EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN-PPE/92-151

15 September 1992

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

DELPHI Collaboration

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 $\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \pm 0.041(sys)$, $\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat) \pm 0.017(sys)$.

(To be submitted to Physics Letters B)

P.Abreu[19], W.Adam[46], T.Adye[34], E.Agasi[28], G.D.Alekseev[13], A.Algeri[12], P.Allen[45], S.Almehed[22], S.J.Alvsvaag[4], U.Amaldi[7], E.G.Anassontzis[3], A.Andreazza[26], P.Antilogus[23], W-D.Apel[14], R.J.Apsimon[34], B.Åsman[41], J-E.Augustin[17], A.Augustinus[28], P.Baillon[7], P.Bambade[17], F.Barao[19], R.Barate[11], G.Barbiellini[43], D.Y.Bardin[13], G.Barker[31], A.Baroncelli[37], O.Barring[22], J.A.Barrio[24], W.Bartl[46], M.J.Bates[34], M.Battaglia[12], M.Baubillier[21], K-H.Becks[48], C.J.Beeston[31], M.Begalli[33], P.Beilliere[6], Yu.Belokopytov[39], P.Beltran[9], D.Benedic[8], A.C.Benvenuti[5], M.Berggren[17], D.Bertrand[2], F.Bianchi[42], M.S.Bilenky[13], P.Billoir[21], J.Bjarne[22], D.Bloch[8], S.Blyth[31], V.Bocci[35], P.N.Bogolubov[13], T.Bolognese[36], M.Bonesini[26], W.Bonivento[26], P.S.L.Booth[20], P.Borgeaud[36], G.Borisov[39], H.Borner[7], C.Bosio[37], B.Bostjancic[40], S.Bosworth[31], O.Botner[44], B.Bouquet[17], C.Bourdarios[17], T.J.V.Bowcock[20], M.Bozzo[10], S.Braibant[2], P.Branchini[37], K.D.Brand[32], R.A.Brenner[7], H.Briand[21], C.Bricman[2], R.C.A.Brown[7], N.Brummer[28], J-M.Brunet[6], L.Bugge[30], T.Buran[30], H.Burmeister[7], J.A.M.A.Buytaert[7], M.Caccia[7], M.Calvi[26], A.J.Camacho Rozas[38], T.Camporesi[7], V.Canale[35], F.Cao[2], F.Carena[7], L.Carroll[20], C.Caso[10], M.V.Castillo Gimenez[45], A.Cattai[7], F.R.Cavallo[5], L.Cerrito[35], V.Chabaud[7], A.Chan[1], Ph.Charpentier[7], L.Chaussard[17], J.Chauveau[21], P.Checchia[32], G.A.Chelkov[13], L.Chevalier[36], P.Chliapnikov[39], V.Chorowicz[21], J.T.M.Chrin[45], M.Ciuchini[37], M.P.Clara[42], P.Collins[31], J.L.Contreras[24], R.Contri[10], E.Cortina[45], G.Cosme[17], G.Cosmo[43], F.Couchot[17], H.B.Crawley[1], D.Crennell[34], G.Crosetti[10], M.Crozon[6], J.Cuevas Maestro[38], S.Czellar[12], E.Dahl-Jensen[27], B.Dalmagne[17], M.Dam[30], G.Damgaard[27], G.Darbo[10], E.Daubie[2], A.Daum[14], P.D.Dauncey[31], M.Davenport[7], P.David[21], W.Da Silva[21], C.Defoix[6], D.Delikaris[7], S.Delorme[7], P.Delpierre[6], P.Del Giudice[37], N.Demaria[42], A.De Angelis[43], M.De Beer[36], H.De Boeck[2], W.De Boer[14], C.De Clercq[2], M.D.M.De Fez Laso[45], N.De Groot[28], C.De La Vaissiere[21], B.De Lotto[43], A.De Min[26], H.Dijkstra[7], L.Di Ciaccio[35], F.Djama[8], J.Dolbeau[6], M.Donszelmann[7], K.Doroba[47], M.Dracos[7], J.Drees[48], M.Dris[29], Y.Dufour[6], L-O.Eek[44], P.A.-M.Eerola[7], R.Ehret[14], T.Ekelof[44], G.Ekspong[41], A.Elliot Peisert[32], J-P.Engel[8], D.Fassouliotis[29], M.Feindt[7], M.Fernandez Alonso[38], A.Ferrer[45], T.A.Filippas[29], A.Firestone[1], H.Foeth[7], E.Fokitis[29], F.Fontanelli[10], K.A.J.Forbes[20], J-L.Fousset[25], S.Francon[23], B.Franek[34], P.Frenkiel[6], D.C.Fries[14], A.G.Frodesen[4], R.Fruhwirth[46], F.Fulda-Quenzer[17], K.Furnival[20], H.Furstenau[14], J.Fuster[7], G.Galeazzi[32], D.Gamba[42], C.Garcia[45], J.Garcia[38], C.Gaspar[7], U.Gasparini[32], Ph.Gavillet[7], E.N.Gazis[29], J-P.Gerber[8], P.Giacomelli[7], R.Gokieli[47], B.Golob[40], V.M.Golovatyuk[13], J.J.Gomez Y Cadenas[7], A.Goobar[41], G.Gopal[34], M.Gorski[47], V.Gracco[10], A.Grant[7], F.Grard[2], E.Graziani[37], G.Grosdidier[17], E.Gross[7], P.Grosse-Wiesmann[7], B.Grossetete[21], S.Gumenyuk[39], J.Guy[34], U.Haedinger[14], F.Hahn[48], M.Hahn[14], S.Haider[28], Z.Hajduk[15], A.Hakansson[22], A.Hallgren[44], K.Hamacher[48], G.Hamel De Monchenault[36], W.Hao[28], F.J.Harris[31], T.Henkes[7], J.J.Hernandez[45], P.Herquet[2], H.Herr[7], T.L.Hessing[20], I.Hietanen[12], C.O.Higgins[20], E.Higon[45], H.J.Hilke[7], S.D.Hodgson[31], T.Hofmokl[47], R.Holmes[1], S-O.Holmgren[41], D.Holthuizen[28], P.F.Honore[6], J.E.Hooper[27], M.Houlden[20], J.Hrubec[46], C.Huet[2], P.O.Hulth[41], K.Hultqvist[41], P.Ioannou[3], D.Isenhower[7], P-S.Iversen[4], J.N.Jackson[20], P.Jalocha[15], G.Jarlskog[22], P.Jarry[36], B.Jean-Marie[17], E.K.Johansson[41], D.Johnson[20], M.Jonker[7], L.Jonsson[22], P.Juillot[8], G.Kalkanis[3], J.Kalkkinen[12], G.Kalmus[34], F.Kapusta[21], M.Karlsson[7], E.Karvelas[9], S.Katsanevas[3], E.C.Katsoufis[29], R.Keranen[12], J.Kesteman[2], B.A.Khomenko[13], N.N.Khovanski[13], B.King[20], N.J.Kjaer[7], H.Klein[7], W.Klempt[7], A.Klovning[4], P.Kluit[28], A.Koch-Mehrin[48], J.H.Koehne[14], B.Koene[28], P.Kokkinias[9], M.Kopf[14], K.Korcyl[15], A.V.Korytov[13], V.Kostioukhine[39], C.Kourkoumelis[3], O.Kouznetsov[13], P.H.Kramer[48], J.Krolikowski[47], I.Kronkvist[22], U.Kruener-Marquis[48], W.Kucewicz[15], K.Kulka[44], K.Kurvinen[12], C.Lacasta[45], C.Lambropoulos[9], J.W.Lamsa[1], L.Lanceri[43], V.Lapin[39], J-P.Laugier[36], R.Lauhakangas[12], G.Leder[46], F.Ledroit[11], R.Leitner[7], Y.Lemoigne[36], J.Lemonne[2], G.Lenzen[48], V.Lepeltier[17], T.Lesiak[15], J.M.Levy[8], E.Lieb[48], D.Liko[46], J.Lindgren[12], R.Lindner[48], A.Lipniacka[47], I.Lippi[32], B.Loerstad[22], M.Lokajicek[13], J.G.Loken[31], A.Lopez-Fernandez[7], M.A.Lopez Aguera[38], M.Los[28], D.Loukas[9], J.J.Lozano[45], P.Lutz[6], L.Lyons[31], G.Maehlum[30], J.Maillard[6], A.Maltezos[9], F.Mandl[46], J.Marco[38], M.Margoni[32], J-C.Marin[7], A.Markou[9], T.Maron[48], S.Marti[45], L.Mathis[1], F.Matorras[38], C.Matteuzzi[26], G.Matthiae[35], M.Mazzucato[32], M.Mc Cubbin[20], R.Mc Kay[1], R.Mc Nulty[20], G.Meola[10], C.Meroni[26], W.T.Meyer[1], M.Michelotto[32], I.Mikulec[46], L.Mirabito[23], W.A.Mitaroff[46], G.V.Mitselmakher[13], U.Mjoernmark[22], T.Moa[41], R.Moeller[27], K.Moenig[7], M.R.Monge[10], P.Morettini[10], H.Mueller[14], W.J.Murray[34], G.Myatt[31], F.Naraghi[21], F.L.Navarria[5], P.Negri[26], B.S.Nielsen[27], B.Nijjhar[20], V.Nikolaenko[39], P.E.S.Nilsen[4], P.Niss[41], V.Obraztsov[39], A.G.Olshevski[13], R.Orava[12], A.Ostankov[39], K.Osterberg[12], A.Ouraou[36], M.Paganoni[26], R.Pain[21], H.Palka[28], Th.D.Papadopoulou[29], L.Pape[7], A.Passeri[37], M.Pegoraro[32], J.Pennanen[12], V.Perevozchikov[39], M.Pernicka[46], A.Perrotta[5], C.Petridou[43], A.Petrolini[10], L.Petrovykh[39], T.E.Pettersen[32], F.Pierre[36], M.Pimenta[19], O.Pingot[2], S.Plaszczynski[17], M.E.Pol[7], G.Polok[15], P.Poropat[43], P.Privitera[14], A.Pullia[26], D.Radojicic[31], S.Ragazzi[26], H.Rahmani[29], P.N.Ratoff[18], A.L.Read[30], N.G.Redaelli[26], M.Regler[46], D.Reid[20], P.B.Renton[31], L.K.Resvanis[3], F.Richard[17], M.Richardson[20], J.Ridky[13], G.Rinaudo[42], I.Roditi[16], A.Romero[42], I.Roncagliolo[10], P.Ronchese[32], C.Ronnqvist[12], E.I.Rosenberg[1], S.Rossi[7], U.Rossi[5], E.Rosso[7], P.Roudeau[17], T.Rovelli[5], W.Ruckstuhl[28], V.Ruhlmann-Kleider[36], A.Ruiz[38], H.Saarikko[12], Y.Sacquin[36], G.Sajot[11], J.Salt[45], J.Sanchez[24], M.Sannino[10], S.Schael[14], H.Schneider[14], B.Schulze[35], M.A.E.Schyns[48], G.Sciolla[42], F.Scuri[43], A.M.Segar[31], R.Sekulin[34], M.Sessa[43], G.Sette[10], R.Seufert[14], R.C.Shellard[33],

