MINOS-doc-3393

A Neural Network Technique for ν_e Identification

Tingjun Yang
Stanford University

August 2, 2007

Abstract

In this memo I describe a neural network based ν_e identification technique that has been developed for the MINOS ν_e appearance oscillation study.

1 Introduction

Several discriminating techniques have been developed in parallel to separate $\nu_\mu \to \nu_e$ signals from the background. The average longitudinal profile of the energy deposition in an EM shower is reasonably well described by a gamma distribution [1]:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}{}^{1} \tag{1}$$

The EM showers are usually condensed while the hadronic showers are usually broader because of the large fluctuations in nuclear interactions. The primary signature of a CC electron neutrino interaction in the MINOS detectors is a shower consistent with the electromagnetic cascade in the absence of a long track. The main background is NC interactions with π_0's. Other background contributions come from high-y ν_μ CC interactions where the muon track is short and invisible, the intrinsic ν_e component in the ν_μ neutrino beams. In the Far Detector, there is one more background source: ν_τ events from $\nu_\mu \to \nu_\tau$ socillations. The purpose of this document is to outline the ANN(Artificial Neural Network) based algorithm I have developed as an option for the ν_e appearance oscillation analysis.

[1] $t = x/X_0$, where X_0 is the radiation length.

2 ANN Input Variables

The following variables in the **NueAna** package have been used in constructing the neural network.

- **fracvars.fract_road** This is defined as the fraction of total visible energy deposited in a 3-strip wide road along the shower development. For ν_e CC interactions, this is peaked towards one because the EM showers are condensed and always have a hard core. In contrast, this fraction for NC and ν_μ CC interactions lies in a much broader distribution.

- **shwfit.par_a** We fit the longitudinal profile of the energy deposition in the shower with Eq. (1). This variable is the fitted parameter a which is related to the rise of the longitudinal shower profile.

- **shwfit.par_b** This is the fitted parameter b which is related to the fall of the longitudinal shower profile.

- **hitcalc.fHitLongEnergy** This is defined as the sum of the magnitudes of the longitudinal projections of the vectors defined by each hit and the vertex.

- **mstvars.e4w+o4w** This is one variable in the MST(Minimum Spanning Tree) analysis. It is defined as the sum of the weights in the MST formed from the larger than average pulseheight hits in the event.

- **subshowervars.PHFracRMSU+PHFracRMSV** This is one variable in the SubShower analysis.

Figure 1 shows the ANN input variables. Red histogram is the signal and blue histogram is the background. The ν_e CC Quasielastic(QE) and Resonance(RES) interactions, in which cases the electromagnetic shower is dominant in the event, are selected as the signals. The NC and ν_μ CC interactions are selected as the background. The following pre-selection cuts were applied:

- Vertex Contained in Fiducial Volume (anainfo.inFiducialVolume==1&&srevent.vtxZ[2]<29)

- Track Length Cut (srtrack.planes<14&&srtrack.trklikePlanes[3]<12)

- Energy Cut (srevent.phNueGeV>1&&srevent.phNueGeV<8)

The track length cut is to remove events with long tracks, which are mainly muons from ν_μ CC interactions. The energy cut is to remove events that have very high/low energy showers.

[2]The cut on vtxZ is only applied in the FD to remove the ν_μ CC backgrounds near the end of the detector.

[3]trklikePlanes is the number of "clean" hit planes in the track, a clean plane is one in which the plane is free from the presence of non-track associated hits

