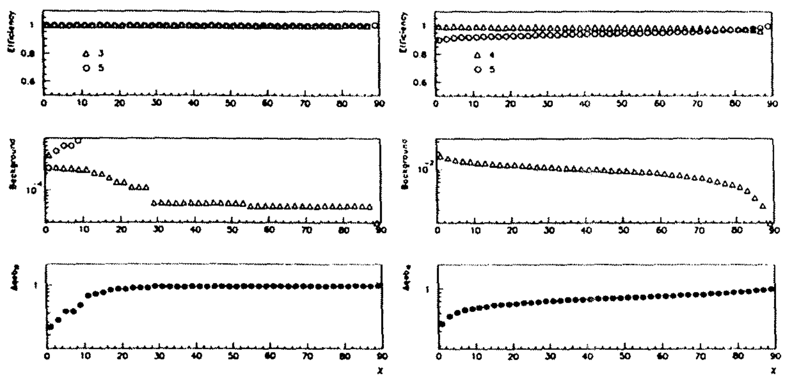


Figure 4.13: Binary Efficiency, Background and Δqeb_{ij} for
a) $i = 2$ $j = 3$;
b) $i = 2$ $j = 4$;
c) $i = 2$ $j = 5$;
d) $i = 3$ $j = 4$.



a) b)

Figure 4.14: Binary Efficiency, Background and Δqeb_{ij} for
a) $i = 3$ $j = 5$;
b) $i = 4$ $j = 5$.

	χ_{min} (deg)	χ_{max} (deg)
χ_{52}	16.0	74.0
χ_{21}	6.00	52.0
χ_{13}	70.0	89.0
χ_{34}	70.0	89.0
χ_{51}	1.72	77.3
χ_{53}	4.73	89.7
χ_{54}	12.8	89.9
χ_{23}	16.1	89.2
χ_{24}	38.4	89.9
χ_{14}	82.4	89.9

Table 4.4: Intervals of scanning on NN output space

One Kind, One Zone

The output space is a five-dimensional space, and according to definitions in previous sections, it can be divided in five zones. Let's take the plane o_i o_j, and define

a 4-dimensional object or hyper-plane $o_i = S.o_j$, which intersects it in a straight line with intercept equal to zero and any slope S. This object divides the plane o_i o_j in two semi-planes, while it divides the whole space in two semi-spaces. It is possible to have ten of these hyper-planes, one for each pair of coordinates, and they can be considered in groups of four to form the boundaries of a zone. By adding the five zones defined in this way, the whole 5-dimensional space is not necessarily recovered, and the intersection of any two of them is not necessarily empty, unless these hyper-planes are defined as follows: Let's take the straight line which passes through the origin of the 5-dimensional space and has direction cosines c_1, c_2, c_3, c_4 and c_5, and assume that it is the intersection of all ten hyper-planes. Then the equations $o_i = S.o_j$ have the solution $o_i = (c_i/c_j).o_j$. Therefore, we have a straight line in the 5-dimensional space, defined through the five cosine c_1, c_2, c_3, c_4, and c_5 which are not all independent since the relation,

$$\sum_i c_i = 1 \tag{4.9}$$

is satisfied. Then, only four degrees of freedom are available to vary the cut in output space.
Finally the five zones are defined by

- zone 1

$$\frac{o_1}{o_2} > \frac{c_1}{c_2},\quad \frac{o_1}{o_3} > \frac{c_1}{c_3},\quad \frac{o_1}{o_4} > \frac{c_1}{c_4},\quad \frac{o_1}{o_5} > \frac{c_1}{c_5}; \tag{4.10}$$

- zone 2

$$\frac{o_2}{o_3} > \frac{c_2}{c_3},\quad \frac{o_2}{o_4} > \frac{c_2}{c_4},\quad \frac{o_2}{o_5} > \frac{c_2}{c_5},\quad \frac{o_2}{o_1} \geq \frac{c_2}{c_1}; \tag{4.11}$$

- zone 3

$$\frac{o_3}{o_4} > \frac{c_3}{c_4},\quad \frac{o_3}{o_5} > \frac{c_3}{c_5},\quad \frac{o_3}{o_1} \geq \frac{c_3}{c_1},\quad \frac{o_3}{o_2} \geq \frac{c_3}{c_2}; \tag{4.12}$$

- zone 4

$$\frac{o_4}{o_5} > \frac{c_4}{c_5},\quad \frac{o_4}{o_1} \geq \frac{c_4}{c_1},\quad \frac{o_4}{o_2} \geq \frac{c_4}{c_2},\quad \frac{o_4}{o_3} \geq \frac{c_4}{c_3}; \tag{4.13}$$

- zone 5

$$\frac{o_5}{o_1} \geq \frac{c_5}{c_1}, \quad \frac{o_5}{o_2} \geq \frac{c_5}{c_2}, \quad \frac{o_5}{o_3} \geq \frac{c_5}{c_3}, \quad \frac{o_5}{o_4} \geq \frac{c_5}{c_4}. \tag{4.14}$$

Efficiency matrix and its inverse

Once a cut is defined in output space, it is possible to compute the efficiency matrix E_{ij} as the probability of having an event of kind j in zone i. A good approximation is the fraction of events j that fall in zone i, that is

$$E_{ij} = \frac{n_{ij}}{N_j} \tag{4.15}$$

where n_{ij} is the number of events of kind j that fall in zone i , while N_j is the total number of events of kind j. For kind 5 , the two-photon final state, we have six sub-kinds, and the probability of having an event of that kind is the weighted sum of probabilities for the six sub-kinds.

$$n_{i5} = \sum_k \frac{m_{ik}}{L_k} \tag{4.16}$$

where m_{ik} is the number of simulated events for sub-kind k that fall in zone i and L_k is the generated luminosity. The total number of events for kind 5 is

$$N_5 = \sum_k \frac{M_k}{L_k} \tag{4.17}$$

where M_k stands for the total numbers of events of sub-kind k.

Then we know the probability of any event of kind j falls in zone i, but we want to know the probability of any event in zone i belongs to kind j, this is related to the inverse of the efficiency matrix.
It is easy to see that the efficiency matrix satisfies the relation (see appendix B)

$$E\mathbf{k} = \mathbf{z} \tag{4.18}$$

where $\mathbf{k}$ is a five dimensional vector whose components are the number of events of each kind, while $\mathbf{z}$ is a five dimensional vector whose components are the number of events in each zone. If K is the space of $\mathbf{k}$ vectors and Z is the space of $\mathbf{z}$ vectors then

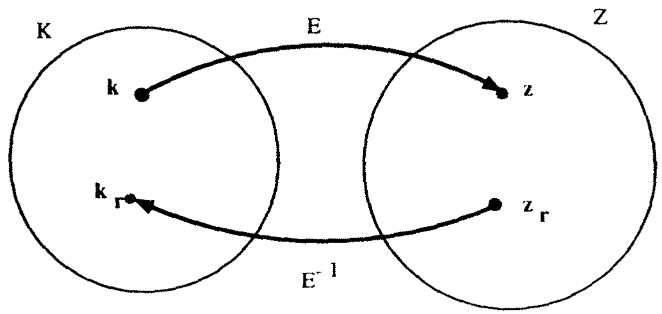


Figure 4.15: Linear transformation E,where the inverse E^{-1} transforms $\mathbf{z}_r$ into $\mathbf{k}_r$

E is a linear transformation that establishes a one to one correspondence between the spaces K and Z (see figure 4.1.4). This transformation is reversible via the inverse of the efficiency matrix E. Assuming that the simulation is accurate, the observer data which are classified into different zones to yield a vector $\mathbf{z}_r$ would imply a produced number of events given by

$$\mathbf{k}_r = E^{-1}\mathbf{z}_r \tag{4.19}$$

Best cut

If the distribution of simulated data in output space matched perfectly that of the real data, it would be unnecessary to select a cut, and in fact it would be totally unnecessary to train a neural net. But this is not the case with our simulated data. To select one of the infinite different cuts in output space we use the criteria of minimum systematic error. The systematic error can be reduced if we reduce the mutual background and if we increase the efficiencies for the different kinds.

