Neural Networks for Higgs Physics

Silvia Tentindo-Repond[b], Pushpalatha C. Bhat[a], Harrison B. Prosper[b]

[a]*Fermi National Accelerator Laboratory* , [b]*Department of Physics, Florida State University, Tallahassee, FL 32306*

Abstract. The main application of neural networks (NN) in Higgs physics so far has been to optimize the signal over background ratio. The positive result obtained imply that the use of NN will lead to a big reduction in the integrated luminosity required for the discovery of the Higgs in RunII. Neural Networks have also been recently used in Higgs physics to set up tagging algorithms to identify the heavy flavor content of jets. Whereas in the previous studies the NN b-tagging methods used are channel-independent, a channel-dependent method has been used in the present work. The signal $p\bar{p} \to WH \to \ell\nu b\bar{b}$ has been studied against the dominant background $p\bar{p} \to Wb\bar{b}$, in an attempt to improve the signal over background ratio by trying to push the invariant mass of the background events further away from the signal. This result would get the equivalent effect of an improved mass resolution.

INTRODUCTION

Higgs Physics

The Higgs boson is the last missing particle in the Standard Model (SM) that would complete the description of almost all phenomena in elementary particle physics in a simple way, and would explain the fundamental mechanism that governs the electro-weak symmetry breaking, the process that provides mass to all fermions.

The SM parameters, the W boson and the top quark masses, are related to the Higgs boson mass. Experimental high precision electroweak measurements constrain the Higgs mass to $M_{H_{SM}} = 107^{+67}_{-45}$ GeV/c^2 and give a 95% confidence level upper limit of 225 GeV/c^2 (1). Some theories like the Minimal Supersymmetric SM (MSSM) predict an upper limit of 130 GeV/c^2.

The low mass of the Higgs boson motivates its search at the Tevatron during Run II. Typical predicted cross sections, for example, for a Higgs mass of 100 GeV/c^{2}are of the order of 1 pb; cross section times branching ratios for the processes studied here are 85 fb (WH) and 3500 fb for the background (W bb). Typical luminosities required at the Tevatron for a signal exclusion at 95% confidence level, or for 3σ or 5σ discovery combining both CDF and D0 data, and all the decay channels, are of the order of respectively : 10, 20 and 40 fb^{-1}, for a 130 GeV/c^{2}Higgs boson.

Dominant decay channels, and detection challenges (b-tagging)

The dominant decay channel for a low mass Standard Model Higgs, up to 130 GeV/c^2, is the $b\bar{b}$ channel, with a branching ratio of about 85%. The $c\bar{c}$ channel and $\tau\tau$ channels are one order of magnitude lower. At higher energies the WW and $t\bar{t}$ decay channels prevail, imposing different detection constraints. For a low mass SM Higgs it is essential to have a good b tagging method.

b-tagging algorithms that enhance the efficiency for detection of the signal $WH \to \ell\nu b\bar{b}$ do nothing to reduce the dominant background $Wb\bar{b}$. A method that discriminates b-tagged jets resulting from Higgs decay, from b-tagged jets not coming from Higgs, is needed. Neural networks can be trained to do this task.

MULTIVARIATE VERSUS TRADITIONAL METHODS IN HIGGS PHYSICS

One of the most important goals of multivariate methods is to improve the ratio of signal over background. This task becomes particularly critical when the signal is very small, and when its properties are very similar to the properties of the background.

In Higgs physics, the first application of multivariate methods has been to maximize the chance to discover the Higgs boson, and to reduce the required luminosity for a 3σ or 5σ discovery (2, 3).

Studies show that in RunII, for a given integrated luminosity, the NN analysis, as compared to the traditional analysis, provides a much higher discovery reach in mass. Typically, for a Higgs mass of 130 GeV/c^2, a factor 2 better in integrated luminosity is reachable.

