Numu CC/Nue CC/ NC Separation with ANN

(Progress Report)

M. Zois, N. Saoulidou*, and G. Tzanakos

Physics Department, University of Athens, Greece

(*) Now at Fermilab

Acknowledgments: C. Andreopoulos, R. Hatcher, Liz Buckley-Geer

MINOS Collaboration Meeting, Sept 20, 2003

Physics Motivation

- $\nu_\mu \neq \nu_\mu$ (Disappearance)
- $\nu_\mu \rightarrow \nu_e$
- $\nu_\mu \rightarrow \nu_s$
- **ν_μ CC / ν_e CC / NC separation very important**
- **Both NearDet and FarDet**

Beam Neutrino Interactions

• ***Near Detector***: **ν_μ CC, ν_e CC, NC, ν-bars**. The a priori probabilities are determined by the **beam.**

• ***Far Detector***: **ν_μ CC, ν_e CC, ν_τ CC, NC, ν-bars** . The a priori probabilities are determined by the **beam, and oscillation.**

• **Event Classification requires the knowledge of the a priori probabilities. This can complicate things in the far detector.**

• **Event classification will lead to signal and background determination for a given process.**

What we have done so far

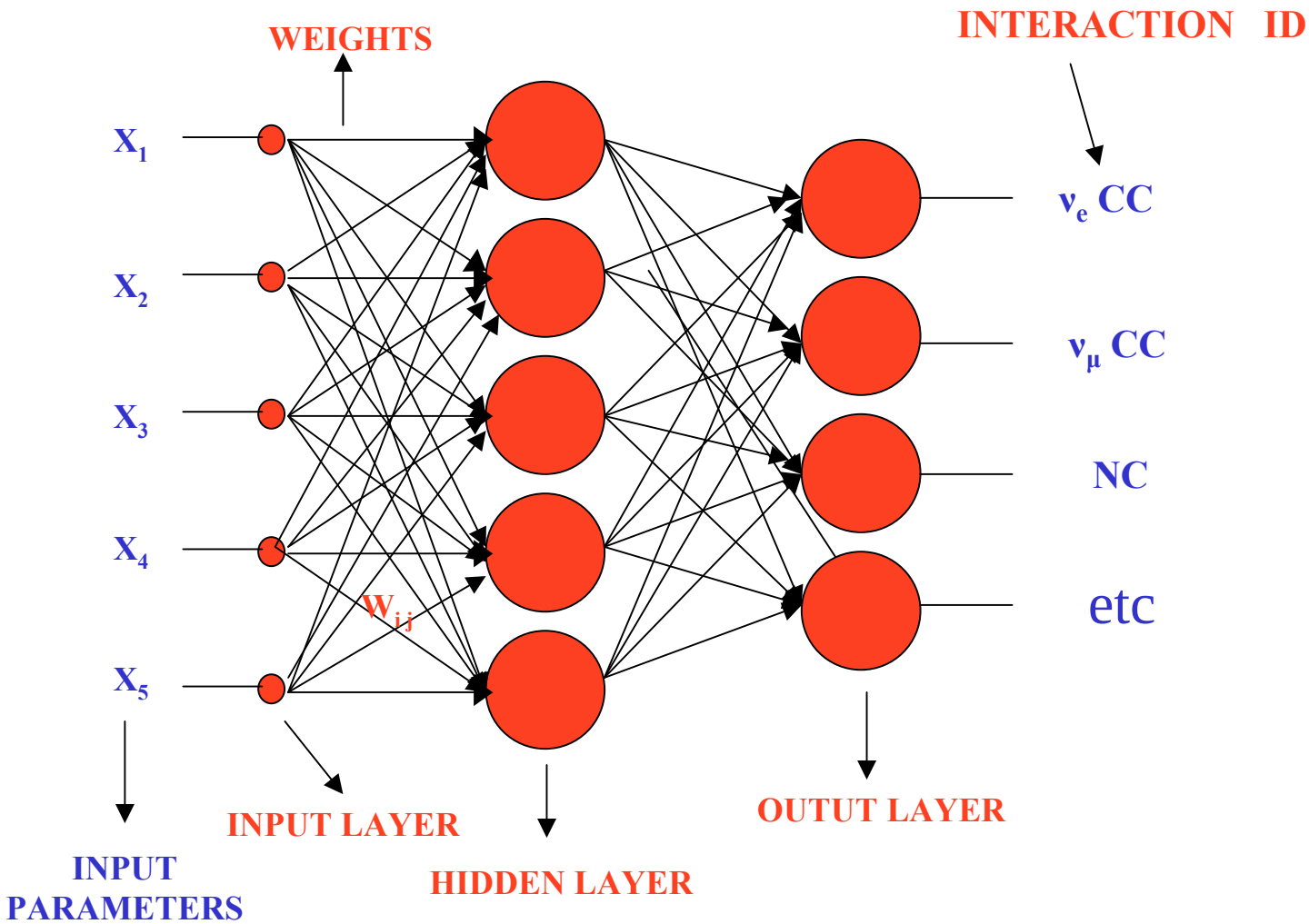
- Generated monoenergetic ν_μ **CC,** ν_e **CC, NC** interactions at the **FARDET** at:

$$E_\nu = 0.5,\ 1,\ 2,\ 3,\ 5\ \text{GeV}$$

- At **each energy** we studied ANNs (**two different packages**) to separate (classify):

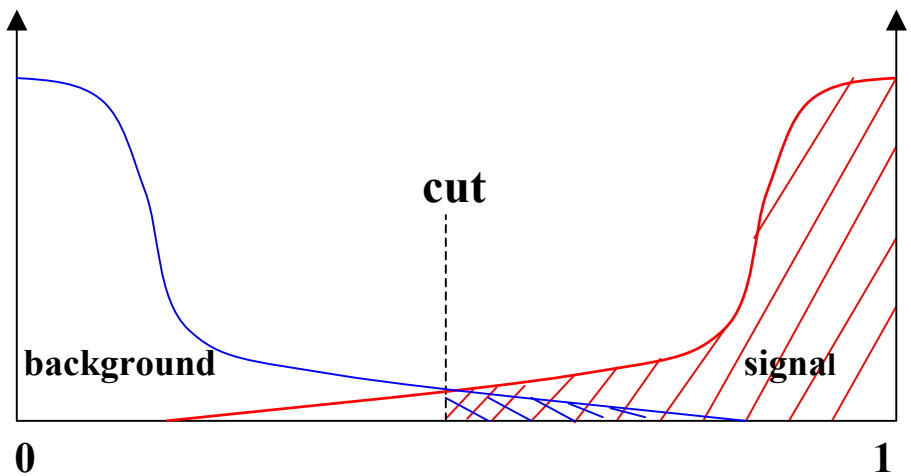
- **ν_μ CC from mixture of (ν_e CC, NC)**
- **ν_e CC from NC**
- **ν_μ CC from NC**
- **We made a comparative study of the performance of the MLPfit and SNNS ANN packages.**

Generic ANN Schematic



ANN Performance Parameters

Network output (selection) function for "background "and "signal" events



S = Total # Signal events

B = Total # Background events

S_C = Signal events above Cut

B_C = Background events above Cut

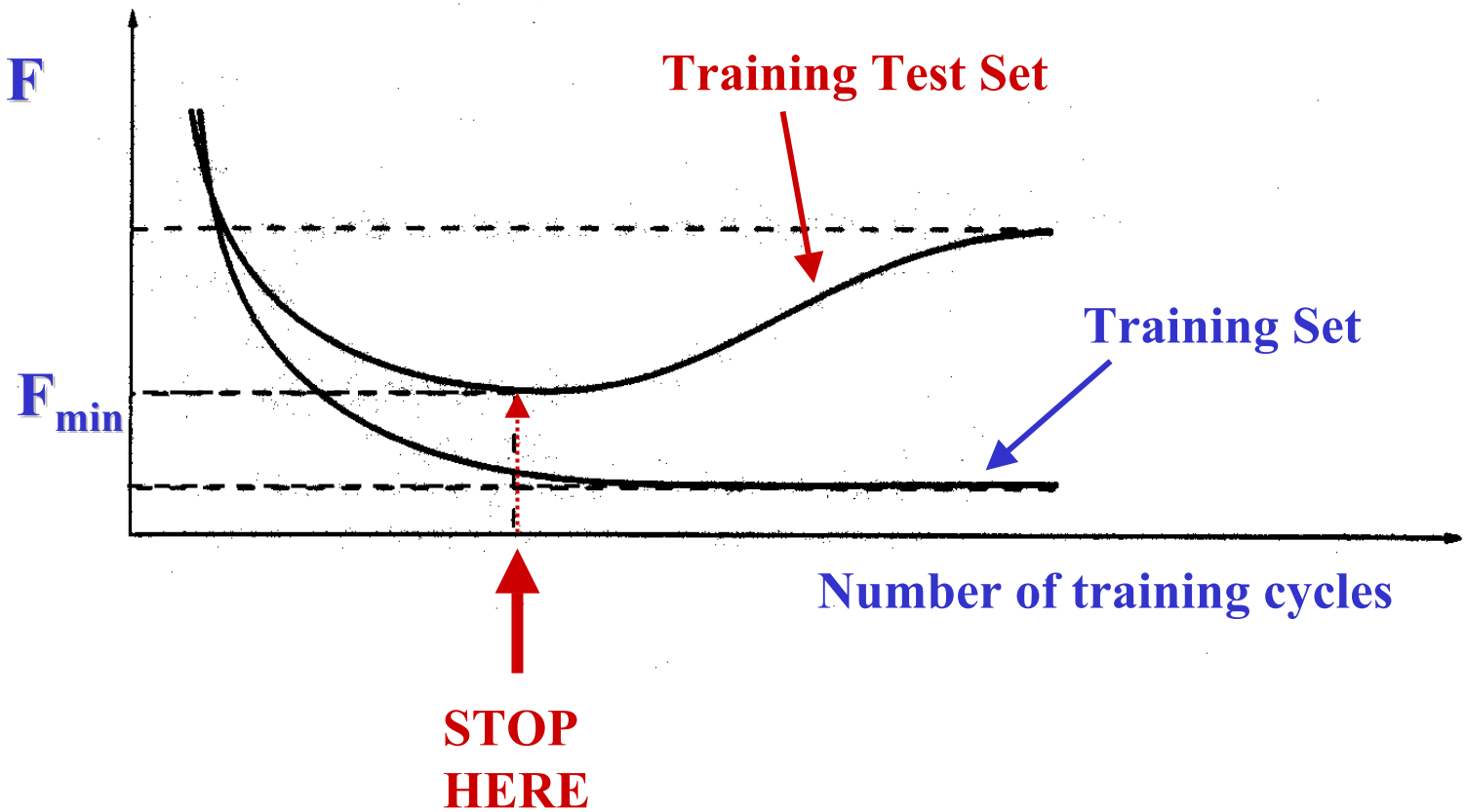
$$\text{efficiency} = \frac{S_C}{S}$$

$$\text{purity} = \frac{S_C}{S_C + B_C}$$

$$\text{contamination} = \frac{B_C}{B}$$

When to Stop Training

F is the "cost function"



Selected ANN Architecture

Experimented with different architectures.
Selected ANN with:

- **6 inputs**
- **3 hidden layers with 12 neurons**
- **1 output**

Input Variables

- **Number of planes:** **"evthdr. planeall.n"**
- **Number of strips:** **"evthdr.nstrip"**
- **Energy (pe)/strip:** **'(evthdr.ph.sigcor/ 100)/evthdr.nstrip'**
- **Longitudinal size:** **'abs(evthdr. plane.beg - evthdr. plane.end)'**
- **Number of digits:** **evthdr.ndigit'**
- **Energy (pe)/plane:** **'(evthdr.ph.pe/100)/ evthdr.planeall.n'**

evthdr.ph.sigcor: in ADC counts (1pe = 100 ADC counts)
evthdr.ph.pe: in ADC counts (1pe = 100 ADC counts)

Procedure

- **Generate MC events: 10K ν_μ + 10K ν_e**
- **Reconstruct**
- **Train the ANN using 1/2 of the sample**
- **Stop training properly to avoid overtraining**
- **Plot the ANN output function (F_{out}) distribution for the total sample**
- **Determine the ANN performance parameters: (efficiency, purity, contamination) versus the output function F_{cut} value**

GMINOS Running:

- **Running twice**

  Numu (‘MFLV 1’ ‘PFLV 14 14 14 14 14 14’)

  Nue (‘MFLV 1’ ‘PFLV 12 12 12 12 12 12’)

- **E_ν: Uniform from E_0-.25 GeV to E_0 +.25 GeV**

  e.g. (‘MFLX 4’ ‘FLXL 0.75 1.25’)

- **Events in far FLS detector**

  (‘SUPR +2’)

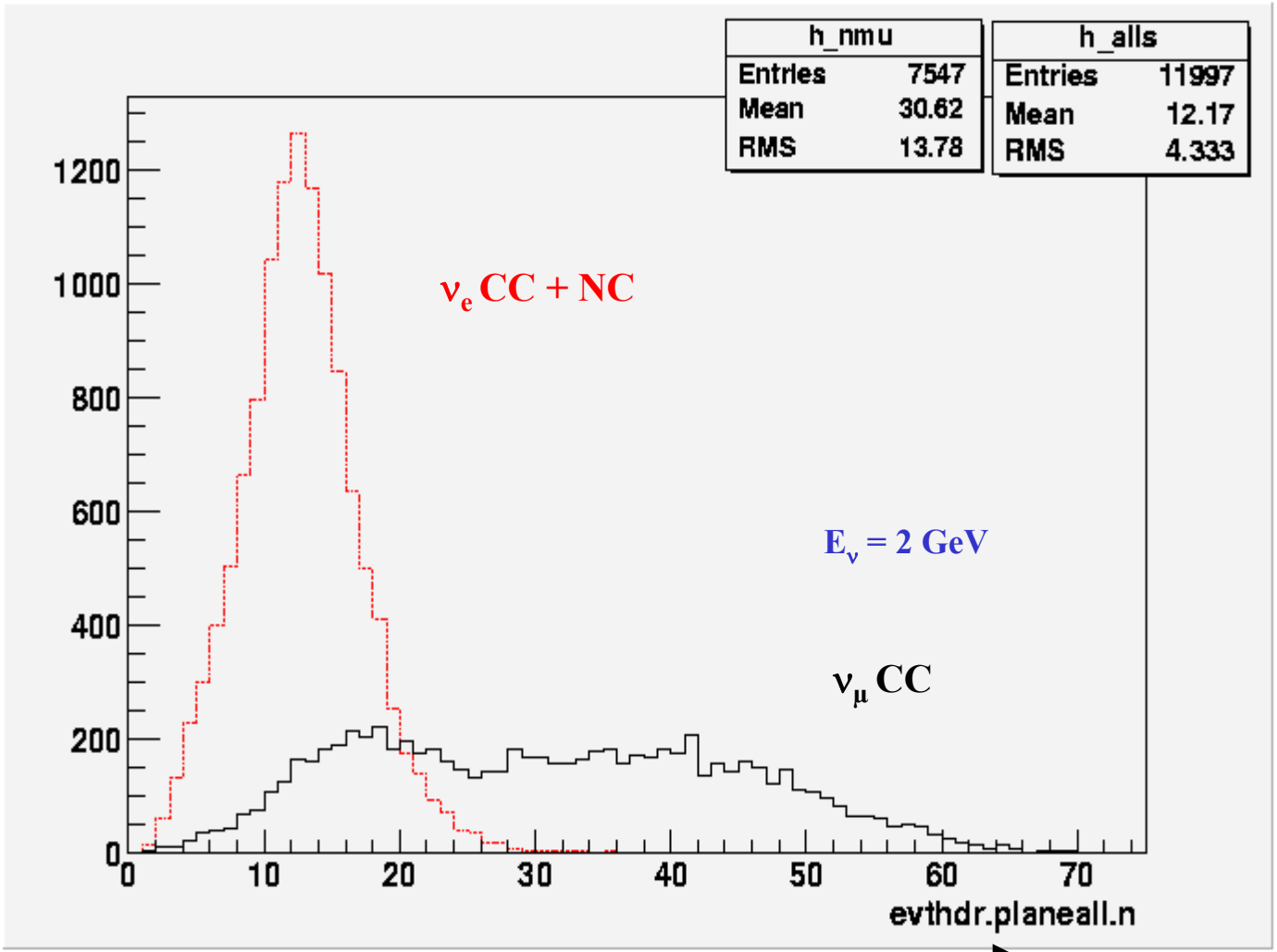
- **Running for 10,000 events each time**

  (‘TRIG 10,000’ ‘FNDE 10,000’)

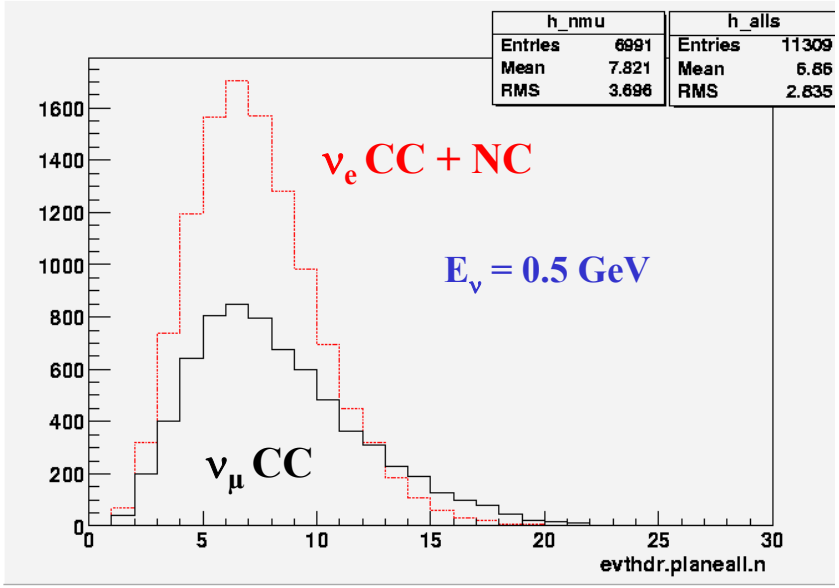
- **Reconstruct the MC events**
- Use **only “digit”** information from the reconstructed events. **No tracks, no showers.** We did not use Reconstructed information (‘trk’ branch at ‘eventsr.root’ tree).

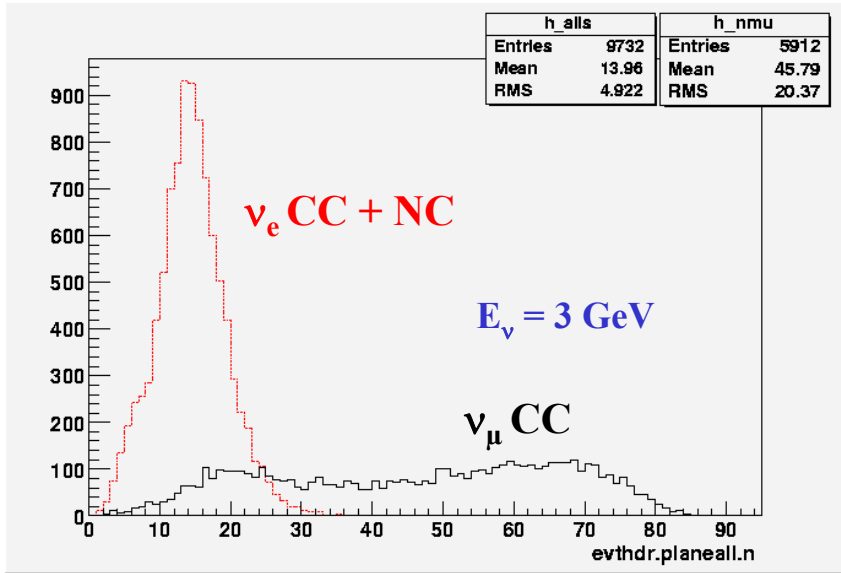
Network for ν_μ CC -(ν_e CC + NC) separation

E_ν (GeV)	ν_μ CC	ν_μ NC	ν_e CC	ν_e NC	NC Total	Grand Total
0.5	6991	2082	7267	1960	4042	18300
1.0	7428	2208	7528	2115	4323	19279
2.0	7547	2223	7586	2188	4411	19544
3.0	5912	1832	5980	1920	3752	15644
5.0	3499	1177	3656	1144	2321	9476