The number of events for kind i in real data is

$$k_i = \sum_j E_{ij}^{-1} z_j \tag{4.20}$$

in order to define efficiency and background an analogy is used with the formula

$$k_i = \frac{z_i - b_i z_i}{e_i} \tag{4.21}$$

where $b_i z_i$ stands for the number of background events and e_i for the efficiency. In terms of the matrix E^{-1}, we have that the efficiency is

$$e_i = \frac{1}{E_{ii}^{-1}} \tag{4.22}$$

while the fractional background is

$$b_i = -\frac{1}{E_{ii}^{-1} z_i} \sum_{j \neq i} E_{ij}^{-1} z_j. \tag{4.23}$$

We now define the relative difference of efficiency over background for a pair of final states as

$$\Delta eob_{ij} = \frac{\left|\frac{e_i}{b_i} - \frac{e_j}{b_j}\right|}{\frac{e_i}{b_i} + \frac{e_j}{b_j}}. \tag{4.24}$$

Note that it is not possible to compute this quantity before measuring z_i.

Finally, the cut which minimizes the quantity

$$\prod_{i<j} \Delta eob_{ij} \tag{4.25}$$

has been chosen as the best cut.

The distributions of the two input variables which are the two most correlated with each output are shown in figures 4.16 through 4.23.

Each figure shows five distributions for each variable, corresponding to the events that fall into the five zones defined by the best cut. It is evident that the NN classification is highly accurate.

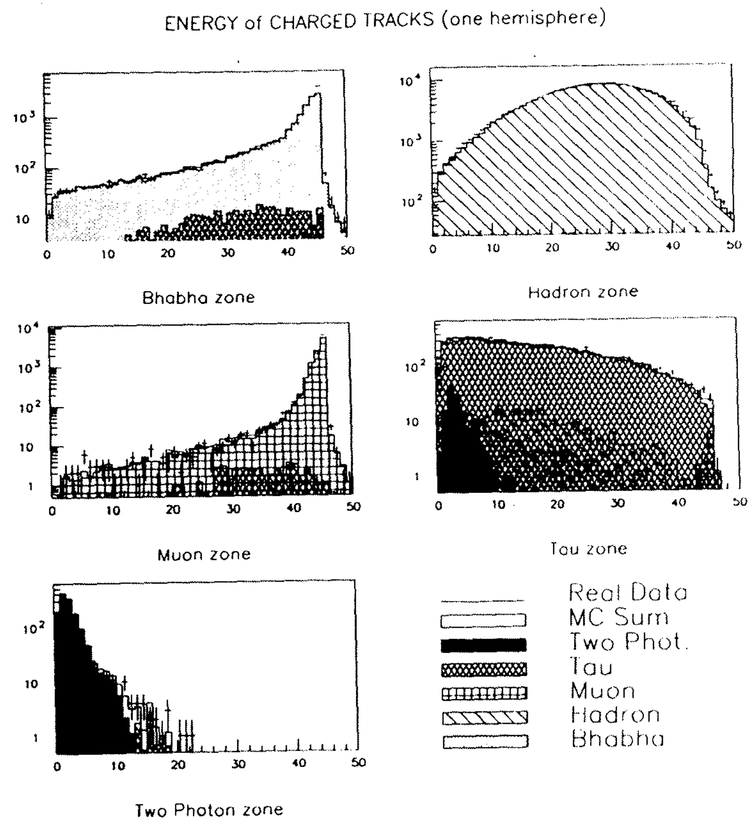


Figure 4.16: Distribution of $E_{ch,h1}$ for each zone of output space. This variable is the most correlated with the output related to tau events, and it is the second most correlated with the outputs related to bhabha and muon events.

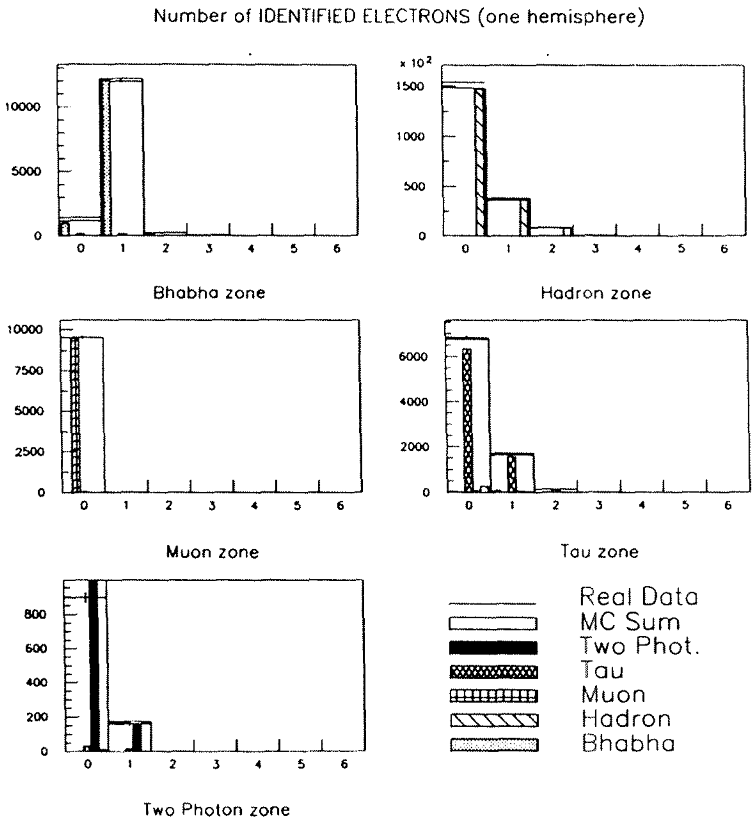


Figure 4.17: Distribution of $n_{e,h1}$ for each zone of output space. This variable is the most correlated with the output related to bhabha events.

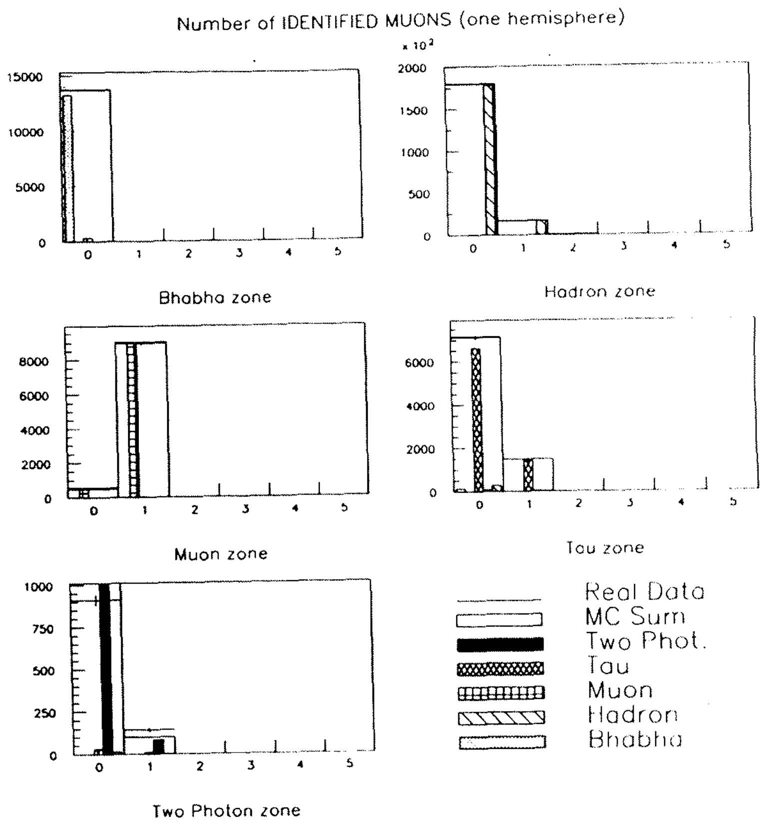


Figure 4.18: Distribution of $n_{\mu,h1}$ for each zone of output space. This variable is the most correlated with the output related to muon events.