I.Siccama[28], P.Siegrist[36], S.Simonetti[10], F.Simonetto[32], A.N.Sisakian[13], T.B.Skaali[30], G.Skjevling[30], G.Smadja[36,23], G.R.Smith[34], R.Sosnowski[7], T.S.Spassoff[11], E.Spiriti[37], S.Squarcia[10], H.Staeck[48], C.Stanescu[37], S.Stapnes[30], G.Stavropoulos[9], F.Stichelbaut[2], A.Stocchi[17], J.Strauss[46], J.Straver[7], R.Strub[8], M.Szczekowski[7], M.Szeptycka[47], P.Szymanski[47], T.Tabarelli[26], S.Tavernier[2], O.Tchikilev[39], G.E.Theodosiou[9], A.Tilquin[25], J.Timmermans[28], V.G.Timofeev[13], L.G.Tkatchev[13], T.Todorov[8], D.Z.Toet[28], O.Toker[12], E.Torassa[42], L.Tortora[37], D.Treille[7], U.Trevisan[10], G.Tristram[6], C.Troncon[26], A.Tsirou[7], E.N.Tsyganov[13], M-L.Turluer[36], T.Tuuva[12], I.A.Tyapkin[21], M.Tyndel[34], S.Tzamarias[7], S.Ueberschaer[48], O.Ullaland[7], V.Uvarov[39], G.Valenti[5], E.Vallazza[42], J.A.Valls Ferrer[45], C.Vander Velde[2], G.W.Van Apeldoorn[28], P.Van Dam[28], M.Van Der Heijden[28], W.K.Van Doninck[2], P.Vaz[7], G.Vegni[26], L.Ventura[32], W.Venus[34], F.Verbeure[2], L.S.Vertogradov[13], D.Vilanova[36], P.Vincent[23], L.Vitale[12], E.Vlasov[39], A.S.Vodopyanov[13], M.Vollmer[48], G.Voulgaris[3], M.Voutilainen[12], V.Vrba[37], H.Wahlen[48], C.Walck[41], F.Waldner[43], M.Wayne[1], A.Wehr[48], M.Weierstall[48], P.Weilhammer[7], J.Werner[48], A.M.Wetherell[7], J.H.Wickens[2], J.Wikne[30], G.R.Wilkinson[31], W.S.C.Williams[31], M.Winter[8], M.Witek[15], D.Wormald[30], G.Wormser[17], K.Woschnagg[44], N.Yamdagni[41], P.Yepes[7], A.Zaitsev[39], A.Zalewska[15], P.Zalewski[17], D.Zavrtanik[40], E.Zevgolatakos[9], G.Zhang[48], N.I.Zimin[13], M.Zito[36], R.Zuberi[31], R.Zukanovich Funchal[6], G.Zumerle[32], J.Zuniga[45]

