Neural network for mass reconstruction of resonance particle with missing energy*

Zhang Zi-Ping (张子平)
(*Department of Modern Physics, University of Science and Technology of China, Hefei 230026*)

Abstract Neural Network can be designed to reconstruct the mass of resonance particle with large missing energy. Taking the Higgs particle search through decaying channel $H^0 \to \tau^+\tau^- \to e\mu x$ and $H^0 \to W^+W^-(ZZ) \to ll\nu\nu$ at LHC collider ($\sqrt{s}$=16 TeV) as examples, neural network correctly reconstructs its mass with right peak position and better width than conventional method. The network also possesses the capability of suppressing background events. This kind of neural network can be widely used in new particle search and precise mass measurement of resonance particle.

Keywords: Neural network, Mass reconstruction, Resonance particles, Higgs search

1 Introduction

Recent years Neural Network (NN) has shown a great power in pattern recognition, particle identification, track reconstruction, etc., and is widely successfully used in high energy physics data analysis[1–5].

In nuclear and particle physics experiment there are varieties of resonance particles whose decayed final states have large energy loss carried away by e.g. neutrinoes, so it is usually very difficult, or even entirely impossible to reconstruct its mass by conventional method. Here we studied the application possibility of NN in mass reconstruction of this kind of resonance particle, by taking the Higgs particle produced at LHC pp collider ($\sqrt{s} = 16$TeV) as an example, because Higgs search will be the most important task of LHC project to reveal the origin of spontaneously breaking symmetry mechanism of standard model, if the ming LEP *II* will not find this particle.

As standard model can not give a prediction of Higgs mass, if it lies in the intermediate mass range, i.e. 80GeV< M_H < 140 GeV, $H^0 \to \tau^+\tau^- \to e\mu x$ is a possible decay channel. Because of the large missing energy carried away by neutrinoes, in order to be able to reconstruct its mass, the machine can only be run at low luminosity to avoid the pile-up effect, in this case since lorentz boosts the final neutrino pair and charged lepton from τ are almost at the same direction, so the charged lepton direction can be taken as the τ's. From the following formula the neutrino pair transverse momentum $p^{\nu}_{T_1}$ and $p^{\nu}_{T_2}$ from each τ can be calculated,

$$p^{\nu}_{T_1} \cdot \vec{u}_{T_1} + p^{\nu}_{T_2} \cdot \vec{u}_{T_2} = \vec{p}_T^{\,miss} \qquad (1)$$

with $\vec{p}_T^{\,miss}$ being the total transverse momentum loss of pp collision, $\vec{u}_{T_1}$ and $\vec{u}_{T_2}$ the unit transverse momentum vectors of charged leptons, then τ's energy and final Higgs mass can be reconstructed. Solution of Eq.1 requires the acolinearity of $\vec{u}_{T_1}$ and $\vec{u}_{T_2}$, i.e. $|\cos(\phi_e - \phi_\mu)|$ cut is needed to select Higgs with big transverse momentum; therefore, the event rate is dramatically decreased[1–5].

In the case of heavy Higgs search ($M_H > 2M_Z$) one important channel is $H^0 \to W^+W^-(ZZ) \to ll\nu\nu$, which has the advantage of 6 times bigger cross section than $H^0 \to (ZZ) \to llll$. But because W and Z are heavy, the final escaped neutrinoes are not colinear with them; therefore, it is entirely impossible to reconstruct this heavy Higgs mass by conventional method. The event signal can be only shown by the broad peak of p_T^Z or transverse mass m_T,

$$m_T^2 = 2 \cdot p_T^Z \cdot p_T^{\nu}(1 - \cos\Delta\phi_Z^{\nu}) \qquad (2)$$

where p_T^{ν} being the total missing energy, $\Delta\phi_Z^{\nu}$ the azimuth between transverse momentum of Z and ν.