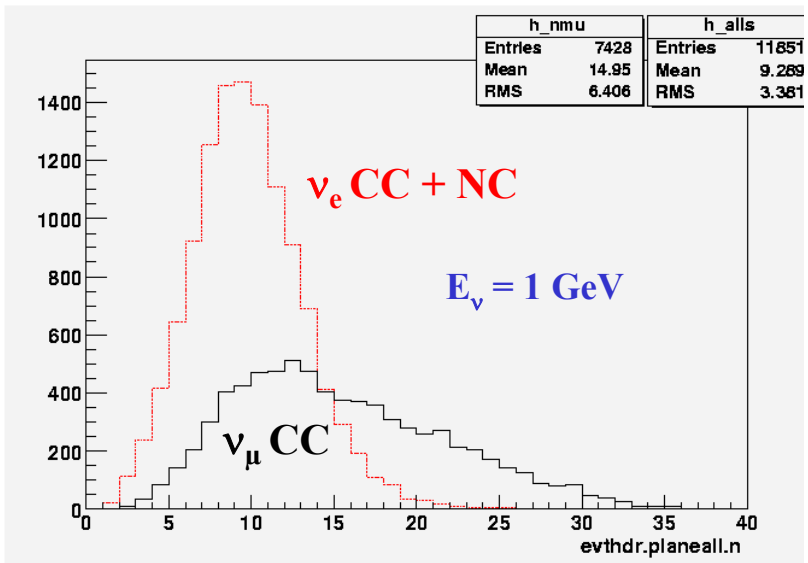


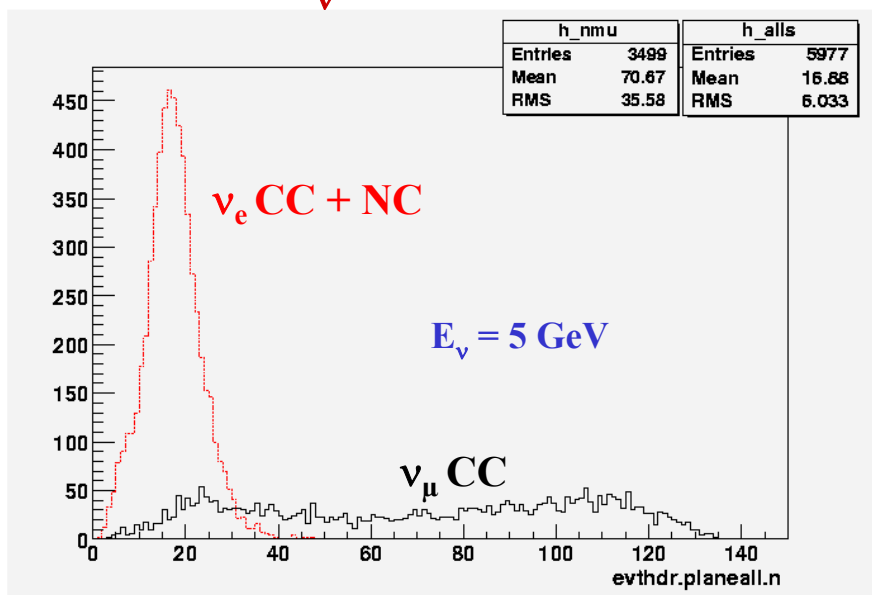
Number of Planes

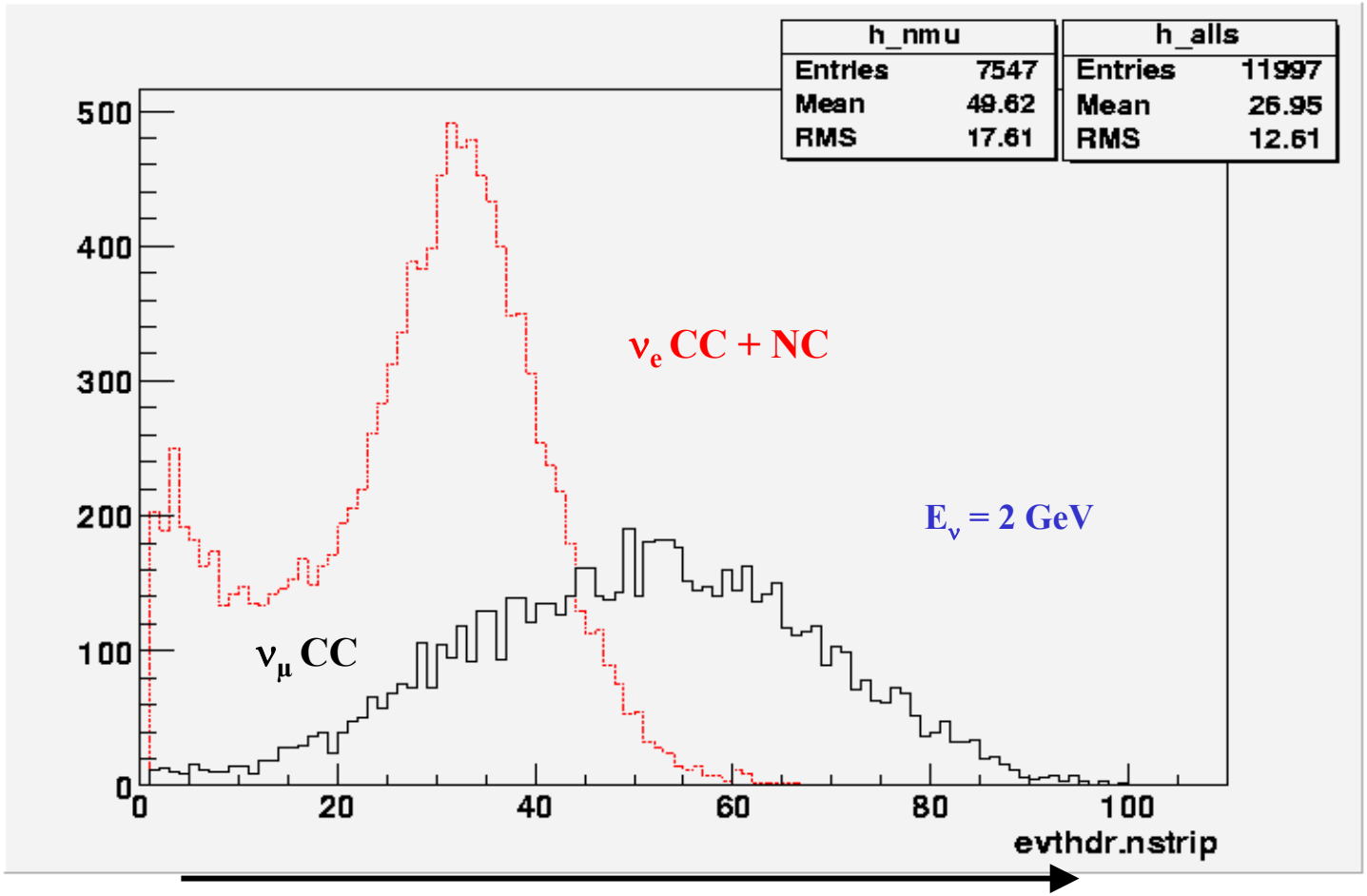




Number of Planes vs E_ν

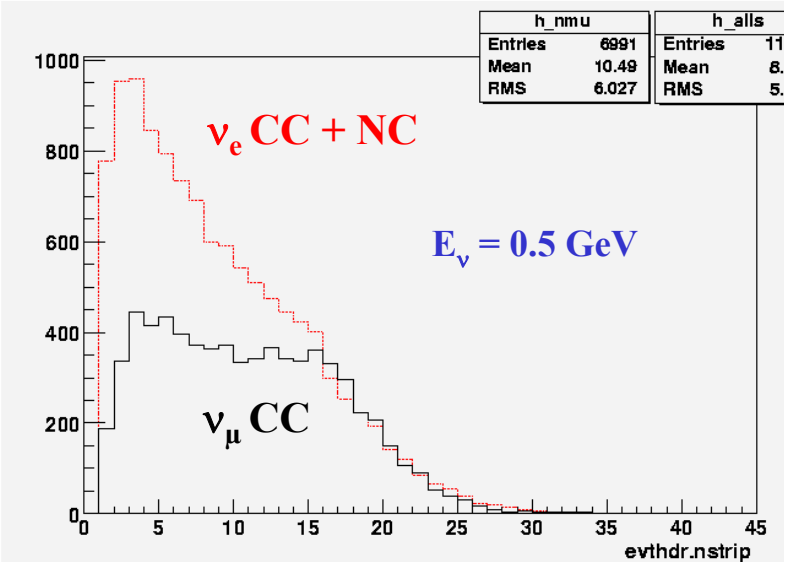


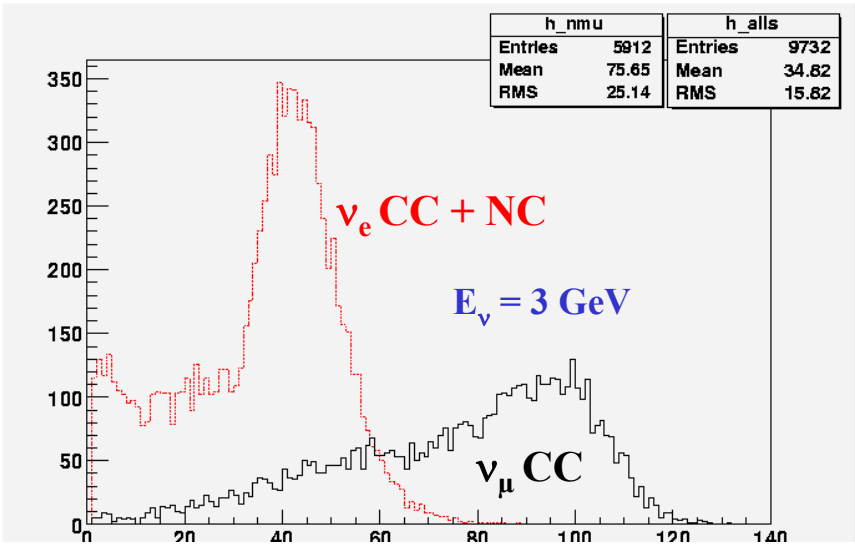


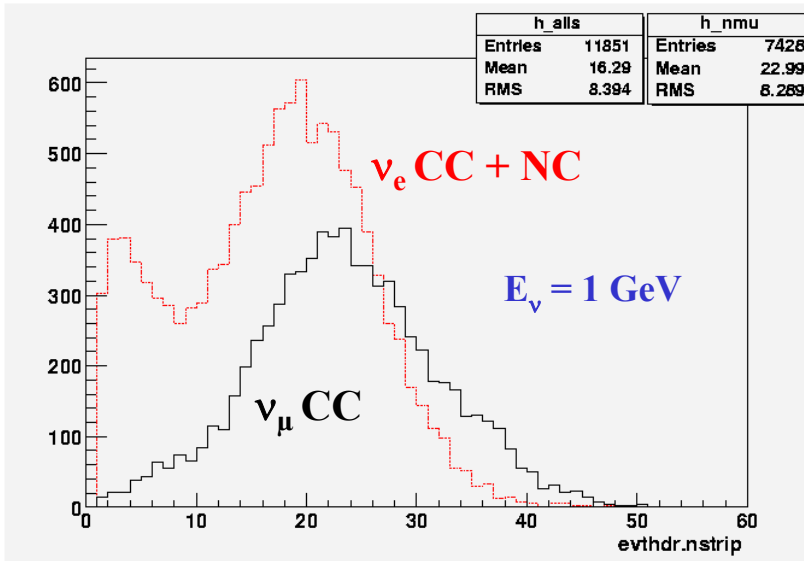


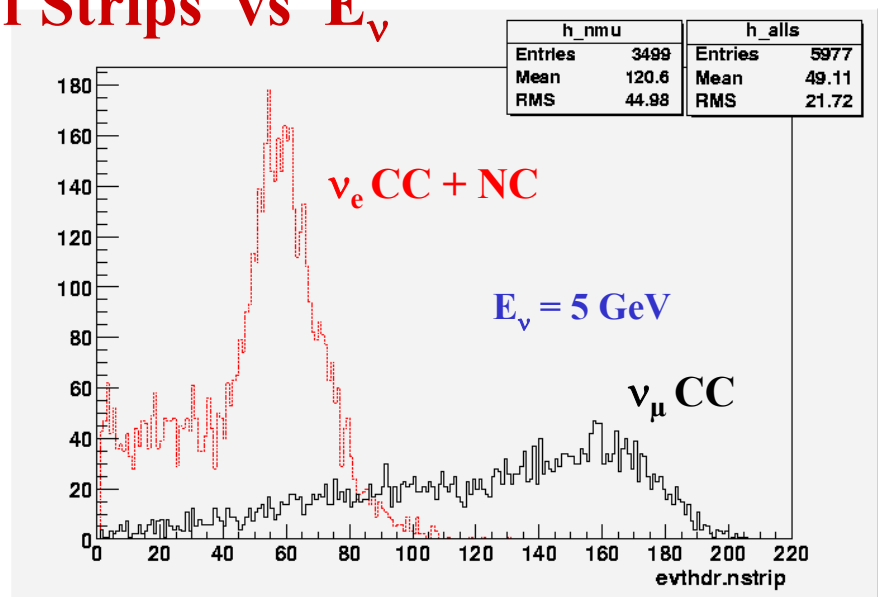
Number of Strips

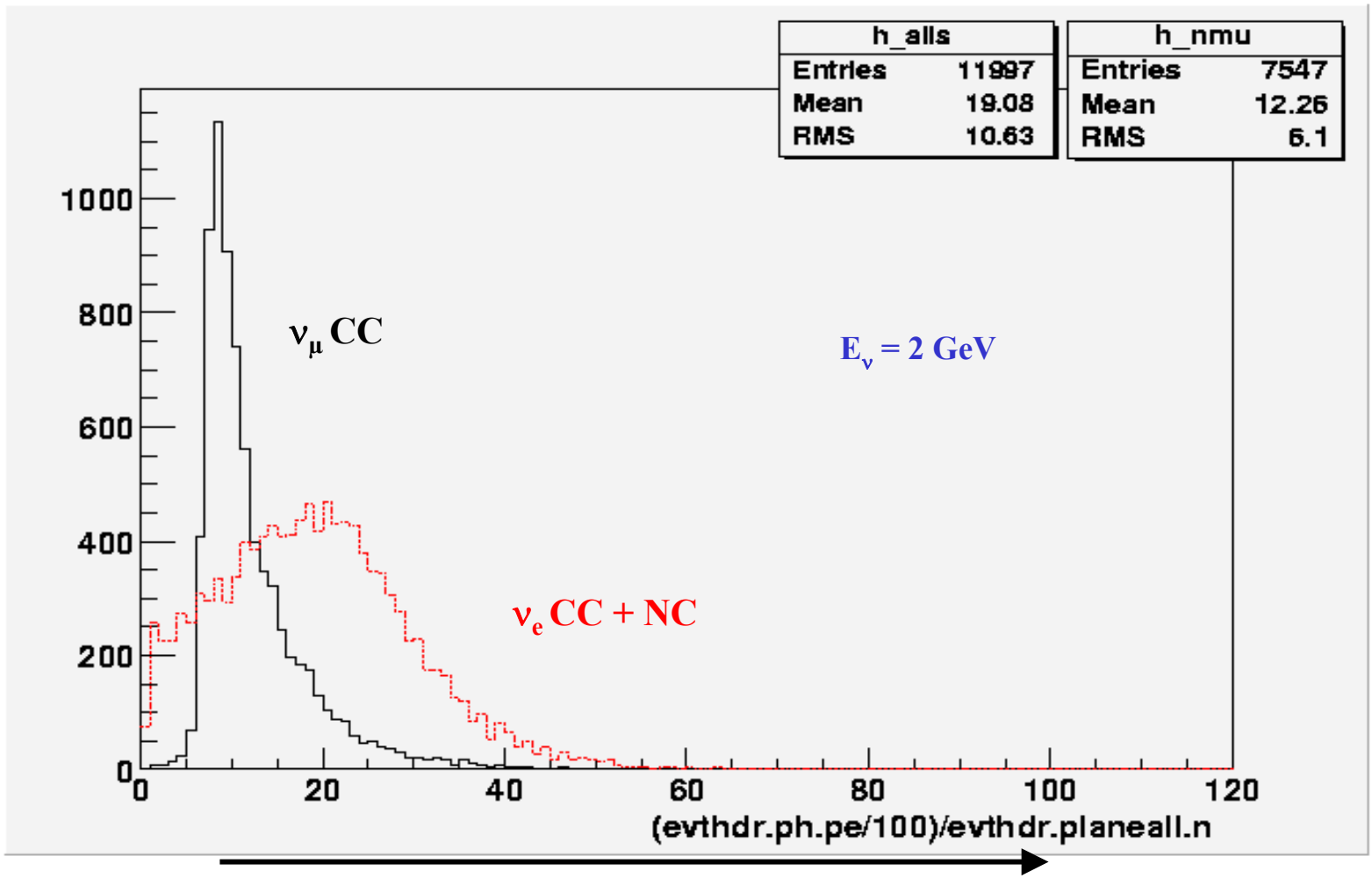
Number of Strips vs E_ν



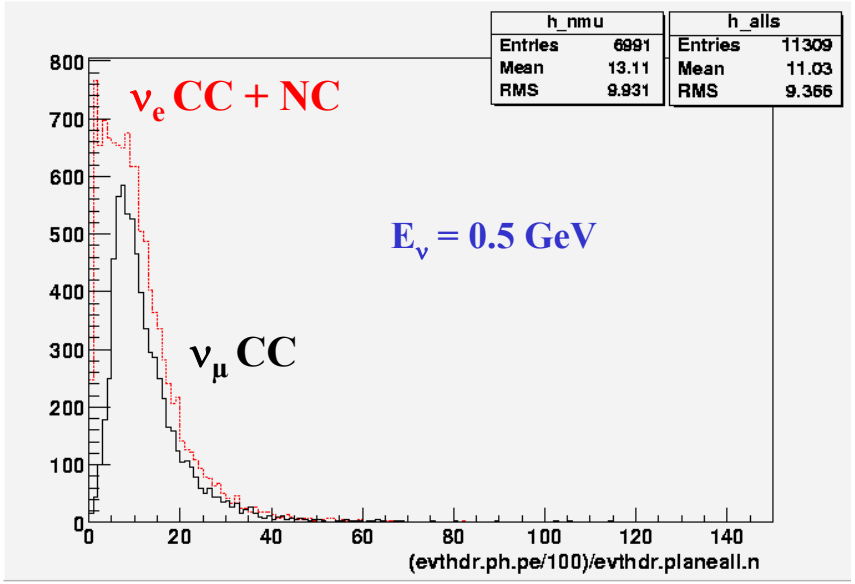


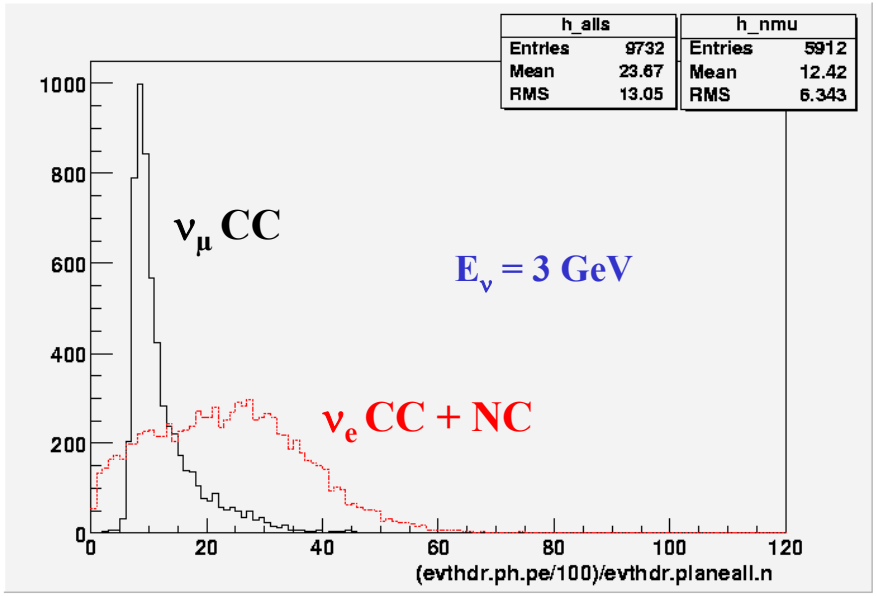




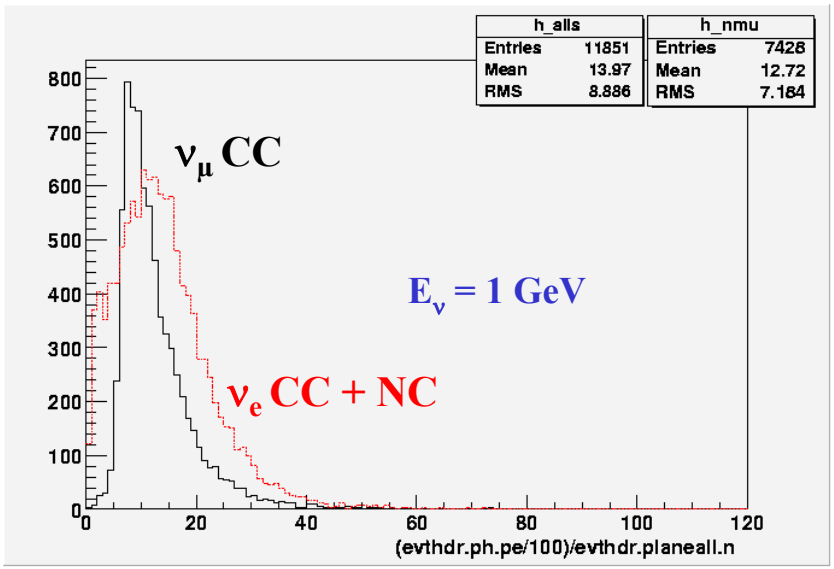


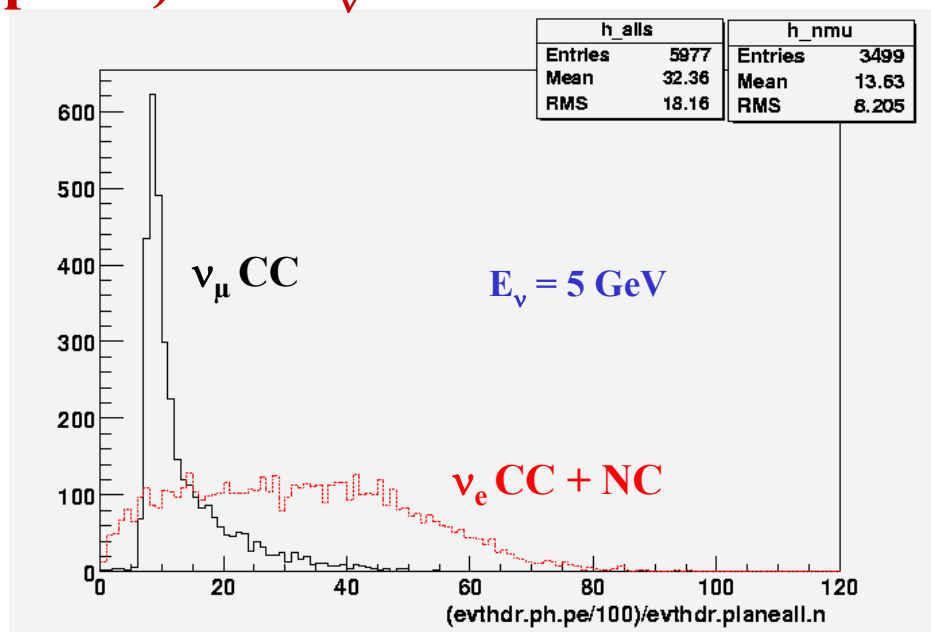
Energy (pe)/plane





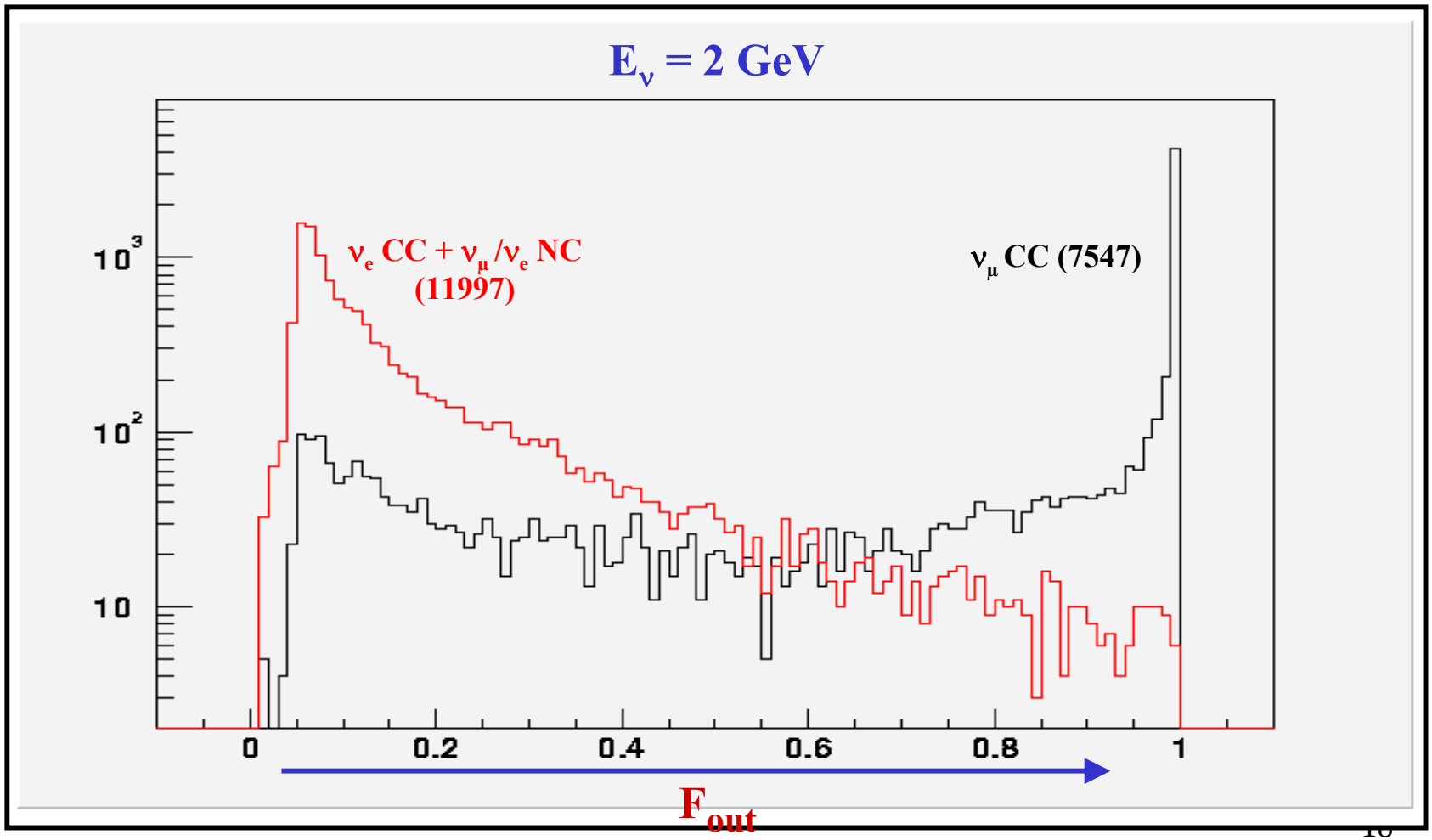
Energy (PE/plane) vs E_ν

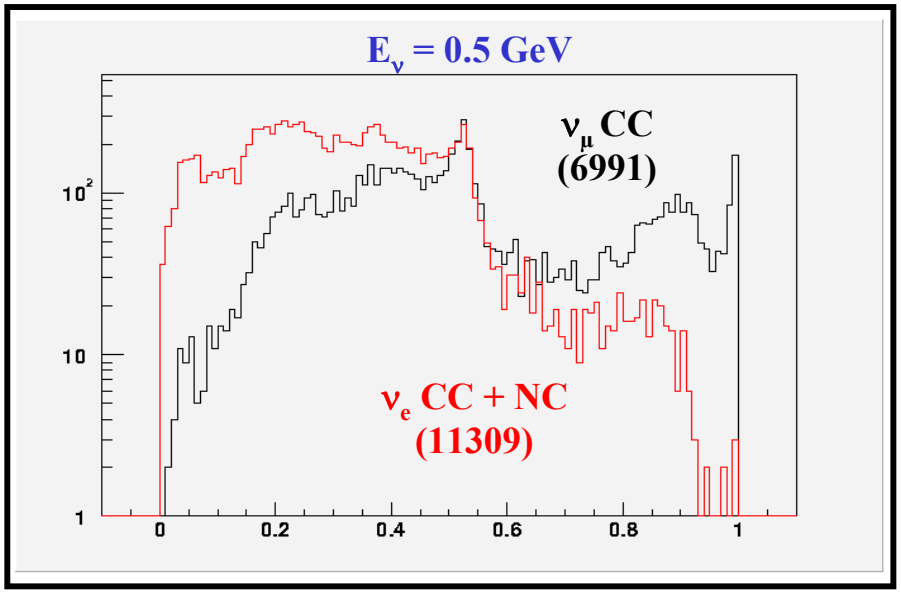


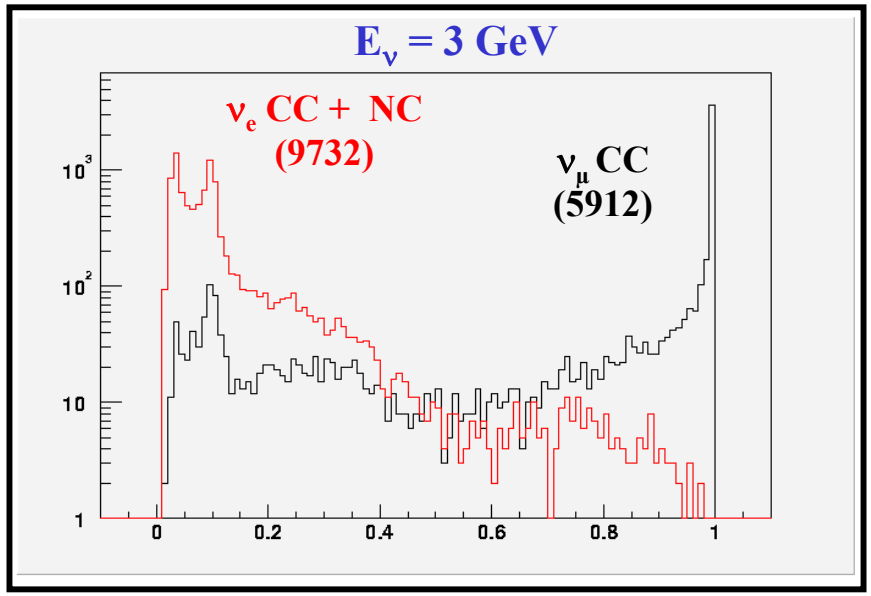


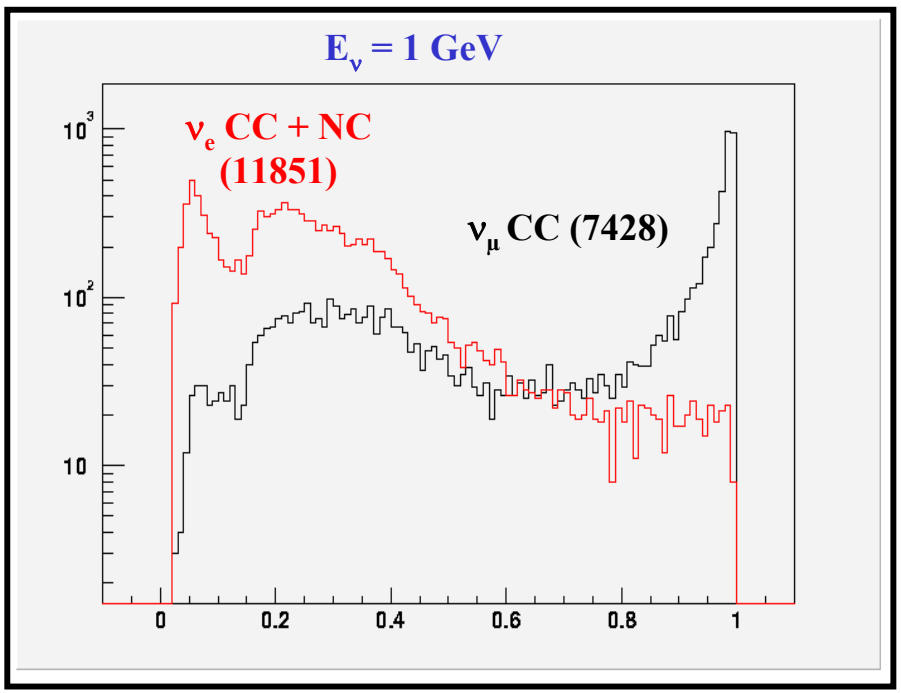
Network for ν_μ CC -(ν_e CC + NC) separation

