Separating $\Upsilon(4S)$ decays from continuum events using a neural network

Fabrizio Salvatore and Glen Cowan

Physics Dept., Royal Holloway, University of London, Egham, Surrey, TW20 0EX

Abstract. We report preliminary results on separating $q\overline{q}$ continuum background events where $q = u, d, s$ or c from $b\bar{b}$ signal events produced in e^+e^- interactions at the $\Upsilon(4S)$ resonance. The analysis uses an Artificial Neural Network (NN) and is based on a Monte Carlo simulation without consideration of detector effects. It was found that for a signal efficiency of 90%, only $\sim$ 16% of the background events survive the NN cut.

1. Introduction

Artificial Neural Networks (NN) have been applied to a variety of problems in high energy physics in order to discriminate between different classes of events. Here the technique is used to separate $b\bar{b}$ events from $\Upsilon(4S)$ decays ('signal') from the continuum background for e^+e^- interactions at a centre-of-mass energy $E_{\rm cm}$ = 10.58 GeV, i.e., at the $\Upsilon(4S)$ resonance. The analysis relies on the different topologies for signal and background at a B Factory: while $b\bar{b}$ events are more 'spherical', uds and $c\bar{c}$ events are more 'jet-like'. This is due to the fact that in the $\Upsilon(4S)$ rest frame, the momenta of the produced B mesons are small and they decay isotropically. On the other hand, events where light quarks ($u\overline{u}$, $d\overline{d}$ or $s\overline{s}$) are produced are characterized by a preferred direction with hadron jets following roughly the quark and antiquark. A similar jet-like structure is also present in $c\overline{c}$ events, although there it is less pronounced.

For this study, two samples of 600k events of the type $e^+e^- \rightarrow q\overline{q}$ have been generated using PYTHIA [1] all at a centre-of-mass energy equal to the mass of the $\Upsilon(4S)$, where the quark flavour q is a mixture of u, d and s in the first set and c in the second. A third data sample of the same size consisted of $\Upsilon(4S)$ decays. The analysis has been performed at the level of the generated particles and no attempt has been made at this stage to simulate detector effects.

2. The shape variables

The variables used to discriminate jet-like from isotropic events are described below. Further information on the variables can be found in [2, 3].

(a) The ratio $R_2 = H_2/H_0$ of the 2nd to the 0th Fox-Wolfram moments. Neglecting particle masses, 4-momentum conservation requires that $H_0 = 1$. For a two-jet event, $H_1 = 0$ and $H_l \sim 1$ for l even and $H_l \sim 0$ for l odd. The R_2 distribution is shown in Fig. 1, where the different behaviour for signal and background events can be observed.

(b) The thrust $T = \sum_i |\hat{T} \cdot \mathbf{p_i}| / \sum_i |\mathbf{p_i}|$, where $\hat{T}$ is the thrust axis of the event [2]. For an isotropic event one has $T = 0.5$ while $T \sim 1$ for a highly directional one, as shown in Figs. 2 (a) and (c).

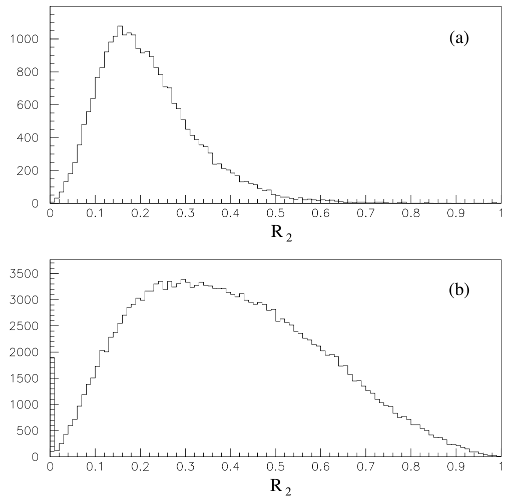

Figure 1. R_2 distribution for (a) Υ(4S) signal and (b) continuum events.

The variables defined above are quite powerful for identifying events where the two-jet topology is well defined and clearly distinguishable from the isotropic case. The signal to background separation decreases, however, when one or more gluons are emitted, which is more likely to happen when light quarks are produced. In this case there are three or more jets, so that the event shape resembles more that of the signal.

To discriminate further between signal and background, the Durham clustering algorithm [4] is applied to group particles into jets. For every pair of particles i and j, the 'distance' y_{ij} is computed as

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

where E_i and E_j are the particles' energies, θ_{ij} is their opening angle and E_{vis} is the total (visible) energy in the event. The pair with the smallest y_{ij} is replaced by a pseudo-particle with four-momentum $p^\mu = p_i^\mu + p_j^\mu$. The procedure is repeated until N jets have been found. In our case $N = 3$ or 4 and the smallest y_{ij} values are called y_3 and y_4, respectively. Once the clustering has been done, the following discriminating variables can be added to the list:

(c) y_3, y_4;

(d) θ_{BZ}, the angle between the highest and second highest energy jet;

(e) $\theta_{\rm KS}$, the angle between the lowest and second lowest energy jet;

(f) $\theta_{\rm NR}$, the angle between the planes defined by the highest and lowest energy jets and by the other two jets;

(g) θ_{34}, the angle between the planes defined by the highest and second highest jets and by the other two jets;

(h) the QCD four-jet matrix element squared, $|M_{QCD}|^2$, averaged over all possible assignments of the jets to the final-state partons [5].

As an example, the distribution of $\ln|M_{QCD}|^2$ is shown in Fig. 2(b) and (d) for signal and background events.