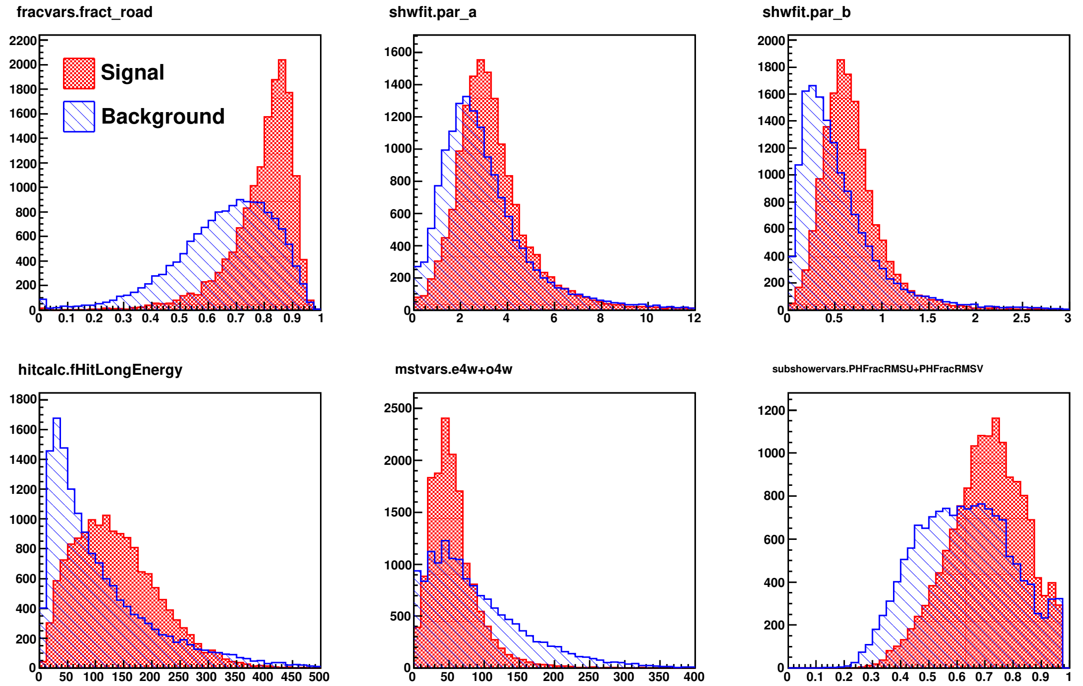

Figure 1: ANN Input Variables. The signals are ν_e CC QE interactions and the backgrounds are NC and ν_μ CC interactions.

3 Constructing the ANN

There is a growing interest in the use of algorithms based on artifical neural networks and related methods in order to perform data analyses in modern high energy physics experiments. An ANN is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. It is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning involves adjustments to the synaptic connections that exist between the neurones.

TMultiLayerPerceptron is a ROOT implementation of ANN. It is a simple feed-forward[4] network based on the MLPfit package. A typical structure of this ANN is shown in Figure 2. We use this ROOT class to build a neural network for the separation of ν_e signals from backgrounds in the ν_e appearance analysis. We use the 6 variables described before as the ANN input variables. We specify the ANN structure to be 6:6:4:1. It has one input layer, two hidden layers and one output layer. The input layer has 6 nodes which correspond to the 6 input variables. The output layer has one node which is the ANN output. One hidden layer has 6 nodes and the other has 4 nodes. We trained the ANN with ν_e CC QE and RES interactions as the signal sample and NC and ν_μ CC interactions as the the background sample. We also build another neural network with 30 input variables, the structure being

[4]Feed-forward ANNs allow signals to travel one way only, from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer.

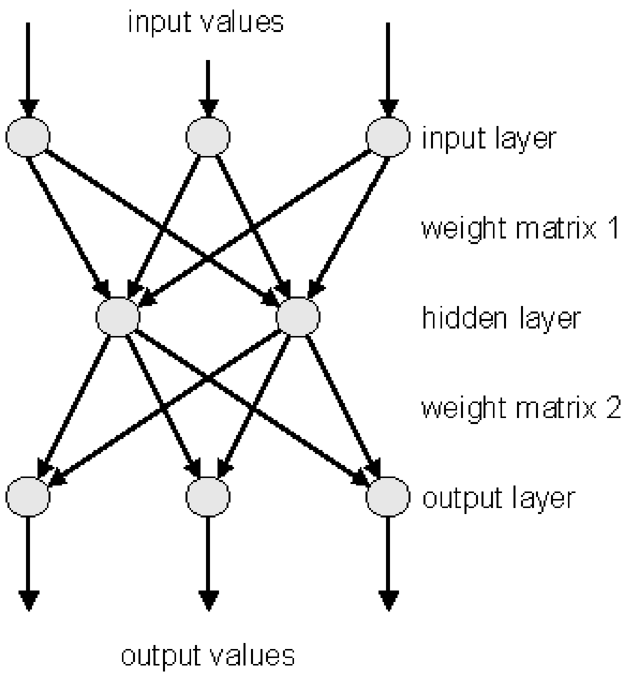

Figure 2: The architecture of an ANN. It is composed of an input layer, a hidden layer and an output layer.