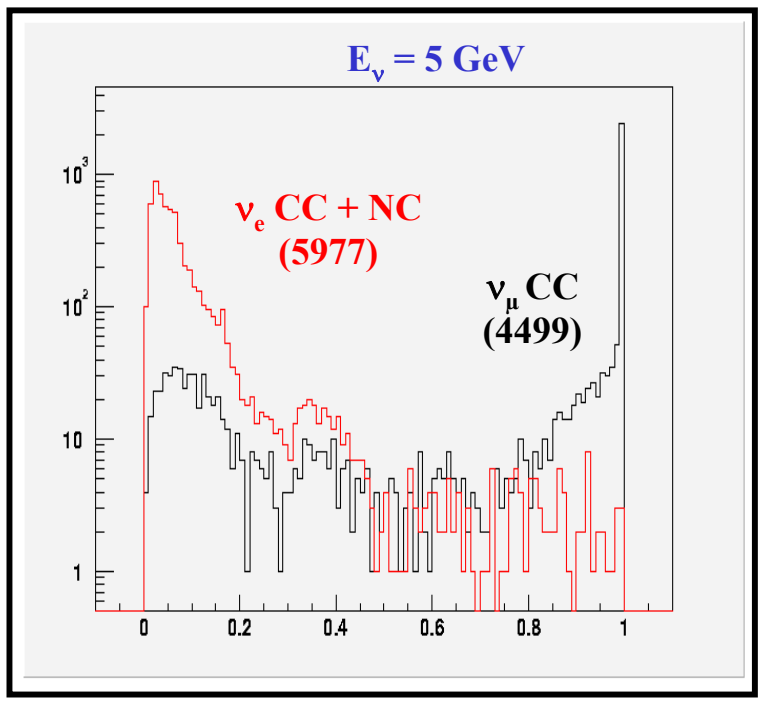
(MC file:19544 events → 7547 ν_μ CC + 11997 ν_eCC+ (ν_μ +ν_e) NC)

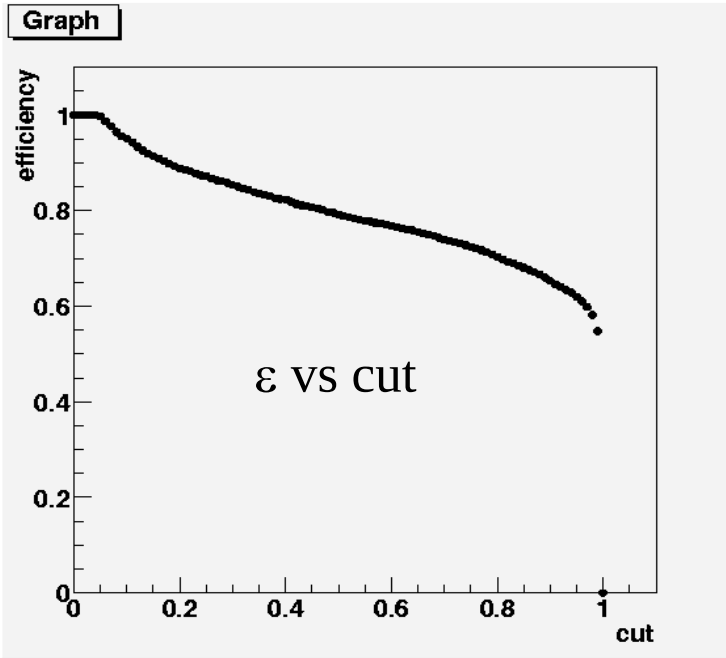


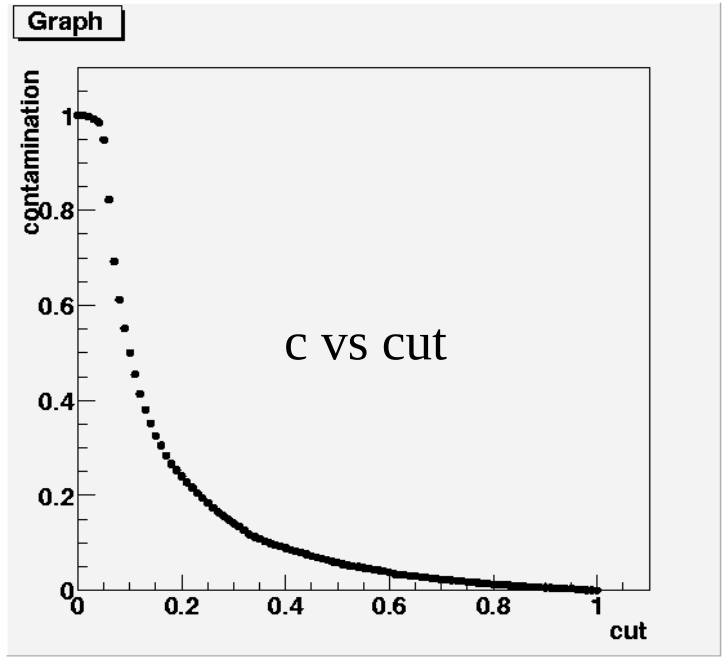

E_ν = 0.5 GeV
ν_μ CC
(6991)
ν_e CC + NC
(11309)
10^2
10
1
0
0.2
0.4
0.6
0.8
1

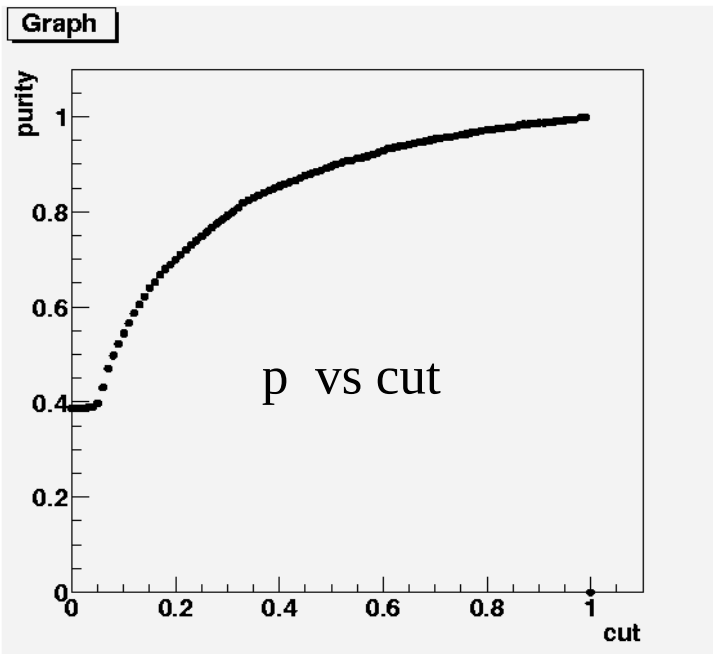

E_ν = 3 GeV
ν_e CC + NC
(9732)
ν_μ CC
(5912)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1


E_ν = 1 GeV
ν_e CC + NC
(11851)
ν_μ CC (7428)
10^3
10^2
10
0
0.2
0.4
0.6
0.8
1


E_ν = 5 GeV
ν_e CC + NC
(5977)
ν_μ CC
(4499)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1



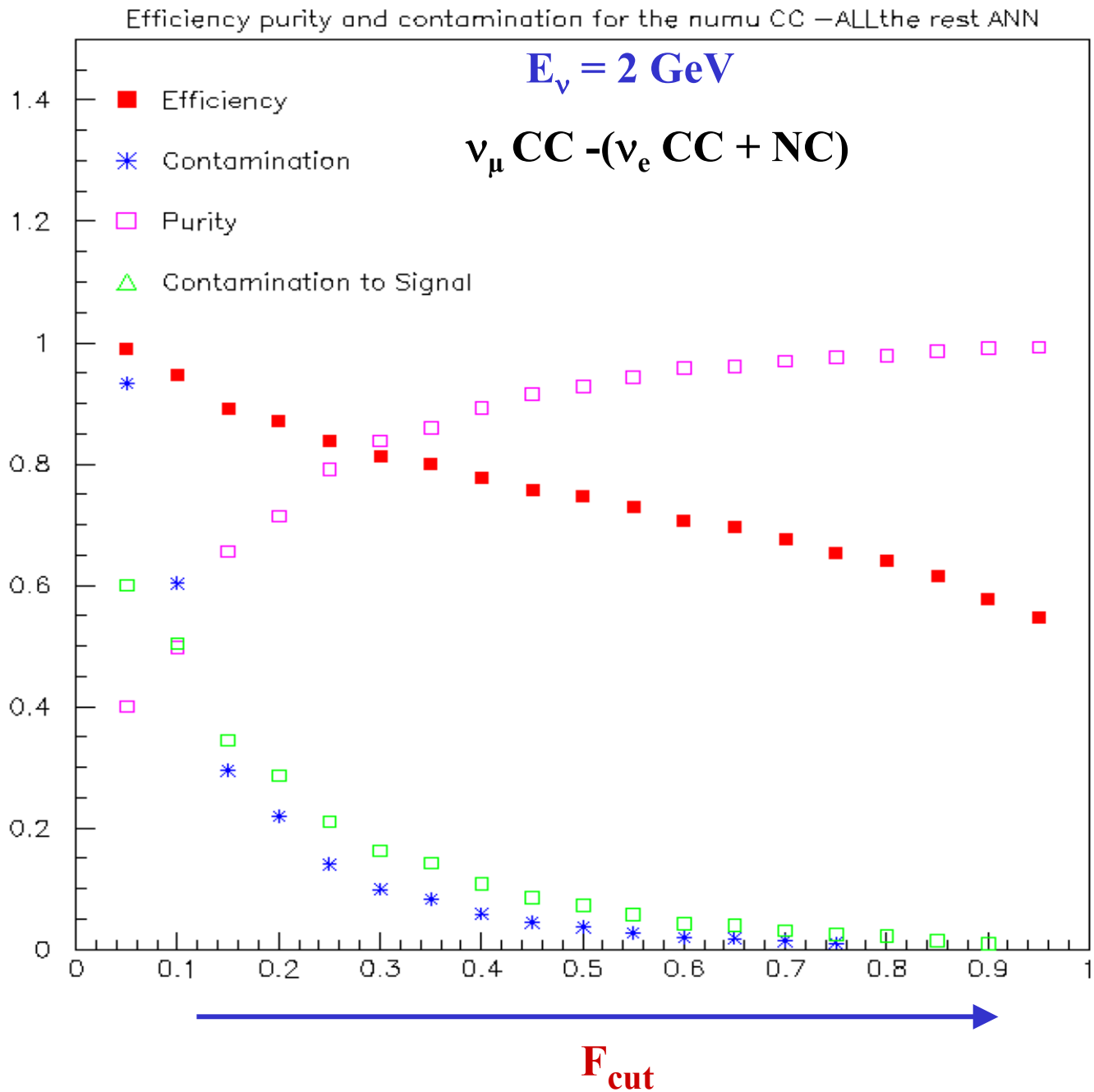




E_ν = 2 GeV

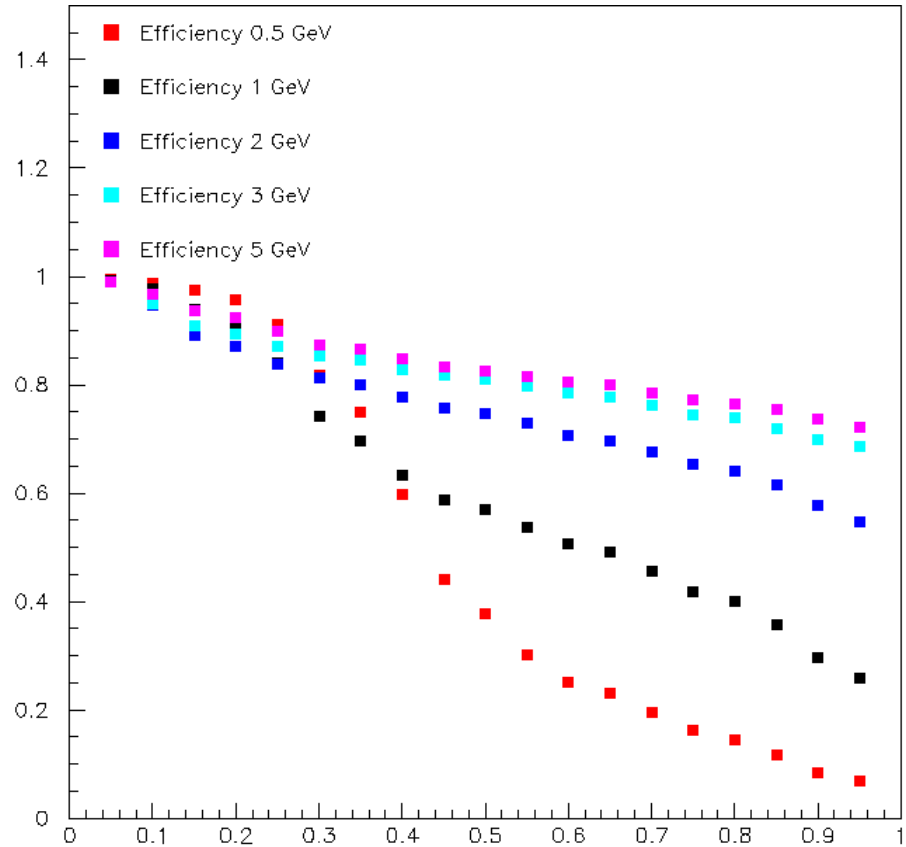
ν_μ CC -(ν_e CC + NC)

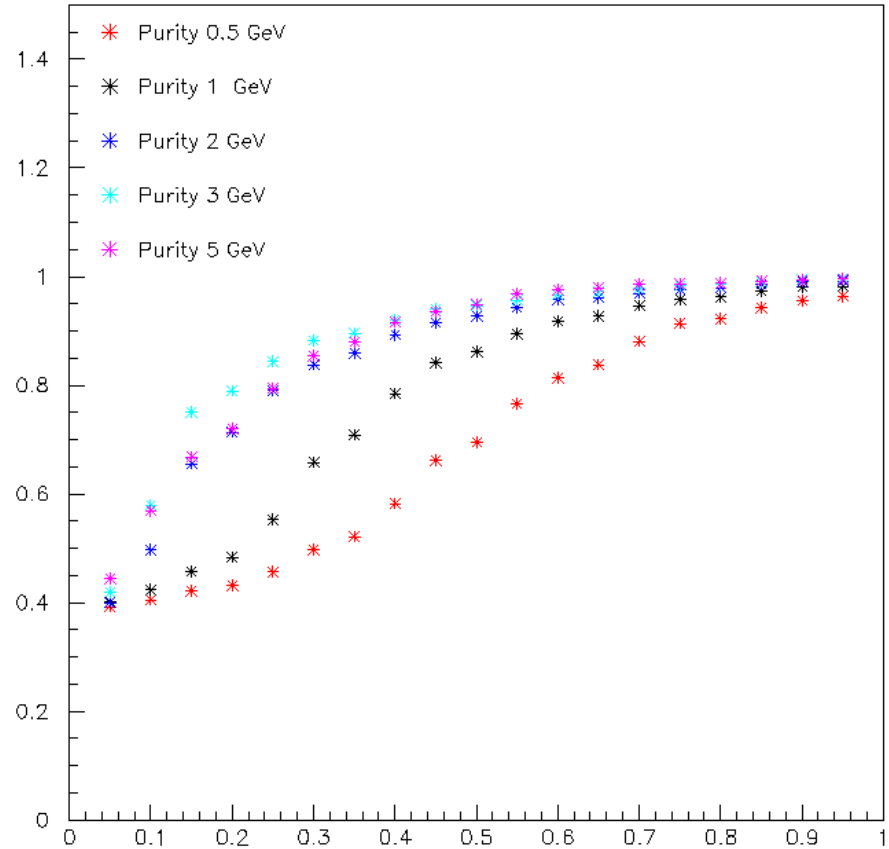
Efficiency (ε)
purity (p)
contamination (c)
vs cut


Efficiency purity and contamination for the numu CC —ALLthe rest ANN
$E_\nu = 2$ GeV
ν_μ CC -(ν_e CC + NC)
Efficiency
Contamination
Purity
Contamination to Signal
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{cut}

Efficiency & Purity vs cut for various energies

ν_μ CC -(ν_e CC + NC)

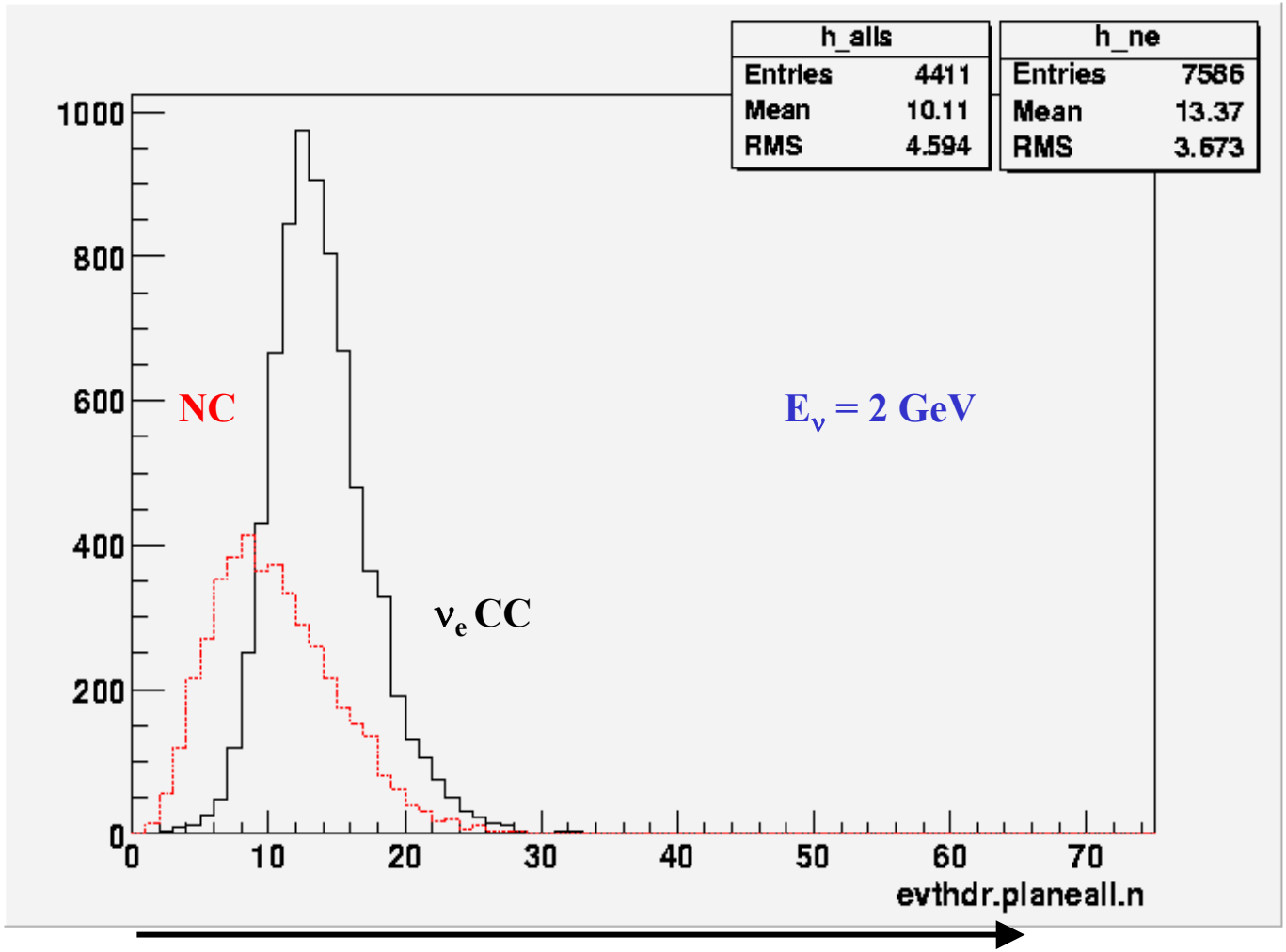




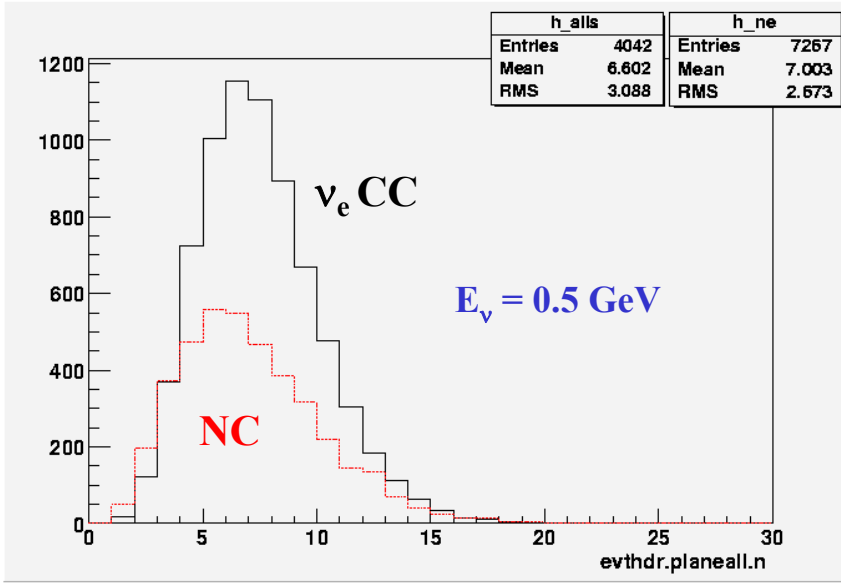
F_{cut}

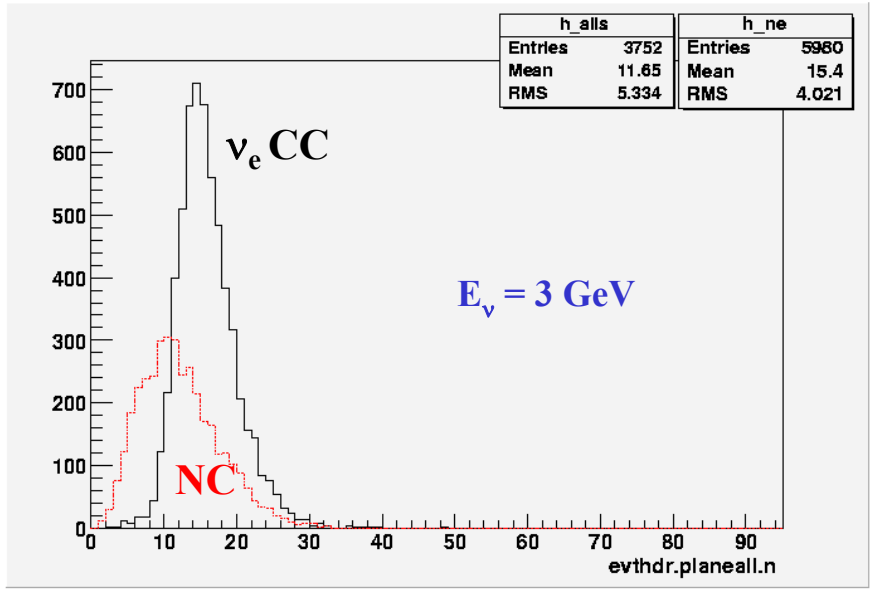
Network for ν_e CC-NC separation

E_ν (GeV)	ν_μ CC	ν_μ NC	ν_e CC	ν_e NC	NC Total	Grand Total
0.5	6991	2082	7267	1960	4042	18300
1.0	7428	2208	7528	2115	4323	19279
2.0	7547	2223	7586	2188	4411	19544
3.0	5912	1832	5980	1920	3752	15644
5.0	3499	1177	3656	1144	2321	9476

