UDINE-92-07-PE-REV
Fermilab Library
0 1160 0019480 7

UNIVERSITA' DI UDINE — ISTITUTO DI FISICA
Via Fagagna 208, I-33100 Udine (Italy)

Udine Report 92/07/PE-Rev.

10 September 1992

Classification of the Decays of the Z^0 into b and c Quark Pairs using a Neural Network

P. Eerola, J. Kalkkinen, G. Cosmo, A. De Angelis,
N. de Groot, M. Los, L. Lyons, E. Torassa, E. Vallazza

Abstract

A classifier based on a Feed-Forward Neural Network has been used for separating a sample of about 123,500 selected hadronic decays of the Z^0, collected by DELPHI during 1991, into three classes according to the flavour of the original quark pair: $u\bar{u}+d\bar{d}+s\bar{s}$ (unresolved), $c\bar{c}$ and $b\bar{b}$. The classification has been used to compute the partial widths of the Z^0 into b and c quark pairs. This gave $\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \pm 0.041(sys)$, $\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat) \pm 0.017(sys)$.

Presented at the Second Workshop on Artificial Neural Networks: From Biology to High Energy Physics, Marciana Marina, 18-26 June 1992. To be published in the Proceedings.

FERMILAB
DEC - 8 1992
LIBRARY

International Journal of Neural Systems, Vol. 3, (Supp. 1992) 000–000
Proceedings of the Neural Networks from Biology to High Energy Physics
© World Scientific Publishing Company

CLASSIFICATION OF THE DECAYS OF THE Z^0 INTO b AND c QUARK PAIRS USING A NEURAL NETWORK

P. Eerola*and J. Kalkkinen†
Research Institute for High Energy Physics, Helsinki, Finland

A. De Angelis and G. Cosmo
Istituto di Fisica dell'Università di Udine and INFN Trieste, Udine, Italy

N. De Groot and M. Los
NIKHEF-H, Amsterdam, The Netherlands

L. Lyons
University of Oxford, Oxford, UK

E. Torassa and E. Vallazza
Universitá di Torino and INFN, Turin, Italy

Received (to be inserted
Revised by Publisher)

A classifier based on a Feed-Forward Neural Network has been used for separating a sample of about 123,500 selected hadronic decays of the Z^0, collected by DELPHI during 1991, into three classes according to the flavour of the original quark pair: $u\bar{u}+d\bar{d}+s\bar{s}$ (unresolved), $c\bar{c}$ and $b\bar{b}$. The classification has been used to compute the partial widths of the Z^0 into b and c quark pairs.

1. Introduction

The difficulties in the classification of hadronic events according to their parent quark flavour can be overcome by utilizing multidimensional variables for the separation[1]. Among the multidimensional classifiers, Feed-Forward Neural Networks[2] appear to be a good candidate for complicated problems such as this. Feed-Forward Neural Networks (NN in the following) can map a set of variables calculated from the event onto a feature space in which the different species are well separated.

The possibility of using a NN for hadronic event classification was explored in Ref. [3], in which the problem of separating Z^0 decays into $b\bar{b}$ pairs was considered. The result of this study was that, in the case of a perfect detector, a separation could be achieved with a higher efficiency than with respect to traditional separation variables[1]. Further studies[4] demonstrated

*Now at CERN, Geneva, Switzerland.
†Visitor at the University of Udine.

that, also in the presence of detector effects, NNs could be a useful tool for the classification of $b\bar{b}$ events, and preliminary results on data have been recently presented[5].

In this paper, it is shown that a NN can be used to classify effectively decays of the Z^0 into $b\bar{b}$ and $c\bar{c}$ pairs. It has been possible to measure, from the data collected by the DELPHI detector[6] at LEP during 1991, the rates of the hadronic decays into $b\bar{b}$ and $c\bar{c}$. The robustness of the separation against systematic uncertainties related to model dependence of the classification has been investigated by varying a wide range of parameters in the main Monte Carlo model used, and by comparing the results obtained with different models. The behaviour of the NN is consistently similar on the Monte Carlo test sample and data. Thus, the NN could be used for extracting flavour probabilities for each event, which offers interesting prospects for subsequent physics analysis.