Another important application of NN in Higgs physics has been to improve the b-tagging efficiency. Along with tagging leptons it will be possible in RunII to do b-tagging by using b lifetime variables. By traditional methods, the efficiency for b-tagging through secondary vertex and impact parameter reconstruction is estimated to be about 45% per single b particle; if two b jets are tagged, this efficiency becomes critically low, and improved methods of tagging are needed. Besides, the method of traditional b-tagging will not be able to discriminate the b component from the c component (about 10% of heavy flavor jets are c-jets for Higgs up to masses of 130 GeV/c^2). A Neural network analysis has been used (9) to combine lifetime variables (tracks consistent with secondary vertex, impact parameter), and kinematic variables (mass, fragmentation). This tagging method can potentially outperform existing tagging algorithms.

Another potential field of application of NN in Higgs physics could be to improve the Higgs invariant mass resolution (P.C.Bhat, APS '99). The di-jet invariant mass variable has proven to be a critical one to discriminate signal from background in Higgs physics. The assumed mass resolution in the recent Run II Susy/Higgs Workshop (5) is 10%.

The mass resolution gets contributions from different processes, at the parton level, particle level and detector level. Improving the mass resolution will be a very challenging task, and the appropriate methods have yet to be worked out to reach the expected resolution. A method that could reduce the background, and at the same time increase the signal-background separation in the mass spectrum, for example, by pushing the background away from the signal, would reach the same goal as to increasing the mass resolution. This is the new application that is attempted here using NN.

CHANNEL-DEPENDENT B-TAGGING

We remark that the methods developed so far (2, 9) for b-tagging with NN are channel-independent , in the sense they are applicable independently of the decay channel studied. In the following we want to investigate the possibility to develop an NN method optimized to a given decay channel. We will call this " channel-dependent b-tagging ", the aim of which is to reduce the background.

Simulation

Here we consider the signal $p\bar{p} \to WH \to \ell\nu b\bar{b}$, and its main background $p\bar{p} \to Wb\bar{b}$.

The WH events were simulated using the PYTHIA program (6) for Higgs boson masses of M_H = 100, 120 and 130 GeV/c^2. The $Wb\bar{b}$ sample was generated using CompHEP (7), with fragmentation done using PYTHIA. To model the expected response of the CDF/DØ Run II detectors we used the SHW program (8).

An initial loose set of cuts on the variables is imposed, based on knowledge of the detector properties, trigger thresholds and kinematic properties of the events of interest:

- the transverse momentum of the isolated lepton $P_T^\ell >$ 15 GeV/c
- the pseudo-rapidity of the lepton $|\eta_\ell| < 2$
- the missing transverse energy in the event $\not{E}_T >$ 20 GeV
- two or more jets in the event with $E_T^{jet} > 10$ GeV and $|\eta_{jet}| < 2$.

NN Analysis

The feed-forward NN package used is MLPfit (version 1.40) (4). The NN configuration chosen here is 3-6-1, 3 input nodes, 6 hidden nodes and one output. One sample (the WH signal data) is used to train the NN, for b-tagged jets (NN output set to 1) and for non b-tagged jets (NN output set to 0).

The set of variables that has been chosen as input to the NN is the following:

- E_T^b – transverse energy of the b-tagged jets.
- N_{trk} – charged track multiplicity in the jets.
- $Width$ – width of jets.

In Fig. 1 are plotted the jet variables used for training the NN, for signal and for background.

The NN function is then used to discriminate b-tagged jets coming from the signal $p\bar{p} \to WH \to \ell\nu b\bar{b}$ from b-tagged jets coming from the background. The hope is that the b-jets in the signal are identified with higher efficiency than those in the background.

In Fig. 2 is plotted the NN output for the most energetic jet : 916 out of 1000 of the most energetic jets are b-jets. (For the second most energetic jet, 787/1000 are b-jets).

These tagged b jets are channel dependent b-tags. The NN output function if applied to the background sample, peaks at 0. The invariant mass plot can be used to see the discrimination power of the NN output . By imposing a N-N cut at 0.4 to jet1 and jet2, 65% (M_{Higgs} = 100 GeV/c^2), 74% (120GeV/c^2) and 58% (130GeV/c^2) signals survive, whereas the background is reduced to 24%.

These studies have just begun, and the method has not been optimized yet. The result of this NN method brings no mass resolution improvement since that is not its purpose. It tries to reduce the background by pushing the remaining background events lower in the invariant mass spectrum.