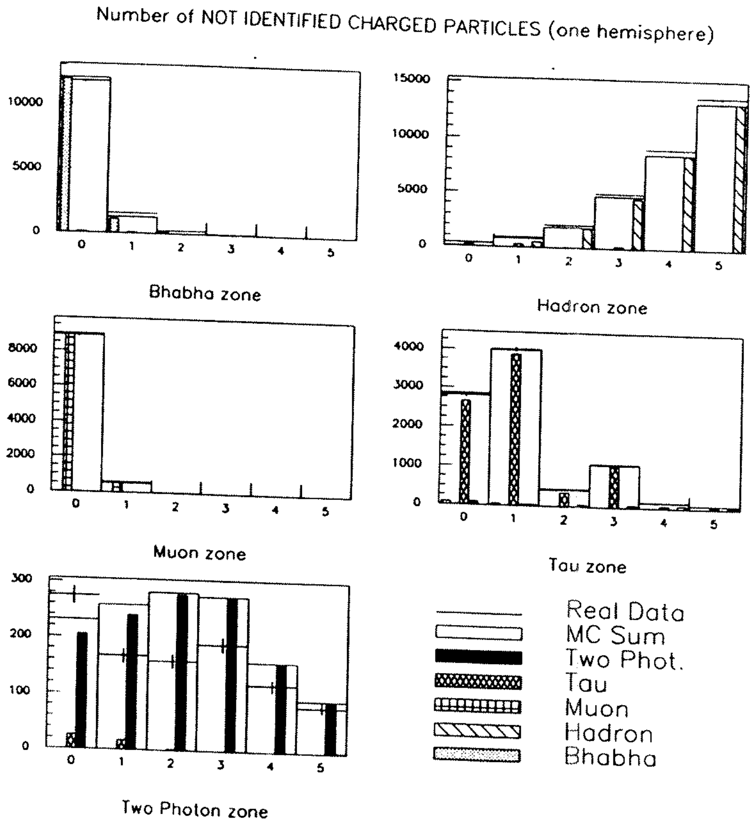


Figure 4.19: Distribution of $n_{ch,h1}$ for each zone of output space. This variable is the most correlated one with the output related to hadron events.

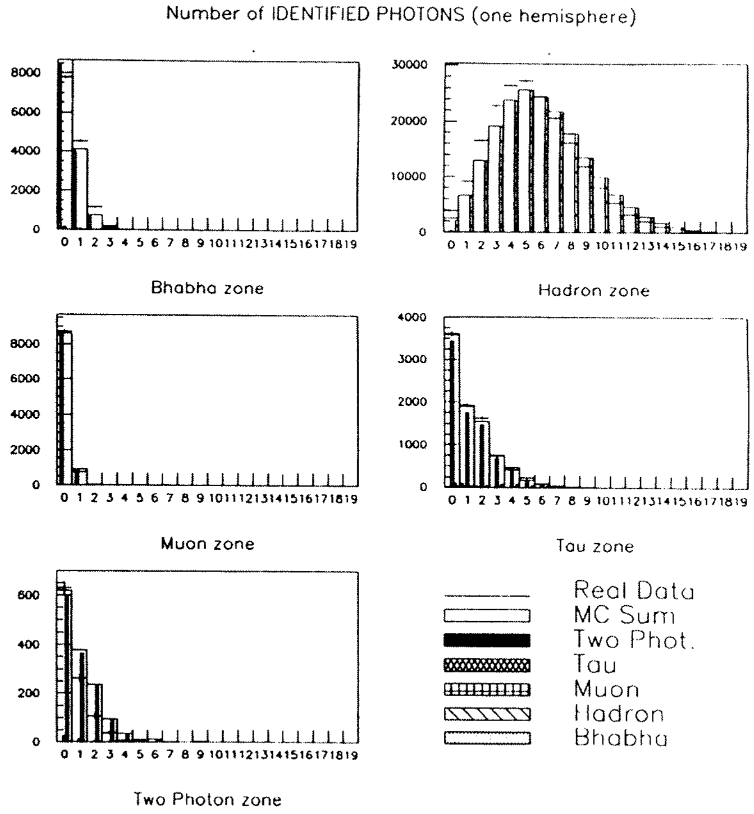


Figure 4.20: Distribution of $n_{\gamma,h1}$ for each zone of output space. This variable is the second most correlated with the output related to hadron events.

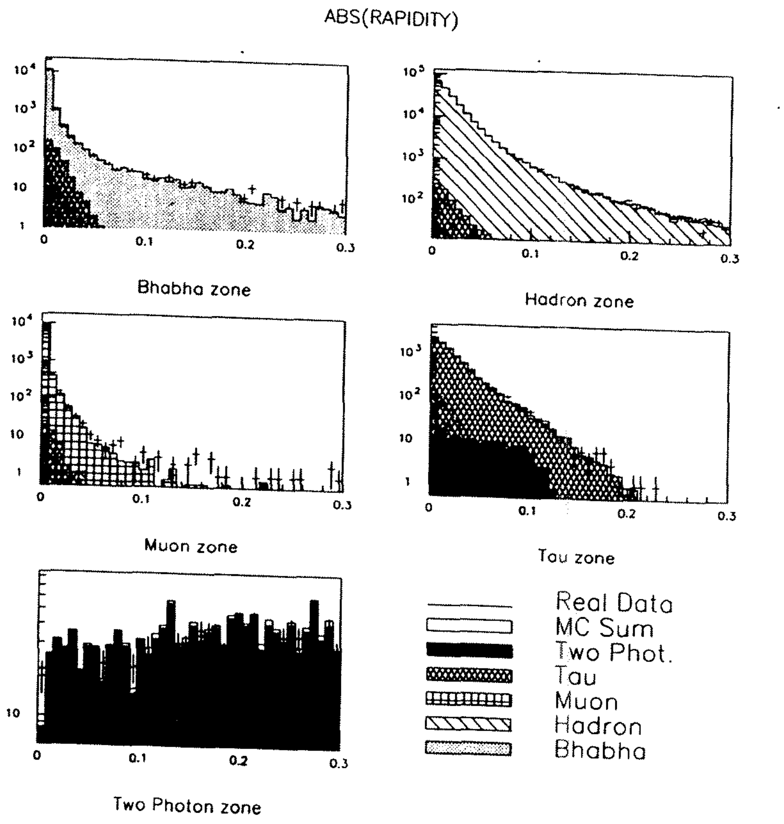


Figure 4.21: Distribution of L for each zone of output space. This variable is the most correlated with the output related to two photon events.

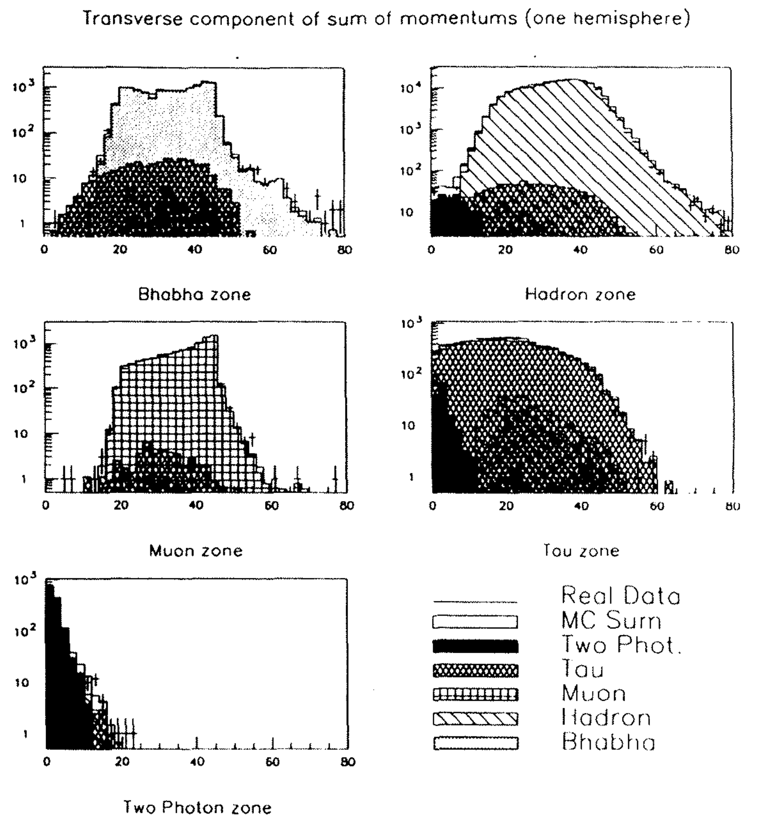


Figure 4.22: Distribution of $P_{t,h1}$ for each zone of output space. This variable is the second most correlated with the output related to two-photon events.