2. Data Selection

The sample of events used in the analysis was collected during 1991 by the DELPHI detector at the LEP e^+e^- collider, operating at center-of-mass energies around the Z^0 peak.

A description of the apparatus can be found in Ref.[6]. Features of the apparatus relevant for the analysis of multihadronic final states (with emphasis on the detection of charged particles) are outlined in Ref.[7], as well as the cuts used for the selection of the hadronic events. The analysis was based on charged particle tracks only.

Strict quality cuts on the detector performance were made, in order to guarantee the reliability of the variables used for the analysis.

A total of 123,475 hadronic decays of the Z^0 satisfying these selections was used in the present analysis. Events due to beam-gas scattering, to $\gamma\gamma$ interactions and to decays into $\tau^+\tau^-$ pairs have been estimated to be less than 0.3% of the sample.

For the study of NN input variables, tuning and testing of the Network, and classification of data, a sample of 255,000 events was generated using the JETSET 7.2 Parton Shower (PS)[8] Monte Carlo program, with the Monte Carlo parameters optimized as in[9]. The generated events were followed through a detailed detector simulation Monte Carlo program DELSIM[10], and processed through the same reconstruction and analysis chain as the data.

3. Variables used for the classification

Nineteen variables were used as an input for the separation. Their choice came from a study of flavour-dependent distributions based on JETSET 7.2 PS, and from the examination of the literature.

The list of the variables used follows[‡]

1. The sphericity $S^{(f)}$ of the first jet, calculated after a boost $\beta = 0.96$ along its axis. The axis of the jet was defined by the sum of the momenta of the particles belonging to it.

2. The directed sphericity $S^{(f)}_{1234}$ of the four most energetic particles in the first jet. For a set of Q particles in a jet, this variable is defined as $S_Q = \sum_Q p_t^2 / \sum_Q |p|^2$, where the $|p|$'s are the momenta in the rest frame of the set Q and the p_t's are their components perpendicular to the original jet direction in the laboratory frame.

[‡]The most energetic jet will be called "first jet", and indicated by the superscript (f); the second most energetic jet will be called "second jet", and indicated by the superscript (s).

3. The directed sphericity $S^{(s)}_{1234}$.

4. The invariant mass $M^{(f)}_{1234}$ of the four most energetic particles in the first jet.

5. The invariant mass $M^{(s)}_{1234}$ of the four most energetic particles in the second jet.

6..9. The products of the homologue directed sphericities for triplets of particles in the first and the second jet, $S^{(f)}_{ijk} \times S^{(s)}_{ijk}$. $(ijk) = (123, 124, 134, 234)$, where the four most energetic particles of the jet are considered.

10..13. The products of the homologue invariant masses for triplets of particles in the first and the second jet, $M^{(f)}_{ijk} \times M^{(s)}_{ijk}$.

14. The momentum of the slowest pion of the jet 1, after a boost along the jet axis corresponding to a D^* energy equal to one half of the beam energy.

15. Same as 14, for the second jet.

16. The sum over the jets of the ratios between the momentum of the leading particle and the momentum of the jet.

17. Sum of the absolute values of the track impact parameters, each one scaled by its error (to reinforce the classification of $b\bar{b}$ events). Tracks with impact parameters greater than 2 mm are omitted because they are likely to come from secondary decays of strange particles.

18. Absolute momentum $|\vec{p}|$ of the most energetic muon (0 if no muons found with momentum greater than 3 GeV/c). Muons were identified as in Ref.[11].

19. $|\vec{p_t}|$ of the most energetic muon with respect to the axis of the closest jet (0 if no muons found with momentum greater than 3 GeV/c).