2 Monte Carlo simulation

Data samples of different processes are

*The Project Supported by National Natural Science Foundation of China
Manuscript received date: 1995–12–20

produced by PYTHIA5.6 generator[9], EHLQ1 (Λ=290MeV) structure function is taken. Assuming that the detector rapidity coverage range is $|\eta| < 3$ and it can precisely measure the momentum of e and μ track with 100% efficiency, A 10GeV/c lower limit cut is put on the transverse momentum of final state charged lepton tracks. Some other parameters and assumptions are as follows:

Electromagnetic calorimeter: resolution is taken as $\sigma/E = 0.15/\sqrt{E} + 0.01$.

Hadron calorimeter: resolution is taken as $\sigma/E = 0.5/\sqrt{E} + 0.02$.

Jets: detector coverage in ϕ range of $[0,2\pi]$ and η range of $|\eta| < 3$ is divided into 100 sub-range seperately giving total 10^4 cells, i.e. the granularity is $\Delta\phi \times \Delta\eta = 0.062 \times 0.06$. Jets are reconstructed by somehow modified LUCELL algorithm, using all the effective hit cells of energy deposit larger than 1.5 GeV, as well as taking the jet cone size as $R \equiv \sqrt{(\Delta\ \phi)^2 + (\Delta\ \eta)^2} = 0.5$. The reconstructed jet's transverse momenta are required to be larger than 10 GeV/c.

The total missing transverse momentum of event $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ can be obtained either from all the reconstructed jets or effective cells.

As for the pile-up effect study, switches in PYTIIIA are turned on to take account of low p_{T} + double and single diffractive processes of total 15 pile-up events, which corresponds to LHC luminosity $\mathcal{L} \sim 1.2\times10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ of 15 ns inter-bunch spacing.

3 Mass reconstruction

3.1 Network structure and training

For this purpose a three layered NN is designed with 13 input neurons, 15 hidden neurons and only one output neuron to give invariant mass information. With M_{H} equally distributed in the range [50GeV, 220GeV], 33000 $\mathrm{H}^0 \rightarrow \tau^+\tau^-$ M.C. events and with M_{H} in the range [200GeV, 750GeV], 15000 $\mathrm{H}^0 \rightarrow \mathrm{W}^+\mathrm{W}^-(\mathrm{ZZ}) \rightarrow ll\nu\nu$ events, are produced as the training sets, respectively.

Of observable physical variables, 13 are selected as the input of this NN, which are: (1) N_{JET}, total number of jets. (2) $p_{\mathrm{T}}^{\mathrm{l}_1}$, transverse momentum of first charged lepton. (3) $p_{\mathrm{T,iso}}^{\mathrm{l}_1}$, transverse momentum of the jet, to which first lepton belongs, substracts $p_{\mathrm{T}}^{\mathrm{l}_1}$. (4,5) $p_{\mathrm{T}}^{\mathrm{l}_2}$, $p_{\mathrm{T,iso}}^{\mathrm{l}_2}$, the same as above for the second lepton's. (6) $|\phi_{\mathrm{l}_1} - \phi_{\mathrm{l}_2}|$, the ϕ-angle difference of the first and second lepton track. (7) $|\eta_{\mathrm{l}_1} - \eta_{\mathrm{l}_2}|$, the rapidity difference of the first and second lepton track. (8) N_{HIT}, the number of effective cells with deposited energy of more than 1.5 GeV in it. (9) $E_{\mathrm{T}}^{\mathrm{total}}$, total transverse momentum of the event. (10) $E_{\mathrm{T}}^{\mathrm{EM}}$, total transverse electromagnetic momentum of the event. (11,12) E_x^{miss}, E_y^{miss}, the x and y components of missing transverse energy calculated from reconstructed jets of the event. (13) $E_{\mathrm{T,max}}^{\mathrm{jet}}$, the largest transverse energy of jet which contains no lepton track. These observables not only more or less carry mass information, but also are able to suppress background events contamination. Each of them is properly nomalized to [0,1]. The target value of output neuron is designed to be $(\lg M_{\mathrm{H}} - 3.91)/1.51$ for intermediate Higgs, and $(\lg M_{\mathrm{H}} - 5.2)/1.6$ for heavy Higgs, respectively corresponding to the M_{H} range to make it in [0,1].