[1] Ames Laboratory and Department of Physics, Iowa State University, Ames IA 50011, USA
[2] Physics Department, Univ. Instelling Antwerpen, Universiteitsplein 1, B-2610 Wilrijk, Belgium
and IIHE, ULB-VUB, Pleinlaan 2, B-1050 Brussels, Belgium
and Faculté des Sciences, Univ. de l'Etat Mons, Av. Maistriau 19, B-7000 Mons, Belgium
[3] Physics Laboratory, University of Athens, Solonos Str. 104, GR-10680 Athens, Greece
[4] Department of Physics, University of Bergen, Allégaten 55, N-5007 Bergen, Norway
[5] Dipartimento di Fisica, Università di Bologna and INFN, Via Irnerio 46, I-40126 Bologna, Italy
[6] Collège de France, Lab. de Physique Corpusculaire, 11 pl. M. Berthelot, F-75231 Paris Cedex 05, France
[7] CERN, CH-1211 Geneva 23, Switzerland
[8] Centre de Recherche Nucléaire, IN2P3 - CNRS/ULP - BP20, F-67037 Strasbourg Cedex, France
[9] Institute of Nuclear Physics, N.C.S.R. Demokritos, P.O. Box 60228, GR-15310 Athens, Greece
[10] Dipartimento di Fisica, Università di Genova and INFN, Via Dodecaneso 33, I-16146 Genova, Italy
[11] Institut des Sciences Nucléaires, Université de Grenoble 1, F-38026 Grenoble, France
[12] Research Institute for High Energy Physics, SEFT, Siltavuorenpenger 20 C, SF-00170 Helsinki, Finland
[13] Joint Institute for Nuclear Research, Dubna, Head Post Office, P.O. Box 79, 101 000 Moscow, USSR.
[14] Institut für Experimentelle Kernphysik, Universität Karlsruhe, Postfach 6980, D-7500 Karlsruhe 1, FRG
[15] High Energy Physics Laboratory, Institute of Nuclear Physics, Ul. Kawiory 26 a, PL-30055 Krakow 30, Poland
[16] Centro Brasileiro de Pesquisas Físicas, rua Xavier Sigaud 150, RJ-22290 Rio de Janeiro, Brazil
[17] Université de Paris-Sud, Lab. de l'Accélérateur Linéaire, Bat 200, F-91405 Orsay, France
[18] School of Physics and Materials, University of Lancaster - Lancaster LA1 4YB, UK
[19] LIP, IST, FOUL - Av. Elias Garcia, 14 - 1°, P-1000 Lisboa Codex, Portugal
[20] Department of Physics, University of Liverpool, P.O. Box 147, GB - Liverpool L69 3BX, UK
[21] LPNHE, Universités Paris VI et VII, Tour 33 (RdC), 4 place Jussieu, F-75230 Paris Cedex 05, France
[22] Department of Physics, University of Lund, Sölvegatan 14, S-22363 Lund, Sweden
[23] Université Claude Bernard de Lyon, 43 Bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France
[24] Universidad Complutense, Avda. Complutense s/n, E-28040 Madrid, Spain
[25] Univ. d'Aix - Marseille II - Case 907 - 70, route Léon Lachamp, F-13288 Marseille Cedex 09, France
[26] Dipartimento di Fisica, Università di Milano and INFN, Via Celoria 16, I-20133 Milan, Italy
[27] Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen 0, Denmark
[28] NIKHEF-H, Postbus 41882, NL-1009 DB Amsterdam, The Netherlands
[29] National Technical University, Physics Department, Zografou Campus, GR-15773 Athens, Greece
[30] Physics Department, University of Oslo, Blindern, N-1000 Oslo 3, Norway
[31] Nuclear Physics Laboratory, University of Oxford, Keble Road, GB - Oxford OX1 3RH, UK
[32] Dipartimento di Fisica, Università di Padova and INFN, Via Marzolo 8, I-35131 Padua, Italy
[33] Depto. de Fisica, Pontificia Univ. Católica, C.P. 38071 RJ-22453 Rio de Janeiro, Brazil
[34] Rutherford Appleton Laboratory, Chilton, GB - Didcot OX11 OQX, UK
[35] Dipartimento di Fisica, Università di Roma II and INFN, Tor Vergata, I-00173 Rome, Italy
[36] Centre d'Etude de Saclay, DSM/DAPNIA, F-91191 Gif-sur-Yvette Cedex, France
[37] Istituto Superiore di Sanità, Ist. Naz. di Fisica Nucl. (INFN), Viale Regina Elena 299, I-00161 Rome, Italy
[38] Facultad de Ciencias, Universidad de Santander, av. de los Castros, E - 39005 Santander, Spain
[39] Inst. for High Energy Physics, Serpukow P.O. Box 35, Protvino, (Moscow Region), CEI
[40] Institut "Jozef Stefan", Ljubljana, Slovenija
[41] Institute of Physics, University of Stockholm, Vanadisvägen 9, S-113 46 Stockholm, Sweden
[42] Dipartimento di Fisica Sperimentale, Università di Torino and INFN, Via P. Giuria 1, I-10125 Turin, Italy
[43] Dipartimento di Fisica, Università di Trieste and INFN, Via A. Valerio 2, I-34127 Trieste, Italy
and Istituto di Fisica, Università di Udine, I-33100 Udine, Italy
[44] Department of Radiation Sciences, University of Uppsala, P.O. Box 535, S-751 21 Uppsala, Sweden
[45] IFIC, Valencia-CSIC, and D.F.A.M.N., U. de Valencia, Avda. Dr. Moliner 50, E-46100 Burjassot (Valencia), Spain
[46] Institut für Hochenergiephysik, Österr. Akad. d. Wissensch., Nikolsdorfergasse 18, A-1050 Vienna, Austria
[47] Inst. Nuclear Studies and, University of Warsaw, Ul. Hoza 69, PL-00681 Warsaw, Poland
[48] Fachbereich Physik, University of Wuppertal, Postfach 100 127, D-5600 Wuppertal 1, FRG

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 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 letter, 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 multi-hadronic final states (with emphasis on the detection of charged particles) are outlined in Ref. [7]. The present analysis relied on the information provided by the central tracking detectors: the Micro Vertex Detector (VD), the Inner Detector (ID), the Time Projection Chamber (TPC), and the Outer Detector (OD). For muon identification, the Barrel and Forward Muon Chambers (MUB and MUF) were used.