Event shape variables and invariant masses discriminate between heavy and light quark events due to the high mass of the b. Variables related to the long lifetime of the b, and kinematics of semileptonic decays were included as well, since they provide clean signatures of $b\bar{b}$ events. All variables were rescaled in such a way that they ranged from 0 to 1.

4. The Neural Network

A NN with 19 nodes in the input layer, one associated with each of the input variables x_i, was used. The input variables define the pattern space P. There were 25 nodes in the hidden layer. The outputs of the three output nodes, forming the vector $\vec{\Theta}$, belong to the feature space F. The components of the output vector were assigned to the three quark classes $u\bar{u}+d\bar{d}+s\bar{s}$ (unresolved), $c\bar{c}$ and $b\bar{b}$.

The network was trained using backpropagation. We refer to[12] for a detailed description of the procedure.

5. Results

The fraction of events, β_k, of each class k ($k = 1...3$), corresponding to $u\bar{u}+d\bar{d}+s\bar{s}$ (unresolved), $c\bar{c}$, and $b\bar{b}$ respectively, were determined from the data in the following way.

For each event, the outputs of the three nodes were projected in a 2-dimensional space (u,v), by normalizing their sum to 1, and plotting them in a Dalitz plot (the distance from each side of an equilateral triangle was proportional to the corresponding output node value) § The results of such a procedure on (a) u, d, s events from simulation, (b) c events from simulation, (c) b events from simulation, and (d) data are displayed in Fig. 1.