30:15:15:1. It gives better signal/background seperation than the one with 6 input variables. Those 30 variables will be listed in the appendix. We will use **ANN6** and **ANN30** to denote the neural network with 6 and 30 input variables respectively.

Figure 3 shows the output of two neural networks, which is a likelihood of being a ν_e CC interaction. We require the ANN output to be greater than **0.5** for an event to be considered as a ν_e candidate.

4 Event Selection

In this section I will show several distributions related to the ANN selection. ANN30 is used unless otherwise specified. The Standard SKZP beam reweighting is applied [2]. Figure 4 shows the ANN PID of ν_e signals and different background components. The signal distribution is scaled up by a factor of 10 for clarity. The following oscillation parameters and exposure are assumed:

$$\Delta m^2 = 2.7 \times 10^{-3} eV^2, sin^2 2\theta_{13} = 0.1, sin^2 2\theta_{23} = 1, \mathrm{POT} = 4 \times 10^{20} \tag{2}$$

Figure 5 shows the visible energy distribution of ν_e-like events after cutting on the ANN PID: **ANN PID>0.5**. E_{vis} is the reconstructed energy assuming the entire shower is electromagnetic. Figure 6 and Figure 7 show the signal selection and background rejection at different cutting stages. Figure 8 shows the ν_e selection efficiency and purity as a function of visible energy with given oscillation parameters. Table 1 shows the data reduction. Figure

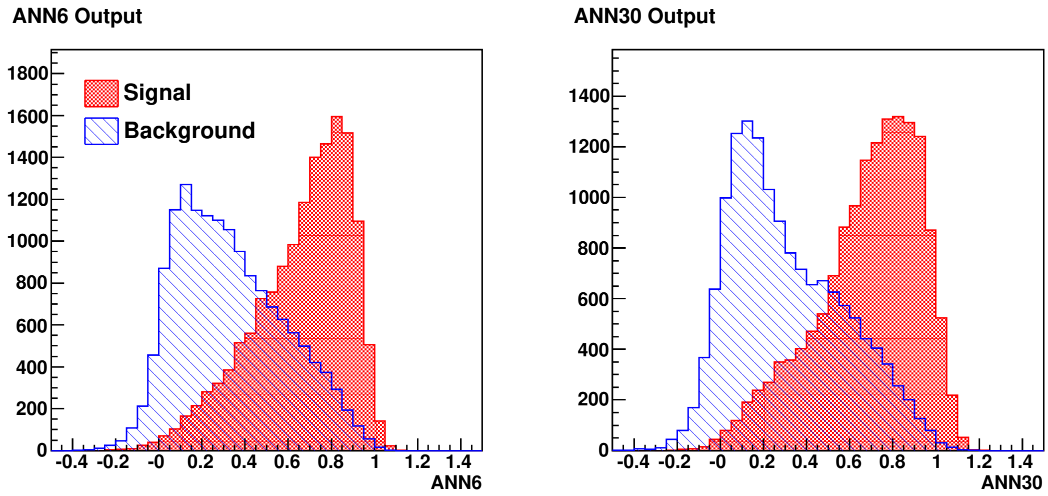

Figure 3: The output of two neural networks - ANN6 and ANN30. Pre-selection cuts are applied. The signals are ν_e CC QE interactions and the backgrounds are NC and ν_μ CC interactions.