The central tracking system of DELPHI covers the region between 25° and 155° in polar angle θ, with reconstruction efficiency near 1. The average momentum resolution for the charged particles in hadronic final states is in the range $\Delta p/p \simeq 0.001p$ to $0.01p$ (p in GeV/c), depending on which detectors are included in the track fit.

The polar angle coverage of the VD is from 42° to 138°. The intrinsic point resolution for single tracks in the transverse plane has been measured to be 8 μm. For tracks with hits in all three layers of the VD, the uncertainty of the track extrapolation to the vertex region is 24 μm for high momentum tracks.

The polar angle acceptance of the muon chambers is $9°-43°$, $52°-128°$ and $137°-171°$. The muon identification algorithm is described in detail in Ref. [8]. The muon track candidate must be associated with hits in the muon chambers, and the identification is

based on a χ^2 fit, where the χ^2 is calculated from the difference between the extrapolated track trajectory and the track element constructed from the hits in the muon chambers. With these criteria, the muon identification efficiency is $(78 \pm 2)\%$ and background due to misidentification is $(1.0 \pm 0.3)\%$ per charged hadron.

Only charged particles fulfilling the following criteria were used in this analysis: (a) momentum, p, larger than 0.1 GeV/c and smaller than 50 GeV/c; (b) impact parameter in the plane transverse to the beam direction, $|r|$, smaller than 5 cm, and along the beam direction, $|z|$, smaller than 10 cm; (c) measured track length above 30 cm; (d) polar angle, θ, between 25° and 155°.

Hadronic events were then selected by requiring that (α) each of the two hemispheres $\cos\theta < 0$ and $\cos\theta > 0$ contained a total energy of the charged particles larger than 3 GeV; (β) the total energy of the charged particles seen in both hemispheres together exceeded 15 GeV; (γ) there were at least 5 charged particles with momenta above 0.2 GeV/c; (δ) the polar angle θ of the sphericity axis was in the range $40° < \theta < 140°$; (ϵ) each of the two most energetic jets[†] in the event had at least four charged particles.

Quality cuts on the detector performance were made, in order to improve the reliability of the variables used for the analysis.

A total of 123,475 events satisfying these selection criteria 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, events were generated by using the JETSET 7.2 Parton Shower (PS) [10] Monte Carlo program, with the Monte Carlo parameters optimized as in Ref. [11]. The generated events were followed through a detailed detector simulation program, DELSIM [12], and processed through the same event reconstruction as the data. A total of 255,000 events satisfying the same selection criteria as the data was used. From this sample, 30,000 events were separated to train the NN, and the remaining events were used as a test sample to obtain the branching fractions (200,000 events), and to check the NN training (100,000 events).

In addition, events were generated with JETSET 7.2 Monte Carlo program with QCD 2^{nd} order generation of the initial state at the parton level (JETSET ME) and the tuning of [11], and simulated with the full detector simulation. This sample, containing 60,000 events, was utilized as an alternative test sample for comparison with the data.

2.1 Variables Used for the Classification

Nineteen variables were used as input for the separation. Their choice came from an examination of flavour dependent distributions based on JETSET PS, and from previous studies [3–5].

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.

[†]The charged particles in the event were clustered in jets according to the JADE/E0 algorithm [9], with $y_{cut} = 0.05$.

[‡]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).

2. The directed sphericity $S^{(f)}_{1234}$ of the four most energetic particles in the first jet. For a set Q of particles in a jet, this variable is defined as

$$S_Q = \frac{\sum_Q p_t^2}{\sum_Q p^2}$$

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.

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 corresponding 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, ordered in decreasing energy.

10..13. The products of the corresponding 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 all 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. All accepted charged particles in the event with impact parameters less than 2 mm were included in the sum. Tracks with large impact parameters were omitted because they are likely to come from secondary decays of strange particles.
18. Absolute momentum p of the most energetic muon (0 if no muons found with momentum greater than 3 GeV/c). Muons were identified as in Ref. [8].
19. 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).

All variables were rescaled in such a way that they ranged from 0 to 1. This was done because a NN is helped when all the input variables have numerical values of roughly the same magnitude.

Event shape variables and invariant masses (1.–13., 16.) discriminate between heavy and light quark events due to the high mass of the b. As an alternative to this set of shape variables, simple kinematical variables like the three momenta of the leading particles, could also be used. In this case, the network should recognize the relevant correlations. Previous studies [13] have indicated that such a network could perform slightly better than a network using shape variables, but the former takes considerably longer to train, and its behaviour is less stable.

Variables related to the long lifetime of the b (17.), and kinematics of semileptonic decays (18., 19.) were included as well, since they provide clean signatures of $b\bar{b}$ events. The muon momentum spectrum is also useful for separation of events originating from c quark pairs. Charm quark event classification was further reinforced by using two variables derived from the characteristic decay of the $D^{*\pm}$ meson into a D^0 and a charged pion, where the pion has a low transverse momentum with respect to the parent $D^{*\pm}$ direction (14., 15.). The data distributions of the network input variables were reasonably reproduced by the JETSET PS with full detector simulation.

3 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 $u\overline{u}+d\overline{d}+s\overline{s}$ (unresolved), $c\overline{c}$ and $b\overline{b}$.

In principle, a single hidden layer is enough to perform any mapping of a continuous function between pattern space P and feature space F. It was checked that by adding a second hidden layer the performance of the network did not improve, but the time needed to train the network increased substantially.

When the number of hidden nodes was reduced from 25 to 18 and further to 15, it was observed that the three output nodes of the NN gave increasingly similar output values, degrading the separation of the classes. It was also checked that by going much beyond 25 hidden nodes one begins to introduce useless nodes to the hidden layer. Therefore, the network with 25 hidden nodes was chosen. A single network with three output nodes was used instead of three single output networks to account automatically for correlations between classes.

In the structure chosen, each node performs a weighted sum of the output values from all the nodes of the previous layer. The node output is computed via a sigmoid function

$$g_T(x) = \frac{1}{1 + e^{-2x/T}}$$

at a "temperature" T. The output o_i of the i-th node of a layer (starting from the second) is then

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

where the sum is made over the nodes of the previous layer.