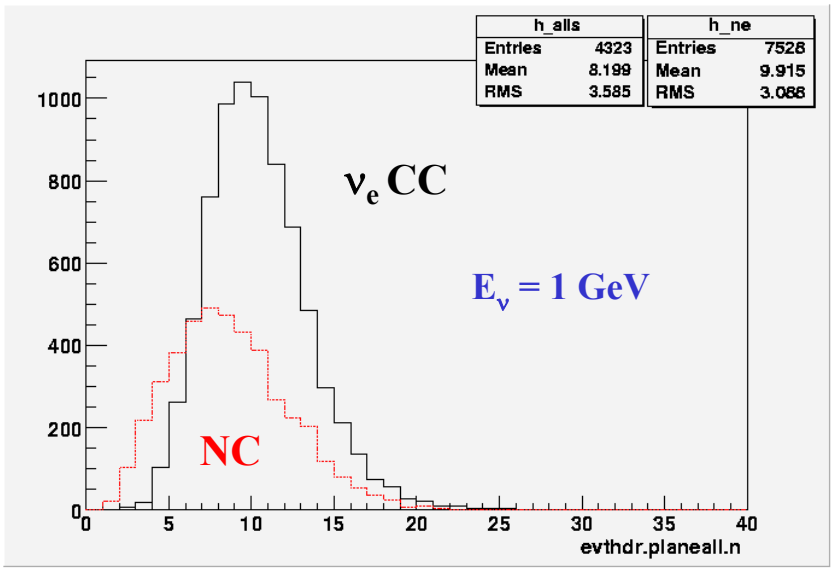


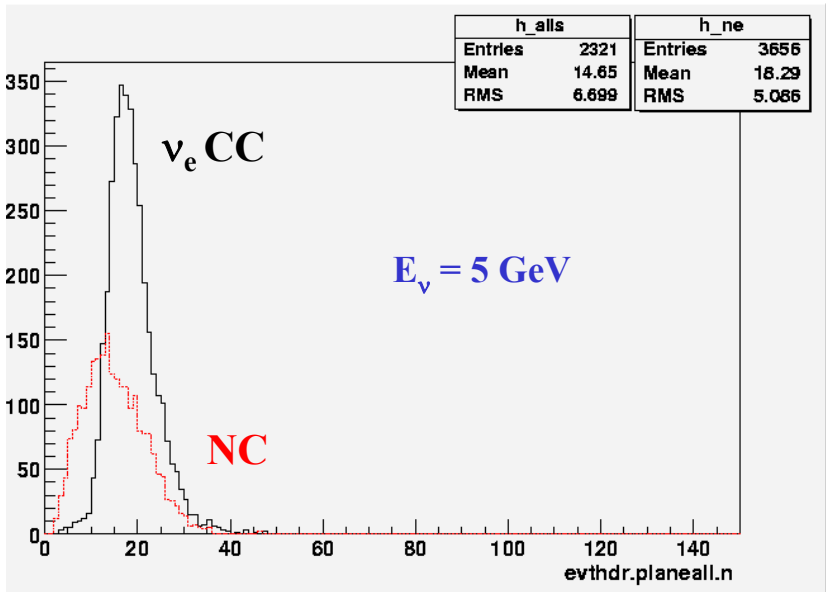
Number of Planes

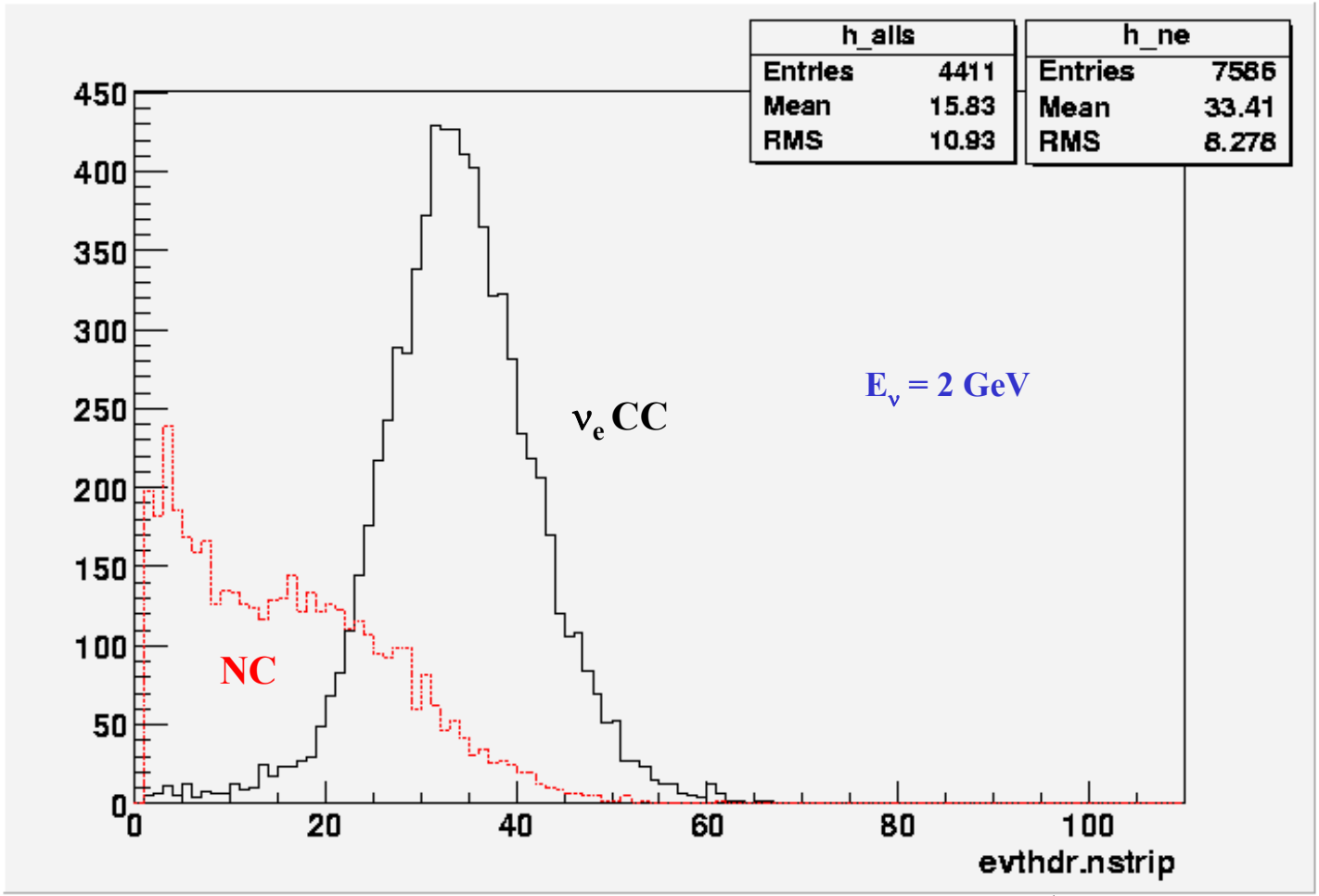




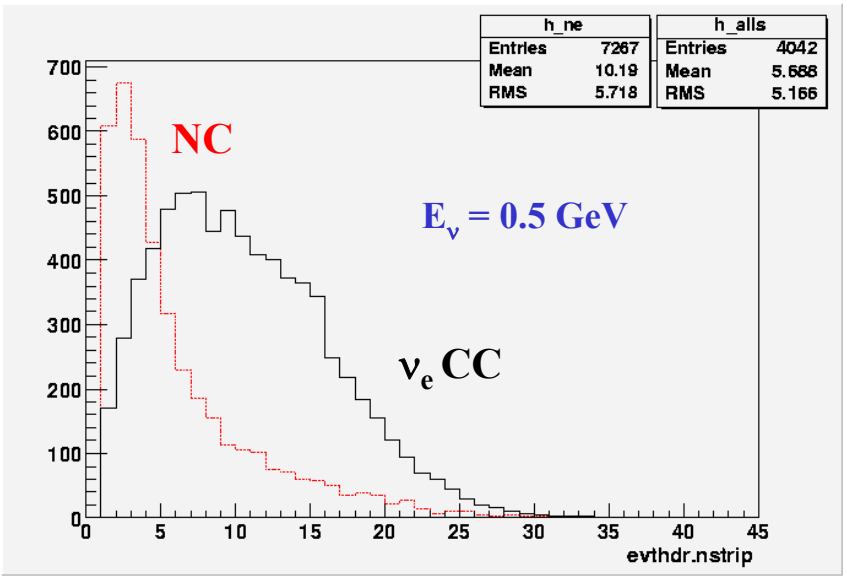
Number of Planes vs E_ν

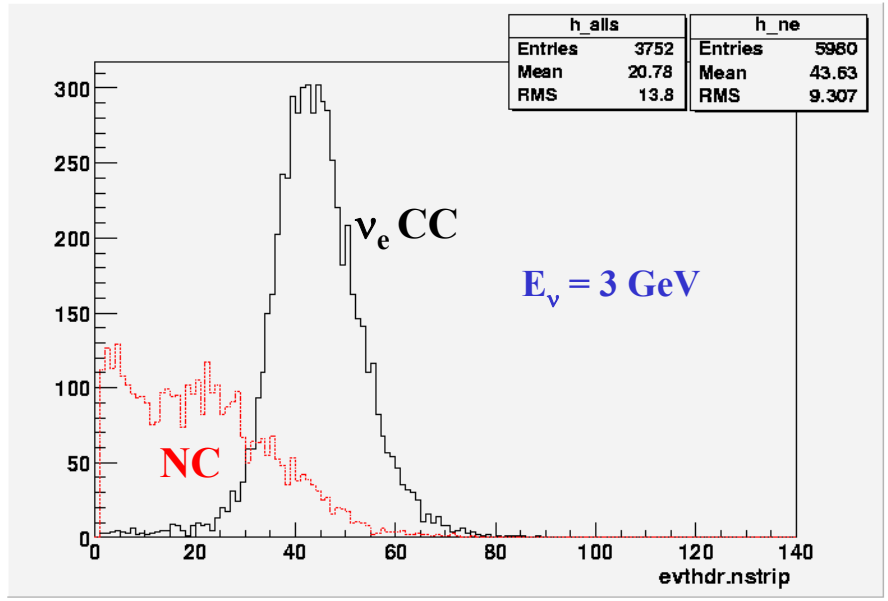




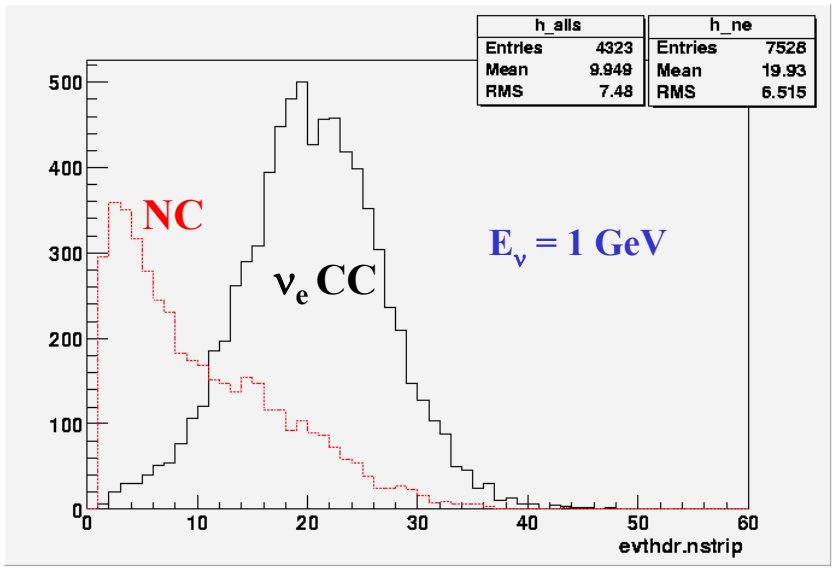


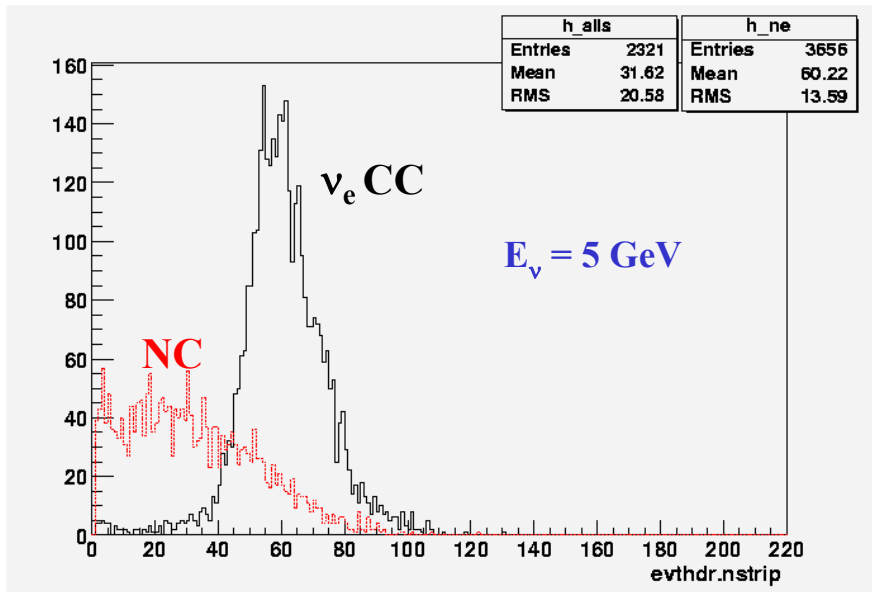
Number of Strips

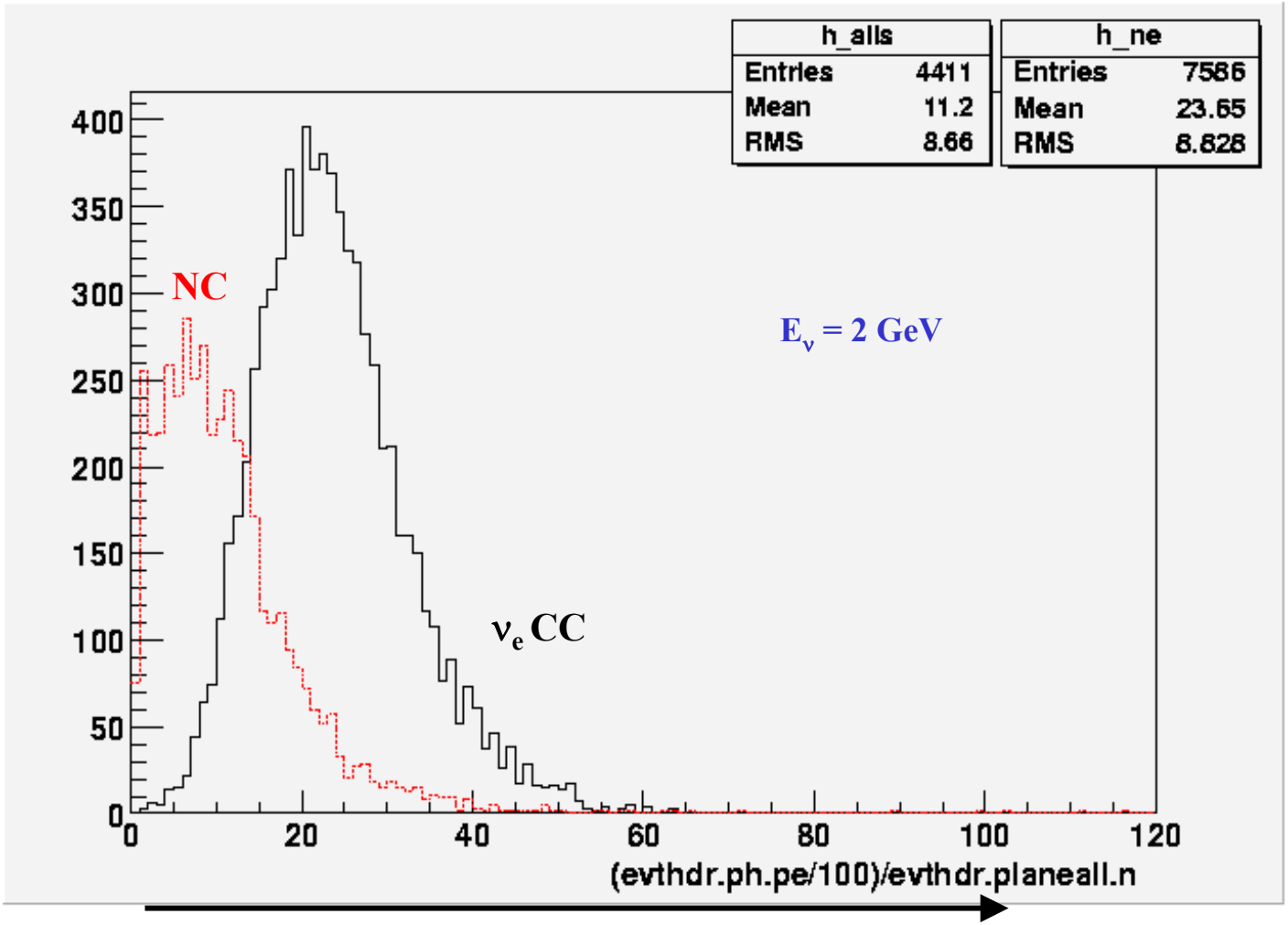




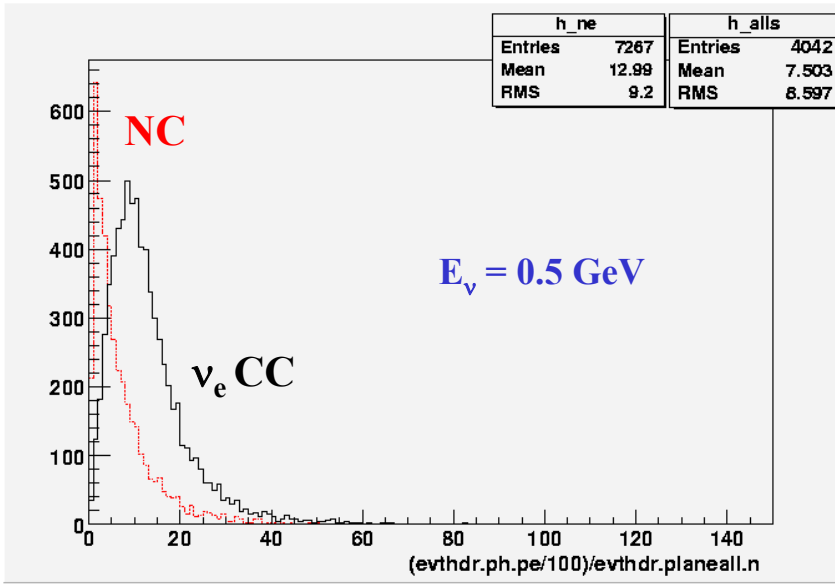
Number of Strips vs E_ν

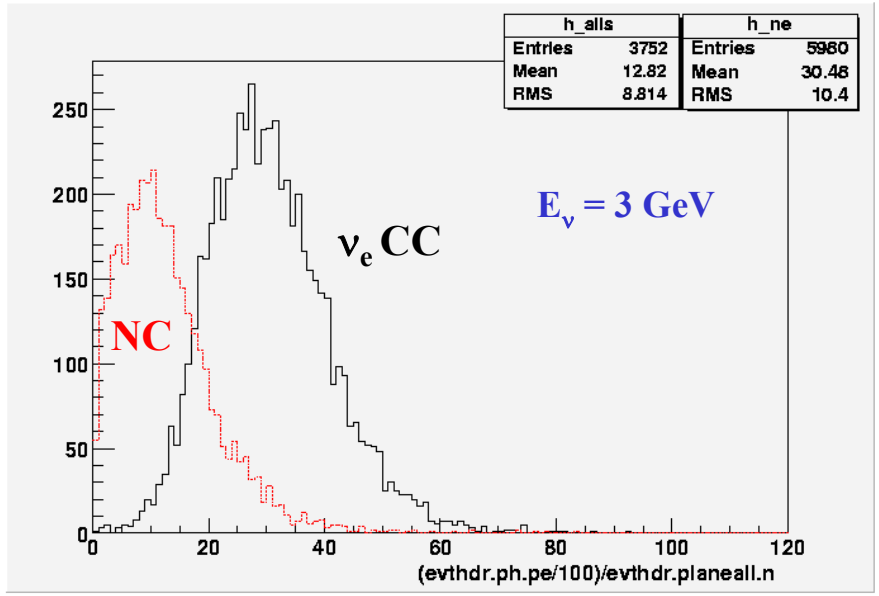




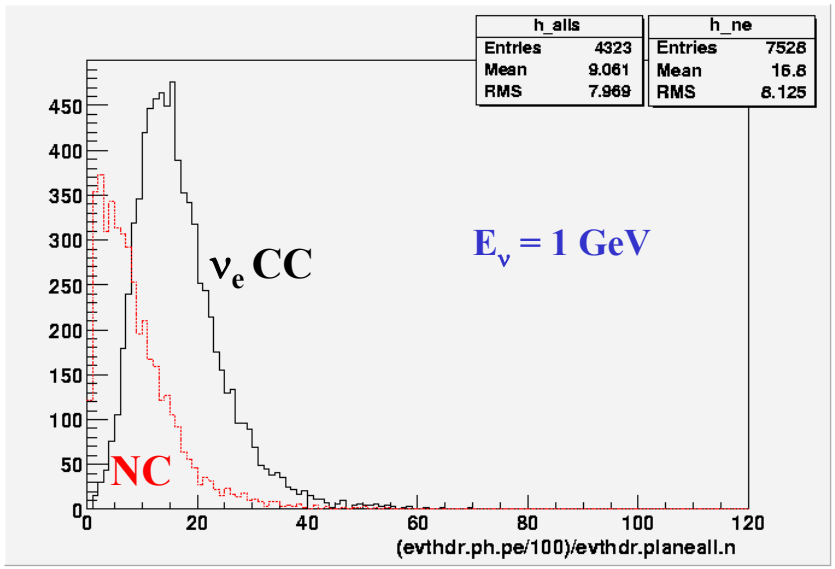


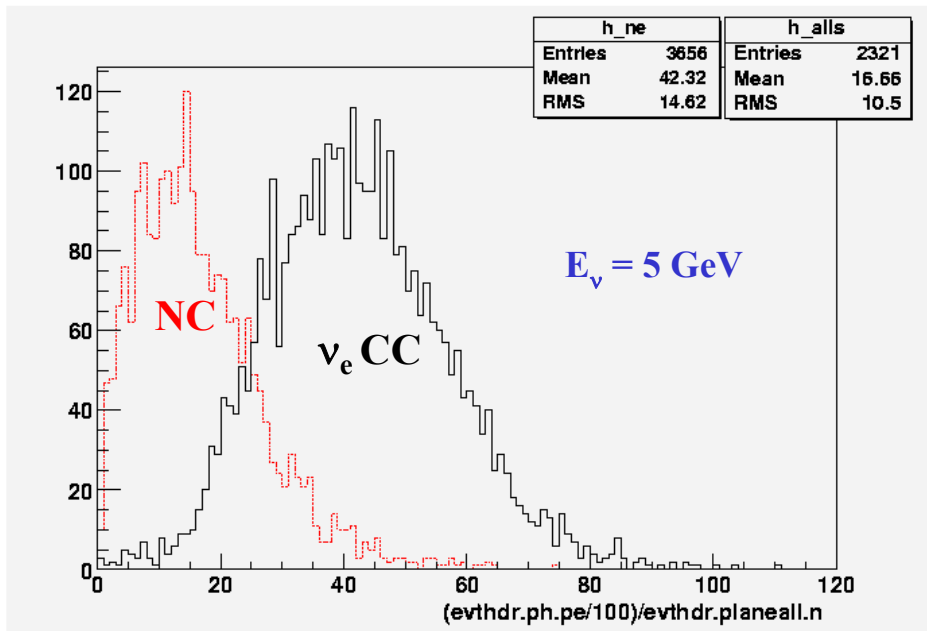
Energy (pe)/plane





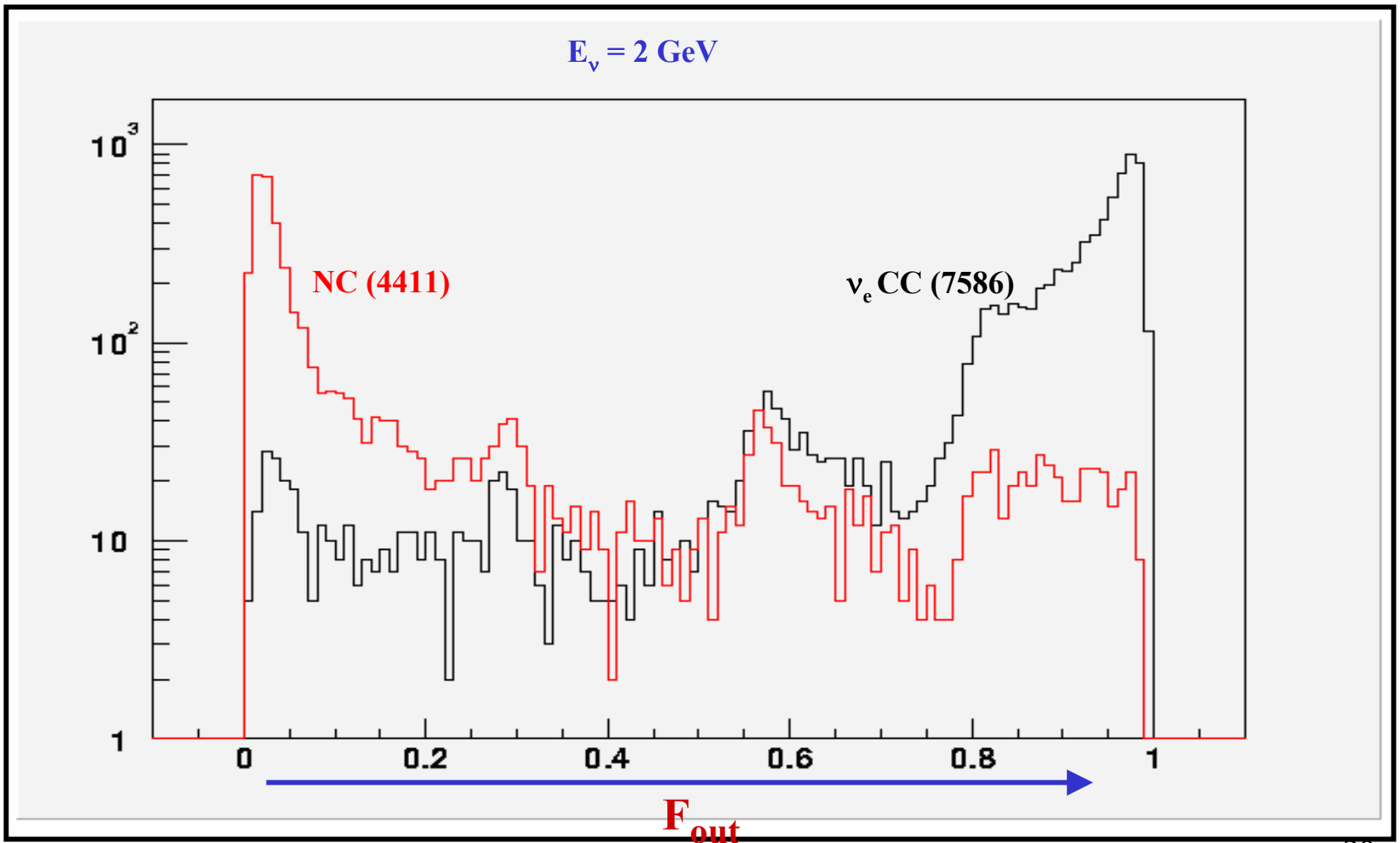
Energy (PE/plane) vs E_ν

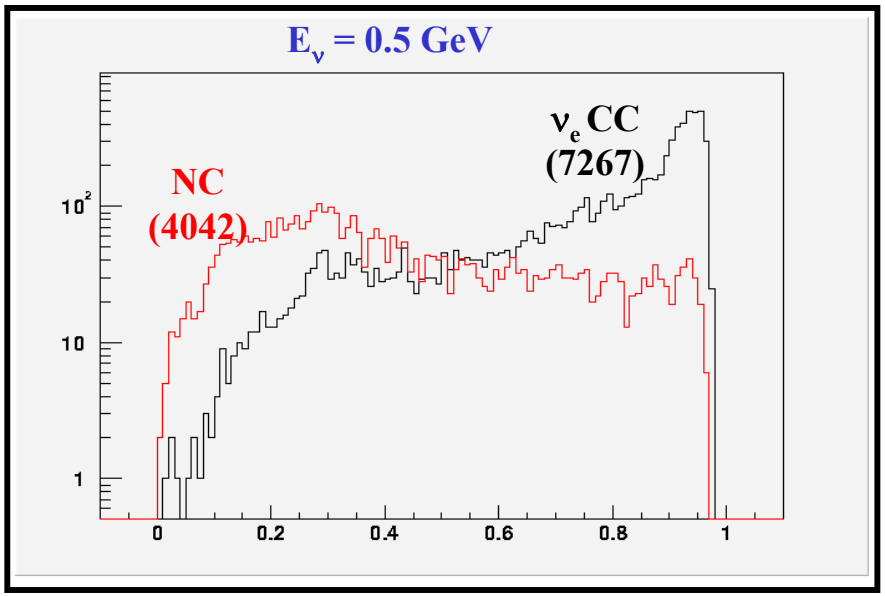


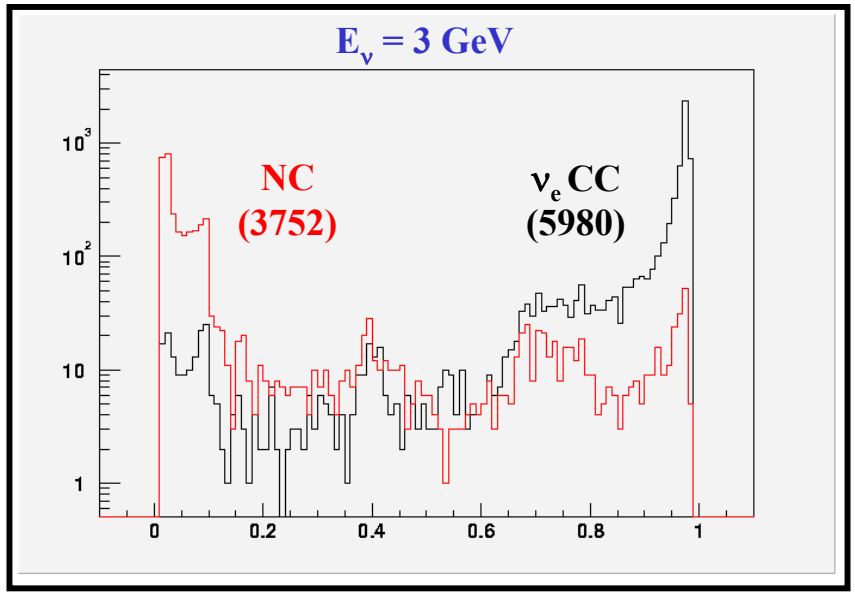


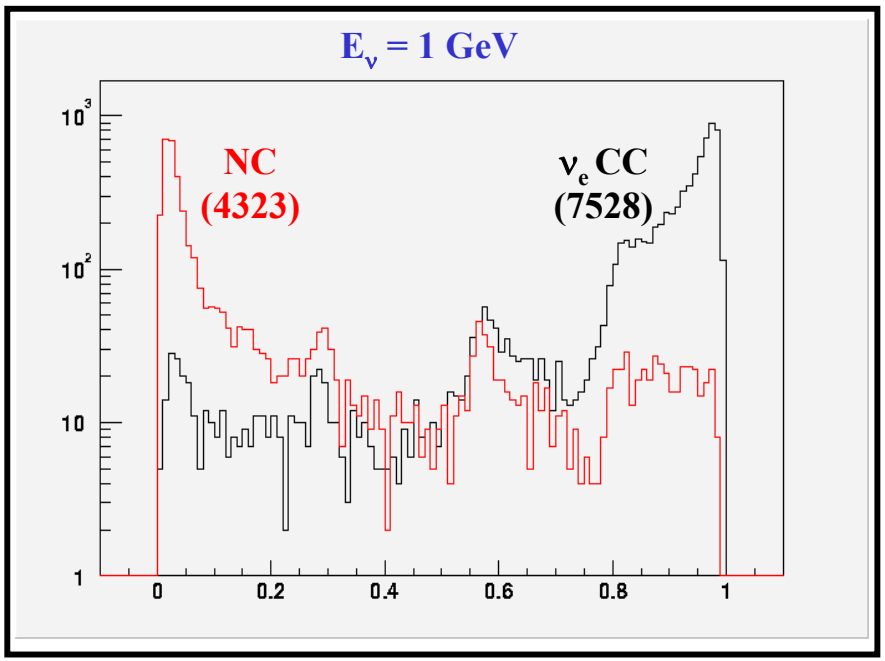
Network for ν_e CC-NC separation

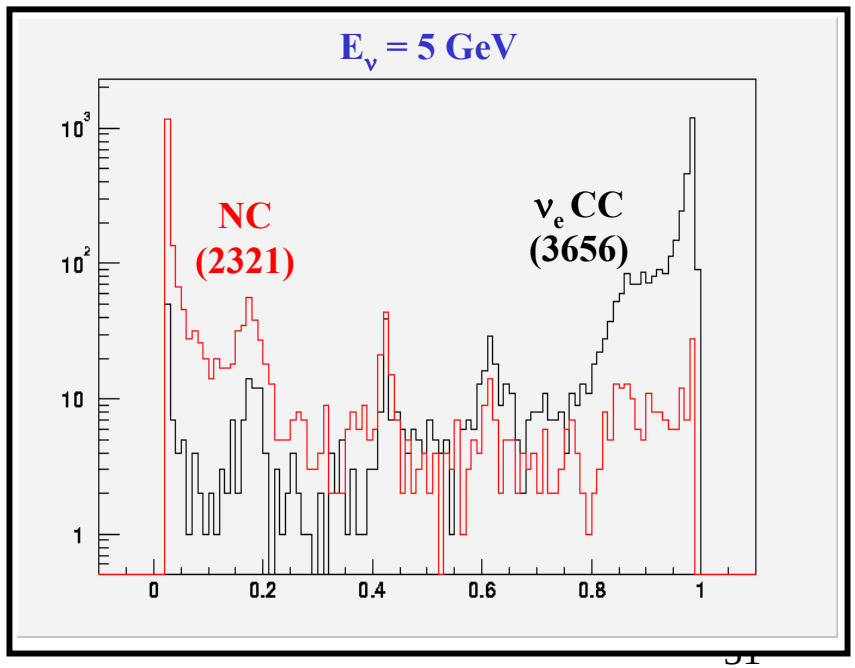