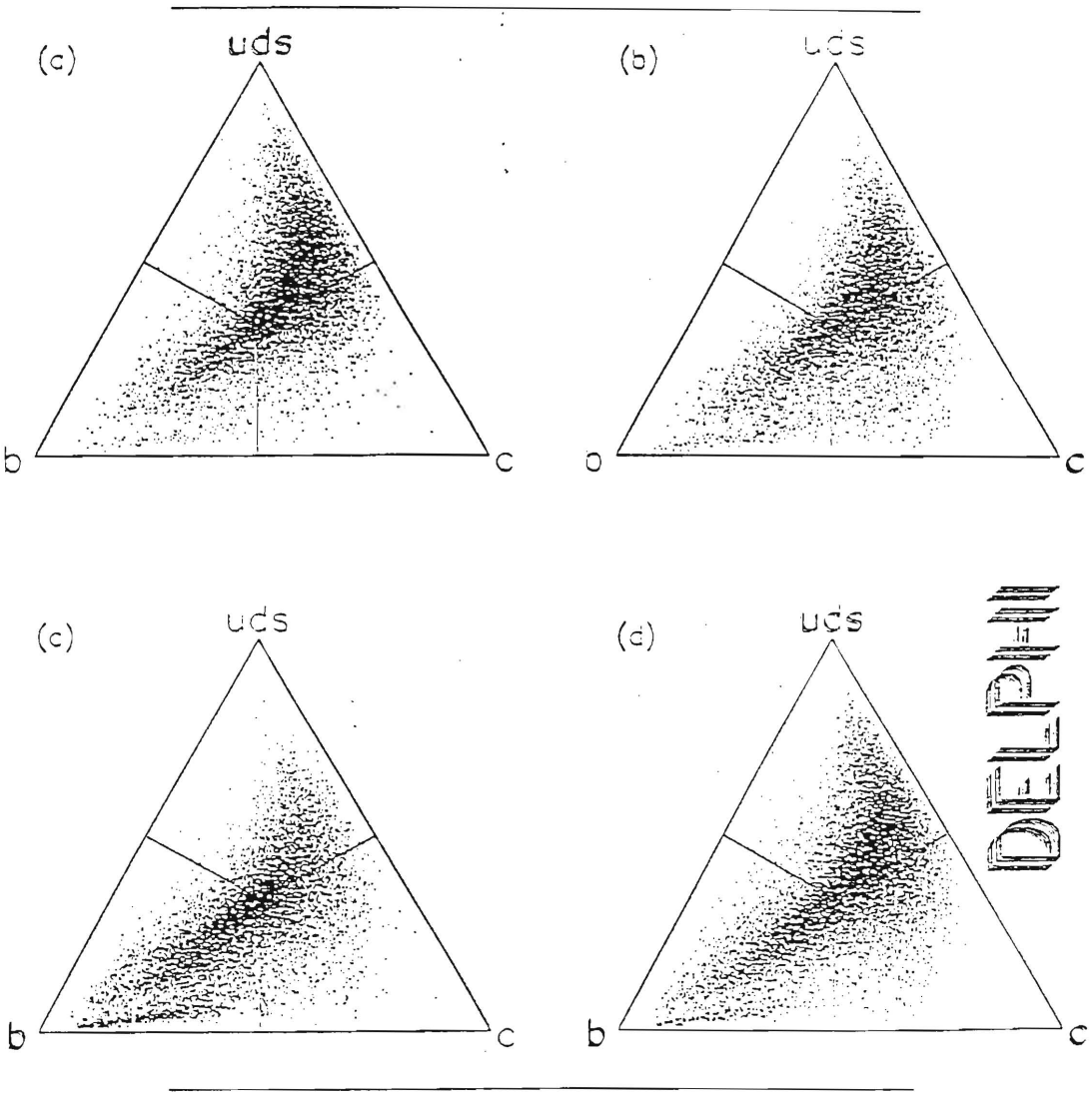

Fig. 1. Dalitz plot (see text) of the network output for simulated $u\bar{u}+d\bar{d}+s\bar{s}$ events (a), $c\bar{c}$ events (b), $b\bar{b}$ events (c), and for the data (d).

The fractions were then obtained by means of a χ^2 fit in the 2 unknown parameters β_2 and β_3 to the form

$$\Re(u,v) = (1-\beta_2-\beta_3)a_1(u,v) + \beta_2 a_2(u,v) + \beta_3 a_3(u,v)$$

where $\Re(u,v)$ is the map of the data through the network into the feature space (Fig. 2(d)), and the $a_k(u,v)$ are the distributions for each class k in the feature space, determined in the

§ v = uds, u = $(uds + 2 \times c)/\sqrt{3}$

test sample. The distributions for the PS test sample are shown in Figures 2(a)–(c). The best fit with the JETSET PS sample, containing 200,000 events, gave χ^2/Number of Degrees of Freedom = 1.0, and correlation coefficient C between β_2 and β_3 was -0.6.

In order to estimate the model dependence of the result, a fit was also performed with a test sample generated with the JETSET 7.2 Matrix Element model (ME), and passed through the full detector simulation. This sample contained 60,000 events. The χ^2/NDF of the fit was 1.1, and correlation coefficient $C = -0.4$.

The average of the two determinations was taken as the result of the measurement. This gave

$$\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \quad (1)$$

$$\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat)\,. \quad (2)$$

In the Dalitz plots (Fig. 1), a separate class of $b\bar{b}$ events appears as a concentrated band. These originate from the events containing identified muons. The muon variables separate $b\bar{b}$ events, and to a lesser extent $c\bar{c}$ events, into two classes which the network correctly recognizes.

Systematic uncertainties can arise from the model used to determine the distributions a_k described above. These distributions come from a Monte Carlo model, which contains several adjustable parameters. The fine tuning of these parameters is, in general, done by assuming that the branching fractions into each flavour are given by the Standard Model. One potential problem is that this may cause "circularity", i.e. our result for the estimated quark fractions could reflect the assumed values when determining the best tuning of the Monte Carlo. To avoid circularity, the range of variation of relevant parameters in JETSET PS, independent of the hadronic branching fractions, was established by studying the dependence of the χ^2 of the rapidity (with respect to the sphericity axis) and aplanarity distributions on the value of the parameters tuned in Ref.[9], allowing free variation of $\Gamma_{c\bar{c}}$ and $\Gamma_{b\bar{b}}$.