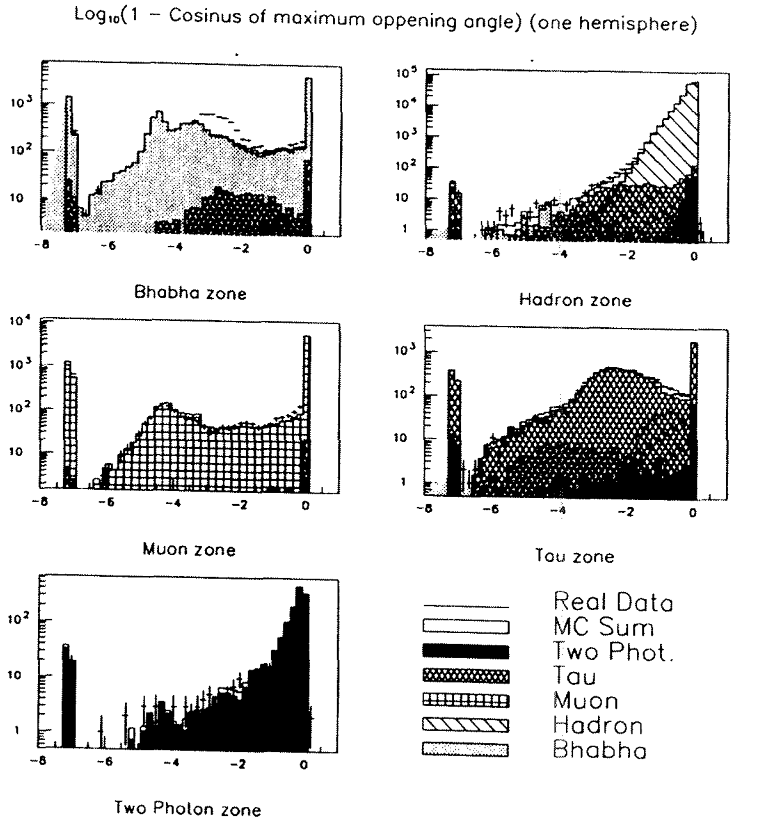


Figure 4.23: Distribution of $cos\theta_{mopa,hi}$ for each zone of output space, amplified around $cos\theta_{mopa,hi} = 1$ by plotting $Log_{10}(1 - cos\theta_{mopa,hi})$. This variable is the second most correlated with the output related to tau events.

4.1.5 Results of NN selection

The efficiency matrix which corresponds to the best cut is

$$E = \begin{pmatrix} 0.9822 & 0.0002 & 0.0029 & 0.0376 & 0.0000 \\ 0.0072 & 0.9980 & 0.0046 & 0.0876 & 0.1181 \\ 0.0001 & 0.0000 & 0.9909 & 0.0049 & 0.0000 \\ 0.0105 & 0.0017 & 0.0015 & 0.8650 & 0.0054 \\ 0.0000 & 0.0001 & 0.0000 & 0.0049 & 0.8766 \end{pmatrix}$$

Then its inverse is

$$E^{-1} = \begin{pmatrix} 1.0186 & -0.0002 & -0.0029 & -0.0442 & 0.0003 \\ -0.0063 & 1.0022 & -0.0045 & -0.1004 & -0.1344 \\ -0.0000 & -0.0000 & 1.0091 & -0.0057 & 0.0000 \\ -0.0123 & -0.0019 & -0.0017 & 1.1569 & -0.0068 \\ 0.0001 & -0.0001 & 0.0000 & -0.0065 & 1.1409 \end{pmatrix}$$

For comparison purposes, the data used for the analysis were chosen to be the set of all 1991 runs at the Z_0 peak, selected by the ALEPH electroweak working group. This ensures an equal data set to start with for all the electroweak cross sections. Out of the 7.9 pb^{-1} of the runs qualified as PERFect and MAYBe by the ALEPH run quality group taken in 1991, 7.5 pb^{-1} have been selected. The main reasons for the rejection of 0.4 pb^{-1} have been DAQ problems, failures of TPC sectors, and failures of the LCAL event builder (bad luminosity).

For all events, the XLUMOK flag is required to be set. This means proper functioning of all parts of ECAL, both sides of LCAL, the TPC (only as far as it is used for the tracking, not dE/dx), the ITC, all parts of HCAL and the trigger. In particular, the single muon trigger, the single charged electromagnetic trigger, the triggers on the ECAL energy in the barrel, in either endcap, in both endcaps, and the LCAL tower coincidence triggers, are required to be enabled.

After applying the cuts defined in section 4.1.3, the set of real data is composed of 233420 events, which are classified by the Neural Net in the five zones, resulting in the number of events in each zone

$$\mathbf{z} = \begin{pmatrix} 13787 \\ 200377 \\ 9566 \\ 8635 \\ 1055 \end{pmatrix}$$

Finally applying equation (4.19), the corrected number of events is obtained.

$$\mathbf{k} = \begin{pmatrix} 13602.7 \\ 199676.0 \\ 9597.88 \\ 9410.66 \\ 1132.76 \end{pmatrix}$$

The vector k is the best estimator of the number of events of each final state kind which fall within the cuts described in section 4.1.3. The acceptance of each process for these cuts is slightly different, and is listed in table 4.5. It should be noted that for the measurement of the observable Q, only the difference in the acceptance of two processes is relevant. Also, the acceptance quoted for Bhabha events is arbitrarily chosen to correspond to the angular range generated in the simulation.

event kind	Number of events	acceptance	Acc. corr. events
Electrons	13602.7	0.5578	24384.7
Hadrons	199676.0	0.8635	231239.6
Muons	9597.88	0.8504	11286.0
Taus	9410.66	0.8335	11290.5

Table 4.5: Number of events corrected by acceptance

4.2 Statistical errors

The statistical error of Q is computed by assuming that the outcome of an e^+e^- collision inside the acceptance, falls in one of the five zones defined above. Therefore, the distribution of events obeys a five-nomial distribution. If n_i is the number of events in zone i, the probability of having the set $n_i : i = 1, ..., 5$ is given by

$$P(n_1, n_2, ..., n_5) = \frac{N!}{n_1! n_2! n_3! n_4! n_5!} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4} p_5^{n_5}. \tag{4.26}$$

Assuming that the total number of events is not a random variable, the covariance matrix of the number of events n_i and n_j is given by

$$cov(n_i, n_j) = \begin{cases} -Np_i p_j & \text{if } i \neq j \\ Np_i(1 - p_i) & \text{if } i = j \end{cases} \tag{4.27}$$

were N is the total number of events. The variance of Q as a function of the numbers n_i, $\sigma(Q(n_i))$, is computed using the approximate formula

$$\sigma^2(Q) = \sum_{i,j} \left.\frac{\partial Q}{\partial n_i}\right|_{n_i=\mu_i} \left.\frac{\partial Q}{\partial n_j}\right|_{n_j=\mu_j} cov(n_i, n_j) \tag{4.28}$$

where μ corresponds to the number of events actually observed. To compute the statistical error of the number of hadronic events, N_{had}, it has been assumed that the quantities z_i are distributed following a Poisson distribution; since in this case the assumption that the total number of events inside the acceptance is a random variable is not valid, then the variance is z_i.

The statistical errors are given in table 4.6

Quantity	Statistical error
Q_μ	0.2157
Q_τ	0.2397
N_{had}	447.6

Table 4.6: Statistical errors

4.3 Systematic errors

The study of systematic errors is divided in two sections. The first one studies errors introduced by the finite statistics of the simulated data sample. The second explores the effects of miss-simulations of the input variable distributions.