(MC file:11997 events → 7586 ν_e **CC** + 4411 **NC**)

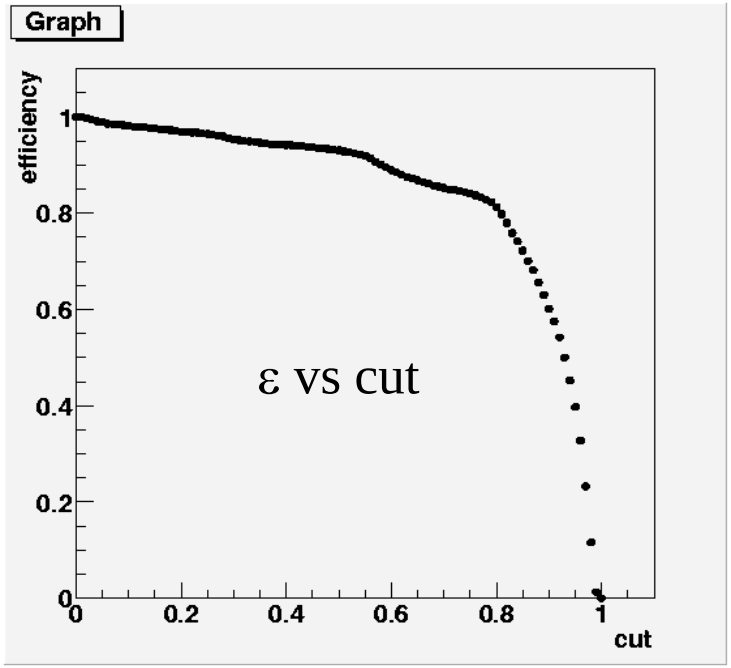


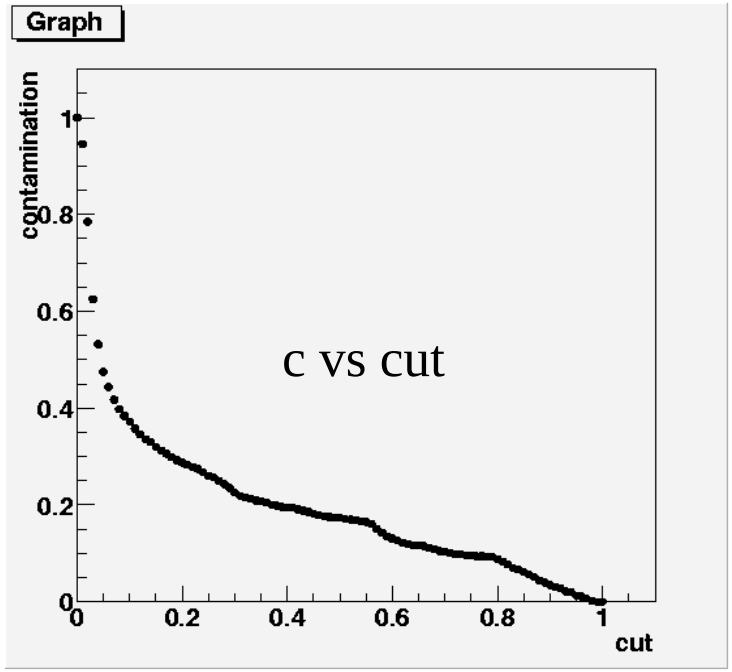

E_ν = 0.5 GeV
NC
(4042)
ν_e CC
(7267)
10^2
10
1
0
0.2
0.4
0.6
0.8
1

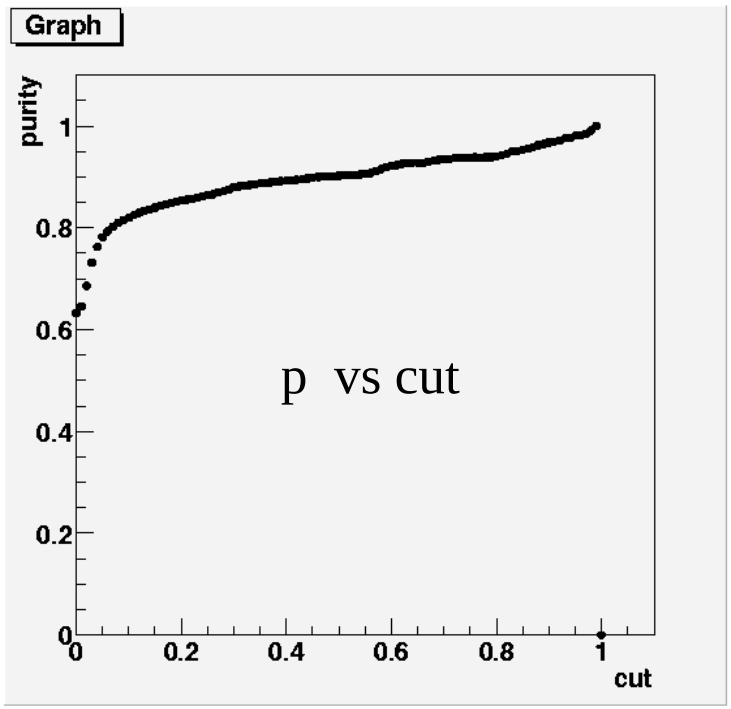

E_ν = 3 GeV
NC
(3752)
ν_e CC
(5980)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1


E_ν = 1 GeV
NC
(4323)
ν_e CC
(7528)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1


E_ν = 5 GeV
NC
(2321)
ν_e CC
(3656)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1



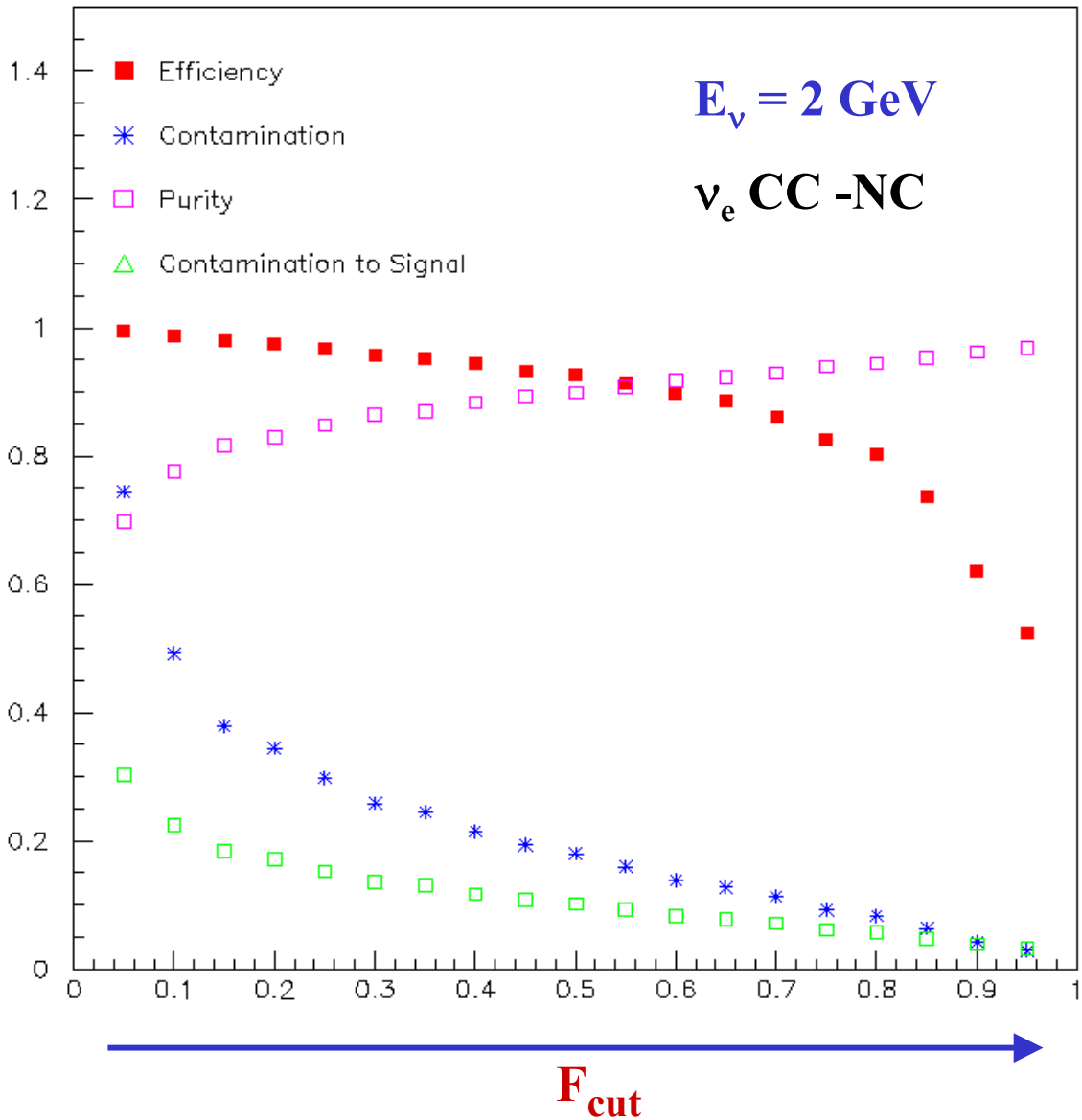




E_ν = 2 GeV

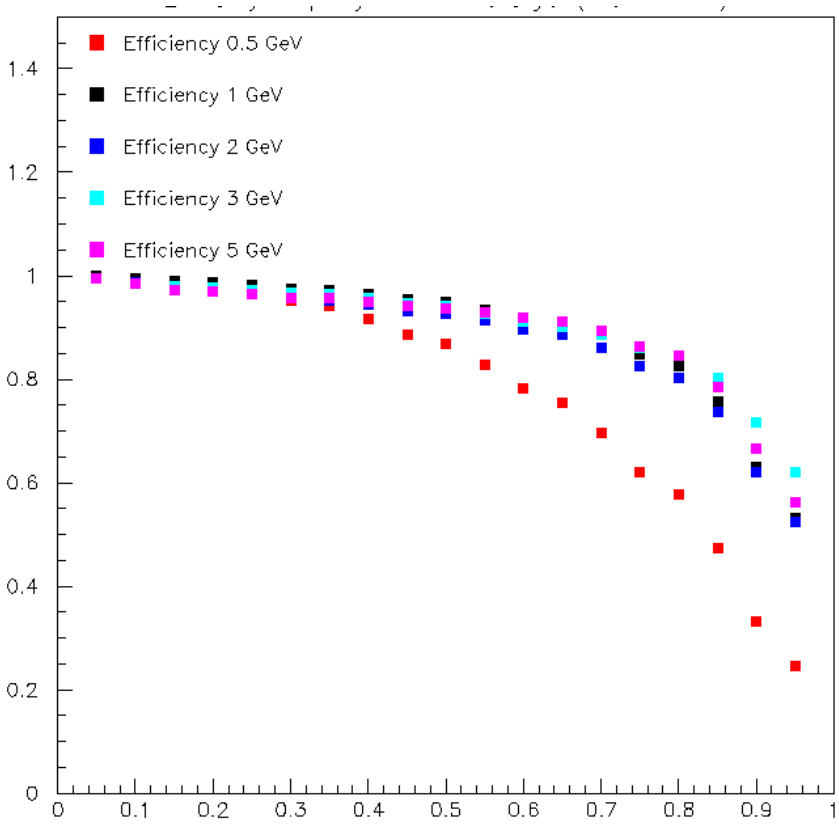
ν_e CC -NC

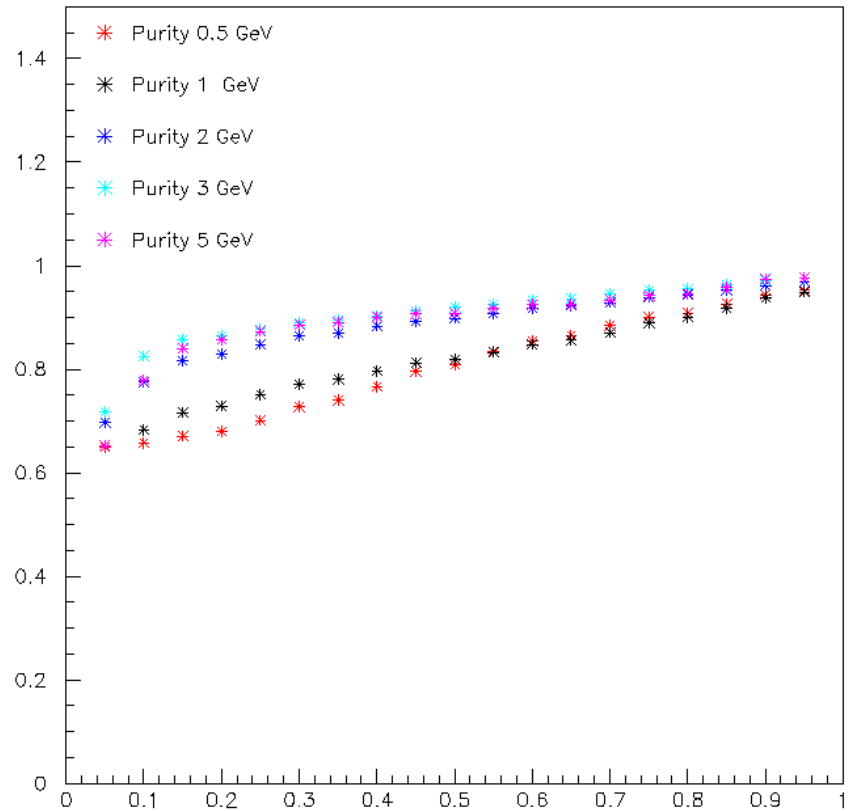
Efficiency (ε)
purity (p)
contamination (c)
vs cut


Efficiency
Contamination
Purity
Contamination to Signal
$E_\nu = 2$ GeV
ν_e CC -NC
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{cut}