The back-propagation training algebra is almost the same as Ref.[10]. The linking weights W_{ij}'s and threshold θ_j's are initialized randomly in [-0.1,0.1]. At the very beginning learning strength η and "momentum" parameter α are taken to be 0.07 and 0.5. As the training goes η is gradually reduced every 10000 loops according to

$$N_{k+1} = N_K \cdot 0.99 \tag{3}$$

but not less than 0.0001. During training the error change

$$\sigma = \sqrt{< (M_{\mathrm{H}}^{\mathrm{NN}}/M_{\mathrm{H}}^{\mathrm{O}})^2 > - < (M_{\mathrm{H}}^{\mathrm{NN}}/M_{\mathrm{H}}^{\mathrm{O}}) >^2} \tag{4}$$

is watched every 30 000 loops, where $M_{\mathrm{H}}^{\mathrm{NN}}$ and $M_{\mathrm{H}}^{\mathrm{O}}$ are the NN output and target values of M_{H} respectively, to make sure it goes down all the way until reaching flat, otherwise to stop it, this is out of the consideration that in a limited number of loops σ may fluctuate but should keep decreasing trend for every enough number of loops. The weights and thresholds are updated for every 10 loops. After first training run if the result is not so optimal, the output weights and thresholds are taken as initial values to restart next run with newly adjusted η and α parameters in order to save CPU time.

3.2 Test and results

At the testing stage of NN intermediate Higgs events of M_H equal 110 GeV and 130 GeV, are seperately fed into the trained NN input to see how is its performance. Fig. 1 gives the results and comparisons with conventional mass reconstruction method in which (a, b) for $M_H = 110$ GeV, (c, d) 130 GeV, (a, c) without $|\phi_e - \phi_\mu|$ cut, (b, d) with this cut. It can be seen that this NN correctly reconstructs the Higgs invariant mass peak position and better width. With $|\phi_e - \phi_\mu| < 0.8$ cut the NN result is consistent with that of conventional method except the later has a peak shift towards low mass end, which can be improved a bit by using the missing transverse energy calculated from all the hit cells as mentioned above. Fig. 2 is the test result by using (a) $t\bar{t}$, (b) W^+W^- and (c) Drell-Yan events as input, which are the backgrounds of $H^0 \to \tau^+\tau^- \to e\mu x$. The Z^0 mass peak is well shown in Drell-Yan's mass spectrum, but no peak in $t\bar{t}$ and W^+W^- ones, so it exhibits the background suppressing ability of this network.