The network training procedure fixes the values of the weights ω_{ij} associated with the node interconnections. The weights can be both positive and negative. If the value of the weight ω_{ij} is zero, then there is no connection between nodes i and j. The aim is to find a mapping of the input pattern space ($x_i \,\epsilon$ P) to the feature space ($\Theta \,\epsilon$ F), such that a good separation of events belonging to a class A from events belonging to the complementary class $\overline{A}$ is obtained. Each class is associated with one output node. Two symmetric target values (1 for class A and 0 for class $\overline{A}$) were used for each of the three output nodes.

In the back propagation learning algorithm the output feature values for the training input events are computed and compared with the desired target values. A squared error function E is computed to quantify the difference between the obtained output $\vec{\Theta}$ and the desired target $\vec{t}$,

$$E = \frac{1}{2} \sum_{events} (\vec{t} - \vec{\Theta})^2\,.$$

This function is minimized by changing ("updating") the weights by an amount computed from the error function by the gradient descent method [2]. The process is controlled by the "learning strength parameter" η and the "momentum" α [2]. Each updating step in the space of weights, computed by gradient descent, is multiplied by η and added to the previous step, multiplied by α. To smooth out fluctuations, weights are updated using the cumulative error from a number of input training events.

For the training of the system, a set of 30,000 simulated events, detailed in section 2, was used. The number of independent events in each class was thus an order of magnitude larger than the number of weights in the network (550 in our case). The weights were updated every 10 events, chosen at random from the three classes $u\bar{u}+d\bar{d}+s\bar{s}$ (unresolved), $c\bar{c}$ and $b\bar{b}$, in such a way that, on the average, there was an equal number of events from each of the three classes[§]. In the following, this sequence of 10 events will be referred to as "update".

Changing the parameters η and α during the training is convenient in order to allow for a fast movement in the space of weights in the early stage of training, and to obtain a controlled approach to the minimum at the later stage. For this reason, the learning and momentum parameters were decreased and increased respectively after every 3,000 updates (an "epoch") according to a rule:

$$\eta_t = \eta_{t-1} \times (\eta_{min}/\eta_{t-1})^{k_\eta}$$

$$\alpha_t = \alpha_{t-1} \times (\alpha_{max}/\alpha_{t-1})^{k_\alpha},$$

where η_{min} and α_{max} are the minimum (maximum) allowed values for the parameters, and subscript t $(t-1)$ refers to the epoch number. Exponents k_η and k_α were set to 0.05 and 0.14, respectively. Given the finite value of the weight change, the gradient descent method might lead to an occasional increase of the error value. In this case, the parameters were reset to their initial values.

The architecture of the network is summarized in Table 1, together with the parameters used in the training phase.

Nodes in the hidden layer	25
T	2.0
α (training)	0.4 - 0.9
η (training)	0.05 - 0.0001

Table 1: *Characteristics of the NN.*

At each step of the learning procedure, an indication of the network performance can be inferred from the error function. A more reliable evaluation is obtained by testing the response of the network on a set of input events independent of the training set. The behaviour of the error function with this new test set shows the error the network makes in generalizing to new data. When the generalization error starts to increase, the NN has 'overlearned' the training sample and its ability to generalize is degraded.

The test sample consisted of about 100,000 simulated events, generated by using JETSET PS (see section 2). The number of events generated in each class corresponded to the Z^0 hadronic branching fractions in the Standard Model. By monitoring the behaviour of the generalization error, the training was stopped after 300,000 updates.

After the network has been trained, its performance can be judged in terms of signal efficiency ε_s (number of events correctly classified as belonging to class q over the total number of events in class q), and purity p (number of correctly classified events in class q over all the events classified as belonging to class q). For example, if the event is classified simply according to the highest output node, one obtains an efficiency for b identification of about 55%, and a purity of about 43%. The efficiency and purity for a single quark class can be improved with a dedicated single output network, but we have chosen to estimate the three quark categories simultaneously. As far as inclusive analyses using single variables are concerned, a better b quark event purity can be obtained by using prompt leptons with a strict cut on the lepton transverse momentum with respect to

[§]It has been verified that equal sample sizes improve the performance of the network [4], and reduce biases.

the jet axis (see for example Ref. [8]), but the efficiency is limited by the semileptonic branching fraction.

One can estimate to which input variables the network is most sensitive by changing one input variable at a time and monitoring the response of the output nodes. The average and the standard deviation of each input variable distribution were calculated from the full data sample. In Fig. 1, the changes of the three output nodes are displayed when each input variable, in turn, was changed by one standard deviation from its average value, while the other inputs remained at their average values. For the muon input variables, the test was performed only when the muon inputs were different from zero.

As expected, the separation is best for $b\bar{b}$ events. The b quark output node shows the strongest gradient in response to the change of the input variables, whereas the light quark output node is least sensitive. The biggest changes to b quark output node are produced by changing the track impact parameter and muon variables, and some of the invariant mass variables. The c quark output node is less sensitive to the lepton variables. The slowest pion of the most energetic jet shows reasonable separation for $c\bar{c}$ events. For example, the sensitivity to this variable is greater than to the sphericity variables for c quark classification. We emphasize that Figure 1 shows the sensitivity of the output to just one input variable at a time. The NN, however, is capable of taking into account correlations between variables.

4 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. 2.

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 test sample. The distributions for the PS test sample are shown in Figures 2(a)–(c). The two-dimensional Dalitz plot was divided into small square bins by overlaying a square on the triangle. The side of the square was equal to the side of the Dalitz triangle. The large square was then divided into 256×256 bins. Bins with too few events were combined with adjacent ones, giving typically a few thousand degrees of freedom. It was checked that this procedure was insensitive to the lower limit on the allowed number of events per bin, and to the way in which the bins were combined. All distributions were normalized to unity.

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.

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

In order to reduce the model dependence of the result, a fit was also performed with the JETSET ME test sample containing 60,000 events. For ME, the χ^2/NDF of the fit was 1.1, and correlation coefficient $C = -0.4$.

The branching fractions obtained are summarized in Table 2. Small corrections have been applied for the different selection efficiencies of the three quark classes, due to the hadronic selection criteria. The statistical errors from the corrections are negligible.

	$1000 \times \Gamma_{c\bar{c}}/\Gamma_h$	$1000 \times \Gamma_{b\bar{b}}/\Gamma_h$
JETSET PS	120 ± 11	238 ± 5
JETSET ME	182 ± 7	226 ± 5
Average	151 ± 8	232 ± 5

Table 2: *Determination of the branching ratios.*

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

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

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

The errors quoted are statistical only. The major contribution to the statistical error comes from the overlap of the Monte Carlo distributions of the three classes. Even though the PS test sample is larger, its statistical error is not smaller, because the distributions of the different quark classes overlap more in PS than in ME. Had the distributions been completely separated, the statistical error arising from the number of data events would have been ± 0.0011 on the $c\bar{c}$ fraction, and ± 0.0013 on the $b\bar{b}$ fraction.