Efficiency & Purity vs cut for various energies

ν_e CC -NC

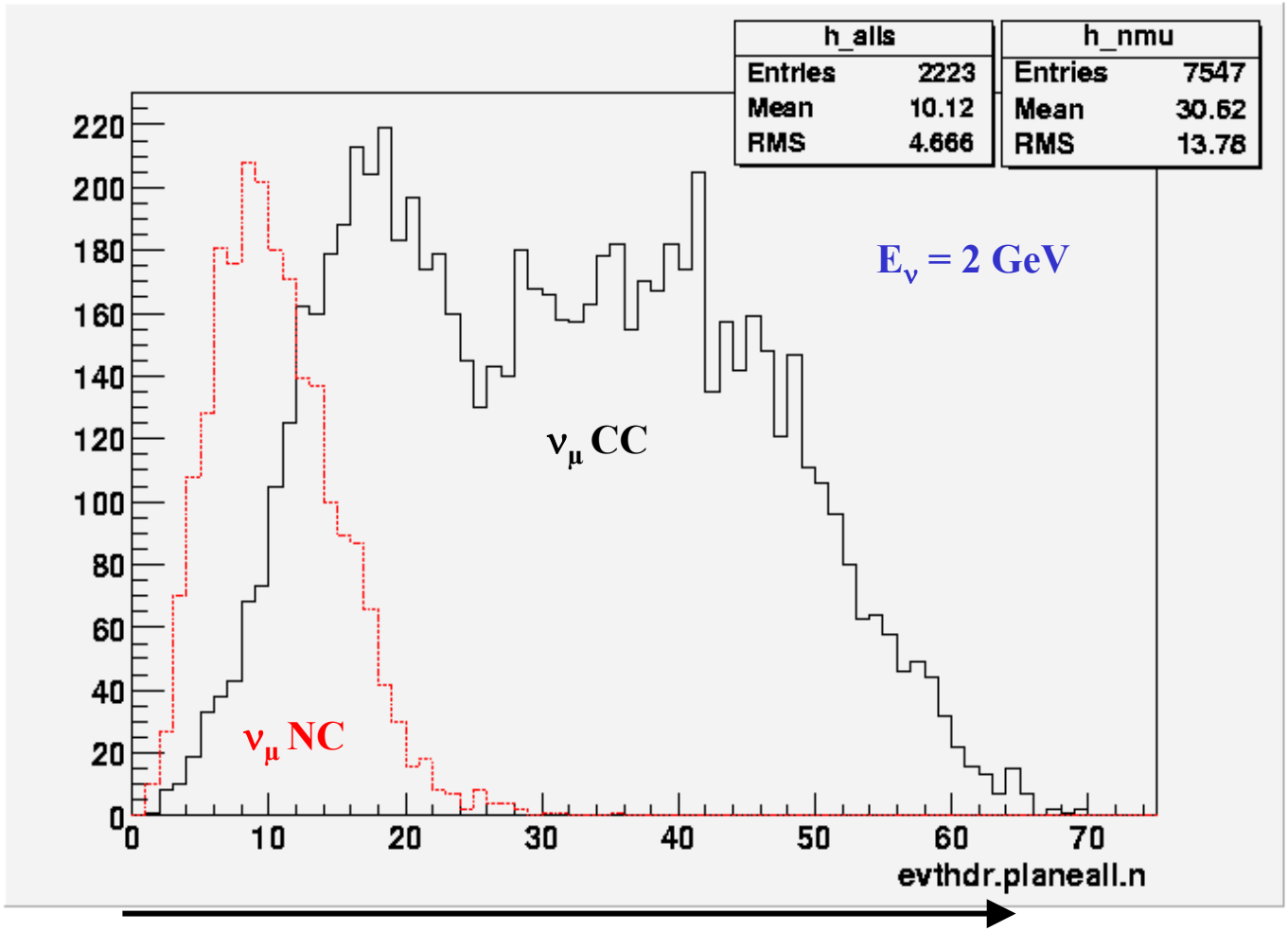




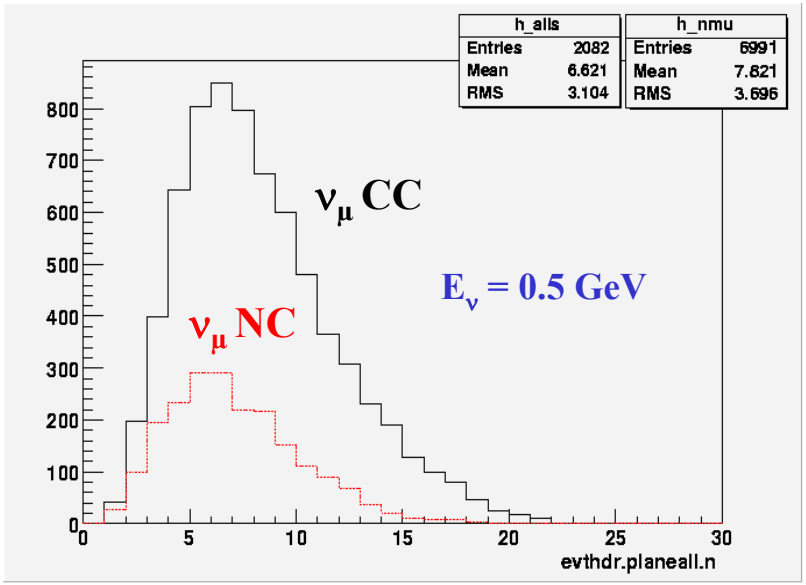
F_{cut}

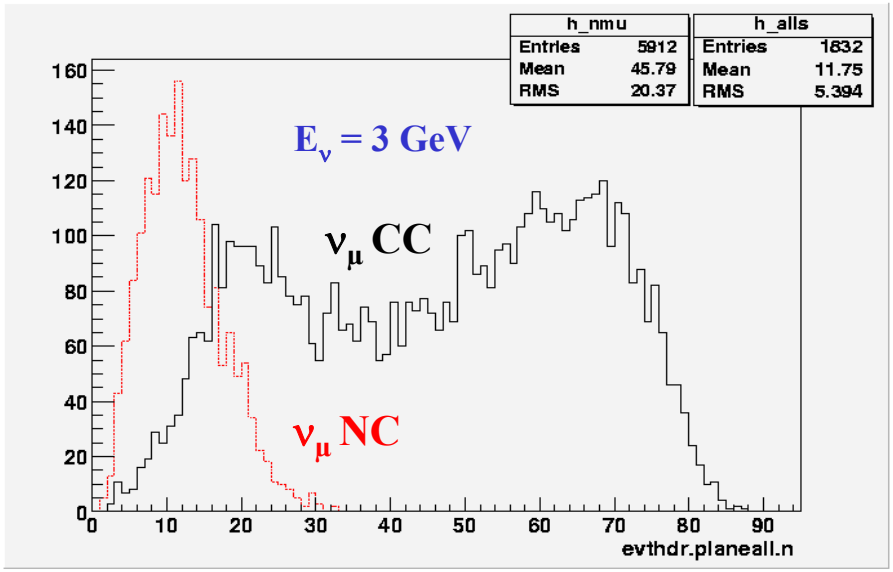
Network for ν_μ CC - ν_μ NC separation

E_ν (GeV)	ν_μ CC	ν_μ NC	ν_e CC	ν_e NC	NC Total	Grand Total
0.5	6991	2082	7267	1960	4042	18300
1.0	7428	2208	7528	2115	4323	19279
2.0	7547	2223	7586	2188	4411	19544
3.0	5912	1832	5980	1920	3752	15644
5.0	3499	1177	3656	1144	2321	9476