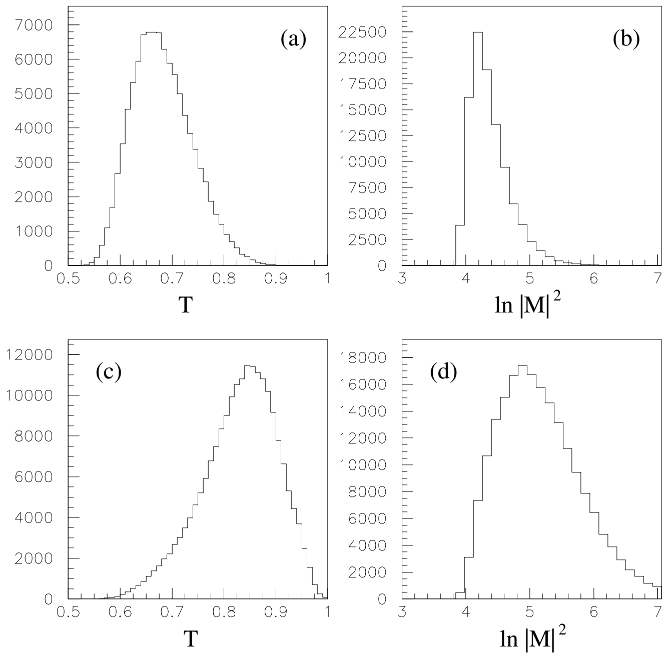

Figure 2. Distribution of (a) thrust and (b) $\ln|M_{QCD}|^2$ (right) for signal events; (c) and (d) show the same for continuum events.

3. The analysis results

All the variables defined in the previous section have been combined in a Neural Network using the package JETNET 3.51 [6]. The network is a 3 layered feed-forward one, with 9 input nodes, 12 hidden nodes and one output. The network has been trained using 100k $b\bar{b}$, 100k uds and 100k $c\bar{c}$ events. Each input variable has been scaled linearly so that it lies in the range [-1,1]; in this way at the start of the training no single variable dominates the inputs to any neuron. The training algorithm is 'error back-propagation'.

After the training, the NN has been used with an independent sample of events, 500k for each type, and the output distributions are shown in Fig. 3. The signal and background curves are very well separated, with continuum background events accumulating around 0 and signal events around 1.

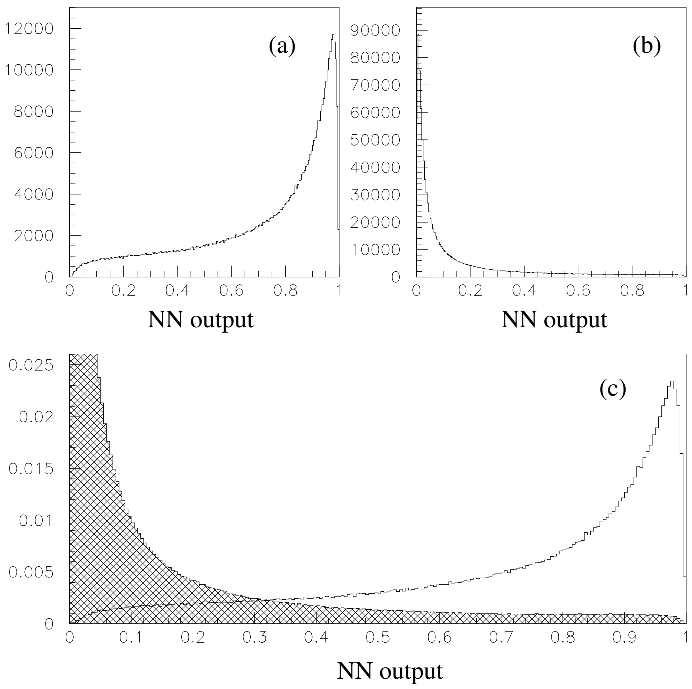

Figure 3. Distributions of the NN output for (a) signal and (b) background events. The distributions are superimposed in (c) after being normalized to unity.

To evaluate the rejection power of the NN we can set a cut on the output such that the efficiency for the signal is 90%; this corresponds to requiring an output greater than 0.35. The fraction of background events surviving the cut is 14% for uds events and 18% for $c\bar{c}$ events.

4. Conclusions

A feed-forward neural network has been applied to separate $q\overline{q}$ continuum events from Υ(4S) decays produced in simulated e^+e^- interactions at the $\Upsilon(4S)$ resonance. Preliminary results are quite encouraging. For a signal efficiency of 90%, only $\sim$ 16% of background events survive the NN cut. More work needs to be done to verify that the results are still valid after a full detector simulation; nevertheless, the NN technique seems a promising option for this kind of analysis.

References

[1] T. Sjöstrand, *Computer Physics Commun.* **82** (1994) 74
[2] P.F. Harrison and H.R. Quinn (eds.), *The BaBar Physics Book: Physics at an Asymmetric B Factory*, SLAC-R-504 (1998) 183
[3] R. Barate *et al.* (ALEPH Collaboration), *Physics Reports* **294** (1998) 1
[4] W. J. Stirling *et al.*, *J. Phys.* **G**: *Nucl. Part. Phys.* **17** (1991) 1567
[5] R. K. Ellis, D. A. Ross and A. E. Terrano, *Phys. Rev. Lett.* **45** (1980) 1225; *Nucl. Phys.* **B178** (1981) 421
[6] C. Peterson, T. Rögnvaldsson, L. Lönnblad *Computer Physics Commun.* **81** (1994) 185