4.3.1 Statistics in simulated data

The systematic error related to statistics in simulated data has been computed in the same way as the statistical error of the measurement. (see section 4.2). The events from each kind of simulated data are assumed to be distributed six-nomially, (eq. 4.26) with the sixth type of event being those that fall outside the acceptance. Therefore, the covariance matrix for the kind k of simulated data is

$$cov(n_{ik}, n_{jk}) = \begin{cases} -Np_{ik}p_{jk} & \text{if } i \neq j \\ Np_{ik}(1-p_{ik}) & \text{if } i = j \end{cases} \tag{4.29}$$

where N is the number of generated events for that kind and p_{ik} is the probability of falling in zone i.
In this case, the partial derivatives of eq. (4.28) have been computed by numerical methods.

4.3.2 Input variable distributions

Comparison of the distributions of input variables for real and simulated data shows that their shapes are well reproduced, but some of their mean values disagree. Therefore, the systematic errors coming from miss-simulations are estimated by shifting the distributions of input variables in simulated data and observing the effect on the value of Q. To compute the amount of shift for an input variable for a kind of simulated data, three steps are followed: First, the mean value of the distribution in the zone belonging to that kind is computed; second, the mean value for real data in the same zone is computed; finally, the distribution of the variable is shifted by the difference between real an simulated data. The shifts applied are summarized in table 4.7, where the following remarks have to be made:

- when shifting E_{ch}, if $n_e + n_\mu + n_{ch} \leq 3$ for an hemisphere then, if $E_{ch} > E_{beam}$ then $E_{ch}$is set equal to E_{beam}, as the Energy flow algorithm saturates the energy of a given track when it is above the beam energy.
- $\Delta n_{ne} = -\Delta n_\gamma$, since the total number of neutral tracks is well simulated and therefore the shift has to be made keeping this number unchanged in order to not overestimate the related systematic error.
- Systematic errors due to variables that become irrelevant in the NN analysis, like n_{l_u} or E_{l_u} has not been quoted.
- For the variables $cos\theta_{h1}$ and $cos\theta_{h2}$ the systematic errors have been estimated by smearing the distributions , i.e. they have been convoluted with a gaussian distribution with a standard deviation of 7×10^{-2} which would be the resolution in $cos\theta_{hi}$.

The systematic errors are summarized in table 4.8, where the error indicated on the first line is due to the finiteness of the simulated data set. Finally, the systematic errors were added quadratically, keeping separate those which produce positive and negative changes in Q. The error coming from the statistics of the simulated data was added on both sides.

<table>
<tr><th rowspan="2">Variable</th><th colspan="5">Shift</th></tr>
<tr><th>bhabha</th><th>hadron</th><th>muon</th><th>tau</th><th>two phot</th></tr>
<tr><td>$E_{ch,h1}$</td><td>0.61</td><td>0.22</td><td>0.16</td><td>0.19</td><td>0.27</td></tr>
<tr><td>$E_{ch,h2}$</td><td>0.61</td><td>0.22</td><td>0.16</td><td>0.19</td><td>0.27</td></tr>
<tr><td>$E_{ne,h1}$</td><td>-0.703</td><td>-1.08</td><td>-0.068</td><td>-0.397</td><td>-0.93</td></tr>
<tr><td>$E_{ne,h2}$</td><td>-0.703</td><td>-1.08</td><td>-0.068</td><td>-0.397</td><td>-0.93</td></tr>
<tr><td>$n_{e,h1}$</td><td colspan="5">-0.0186</td></tr>
<tr><td>$n_{e,h2}$</td><td colspan="5">-0.0186</td></tr>
<tr><td>$n_{\mu,h1}$</td><td colspan="5">9.5×10^{-3}</td></tr>
<tr><td>$n_{\mu,h2}$</td><td colspan="5">9.5×10^{-3}</td></tr>
<tr><td>$n_{ch,h1}$</td><td colspan="5">5×10^{-3}</td></tr>
<tr><td>$n_{ch,h2}$</td><td colspan="5">5×10^{-3}</td></tr>
<tr><td>$n_{\gamma,h1}$</td><td>0.1192</td><td>-0.541</td><td>-7.8e-3</td><td>-0.058</td><td>-0.362</td></tr>
<tr><td>$n_{\gamma,h2}$</td><td>0.1192</td><td>-0.541</td><td>-7.8e-3</td><td>-0.058</td><td>-0.362</td></tr>
<tr><td>$n_{ne,h1}$</td><td>0.0538</td><td>0.43</td><td>0.0146</td><td>0.077</td><td>-0.106</td></tr>
<tr><td>$n_{ne,h2}$</td><td>0.0538</td><td>0.43</td><td>0.0146</td><td>0.077</td><td>-0.106</td></tr>
<tr><td>L</td><td colspan="5">-4×10^{-4}</td></tr>
<tr><td>$P_{t,h1}$</td><td></td><td>-.4</td><td></td><td></td><td></td></tr>
<tr><td>$P_{t,h2}$</td><td></td><td>-.4</td><td></td><td></td><td></td></tr>
<tr><td>Y</td><td colspan="5">3.7×10^{-3}</td></tr>
<tr><td>M</td><td colspan="5">-3.8×10^{-3}</td></tr>
<tr><td>N</td><td colspan="5">-1.19×10^{-3}</td></tr>
<tr><td>S_{h1}</td><td>4.85e-3</td><td colspan="4">-6.2×10^{-3}</td></tr>
<tr><td>S_{h2}</td><td>4.85e-3</td><td colspan="4">-6.2×10^{-3}</td></tr>
<tr><td>$cos\theta_{mopa,h1}$</td><td colspan="5">1.5×10^{-3}</td></tr>
<tr><td>$cos\theta_{mopa,h2}$</td><td colspan="5">1.5×10^{-3}</td></tr>
<tr><td>$\Delta\phi$</td><td colspan="5">-0.225</td></tr>
<tr><td>W^2_{h1}</td><td colspan="5">-14</td></tr>
<tr><td>W^2_{h2}</td><td colspan="5">-14</td></tr>
</table>

Table 4.7: Shifts applied to input variables to compute systematics errors

Variable	ΔQ_μ	ΔQ_τ	ΔN_{had}
stat mc	0.01284	0.05844	29.33
$E_{ch,h1}$	-0.00096	0.05945	15.95
$E_{ch,h2}$	0.00152	0.09898	33.08
$E_{ne,h1}$	0.00029	0.00759	3.70
$E_{ne,h2}$	-0.00509	-0.11318	-49.17
$n_{e,h1}$	-0.00018	0.01816	-1.09
$n_{e,h2}$	-0.00029	0.01208	-2.66
$n_{\mu,h1}$	0.02257	0.00493	-3.97
$n_{\mu,h2}$	0.02086	-0.00123	-4.42
$n_{ch,h1}$	-0.00008	-0.00156	-1.00
$n_{ch,h2}$	-0.00002	-0.00002	-0.25
$n_{\gamma,h1}$	0.00244	0.01460	14.02
$n_{\gamma,h2}$	-0.00091	0.0112	4.58
$n_{ne,h1}$	0.00023	0.01271	3.25
$n_{ne,h2}$	0.00014	0.00841	2.08
L	-0.00119	-0.01182	-4.48
$cos\theta_{h1}$	0.00000	0.00000	0.02
$cos\theta_{h2}$	0.00000	0.00000	0.02
$P_{t,h1}$	-0.00072	-0.01402	-8.83
$P_{t,h2}$	-0.00019	-0.00165	-2.19
Y	-0.00182	-0.02135	-12.17
M	-0.00257	-0.03501	-21.28
N	-0.00213	-0.02705	-15.89
S_{h1}	-0.00365	0.00657	-2.77
S_{h2}	-0.00197	-0.00704	-8.39
$cos\theta_{mopa,h1}$	-0.00083	-0.00735	-4.94
$cos\theta_{mopa,h2}$	-0.00073	-0.01330	-8.94
$\Delta\phi$	-0.00090	0.03749	10.88
W^2_{h1}	-0.00141	-0.00878	-6.78
W^2_{h2}	-0.00106	-0.00999	-2.95
Total+	0.03343	0.13901	50.72
Total−	-0.01521	-0.13942	-67.20

Table 4.8: Systematics errors. Total+ is the quadratic combination of positive errors, while Total− is the quadratic combination of negative errors.