In the Dalitz plots (Fig. 2), 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. Figure 3 shows the Dalitz plot for events which contained a muon with a transverse momentum of at least 1 GeV/c.

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 as follows. A set of 100,000 events was generated using the parameters of Ref. [11] ("central" simulation in the following), retaining the charged stable particles only. For each of the parameters, an optimization was then performed 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. [11], allowing free variation of $\Gamma_{c\bar{c}}$ and $\Gamma_{b\bar{b}}$. The two extreme values of the range were conservatively taken when the χ^2 increase corresponded to at least four standard deviations with respect to the central tuning[‖].

[‖]In common with most other estimates of the uncertainties in the Monte Carlo parameters, our error estimates are derived from the diagonal elements of the inverse error matrix, rather than those of the error matrix itself. This underestimate is partially compensated by the fact that we allow χ^2 to increase by a large amount as compared with its best value.

The relevant parameters for the parton phase of JETSET PS are the QCD parameter Λ, and the cutoff parameter Q_0 of the parton evolution. The range of variation of Λ was established to be between 0.28 and 0.32 GeV, but conservatively the parameter was allowed to vary between 0.25 and 0.32 GeV. The value of Q_0 was found to be between 0.6 and 1.4 GeV. The fragmentation in JETSET PS is governed by the Lund symmetric fragmentation function with two parameters a and b, of which essentially only one is a free parameter. The transverse momentum of primary hadrons is parametrized by a Gaussian of width σ_q. The range of variation of the a parameter was found to be between 0.14 and 0.26 (the b parameter was fixed to 0.34 GeV^{-2}), and σ_q lay between 355 MeV/c and 415 MeV/c.

Other sources of systematic uncertainty due to free parameters in JETSET PS considered were: choice of the fragmentation function, lifetime of the b quark, semileptonic branching fraction of b and c quarks, and branching fraction of $\mathrm{D}^{*\pm}$ meson into $\mathrm{D}^0\pi^\pm$.

The systematic uncertainty from fragmentation was checked by using the Peterson fragmentation scheme instead of the Lund symmetric fragmentation for the b and c quarks. The ϵ_b parameter was allowed to vary between 0.002 and 0.009, and the ϵ_c parameter was allowed to vary between 0.010 and 0.080. The two parameters were varied simultaneously in the same direction. The parameters a and b were kept fixed at the central values 0.20 and 0.34 GeV^{-2}. For the lifetime of b, a range of variation between 1.20 ps and 1.35 ps was allowed. A range of variation between 0.090 and 0.102 was allowed for the semileptonic branching fraction of c, and between 0.100 and 0.123 for the semileptonic branching fraction of b. The two fractions were varied simultaneously in the same direction. It was checked that when the fractions were varied incoherently, the results remained within the range estimated by varying the fractions as above. The branching ratio, $\mathrm{Br}(\mathrm{D}^{*\pm} \to \mathrm{D}^0\pi^\pm)$, was varied between 0.50 and 0.65.

In the study of systematic uncertainties, to save computer time, a fast detector simulation was used. The branching fractions were obtained by considering the relevant Monte Carlo simulation as the test sample and fitting the data $\Re(u, v)$ with the distributions obtained from this test sample. The results are summarized in Table 3. The χ^2/NDF of the fits lies between 1.0 and 1.1.

The total systematic uncertainty due to the variation of input parameters in the JETSET PS was estimated by combining, in quadrature, the individual contributions. These were taken as the half differences between the branching fractions obtained with each pair of extreme values allowed for the parameters, and subtracting, in quadrature, the contribution expected from the statistical error due to the finite simulation statistics (± 0.003 for $c\bar{c}$, ± 0.002 for $b\bar{b}$). This gave

$$\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.023(sys, param) \quad (3)$$

$$\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.010(sys, param). \quad (4)$$

	$1000 \times \Delta(\Gamma_{c\bar{c}}/\Gamma_h)$	$1000 \times \Delta(\Gamma_{b\bar{b}}/\Gamma_h)$
$\Lambda = 250$–320 MeV	± 5.5	∓ 5
$Q_0 = 0.6$–1.4 GeV	± 11	∓ 1.5
$\sigma_q = 355$–415 MeV/c	± 18.5	± 4
$a = 0.14$–0.26	∓ 7.5	∓ 6.5
$\epsilon_b = 0.002$–0.009, $\epsilon_c = 0.010$–0.080	± 1.5	∓ 5.5
$\mathrm{Br}(\mathrm{D}^{*\pm} \to \mathrm{D}^0\pi^\pm) = 0.50$–$0.65$	∓ 2.5	± 0
$\tau_b = 1.20$–1.35 ps	∓ 1.5	∓ 1
$\mathrm{Br}(c \to \ell\mathrm{X}) = 9.0$–$10.2$ %, $\mathrm{Br}(b \to \ell\mathrm{X}) = 10.0$–$12.3$ %	± 2.5	∓ 3
Total, JETSET PS	± 23	± 10

Table 3: *Summary of the systematic effects from JETSET PS.*

A comparison with HERWIG [14] was also performed. The default parameters of the version 5.4 were taken, and the fast simulation was used. The results from the fit of the data using HERWIG as the test sample, corrected for the difference between fast and full simulation, were: $\Gamma_{c\bar{c}}/\Gamma_h = 0.134 \pm 0.005$, $\Gamma_{b\bar{b}}/\Gamma_h = 0.223 \pm 0.003$. The χ^2/NDF of the fit was 1.0.

The systematic uncertainty associated with the choice of model was estimated by calculating the standard deviation of the branching fraction determinations obtained with JETSET PS, JETSET ME (Table 2), and HERWIG. The obtained values were

$$\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.033(sys, model) \quad (5)$$
$$\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.008(sys, model)\,. \quad (6)$$

Systematic uncertainties can also emerge from imperfections in the detector simulation. This was checked separately for muon identification and for the impact parameters. By varying the muon identification efficiency, ϵ_μ, and the background due to misidentification, b_μ, as measured in Ref. [8] ($\epsilon_\mu = (78 \pm 2)\%, b_\mu = (1.0 \pm 0.3)\%$ per charged hadron), systematic uncertainties of 0.009 and 0.007 were obtained for the $c\bar{c}$ and $b\bar{b}$ hadronic branching fractions, respectively. The muon background fraction was the major contribution to the systematic uncertainty. The impact parameter sum of the $b\bar{b}$ events was allowed to vary $\pm 10\%$, resulting in an uncertainty of 0 and 0.008 for the $c\bar{c}$ and $b\bar{b}$ hadronic branching fractions, respectively. The systematic uncertainty was estimated in the same way as the uncertainty due to JETSET PS parameters, i.e. the estimated statistical error was subtracted in quadrature from the half difference between each pair of extreme values.