	ν_e^{osc}	background	CC	NC	ν_e^{beam}	ν_τ	FOM
Fid. Vol. Cut	25.6	1741.2	1193.0	509.5	23.7	15.0	
Pre-selection	21.8	306.5	60.3	232.6	6.9	6.7	
ANN6 PID>0.5	13.3	78.5	13.2	58.4	4.1	2.7	1.51
efficiency,bkgd rej.	0.519	0.045	0.011	0.115	0.114	0.101	
ANN30 PID>0.5	13.4	74.5	12.1	55.7	4.1	2.6	1.55
efficiency,bkgd rej.	0.523	0.043	0.010	0.109	0.110	0.103	

Table 1: Summary of data reduction for $\Delta m^2 = 2.7 \times 10^{-3} eV^2, sin^2 2\theta_{13} = 0.1, sin^2 2\theta_{23} = 1$ and POT $= 4 \times 10^{20}$. The standard SKZP beam reweighting is applied. FOM=sig/sqrt(background)

of Merit (FOM) is calculated as the number of signal events divided by the sqare root of the number of total background events.

5 Summary

The ANN technique provides a useful method of selecting ν_e CC interactions for the ν_e appearance analysis. It can select 52% of the signals while rejecting 99% of the ν_μ CC and 89% of the NC backgrounds.

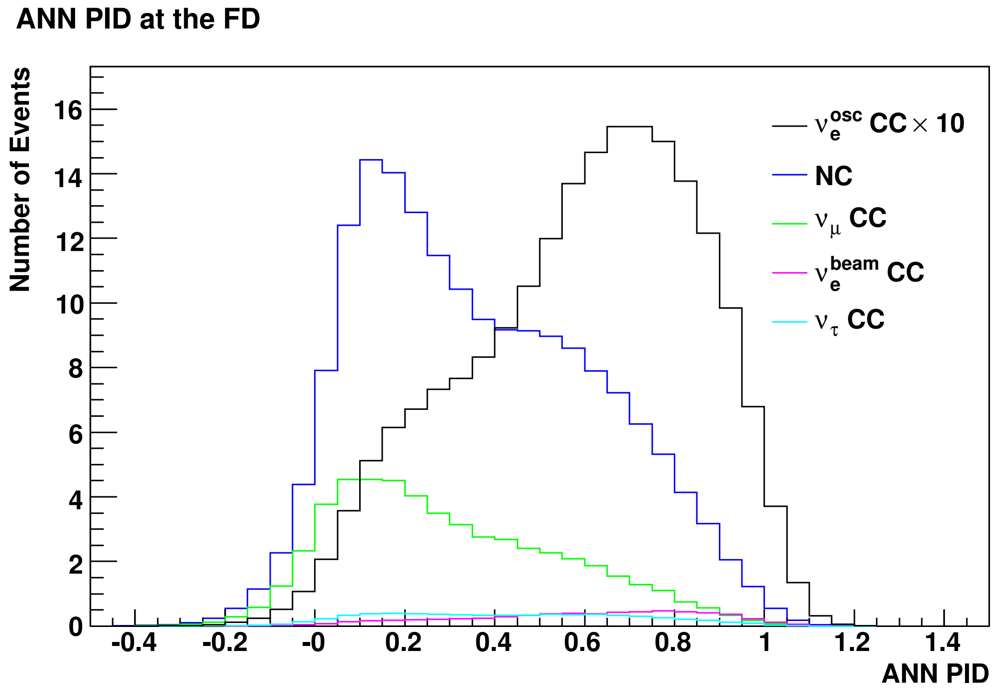

Figure 4: The ANN PID of signals and different background components. The signal distribution(black curve) is scaled by a factor of 10 for clarity. Pre-selection cuts are applied.