4.4 Results Q_μ, Q_τ and N_{had}

In summary, the corrected number of events and the estimated acceptance for each final state (table 4.5), together with the statistical and systematic error estimates (table 4.6 and 4.8), yield the experimental results:

$$Q_\mu = 20.489 \pm 0.216(stat.) + 0.034 - 0.015(syst.)$$

$$Q_\tau = 20.481 \pm 0.240(stat.) + 0.127 - 0.128(syst.)$$

$$N_{had} = 231239.6 \pm 447.6(stat.) + 50.7 - 67.2(syst.)$$

The luminosity measuered for this this set of data is $L = 7546.4 \pm 16.9 nb^{-1}$ then the total cross section for hadron events is

$$\sigma_h = 30.64 \pm 0.06(stat.) + 0.01 - 0.01(syst.) \pm 0.07(lumi.)$$

Chapter 5

Results and Conclusions

Three different quantities have been measured in this work; two of them are Q_μ and Q_τ defined by eqs. (2.1) and (2.2), while the third one is the number of hadronic events N_{had}, which has been combined with the measured luminosity to give the hadronic cross section σ_h. The luminosity of $7.5 pb^{-1}$ corresponds to the 1991 period of ALEPH experiment.

The experimental results are summarized in table 5.1

Q_μ	$20.489 \pm 0.216(stat.) + 0.034 - 0.015(syst.)$
Q_τ	$20.481 \pm 0.240(stat.) + 0.127 - 0.128(syst.)$
$\sigma_h(nb)$	$30.64 \pm 0.06(stat.) + 0.01 - 0.01(syst.) \pm 0.07(lumi.)$

Table 5.1: Experimental results for Q_μ, Q_τ and σ_h

5.1 On R

The theoretical formulas for Q_μ , Q_τ and σ_h given in section 2.3 have been fitted to the experimental values by varying the four parameters M_Z, Γ_Z, σ_h^0 and R. The parameters M_Z and Γ_Z have been very loosely constrained. The results from the fit are summarized in table 5.2 . The confidence level is $C.L. = 51\%$. As explained in section 2.3.3, the sensitivity to M_Z and Γ_Z is practically null. The measurement of σ_h is used to constrain σ_h^0, which has a small correlation to R. As expected, the value of R is well determined by the measurement.

	$M_Z(GeV)$	$\Gamma_Z(GeV)$	$\sigma_h^0(nb)$	R
Estimation	91.0 ± 10.0	2.5 ± 2.4	41.0 ± 10.0	
Fit	91.2 ± 2.5	2.5 ± 2.0	41.6 ± 6.7	$20.68 \pm 0.16(stat) \pm 0.06(syst.)$

Table 5.2: R from fit of Q_μ Q_τ and σ_h

5.2 On α_s

To obtain α_s from the simultaneous fit of Q_μ, Q_τ and σ_h, we have replaced R for its development in terms of α_s/π. The quotient of the hadronic partial width without QCD corrections over the leptonic partial width, R_0, has been assumed, then α_s can be determined with an error $\Delta\alpha_s \approx \pi\Delta R/R$

Taking $R_0 = 19.943$ (ref. [8]) the results from the fit are summarized in table 5.3. The confidence level is $C.L. = 51\%$

	$M_Z(GeV)$	$\Gamma_Z(GeV)$	$\sigma_h^0(nb)$	α_s
Estimation	91.0 ± 10.0	2.5 ± 2.4	41.0 ± 10.0	
Fit	91.2 ± 1.3	2.5 ± 0.6	41.6 ± 5.3	$0.109 \pm 0.023(stat) \pm 0.004(syst)$

Table 5.3: α_s from fit of Q_μ Q_τ and σ_h

5.3 Conclusions

A new method for classifying final states in the process $e^+e^- \to Z_0 \to f\bar{f}$, where f is a charged fermion, has been developed. The method simultaneously maximizes the efficiencies and minimizes the background for all the species. In addition, the Neural Net can utilize information from all the ALEPH subdetectors and also exploit their correlations. The method yields measurements with small systematic errors.

The results presented for the 1991 ALEPH data set compare well with the results of the published traditional analysis. It is desirable to compare the value of α_s obtained from R to those from event shape analyses, at precisions of 10%. The low systematics of the method presented will allow this comparison to be performed when a data sample of about four million Z_0 decays becomes available in the near future.

Appendix A

Learn and Grow Neural Net training algorithm

Neural Networks have become a very useful tool for event classification and more generaly, in pattern recognition problems where the number of variables to consider is large. The "back propagation" method of training has become popular to solve these kinds of problems. In High Energy Physics, especially in the last decade, the use of Neural Nets has increased, from practically nothing until several dozen data analyses published per year. This work consists in the development of a new training method based on the "back propagation" method, which has been modified in order to have an independent temperature for each neuron. In addition, an alghorithm of growth has been developed (i.e. the change of the arquitecture during the training proccess). Since the training is stochastic (i.e. the parameters are modified for each event showed to the NN), the mix of the different kinds of events to be classified is performed on-line, in order to optimize the NN to the actual proportions of each kind that will be found in real data, as well as to increase the amount of events for the kinds that are least well known for the NN.

The algorithm and the special features described above has been coded in a program called Master In Neuron Developement (MIND).

A.1 Neuron response and NN Arquitecture

A Neural Net is a set of formally identical functions (neurons) whose arguments are linear combinations of the values of the other functions (axons). The coefficients of these linear combinations are called weights (activation level of the axon), the value of each function or activation function is called the neuron output while its argument is the input. In the arquitecture of "forward feed net" the information flows in only one way and the net topology is such that none of the neurons are connected in such a way that its output is reflected in its input, or in other words the value of one function is independent of itself.

The NN is composed of several layers. If the information flow is followed, the first one is the input layer, where the number of neurons is equal to the number of variables. The last layer is the output layer, and its number of neurons is equal to the number of kinds to distinguish. In between are found the hidden layers, whose quantity and number of neurons can be varied. In order to determine the hidden layer topology, an algorithm of "learn-and-grow" has been developed.

A graphical representation of a feed forward Neural Net can be seen in figure A.1. Eeach neuron is represented by a circle, and its connections are represented by bars with widths proportional to the weight of the connection.

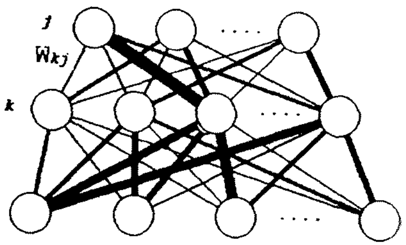


Figure A.1: Grafical representation of a Neural Net. The circles represent the neuron, while the straigth lines represent the weigths W_{kj}. Each neuron has associated a temperature which is the inverse of the parameter β and a chemical potential μ.

The activation function has to be limited by two assymptotes or saturation values, typically 1 and 0, and it has to have a range where the argument is aproximately linear, which is called the sensitivity range. A function that fulfills these requirements is the sigmoid function

(A.1) $$O(I) = \frac{1}{1 + e^{\beta(I-\mu)}}$$

where the input I is the linear combination of previous layer outputs using weights W_{ij}. Therefore, the input for neuron i is

(A.2) $$I_i = \sum_j W_{ij}.O_j$$

μ is called the "chemical potencial" in analogy with the distribution in energy of occupation numbers in a Fermi-Dirac gas, and β is called the inverse of the temperature for the same reason (see figure A.2).

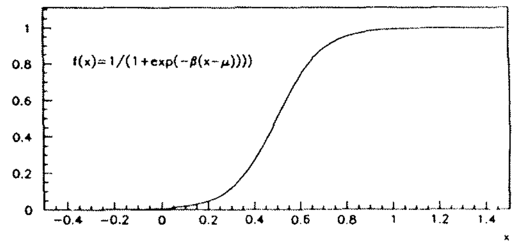


Figure A.2: Sigmoid function with $\beta = 10$ and $\mu = 0.5$.

The Net training process, consists of varying the parameters W_{ij}, β and μ in order to minimize the energy of the sample, defined as the sum of the individual event energies

(A.3) $$E = \frac{1}{2} \sum_{sample} \sum_{j=1}^{N_k} (O_j - d_j)^2$$

where j runs from 1 to N_k, being N_k the number of neurons in the output layer. O_j is the output value of the output neuron j, while d_j is the desired output for the same neuron. For an event of kind i, d_j takes the value $d_j = \delta_{ij}$ (i.e. $d_j = 1$ if $i = j$, and $d_j = 0$ if $i \neq j$).