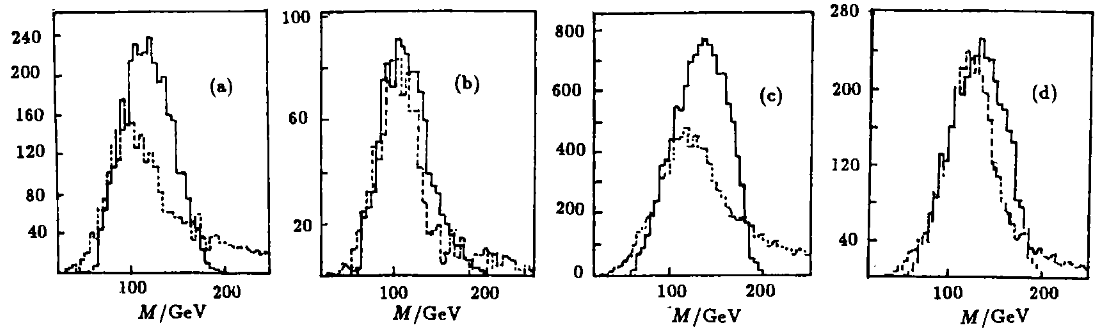

Fig.1 Intermediate M_H reconstructed by NN, dotted lines are results by conventional method (a) $M_H = 110$ GeV, no $|\cos(\phi_e - \phi_\mu)|$ cut; (b) $M_H = 110$ GeV, with $|\cos(\phi_e - \phi_\mu)| < 0.8$ cut; (c) $M_H = 130$ GeV, no $|\cos(\phi_e - \phi_\mu)|$ cut; (d) $M_H = 130$ GeV, with $|\cos(\phi_e - \phi_\mu)| < 0.8$ cut

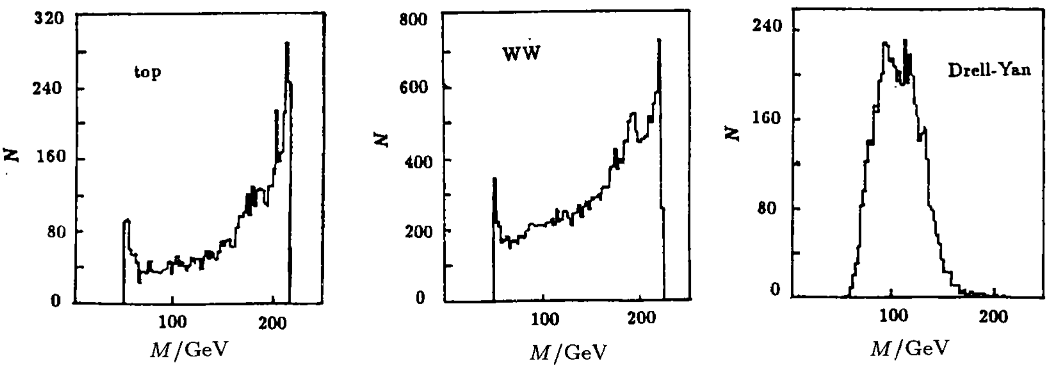

Fig.2 Invariant mass of (a) $t\bar{t}$, (b) W^+W^- and (c) Drell-Yan reconstructed by NN of intermediate Higgs

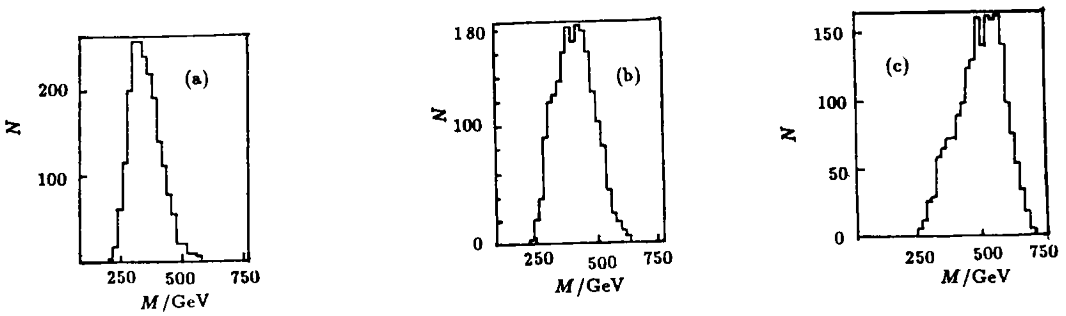

Fig.3 Heavy M_H reconstructed by NN, (a) $M_H = 300$ GeV, (b) $M_H = 400$ GeV, (c) $M_H = 500$ GeV

The same test procedure is done for heavy Higgs. Fig. 3 is the results of $M_{\mathrm{H}} = 300\,\mathrm{GeV}$, 400 GeV and 500 GeV without any cuts. It shows the correct peak position, and their width reasonably increases as M_{H} goes up. The asymmetries of 300GeV and 500GeV mass spectra come from asymmetry of training sample around these two mass points, which is also reasonable. Fig. 4 shows the test result of heavy Higgs background events (a) $t\bar{t}$ and (b) W^+W^-, *which is consistent with that of Fig.2.*