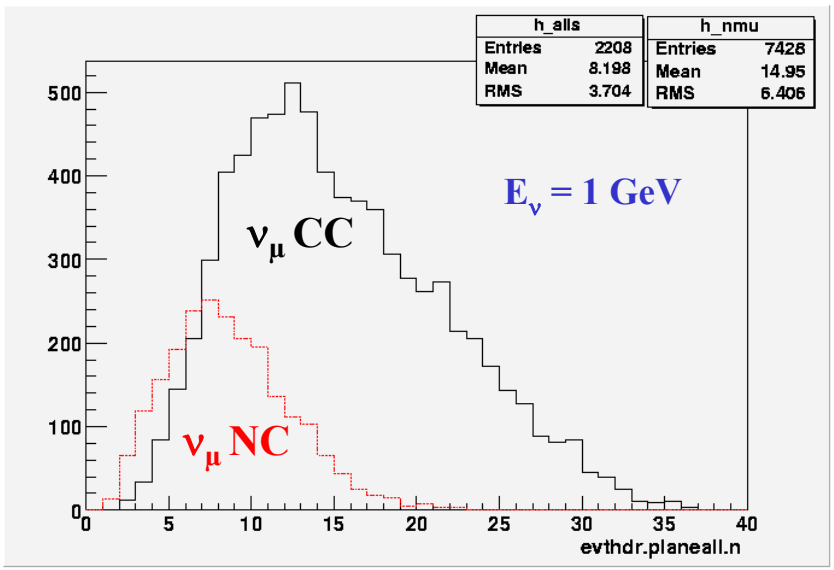


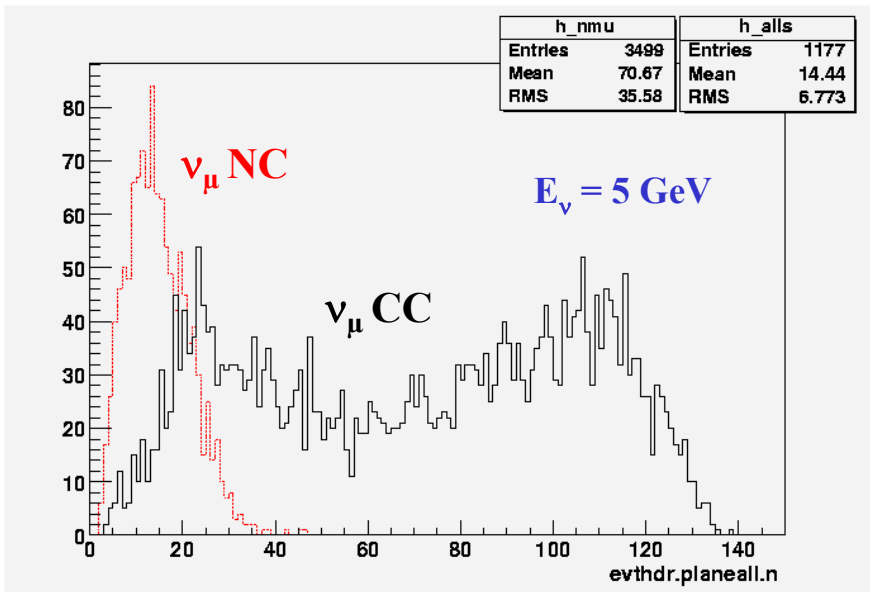
Number of Planes

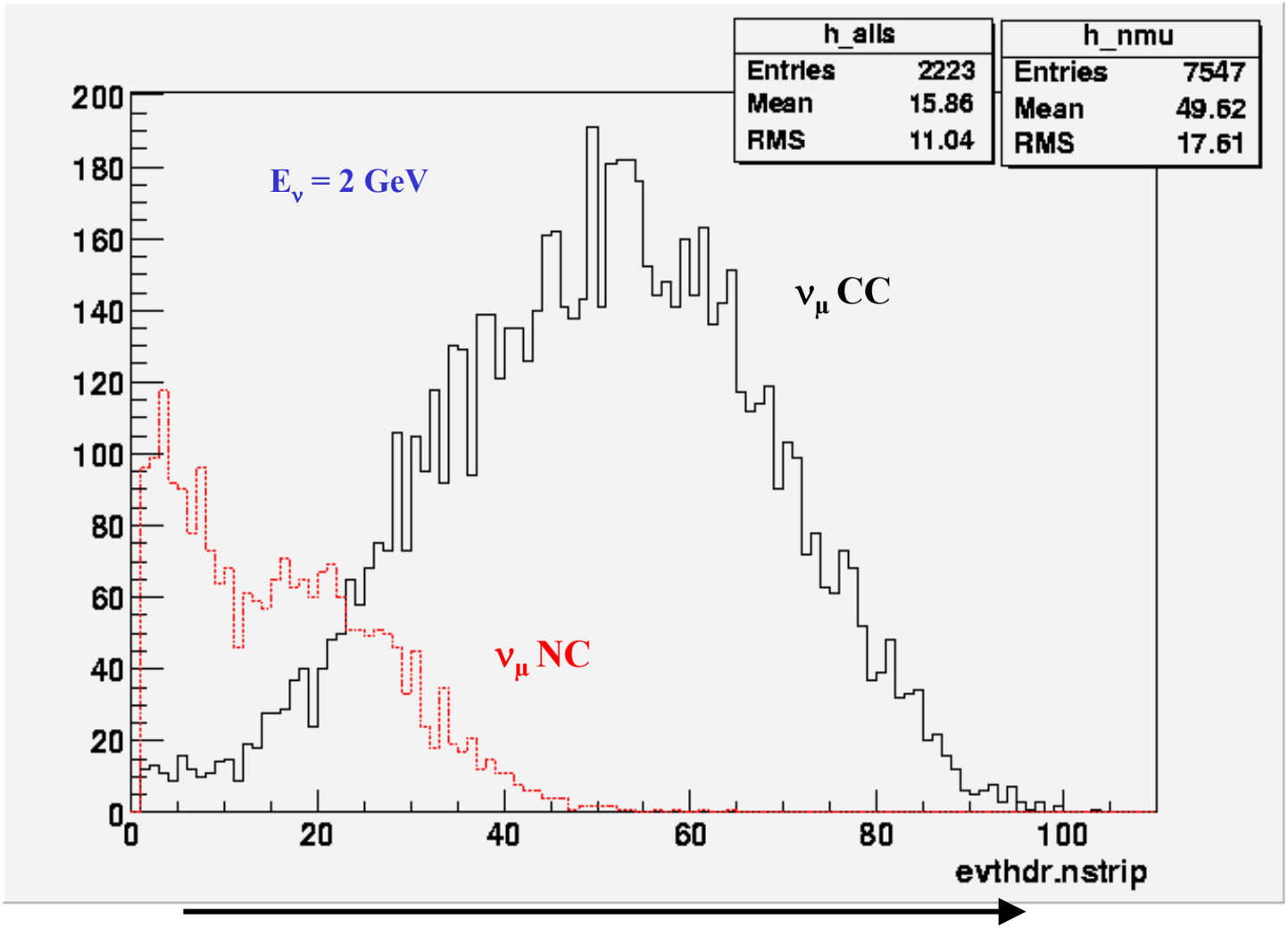




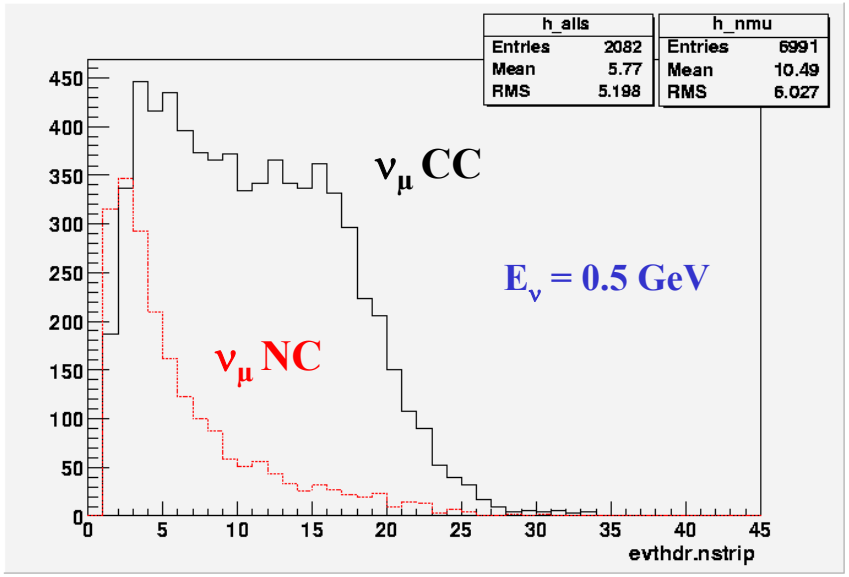
Number of Planes vs E_ν

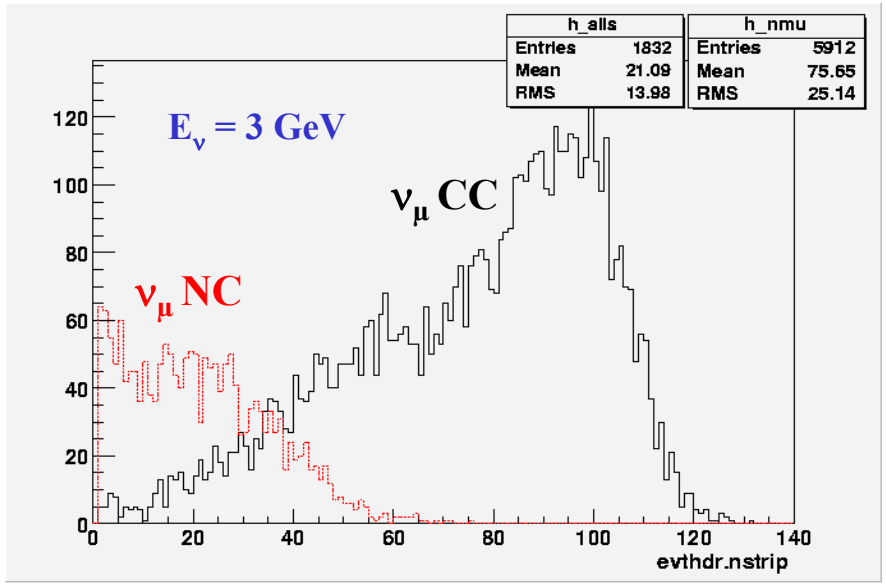




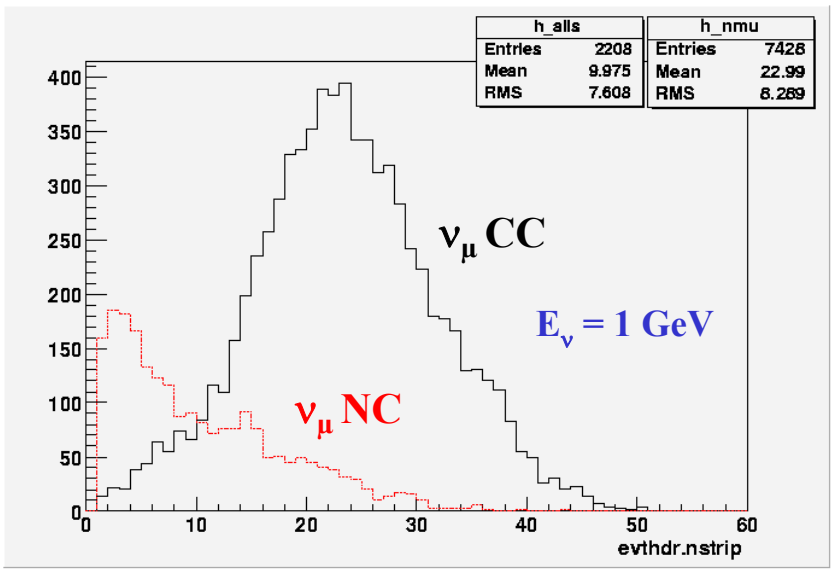


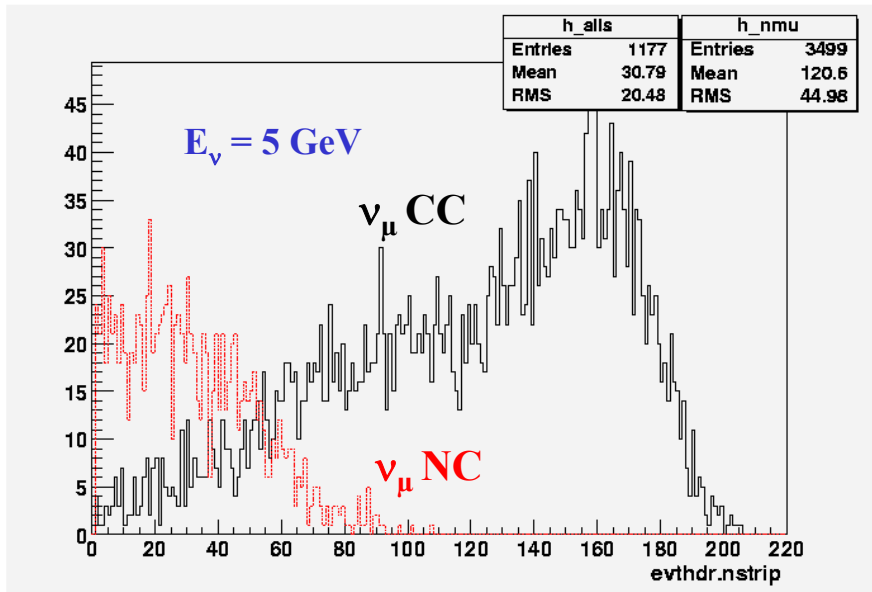
Number of Strips

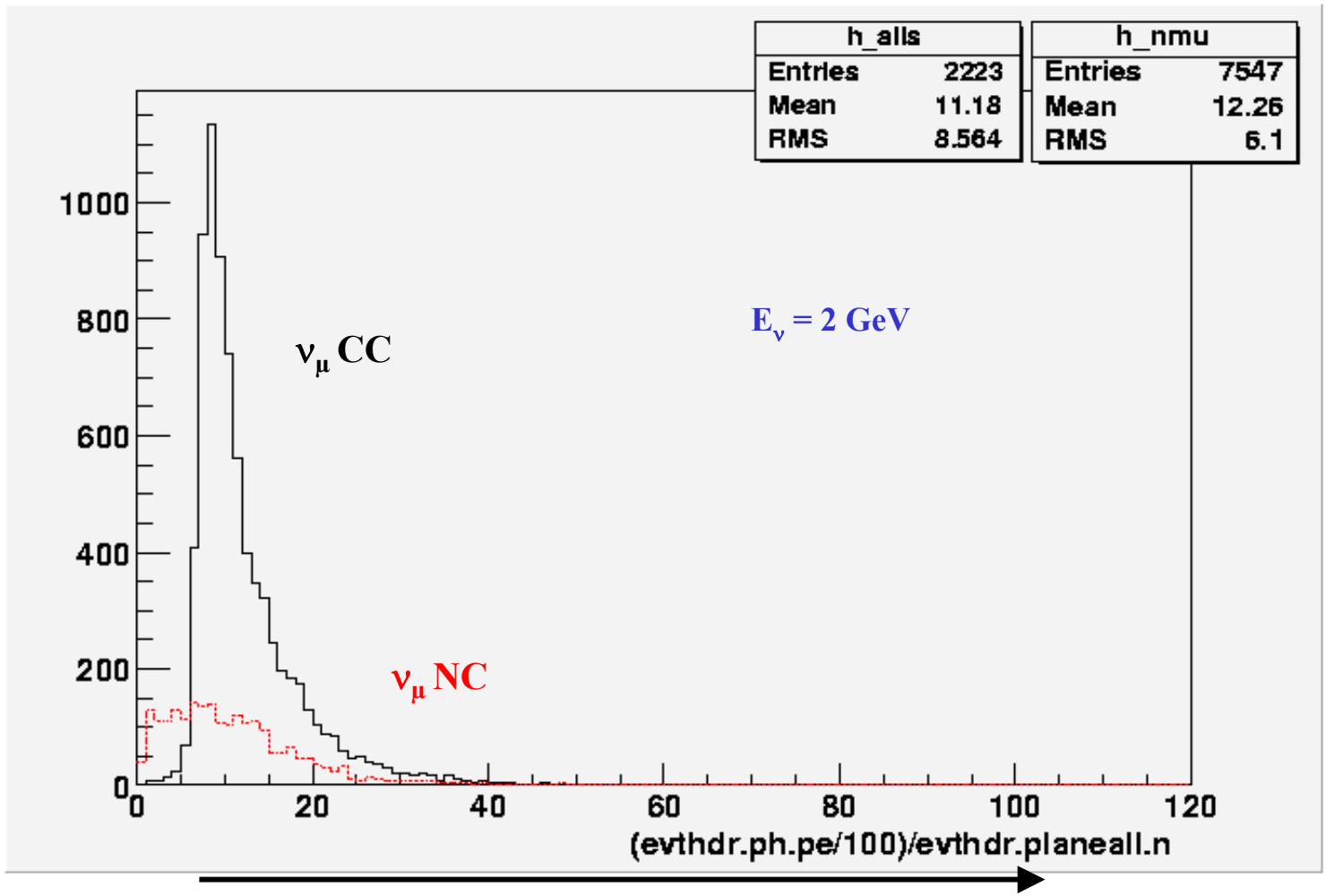




Number of Strips vs E_ν

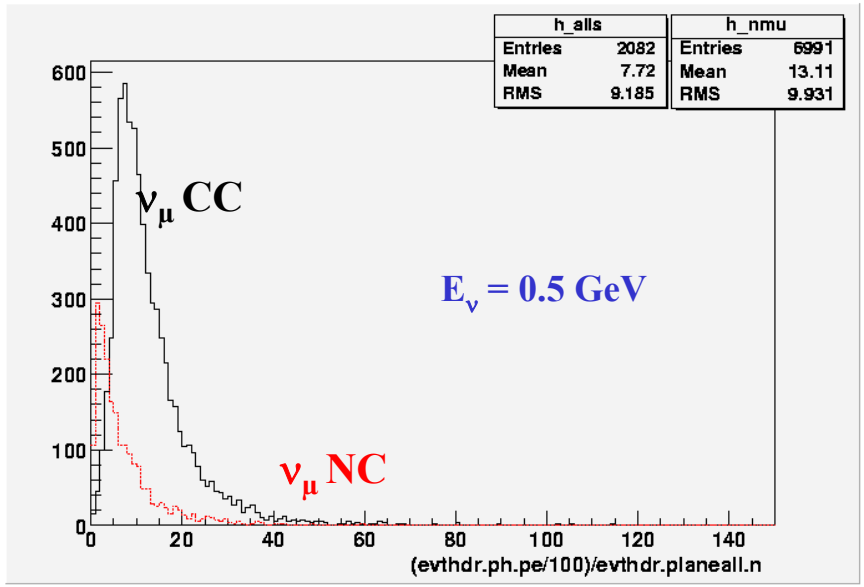


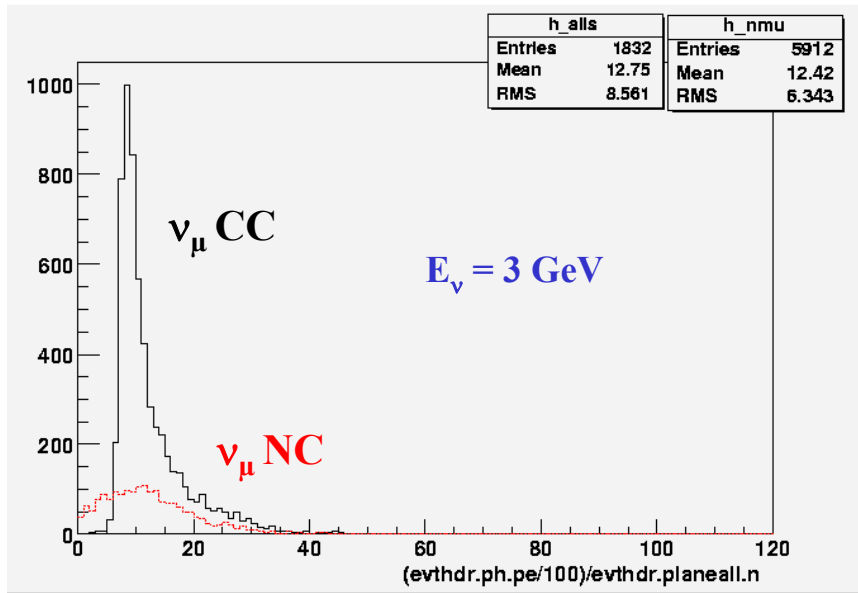


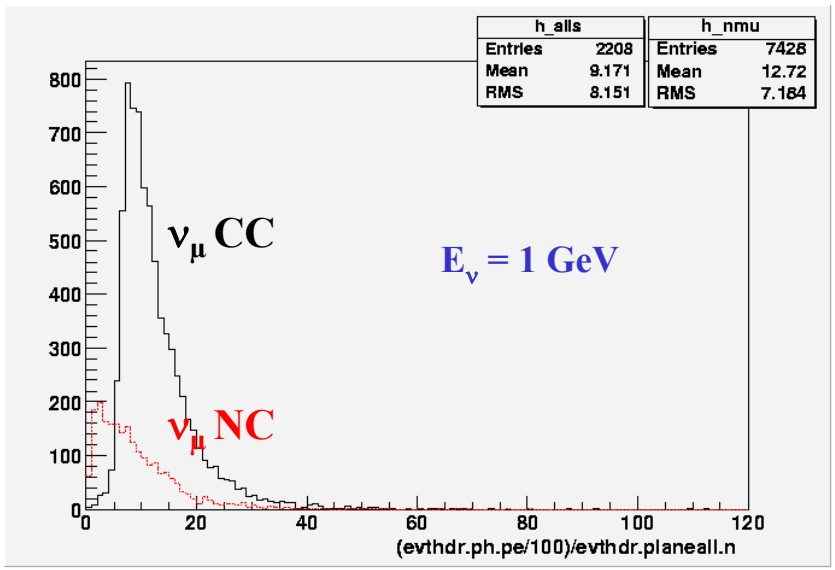


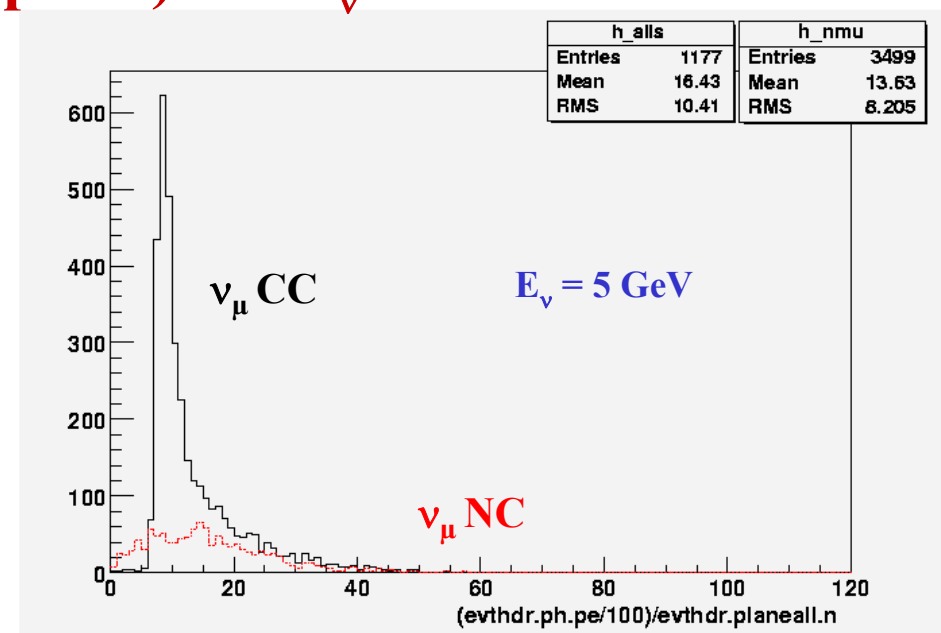
Energy (pe)/plane

Energy (PE/plane) vs E_ν



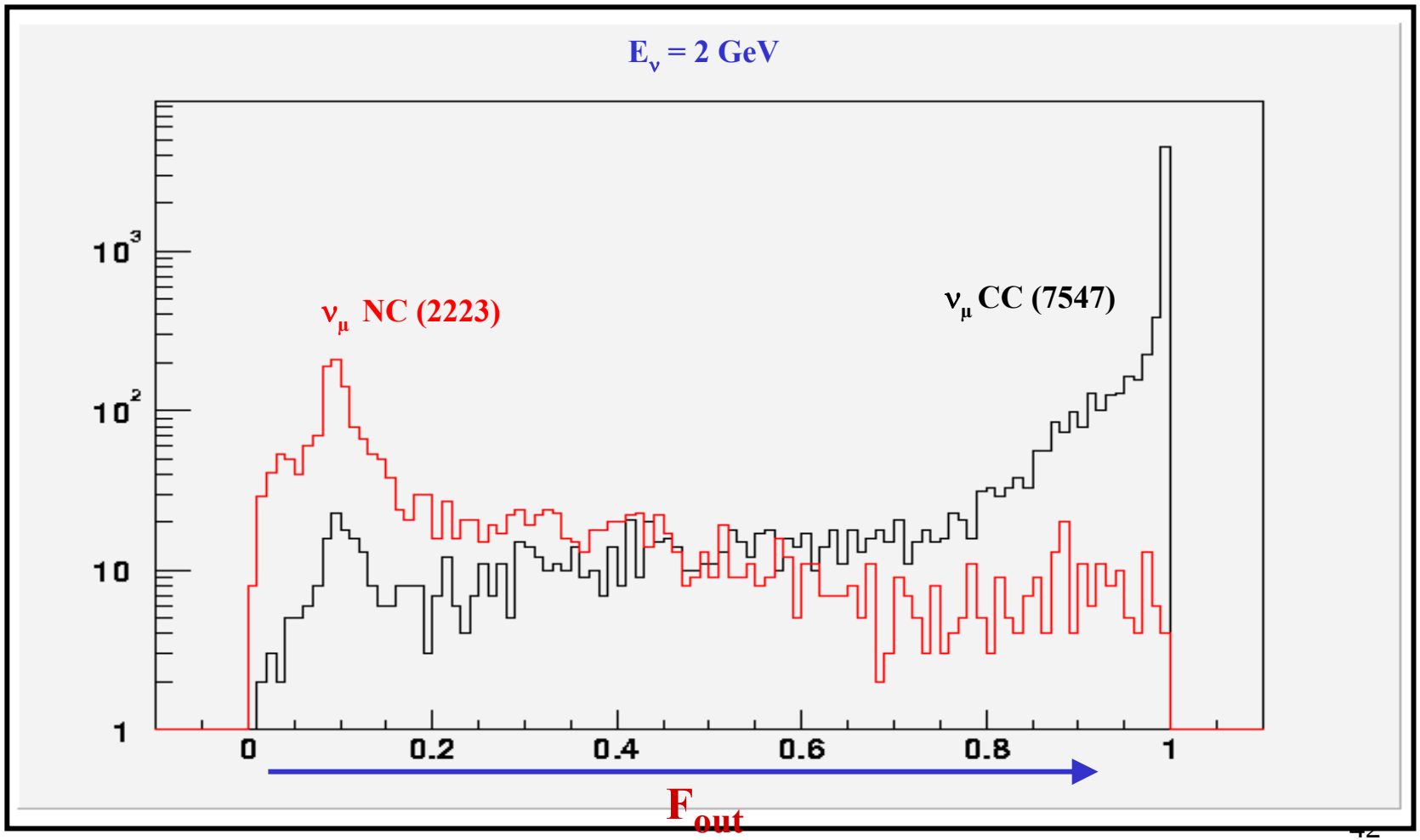


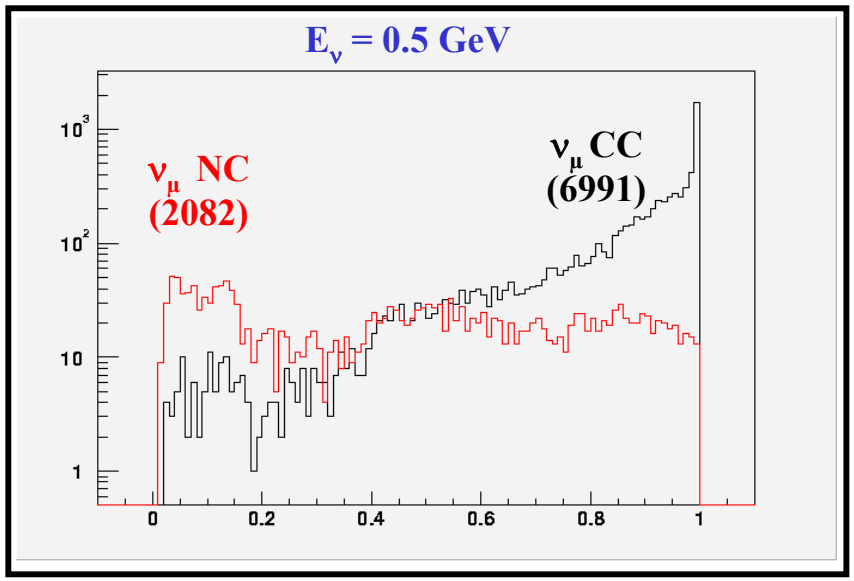


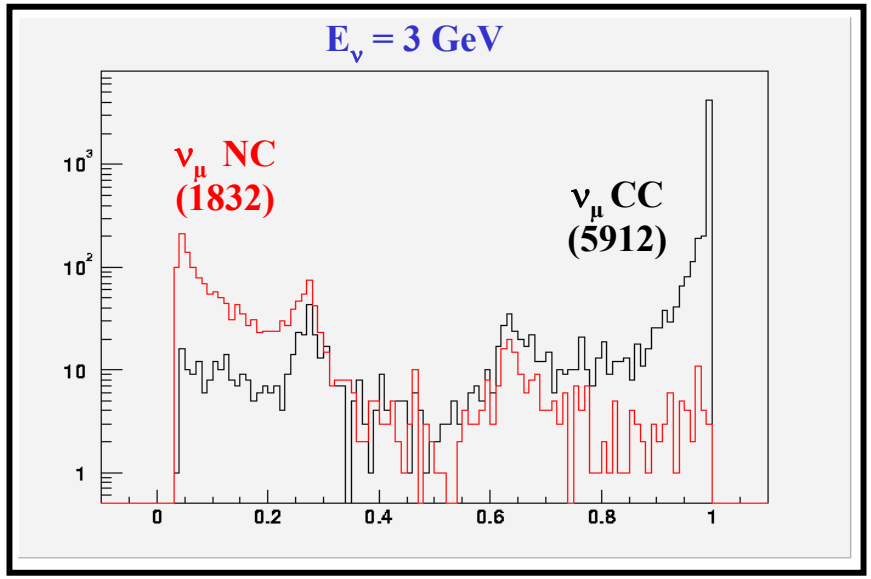


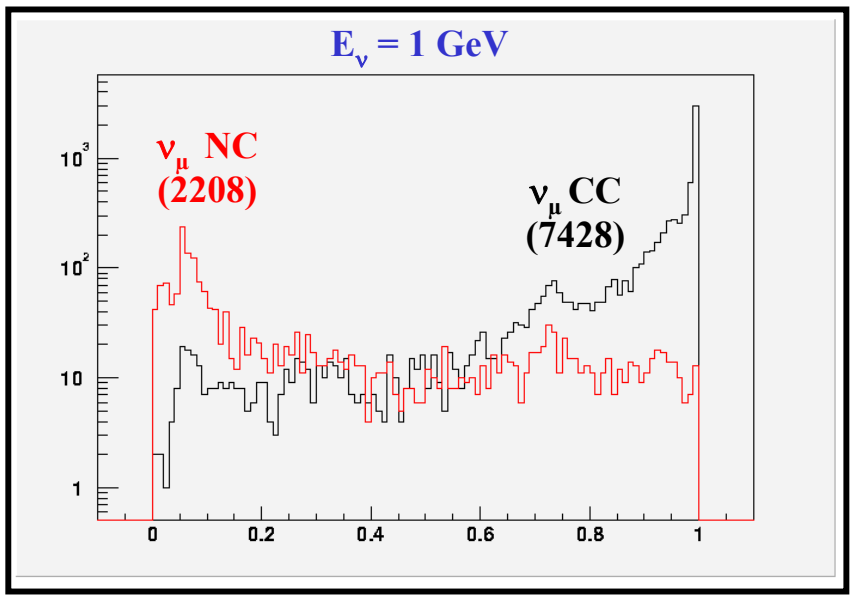
Network for ν_μ CC- ν_μ NC separation

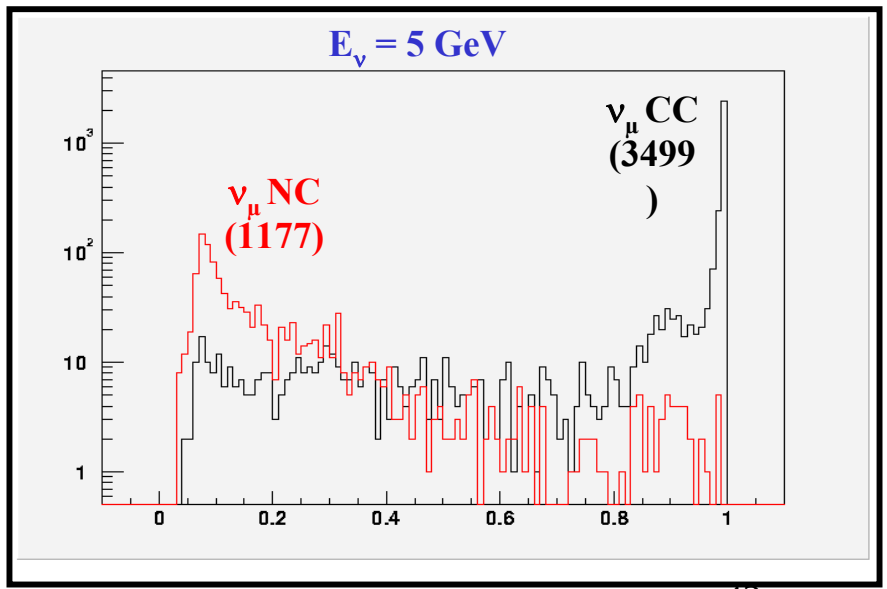
(MC file:9770 events → 7547 ν_μ CC + 2223 ν_μ NC)

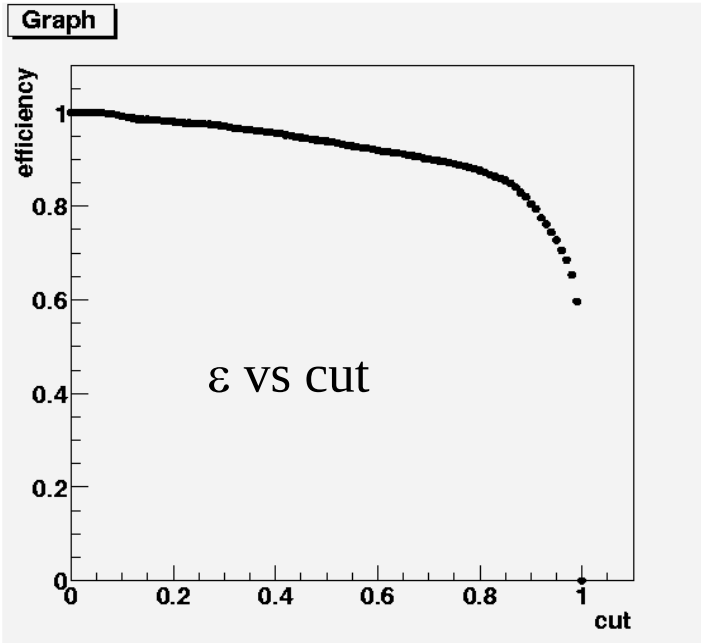


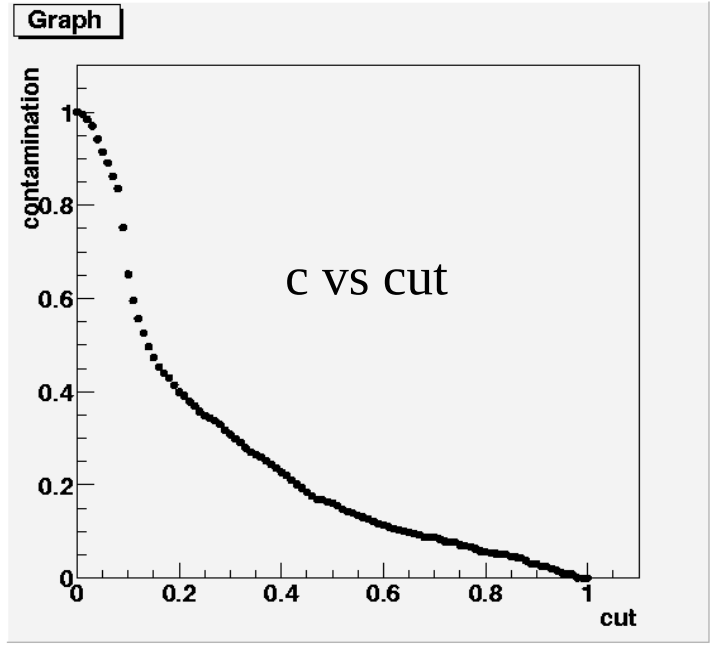

E_ν = 0.5 GeV
ν_μ NC
(2082)
ν_μ CC
(6991)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1

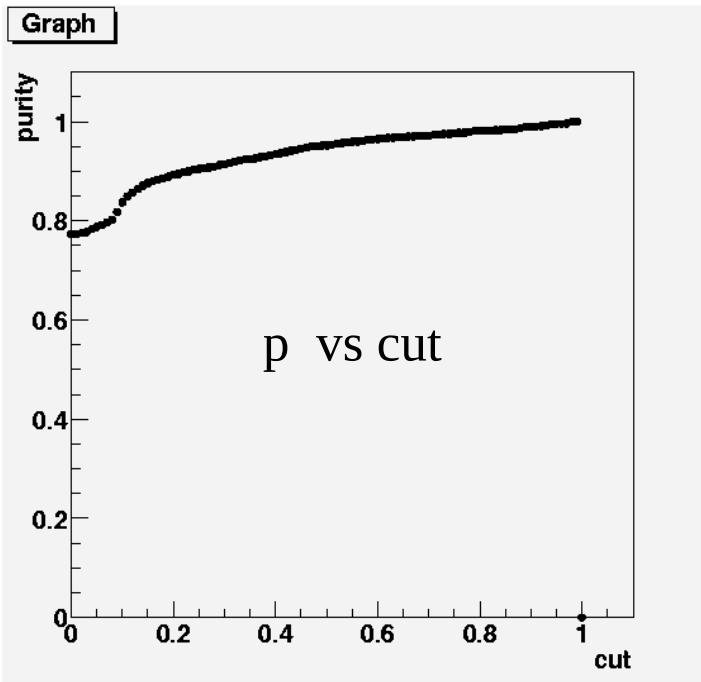

E_ν = 3 GeV
ν_μ NC
(1832)
ν_μ CC
(5912)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1


E_ν = 1 GeV
ν_μ NC
(2208)
ν_μ CC
(7428)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1