A.2 Extended Back Propagation method

A stochastic method is used to minimize the energy of the sample. It consists of varying the Net parameters W_{ij}, β_j and μ_j event by event, along a direction opposite to the energy gradient. Given eqs. (A.1) and (A.2), the output O_i of the neuron i is a function of the weights W_{ji} and the parameters β_i and μ_i

(A.4) $$O_i = O_i\left(\beta_i(I_i - \mu_i)\right)$$

Then the energy of the sample is a functional,

(A.5) $$E = E(W_{jk}, \beta_k, \mu_k; I_{n_i}, d_{n_o})$$

where I_{n_i} is the value of the input variable number n_i while d_{n_o} is the dessired value of the output neuron number n_o. The components of the energy gradient are given by

(A.6) $$\begin{aligned} \frac{\partial E}{\partial W_{lk}} &= \Delta_k \frac{\partial O_k}{\partial a_k} \beta_k O_l \\ \frac{\partial E}{\partial \mu_k} &= -\Delta_k \frac{\partial O_k}{\partial a_k} \beta_k \\ \frac{\partial E}{\partial \beta_k} &= \Delta_k \frac{\partial O_k}{\partial a_k} (I_k - \mu_k), \end{aligned}$$

where

(A.7) $$\Delta_k = \begin{cases} O_k - d_k & \text{if } k \text{ is an output neuron,} \\ \sum_m W_{km} \beta_m \frac{\partial O_m}{\partial a_m} \Delta_m & \text{otherwise.} \end{cases}$$

and

(A.8) $$a_k = \beta_k (I_k - \mu_k)$$

Note that these equations are valid for the family of functions whose argument are a combination of β_k, I_k, μ_k, like in a_k. In particular if we choose the sigmoid (A.1) as the activation function, its derivatives are given by

(A.9) $$\frac{\partial O_m}{\partial a_m} = O_m (1 - O_m).$$

A.3 Taking a step

The gradient A.6 is computed for each event and the changes in Net parameters for event i are given by

(A.10) $$(\Delta W_{lk})_i = s \frac{\frac{\partial E}{\partial W_{lk}}}{|\nabla E|} + I (\Delta W_{lk})_{i-1}$$

(A.11) $$(\Delta \mu_k)_i = s \frac{\frac{\partial E}{\partial \mu_k}}{|\nabla E|} + I (\Delta \mu_k)_{i-1}$$

(A.12) $$(\Delta \beta_k)_i = s \frac{\frac{\partial E}{\partial \beta_k}}{|\nabla E|} + I (\Delta \beta_k)_{i-1}$$

The step s is computed for event i by the equation

(A.13) $$s = s_0 \frac{E_i}{\langle E \rangle} \frac{1}{|\nabla E|}$$

where E_i is the event energy and $\langle E \rangle$ is the sample mean energy.

The inertia I is is the fraction of the previous step that is being vectorialy added, which is given by

(A.14) $$I = I_0 \frac{|E_i - \langle E \rangle|}{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}$$

A.4 Simulated data treatement

A.4.1 rescaling

The input variables are transformed in order to have a range in the interval $[0, 1]$ for the input neurons. The transformation for input variable x is

(A.15) $$x' = \frac{x - \langle x \rangle}{6\sigma_x} + \frac{1}{2}$$

with

(A.16) $$\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$$

where the mean values $\langle x \rangle$ and $\langle x^2 \rangle$ are taken over the full sample. This transformation maps the interval $[-3\sigma_x, +3\sigma_x]$ to the interval $[0, 1]$

A.4.2 On-line mixing

The sample of simulated data is constituted for events of different kinds but not necessarily in the proportion that would be found in real data. The mixing of these kinds is performed during the training process. This feature allows to modify the proportions to optimize the final parameters. The mixing is performed to have the maximum eficiency for each kind and the minimum background (see next section).

A.4.3 sampling the sample

The parameters of the minimization process (s_0, l_0) have to be changed, when the minimization problem changes its scale. The input sample is divided in sets of events, which have to be large enough to have statistical significance, but they small enough to allow a change on minimization parameters to prevent divergences. That set of events constitutes a Loop.

A.5 On-line Efficiency and Background

During the minimization process, efficiences and backgrounds can be computed assuming a "central" cut, meaning that an event is classified as belonging to a kind when the corresponding output is the largest one. Then an efficiency matrix is obtained which has to be corrected by factors to have the estimated proportions in real data. This permits to compute efficiencies and backgrounds that one would have if the sub-sample had the actual proportions. Therefore we have a quantity to evaluate the qualily of the classification, that is the quotient of eficiency over background. Depending on this quality, the on-line mixing is modified to increase the number of events of each kind inversely proportionally to that quotient.

A.6 Slope as index of activity

At the end of each Loop, the energy mean value of each kind is computed, and the mean value of these numbers is calculated to yield a global mean value. Then a straight line is fitted over previous global mean values as a function of the number of loops. Therefore the slope of this straight line is considered an index of variation,

which is used to modify the parameters of the minimization, or to decide on a mutation as will be described in the next section.

A.7 Changing minimization parameters

The minimization parameters are modified depending on the slope. When the slope is negative but inside a given range, s_0 is increased by 10%; when the slope is positive and inside that range, s_0 is unchanged but when the slope is above that range s_0 is decreased by 10%.

A.8 Mutation of a NN

With a given arquitecture, it is possible to reach a certain level of separation of the diferent kinds. Once that point has been reached it is desirable to continue with a new arquitecture, normally the most simple extension, without loosing what has been previously learned. The parameter that controls the mutation of the NN is the slope h defined previously. The conditions to mutate are fulfilled when the absolute value of the slope falls below a threshold (typically :10^{-5}). When the NN has only two layer (i.e. input an output) the mutation increases the NN in one layers, adding a hidden layer with the same number of neurons as

the output layer. In turn, when the NN has three or more layers, the mutation only increases in one neuron one of the hidden layers, begining with the layer closest to the output. If the previous layer has the same number of neurons, the mutation modifies that layer instead. In others words, going from the input to the output, the Net never has two consecutive layers one larger than the other, with the exception of the input layer in relation to the first hidden layer. When the mutation results in adding a layer, the old output layer becomes the last hidden one, and a one to one correspondence is set between neurons of these two layers. Therefore, the new weights and β parameter are set to one if they connect two corresponding neurons, while μ is set to .5; the rest of the new weights are set to zero. When the mutation results in adding one neuron, its parameters are set randomly in the interval $[0,1]$.

Appendix B

Efficiency matrix and its inverse

Given, the number of events of kind j that fall in zone i, n_{ij}, the elements of the efficiency matrix E, are defined as

(B.1) $$E_{ij} = \frac{n_{ij}}{k_j}$$

where k_j is the total number of events of kind j, defined as

(B.2) $$k_j = \sum_m n_{mj}.$$

Let's define the vector whose components are the total number of events that fall in zone i

(B.3) $$z_i = \sum_l n_{il}$$

and let's compute the element i of the vector $E\mathbf{k}$

(B.4) $$\begin{aligned}\sum_l E_{il} k_l &= \sum_l \frac{n_{il}}{k_l} k_l \\ &= \sum_l n_{il} \\ &= z_i.\end{aligned}$$

Therefore the relationship

(B.5) $$E\mathbf{k} = \mathbf{z}$$

Is valid. Converselly, the vector $\mathbf{k}$ can be computed using the inverse of E,

(B.6) $$\mathbf{k} = E^{-1}\mathbf{z}.$$

Bibliography

[1] S.L. Glashow, Nucl. Phys. **B22** (1961) 579.
S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264.
A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm (Almqvist and Wiksell, Stockholm) (1968) 367.
S.L. Glashow, J. Iliopoulos, L. Maiani, Phys. Rev. **D2** (1970) 1285.

[2] Z Physics at LEP I, G. Altarelli et al. (eds.), CERN 89-08, 1989.