3.3 Mass reconstruction for pile-up events

As already mentioned that the crucial problem for $\mathrm{H}^0 \to \tau^+\tau^-$ search at high luminosity comes from the impossibility of Higgs mass reconstruction by conventional method because of pile-up effect. Here NN method, in principle the same as before, is adopted to see if it helps. Using a training sample of M_{H} equally distributed in [50 GeV,180 GeV] with the pile-up parameters given in Section 2, test results for $M_{\mathrm{H}} = 100\,GeV$ and 130 GeV events are *shown in Fig.5, from which the correct peak* position can be seen.

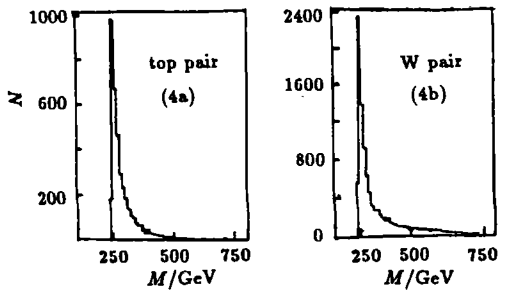

Fig.4 Invariant mass of (a) $t\bar{t}$, and (b) W^+W^- reconstructed by NN of heavy Higgs

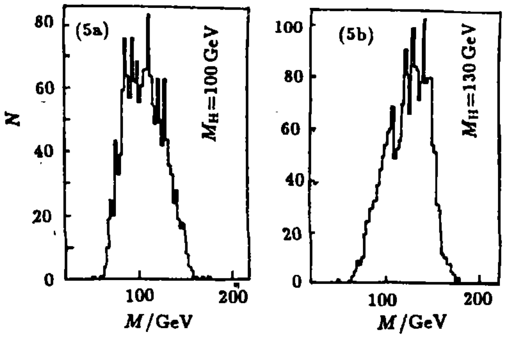

Fig.5 Invariant mass of (a) M_{H}=100 GeV, (b) M_{H}=130 GeV reconstructed by NN of intermediate Higgs for pile-up events

4 Conclusion and discussion

Based on M.C. simulation data, this study demonstrates that NN method can be successfully applied to mass reconstruction of resonance particle, even there is large energy loss in its final state, which has special application importance in new particle search and precise mass measurement. In real data analysis the consistency of data and Monte Carlo have to be checked carefully. From the deep study of detector responses better physical variables representing events properties can be chosen to improve the NN performance.

At high luminosity run with existence of pile-up effect, NN method has particular significance as higher event rate may make the Higgs search in intermediate mass range through $\mathrm{H}^0 \to \tau^+\tau^- \to \mathrm{e}\mu x$ become possible.

References

1 Bellantoni L, Conway J S, Jacobsen J E *et al.*, CERN-PPE/91-80,1991

2 Innocente V, Wang Y F, Zhang Z P. Nucl Instr and Meth, 1992; A323:647

3 Bortolotto C, Angelis A De, Lanceri L. Nucl Instr and Meth, 1991; A306:459

4 Lönnblad L, Peterson C, Rögnvaldsson T. Phys Lett, 1992; B278:181

5 Joswig M, Kolanoski H, Kublun S *et al.*. DESY 93-167, preprint

6 Lella L D. CERN 90-10/ECFA 90-133, 1990; II:530

7 Bos K, Anselmo F, Eijk B. CERN 90-10/ECFA 90-133:538

8 Kunszt Z, Zwirner F. CERN 90-10/ECFA 90-133:578

9 Sjöstrand T. PYTHIA5.6 and JETSET7.3 physics and mannual, CERN-TH. 6488/92, 1992

10 Lönnblad L,Peterson C, Rögnvaldsson T. Nucl Phys, 1991; B349:675