The systematic uncertainty in the branching fraction determinations, due to the free parameters in the simulation in the range established as described in the previous paragraph, were estimated to be: $\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.023$, $\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.010$.

Other sources of systematic uncertainties considered were: detector modelling in the simulation, and the model used for generating the initial state (JETSET PS, JETSET ME or HERWIG[13]). This gave $\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.009$, $\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.011$ from the detector modelling, and $\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.033$, $\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.008$ from the model dependence.

6. Conclusions

By combining in quadrature the systematic uncertainties, the final results were

$$\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \pm 0.041(sys) \quad (3)$$

$$\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat) \pm 0.017(sys)\,. \quad (4)$$

The results are consistent with the Standard Model, which gives in the Born approximation $\Gamma_{c\bar{c}}/\Gamma_h = 0.171$, $\Gamma_{b\bar{b}}/\Gamma_h = 0.217$.

The behaviour of the Neural Network has been investigated against a wide range of systematic uncertainties. The Neural Network has been found to be able to generalize consistently

to data, and thus, it has been demonstrated that a Neural Network could reliably be used for assigning events with a probability of coming from the hadronization of a $b\bar{b}$ or of a $c\bar{c}$ pair.

References

1. R. Marshall, *Zeit. Phys.* **C26** (1984) 291.
2. D.E.Rumelhart, G.E.Hinton and R.J.Williams, "Learning Internal Representations by Error Propagation", in D.E.Rumelhart and J.L.McLelland (Eds.), "Parallel Distributed Processing: Explorations in the Microstructure of Cognition (Vol.1)", MIT Press (1986).
3. L. Lönnblad et al., *Nucl. Phys.* **B349** (1991) 675.
4. C. Bortolotto, A. De Angelis and L. Lanceri, *Nucl. Instr. Meth.* **A306** (1991) 459;
   L. Bellantoni et al., *Nucl. Instr. Meth.* **A310** (1991) 618.
5. C. Bortolotto et al., "A Measurement of the Partial Hadronic Widths of the Z^0 using Neural Networks", in O. Benhar, C. Bosio, P. Del Giudice and E. Tabet (Eds.), "Neural Networks: From Biology to High Energy Physics", p. 445, ETS Editrice, Pisa (1991). Proceedings of the Workshop held on Isola d'Elba (Italy), June 1991;
   P. Henrard (ALEPH), presented at the 4th Symposium on Heavy Flavour Physics, Orsay, June 1991;
   B. Brandl (ALEPH), Heidelberg preprint HD-IHEP 92-01, presented at the Second International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, L'Agelonde, January 1992.
6. P. Aarnio et al. (DELPHI), *Nucl. Instr. Meth.* **A303** (1991) 233.
7. P. Aarnio et al. (DELPHI), *Phys. Lett.* **240B** (1990) 271.
8. T. Sjöstrand, *Comp. Phys. Comm.* **27** (1982) 243, ibid. **28** (1983) 229;
   T. Sjöstrand and M. Bengtsson, *Comp. Phys. Comm.* **43** (1987) 367.
9. W. de Boer et al., IEKP-KA/91-07, Karlsruhe, June 1991.
10. DELSIM User Manual, DELPHI 87-96 PROG-99, Geneva, July 1989;
    DELSIM Reference Manual, DELPHI 87-98 PROG-100, Geneva, July 1989.
11. P. Abreu et al. (DELPHI), CERN-PPE/92-79, May 1992.
12. P. Abreu et al. (DELPHI), "Classification of the Hadronic Decays of the Z^0 into b and C Quark Pairs Using a Neural Network", CERN-PPE/92-151 (September 1992), submitted to Phys. Lett. B.
13. G. Marchesini and B.R. Webber, *Nucl. Phys.* **B238** (1984) 1.