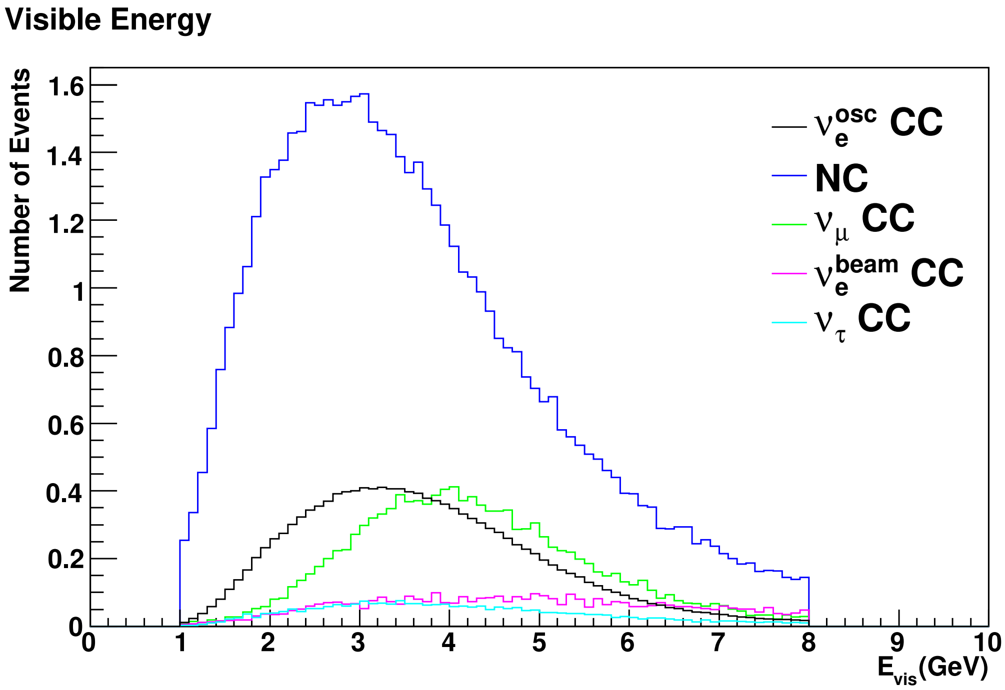

Figure 5: The visible energy distribution of ν_e-like events after making a ANN PID cut at 0.5.

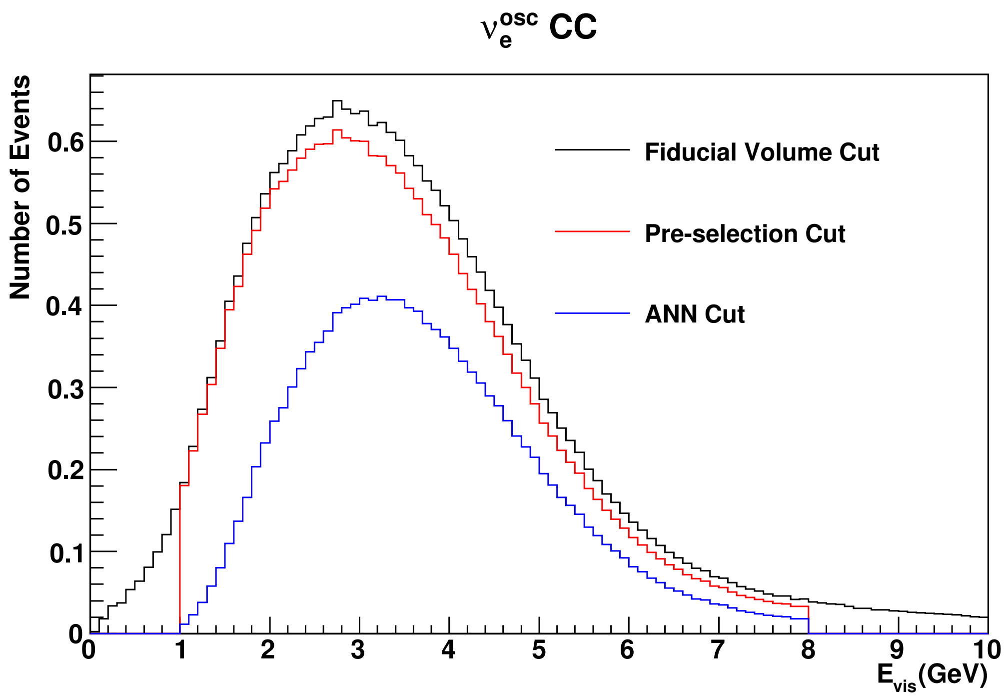

Figure 6: Signal selection.

References

[1] E. Longo and I. Sestili, Nucl. Instrum. Meth. **128**, 283 (1975) [Erratum-ibid. **135**, 587 (1976)].