By adding in quadrature the contributions due to detector modelling, systematic uncertainties

$$\Delta(\Gamma_{c\bar{c}}/\Gamma_h) = \pm 0.009(sys, det) \quad (7)$$
$$\Delta(\Gamma_{b\bar{b}}/\Gamma_h) = \pm 0.011(sys, det) \quad (8)$$

were obtained.

By using the measurements (1) and (2), and combining in quadrature the systematic uncertainties from (3–8), the final results were

$$\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \pm 0.041(sys) \quad (9)$$
$$\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat) \pm 0.017(sys)\,. \quad (10)$$

The same NN was also used to determine the branching fractions when the muon inputs were removed both from data and the Monte Carlo test sample. The average of the determinations by using PS and ME models as test samples was

$$\Gamma_{c\bar{c}}/\Gamma_h = 0.153 \pm 0.009(stat) \pm 0.044(sys) \quad (11)$$
$$\Gamma_{b\bar{b}}/\Gamma_h = 0.230 \pm 0.006(stat) \pm 0.015(sys)\,. \quad (12)$$

The systematic uncertainty was obtained from the estimates (3,4) and (7,8) by removing the sources of systematic uncertainties related to semileptonic decays, and including the model dependence as the standard deviation of the results without muons, obtained from PS, ME and HERWIG test samples.

The LEP collaborations have recently measured the product of the hadronic branching fraction of the Z^0 into c and b quark pairs times the inclusive semileptonic branching fraction of the hadrons produced from these quarks [15]. The average of their results is $\mathrm{Br}(c \to \ell \mathrm{X}) \times \Gamma_{c\bar{c}}/\Gamma_h = 0.0156 \pm 0.0036$, $\mathrm{Br}(b \to \ell \mathrm{X}) \times \Gamma_{b\bar{b}}/\Gamma_h = 0.0233 \pm 0.0008$. Combining these results with the measurements (11,12), the semileptonic branching fractions of c and b quarks were determined to be:

$$Br(c \to \ell X) = (10 \pm 4)\% \quad (13)$$
$$Br(b \to \ell X) = (10.1 \pm 0.8)\%\,. \quad (14)$$

5 Conclusions

By using a classifier based on a Feed-Forward Neural Network, the hadronic branching fractions of the Z^0 into c and b quark pairs have been determined to be

$$\Gamma_{c\bar{c}}/\Gamma_h = 0.151 \pm 0.008(stat) \pm 0.041(sys)$$
$$\Gamma_{b\bar{b}}/\Gamma_h = 0.232 \pm 0.005(stat) \pm 0.017(sys)\,.$$

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.

Acknowledgements

We are greatly indebted to our technical collaborators and to the funding agencies for their support in building and operating the DELPHI detector, and to the members of the CERN-SL Division for the excellent performance of the LEP collider.

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] P. Abreu *et al* (DELPHI), CERN-PPE/92-79, May 1992.

[9] W. Bartel *et al* (JADE), *Zeit. Phys.* **C33** (1986) 23.

[10] T. Sjöstrand, Comp. Phys. Comm. **27** (1982) 243, ibid. **28** (1983) 229;
T. Sjöstrand and M. Bengtsson, Comp. Phys. Comm. **43** (1987) 367.

[11] W. de Boer *et al*, IEKP-KA/91-07, Karlsruhe, June 1991.

[12] DELSIM User Manual, DELPHI 87-96 PROG-99, Geneva, July 1989;
DELSIM Reference Manual, DELPHI 87-98 PROG-100, Geneva, July 1989.

[13] M. Los and N. De Groot, "B Tagging in DELPHI with a Feed-Forward Neural Network", in O. Benhar, C. Bosio, P. Del Giudice and E. Tabet (Eds.), "Neural Networks: From Biology to High Energy Physics", p. 459, ETS Editrice, Pisa (1991). Proceedings of the Workshop held on Isola d'Elba (Italy), June 1991;
P. Branchini, M. Ciuchini and P. Del Giudice, "B Tagging with Neural Networks: An Alternative use of Single Particle Information for Discriminating Jet Events", INFN Sanitá preprint INFN-ISS 92/1, presented at the Second International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, L'Agelonde, January 1992;
G. Bahan and R. Barlow, "Identification of b Jets using Neural Networks", Manchester preprint MAN/HEP/92/1.

[14] G. Marchesini and B.R. Webber, *Nucl. Phys.* **B238** (1984) 1.

[15] D. Decamp *et al* (ALEPH), *Phys. Lett.* **244B** (1990) 551;
B. Adeva *et al* (L3), *Phys. Lett.* **261B** (1991) 177;
M.Z. Akrawy *et al* (OPAL), *Phys. Lett.* **263B** (1991) 311;
P. Abreu *et al* (DELPHI), to be published in *Zeit. Phys. C.*.

Figure Captions

1. The change of the output value of the $u\bar{u}+d\bar{d}+s\bar{s}$ output node (a), the $c\bar{c}$ output node (b), and the $b\bar{b}$ output node (c), when each input variable is varied by one standard deviation from its average value.
2. 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).
3. Dalitz plot (see text) of the network output for data events with a high transverse momentum muon (p_t larger than 1 GeV/c).