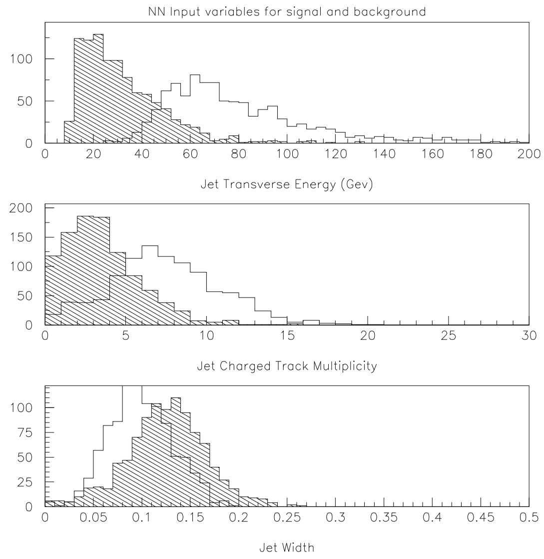

FIGURE 1. Jet variables used as input for training NN in channel-dependent b-tagging studies for signal (unshaded) and background (shaded)

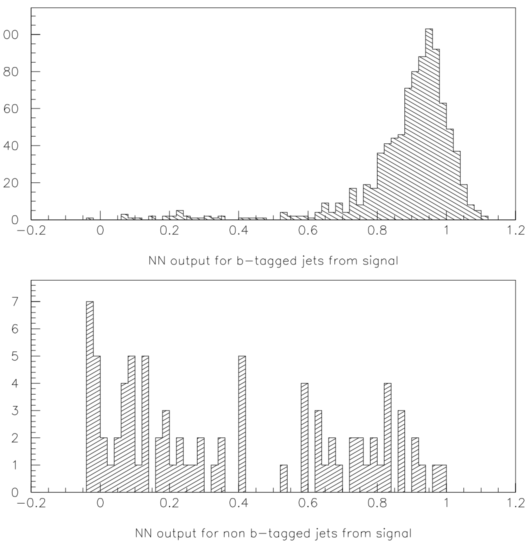

FIGURE 2. NN output distribution for most energetic b-tagged jet and most energetic non b-tagged jet

SUMMARY

Neural networks methods in Higgs physics have been used to maximize discovery potential. They have been used as a b-tagging algorithm and specifically as a discriminator of charm and bottom content in heavy flavor jets. Channel - dependent b-tagging is being studied and has been applied to the channel $p\bar{p} \to WH \to \ell\nu b\bar{b}$ against the dominant background $p\bar{p} \to Wb\bar{b}$, and better separation between signal and background has been obtained. We plan to continue the systematic study of these NN methods, as well as those that may improve the di-jet invariant mass resolution.

REFERENCES

1. LEP Electroweak Group, http://www.cern.ch/LEPEWWG /plots/summer99.
2. Bhat, P. C., Gilmartin, R., Prosper, H. B., *Phys.Rev.D, 62, 074022(2000).*
3. Bhat, P. C., these proceedings.
4. MLPfit v1.40, Schwindling, J., Mansoulie, B. http://home.cern.ch/ schwind/MLPfit.htlm
5. Run II Higgs Working Group of the Run II SUSY/Higgs workshop, http://fnth37.fnal.gov/higgs.html; T. Han, A. S. Turcot and R. Zhang, *Phys. Rev.* D **59**, 093001 (1999).
6. PYTHIA, T. Sjöstrand, *Comp. Phys. Comm.* **82***, 74 (1994).*
7. CompHEP, A. S. Belyaev, A. V. Gladyshev and A. V. Semenov, e-print hep-ph/9712303; E.E. Boos *et al.*, e-print hep-ph/9503280.
8. SHW v2.0, Conway, J. - available at http://www.physics.rutgers.edu/∼jconway/soft/shw/shw.html (unpublished).
9. Wolinski, D., Amidei, D., Demina, R.- CDF/ANAL/EXOTIC/CDFR/5051 (Sept 1999).