[3] H.Buckhardt *et al.*Particles and Fields43 (1989) 497 In ' Polarization at LEP' , CERN 88-06, (1988) 145.

[4] W. Beenakker and W. Hollik, Z.Phys. **C40** (1988) 141.

[5] A.A. Akhunov *et al.*Nucl. Phys.**B276** (1986) 1.

[6] F.A. Berends and R.Kleiss, Nucl. Phys.**B260**(1985) 32.

[7] P.kalyniak,J.N. Ng and P. Zakarauskas, Phys. Rev.**D29** (1984) 502; P.kalyniak,J.N. Ng and P. Zakarauskas, Phys. Rev.**D30** (1984) 123; E.Franco in Physics at LEP, CERN 86-02, p.187 ed. by J. Ellis and R.Peccei.

[8] T.Hebbeker, M.Martinez, G. Passarino, and G.Quast CERN-PPE/94-44

[9] T.Hebbeker, at the 1991 Joint Int.Lepton Photon Symposium at High Energies abd Europen Physical society Conference on High Energy, Geneva, 25 July - 2 August 1991.

[10] *LEP Design Report*, CERN-LEP/84-01 (1984).

[11] B. Adeva *et al.*, L3 Collaboration, *The L3 Detector at LEP*, Nucl. Inst. and Meth. **A289** (1991) 35.

[12] K. Ahmet *et al.*, OPAL Collaboration, *The OPAL Detector at LEP*, Nucl. Inst. and Meth. **A305** (1991) 275.

[13] P. Aarnio *et al.*, DELPHI Collaboration, *The DELPHI Detector at LEP*, Nucl. Inst. and Meth. **A303** (1991) 233.

[14] D. Decamp *et al.*, ALEPH Collaboration, *A Detector for Electron-Positron Anihilations at LEP*, Nucl. Inst. and Meth. **A294** (1990) 121.

[15] D. Decamp *et al.*, ALEPH Collaboration, *The ALEPH Handbook*, ALEPH internal report, ALEPH 89-77 (1989).

[16] P. Mató, *Data Processing for Large e^+e^- Experiments*, Ph.D. thesis, Univ. of Barcelona (1990).

[17] M. Delfino *et al.*, *The ALEPH Event Reconstruction Facility: Parallel Processing using Workstations*, Comp. Phys. Comm. **57** (1989) 401.

[18] A. Pacheco, *Measurement of the Electroweak Parameters from Z Decays into Lepton Pairs with the ALEPH Detector at the LEP e^+e^- Collider*, Ph.D. thesis, Univ. Autònoma of Barcelona (1990).

[19] J. Knobloch, *JULIA Users and Programmers guide*, ALEPH internal report, ALEPH 90-115, SOFTWR 90-11 (1990).

[20] D. Decamp *et al.*, ALEPH Collaboration, *Measurement of electroweak parameters from Z decays into fermion pairs*, Z. Phys. C-Particles and Fields **48** (1990) 365.

[21] D. Decamp *et al.*, ALEPH Collaboration, *Improved measurements of electroweak parameters from Z decays into fermion pairs*, Z. Phys. C Particles and Fields **53** (1992) 1.

[22] GALEPH, ALEPH internal report, ALEPH 86-119 (1986).

[23] S. Jadach *et al.*, *Z Physics at LEP I*, CERN Report CERN/89-08, Vol. 3 (1989) 69.
S. Jadach *et al.*, CERN-TH961-90, to be published in Comput. Phys. Commun.

[24] M. Böhm, A. Denner, W. Hollik, Nucl. Phys. **B304** (1988) 687.
F.A. Berends, R. Kleiss, W. Hollik, Nucl. Phys. **B304** (1988) 712.

[25] T. Sjostrand, *The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics*, LU TP 85-10.

[26] F.A. Berends *et al.*, Phys. Lett. **B148** (1984) 489.

[27] F.A. Berends *et al.*, Nucl. Phys. **B253** (1985) 421.

[28] B. Bloch-Devaux *et al.*, *KINGAL Users Guide*, ALEPH internal report, ALEPH 87-53, SOFTWR 87-7 (1988).

[29] R. Brun *et al.*, *GEANT3 Users Guide*, CERN DD/EE/84-1 (1987).

[30] H.Fesefeldt, *GHEISHA*, Aachen 3rd Inst.

[31] G. Batignani *et al.*, *The ALEPH Silicon Vertex Detector*, Nucl. Phys. **B** (Proc. Suppl.) **23A** (1991) 291.

[32] R. Settles, *ALEPH in 1991*, ALEPH internal report, ALEPH 92-68 (1992).

[33] G.J. Barber *et al.*, *Performance of the Three-Dimensional Readout of the ALEPH Inner Tracking Chamber*, Nucl. Inst. and Meth. **A279** (1989) 212.

[34] W.B Atwood *et al.*, *Performance of the ALEPH Time Projection Chamber*, Nucl. Inst. and Meth. **A306** (1991) 446.

[35] P. Némethy *et al.*, *Gated Time Projection Chamber*, Nucl. Inst. and Meth. **A297** (1983) 273.

[36] B. Jost *et al.*, *A Multi-Function FASTBUS Front-End for VAXBI computers*, IEEE Trans. Nucl. Sci. NS-36 (1989) 1452.

[37] H. Videau, *The ALEPH Pictorial Electromagnetic Calorimeter - A High Granularity Gaseous Calorimeter Operated in a High Magnetic Field*, Invited talk at the 2nd Pisa Meeting on advanced detectors, Castiglione, Italy, June 1983.

[38] D. Schlatter *et al.*, *ALEPH in Numbers*, ALEPH internal report, ALEPH 93-58 (1993).

[39] J. Knobloch and P. Norton, eds., ALEPH internal report, ALEPH 88-46 (1988).

[40] M. Consoli, W. Hollik *et al.*, in G. Altarelli, et al. (eds.), CERN 89-08, 1989.

[41] F.A. Berends, G. Burgers and W.L. van Neerven, Nucl. Phys.**B297** (1988) 429, E Nucl. Phys.**B304** (1988) 921.

[42] O. Nicrosini and L. Trentadue Phys. Lett.**B196** (1987) 551.
R. Kleiss and C. Verzegnassi, Z Physics at LEP I, Vol. 1: Standard Physics, CERN 89-08 (1989) 7.

[43] M. Martinez *et al.*, Z. Phys. C–Particles and Fields **49** (1991) 645.

[44] S. Jadach, J.H. Kühn, R.G. Stuart Z. Wąs Z. Phys. C–Particles and Fields **38**, (1988) 609.

[45] R. Miquel, *Radiative corrections to single photon processes*, Ph.D. thesis, Univ. Autònoma of Barcelona (1989).

[46] F.A. Berends *et al.*, in G. Altarelli, R. Kleiss and C. Verzegnassi, Z Physics at LEP I, Vol. 1: Standard Physics, CERN 89-08 (1989) 105.

[47] M. Grecco, Riv. Nuovo Cimento **11** (1988) 1.
M. Caffo, E. Remiddi, F. Semeria, in G. Altarelli, R. Kleiss and C. Verzegnassi, Z Physics at LEP I, Vol. 1: Standard Physics, CERN 89-08 (1989) 171.

[48] M. Martinez *et al.*, Z. Phys. C–Particles and Fields **49** (1991) 645.

[49] Z. Wąs, S. Jadach, Phys. Rev. **D41** (1990) 1425.

[50] D. Levinthal, F. Bird, R.G. Stuart, B.W. Lynn, Z. Phys. C–Particles and Fields **53** (1992) 617.

[51] The LEP Collaborations: ALEPH, DELPHI, L3 and OPAL, Phys. Lett. **B276** (1992) 247.

[52] D. Decamp *et al.*, ALEPH Collaboration, *Searches for the standard Higgs boson*, Phys. Lett.**B246** (1990) 306.

[53] G. Capon, D. Cinabro: Program package QMUIDO on the CERNVM ALEPH UPHY disk, serves to identify muons with any ALPHA programs; unpublished. We accept only positive muon flags, i.e. the track must match an existing muon pattern in the HCAL or the muon chambers better than any other track candidate.

[54] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438.
G. Altarelli and G. Parisi, Nucl. Phys. **B126**, (1977) 298.