Fig. 1

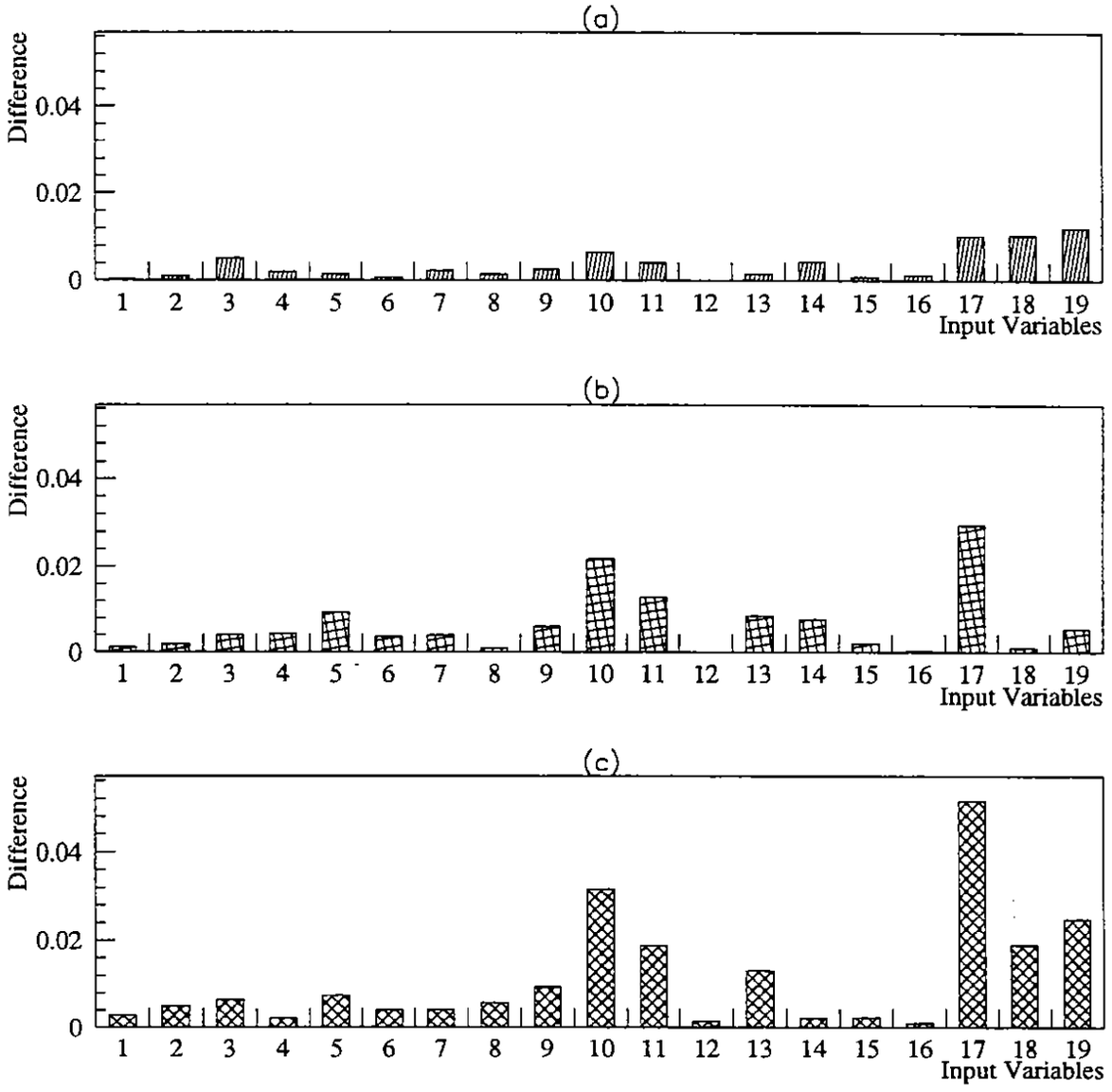
(a)
Difference
0.04
0.02
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Input Variables
(b)
Difference
0.04
0.02
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Input Variables
(c)
Difference
0.04
0.02
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Input Variables

Fig. 2

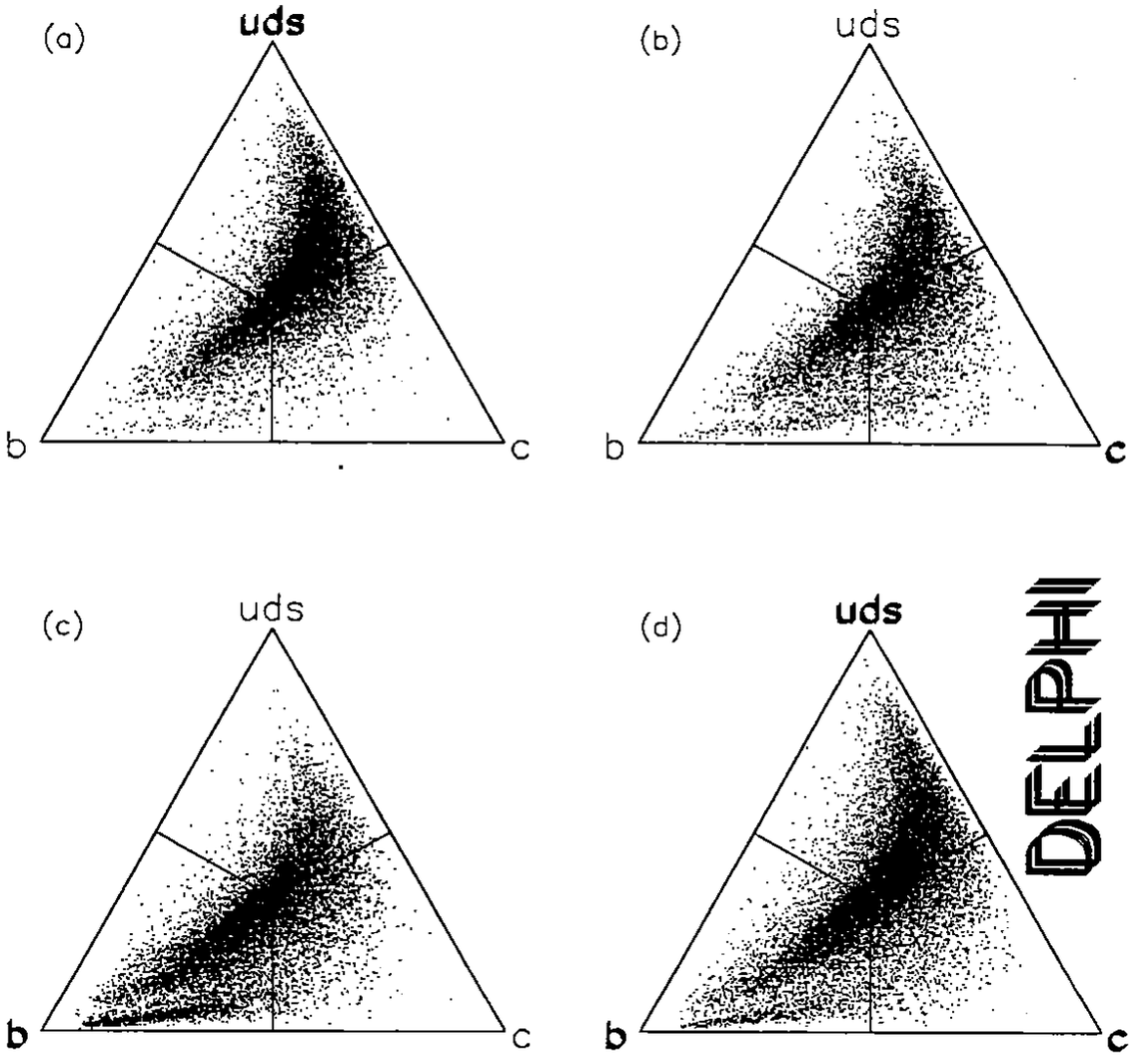

Fig. 3