[2] MINOS-doc-2965-v1

A 30 ANN Input Variables

- Shower Fit Variables(4)
  - shwfit.par_a - Fitted parameter a(shower rise).
  - shwfit.par_b - Fitted parameter b(shower fall).
  - shwfit.uv_molrad_peak - Moliere radius from peak.
  - shwfit.uv_rms - RMS in coordinate.
- Minimal Spanning Tree(MST) Variables(5)
  - mstvars.e4w, mstvars.o4w - The sum of the weights in the mst formed from the larger than average pulseheight hits in the event (actually the largest primary mst found) in the even or odd view.

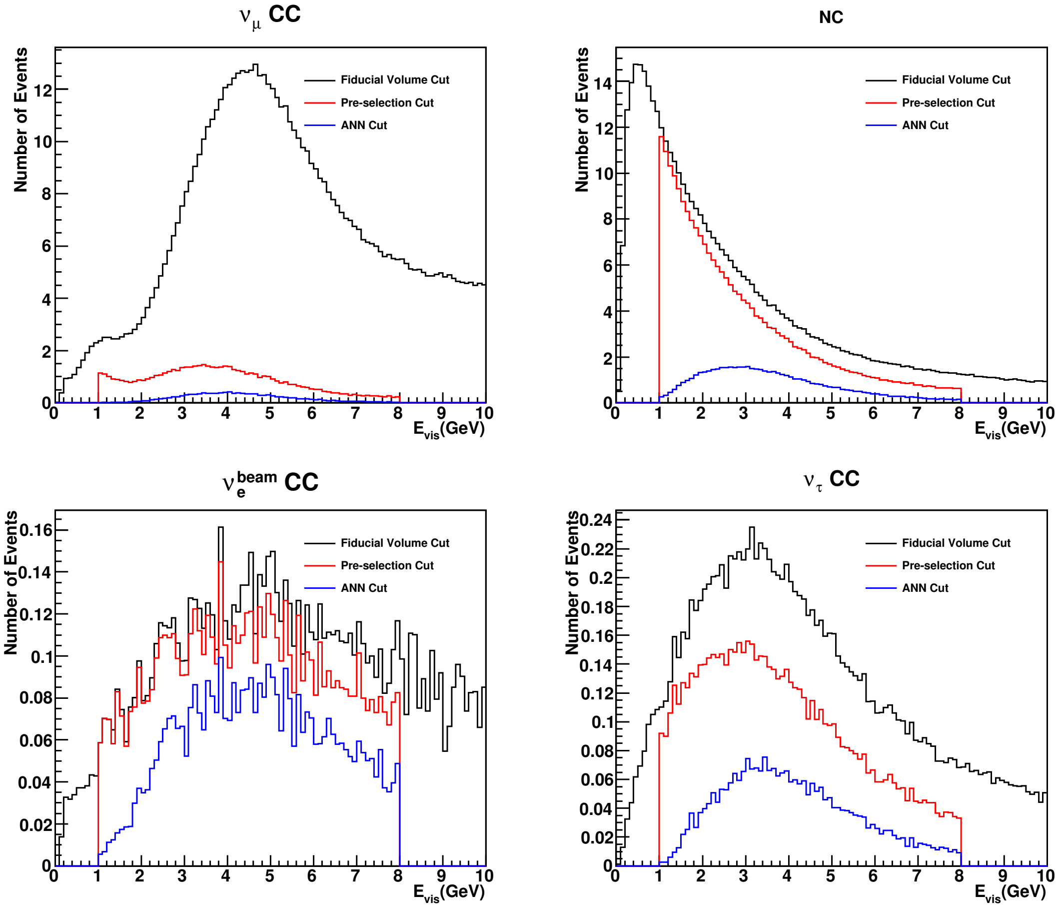

Figure 7: Background rejection.

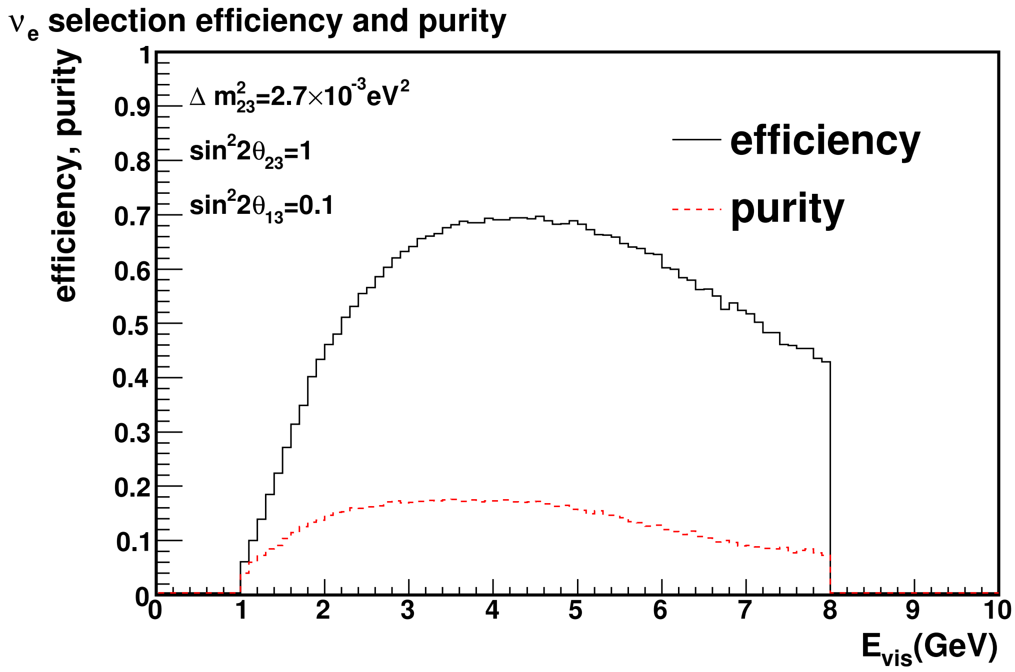

Figure 8: ν_e selection efficiency and purity.

- mstvars.osmtot - The totat pulseheight in mips that's picked up and put in any mst (in the odd view).
- mstvars.eeprob, mstvars.oeprob - A likelihood of being a nue.

- 3D Hit variables(6)
  - hitcalc.fHitLongEnergy - Sum of the magnitudes of the longitudinal projections of the vectors defined as above.
  - hitcalc.fHitTransCMEnergyRatio - Ratio of hitcalc.fHitTransCMEnergy to hitcalc.fHitTotalEnergy.
  - hitcalc.fHitTransEnergyRatio - Ratio of hitcalc.fHitTransEnergy to hitcalc.fHitTotalEnergy.
  - hitcalc.fHitLongEnergyRatio - Ratio of hitcalc.fHitLongEnergy to hitcalc.fHitTotalEnergy.
  - hitcalc.fHitFarAngle - Angle between the straight line joining the farthest hit with the vertex and the Z axis, along the incoming neutrino direction.
  - hitcalc.fHitPeakAngle - Angle between the straight line joining the highest energy hit with the vertex and the Z axis.

- Angular Clustering Variables (determined from clustering 3D hits)(2)

- angcluster.fACluRmsShwAxis - RMS value of the radial distribution of the hits belonging to the primary shower/cluster around the primary shower direction.
- angcluster.fACluRmsZAxis - RMS value of the radial distribution of the hits belonging to the primary shower/cluster around the Z axis.

- Fractional Energy Variables(7)
  - fracvars.fract_n_planes(n=2,4,6) - The fraction of the observed total energy which is deposited in n consecutive planes with the highest pulse height.
  - fracvars.fract_n_counters(n=4,8,12) - The fraction of the observed total energy which is deposited in n strips with the highest pulse height.
  - fracvars.fract_road - The fraction of the observed total energy which is deposited in a narrow road.
- SubShower Variables(6)
  - subshowervars.E2to1U
  - subshowervars.E2to1V
  - subshowervars.PHFracRMSU
  - subshowervars.PHFracRMSV
  - subshowervars.PHAvgDevU
  - subshowervars.PHAvgDevV