E_ν = 5 GeV
ν_μ CC
(3499
)
ν_μ NC
(1177)
10^3
10^2
10
1
0
0.2
0.4
0.6
0.8
1



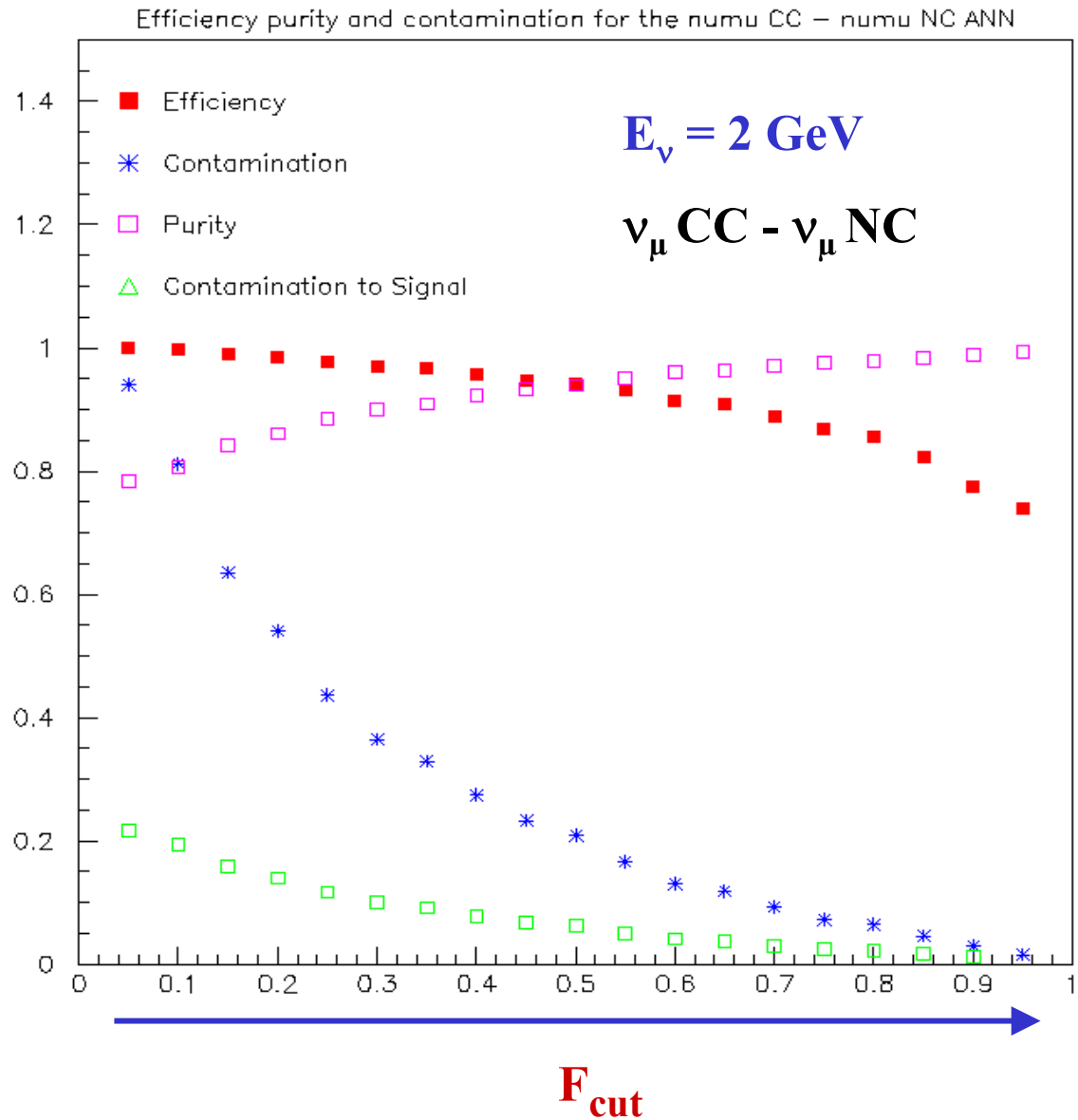




E_ν = 2 GeV

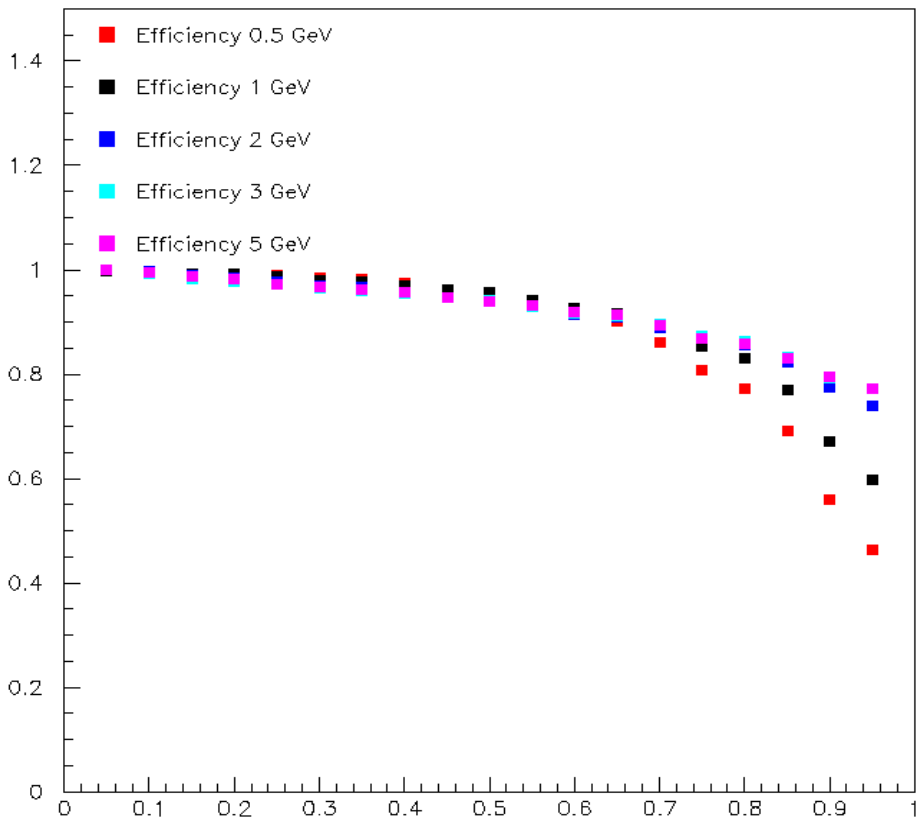
ν_μ CC - ν_μ NC

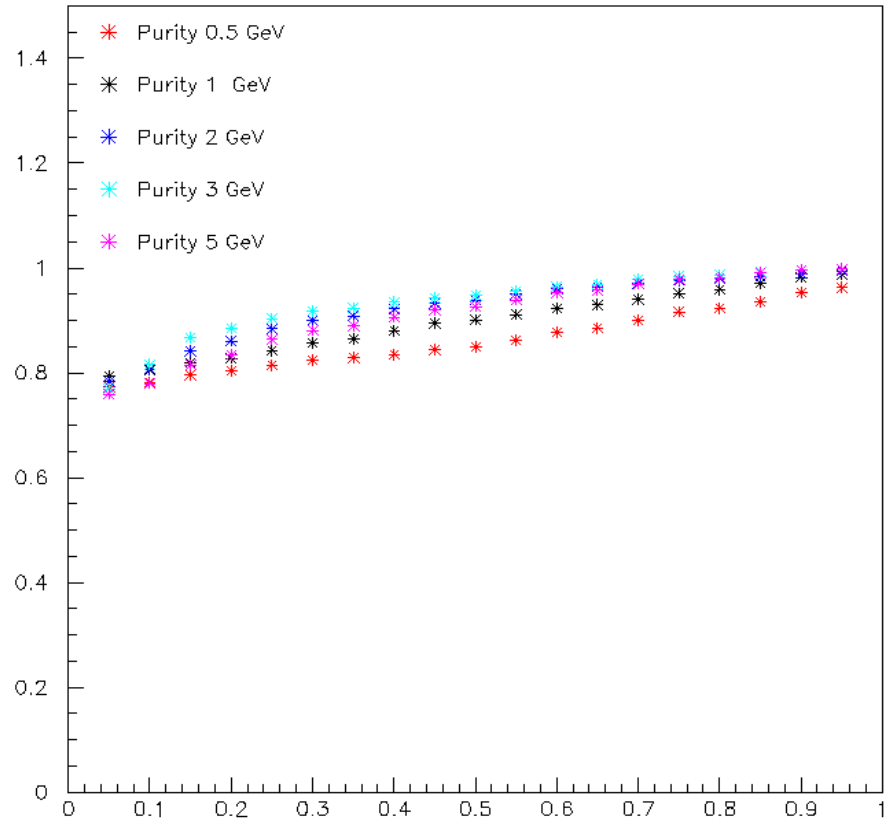
Efficiency (ε)
purity (p)
contamination (c)
vs cut


Efficiency purity and contamination for the numu CC – numu NC ANN
Efficiency
Contamination
Purity
Contamination to Signal
$E_\nu = 2$ GeV
ν_μ CC - ν_μ NC
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{cut}

Efficiency & Purity vs cut for various energies

ν_μ CC - ν_μ NC





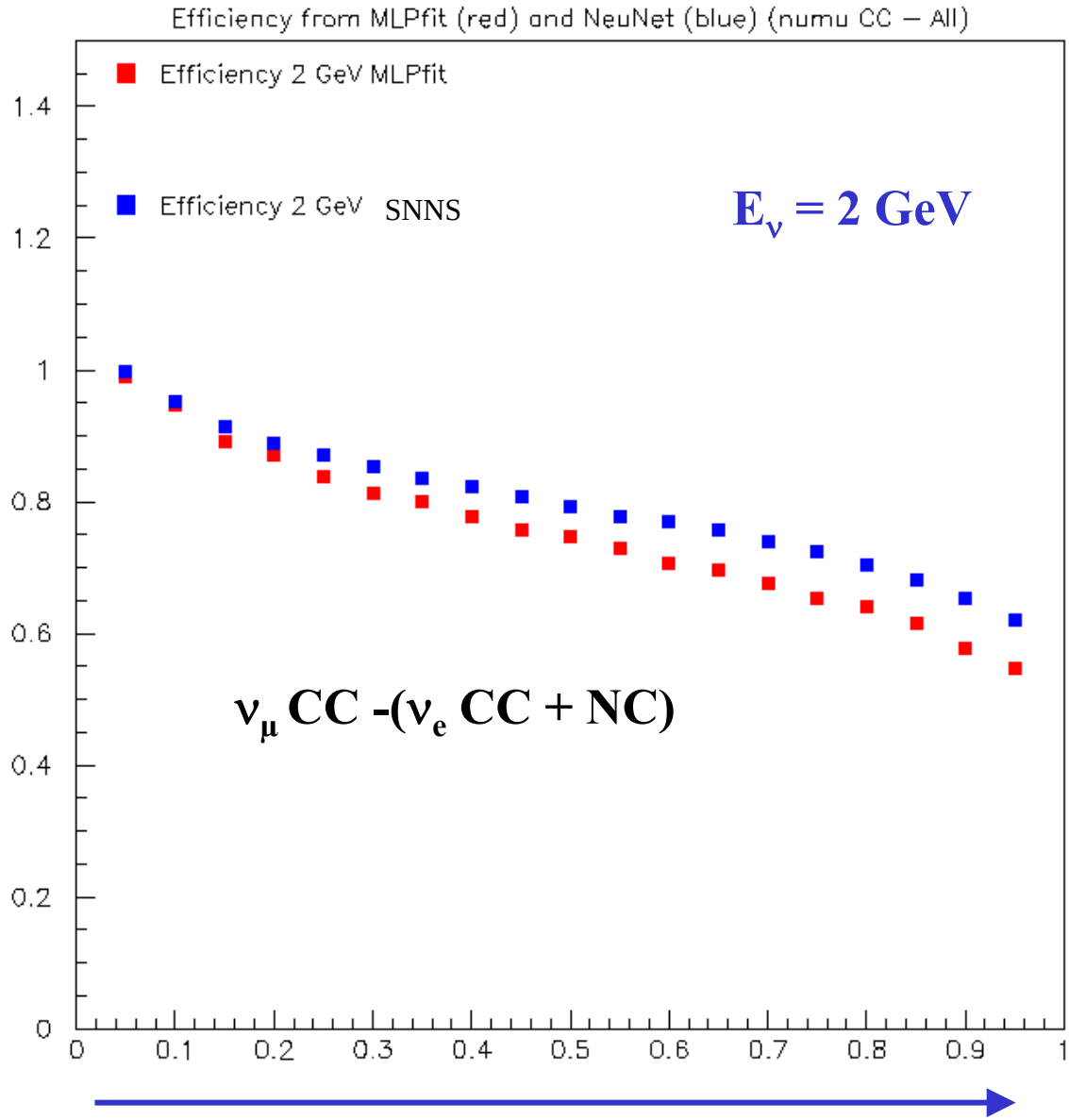
F_{cut}

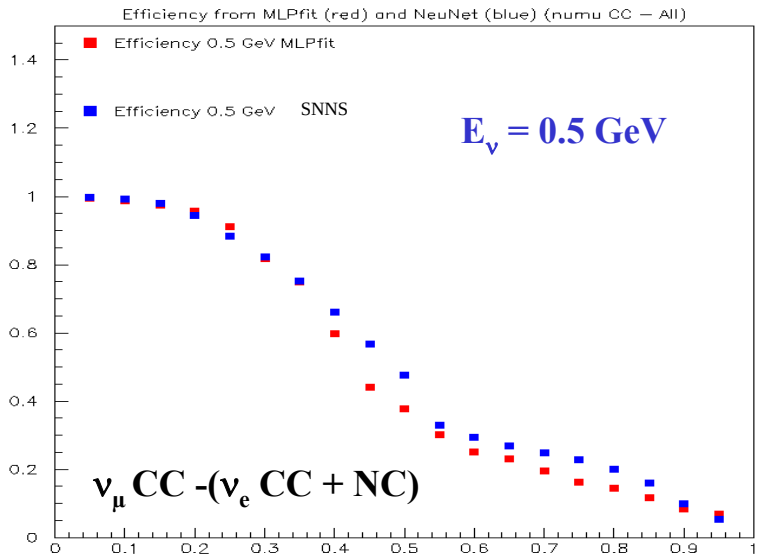
Comparison between MLP and SNNS

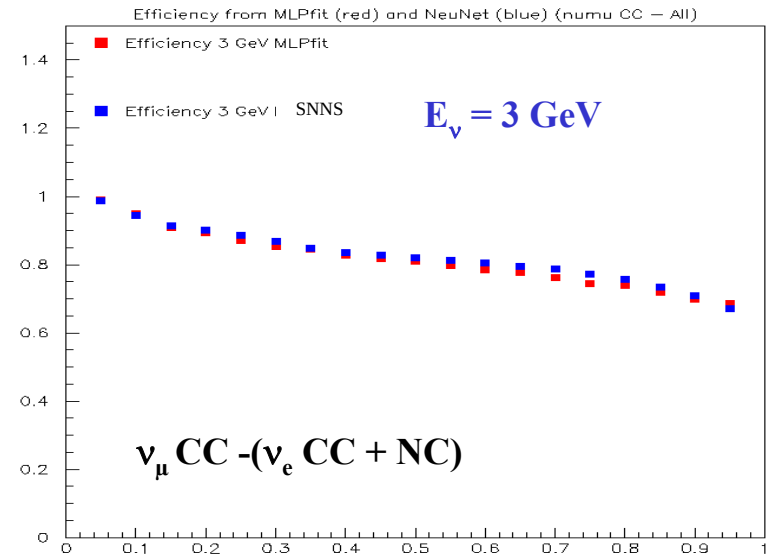
- **Run the 3 ANN with both MLPfit and SNNS packages.**

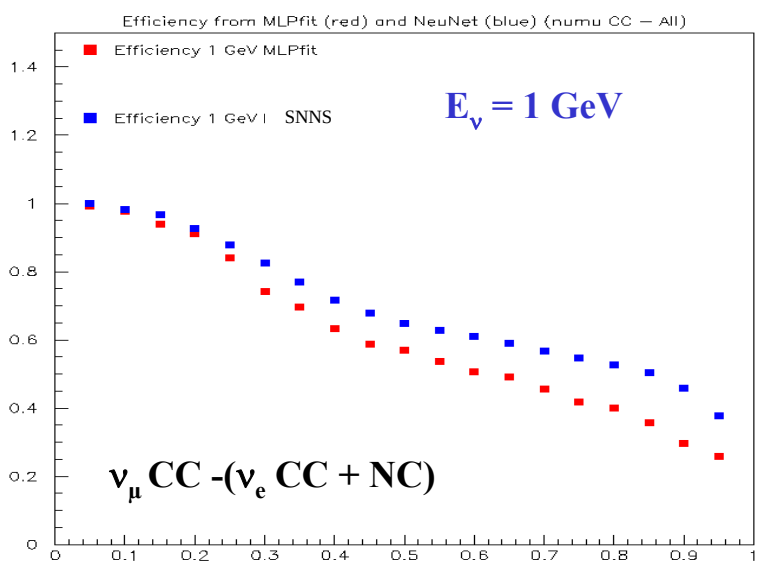
- **For each ANN we tried to train with the necessary number of "epochs" (training cycles) just to achieve optimum training and avoid overtraining.**

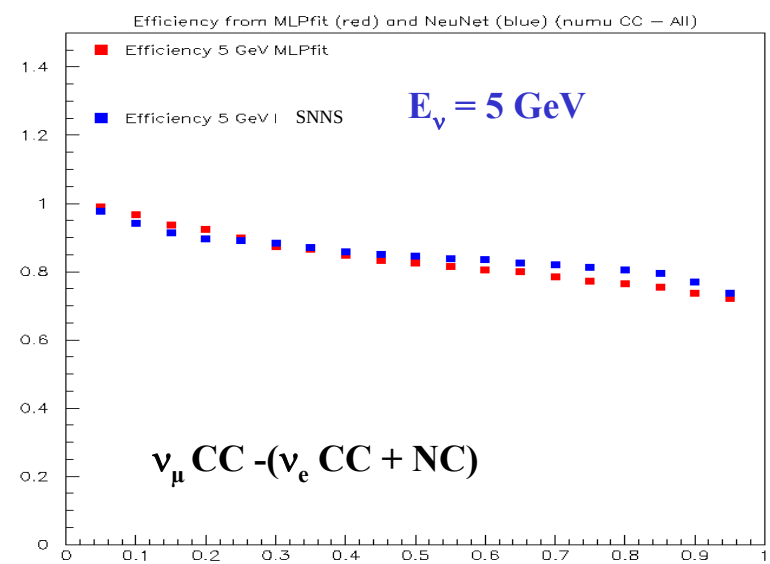
- Studied **efficiency** and **purity** versus $\mathbf{f_{out}}$ for all **energies.**

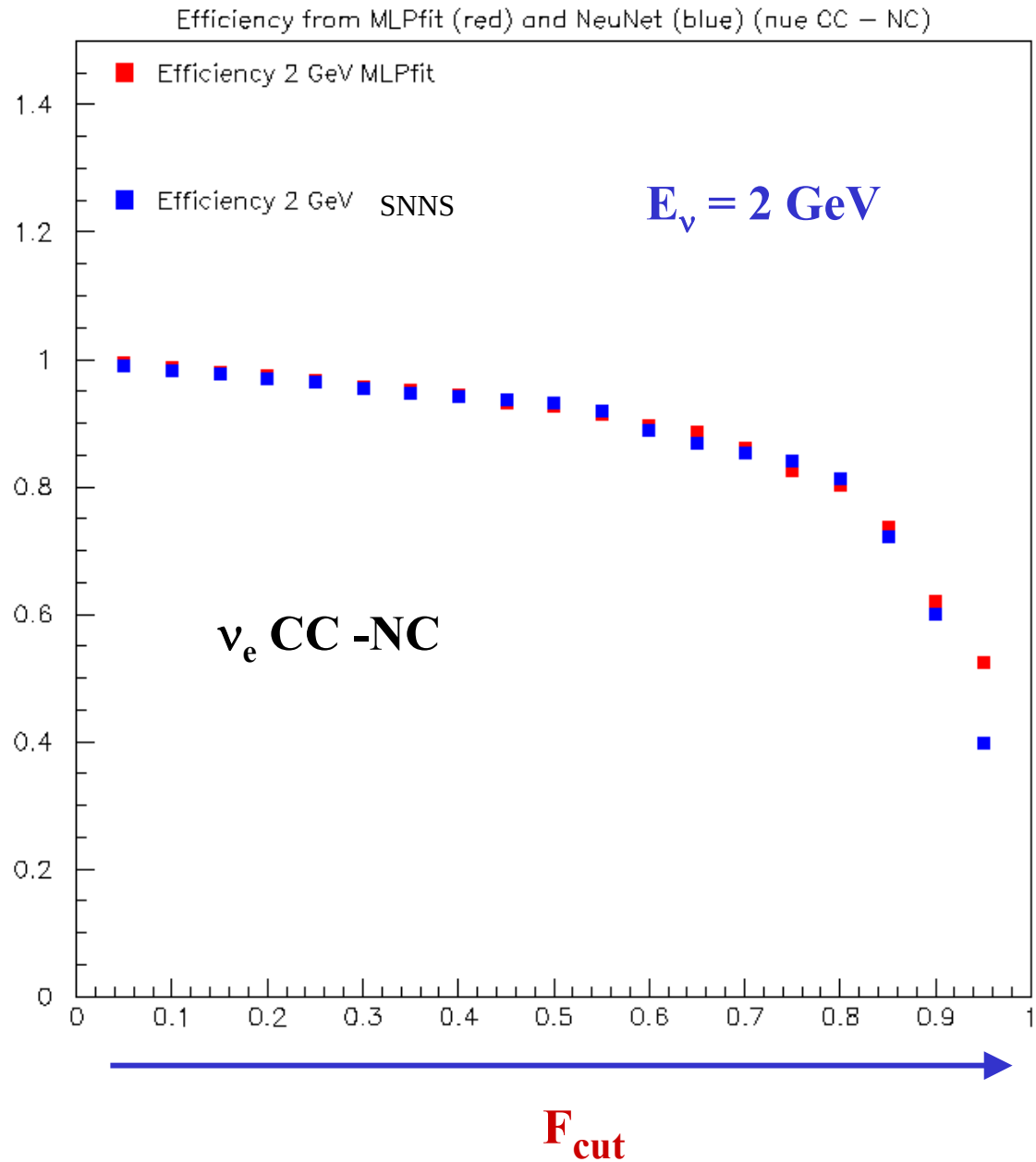

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – All)
Efficiency 2 GeV MLPfit
Efficiency 2 GeV SNNS
$E_\nu = 2$ GeV
ν_μ CC -(ν_e CC + NC)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{cut}

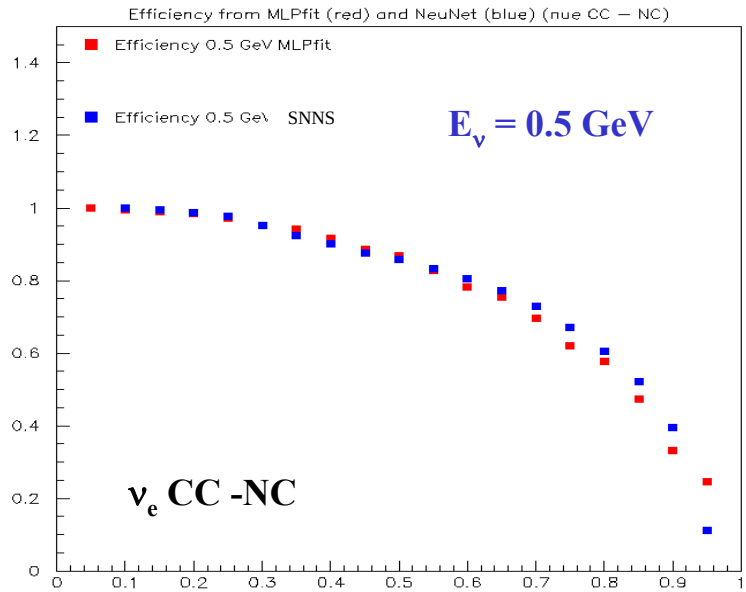

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – All)
Efficiency 0.5 GeV MLPfit
Efficiency 0.5 GeV SNNS
$E_\nu = 0.5$ GeV
ν_μ CC -(ν_e CC + NC)

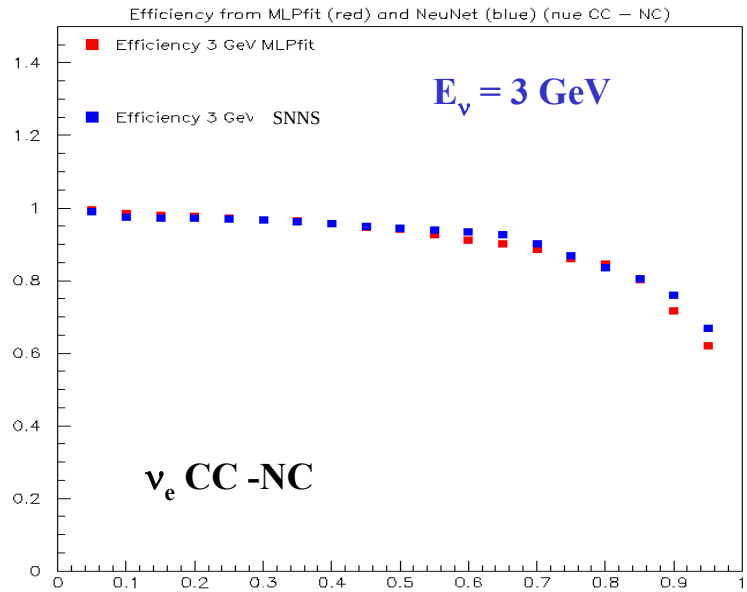

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – All)
Efficiency 3 GeV MLPfit
Efficiency 3 GeV SNNS
$E_\nu = 3$ GeV
ν_μ CC -(ν_e CC + NC)

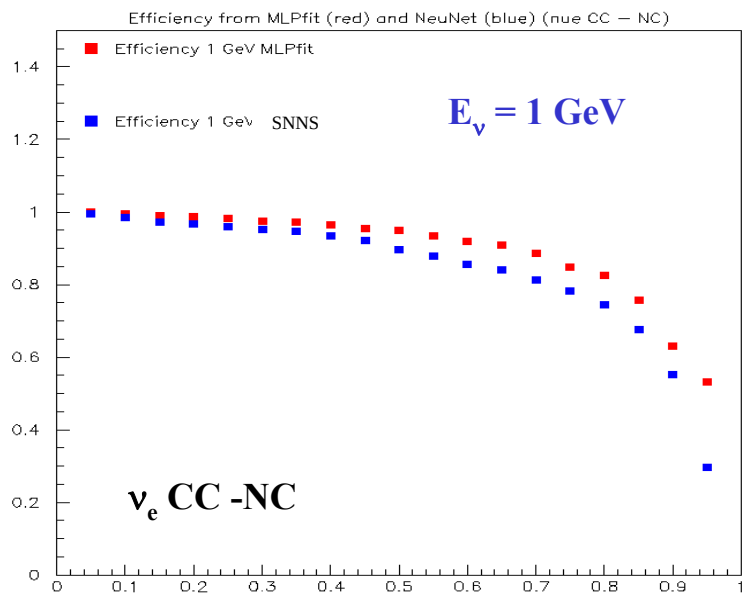

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – All)
Efficiency 1 GeV MLPfit
Efficiency 1 GeV SNNS
$E_\nu = 1$ GeV
ν_μ CC -(ν_e CC + NC)

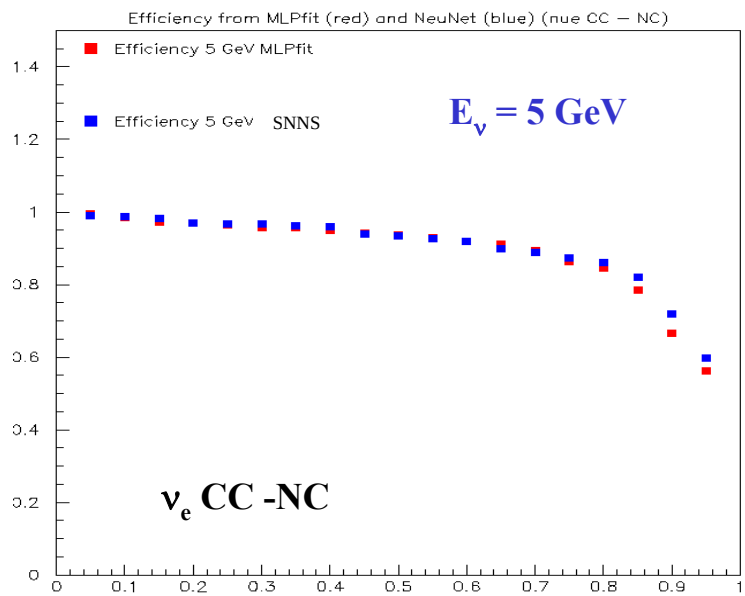

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – All)
Efficiency 5 GeV MLPfit
Efficiency 5 GeV SNNS
$E_\nu = 5$ GeV
ν_μ CC -(ν_e CC + NC)

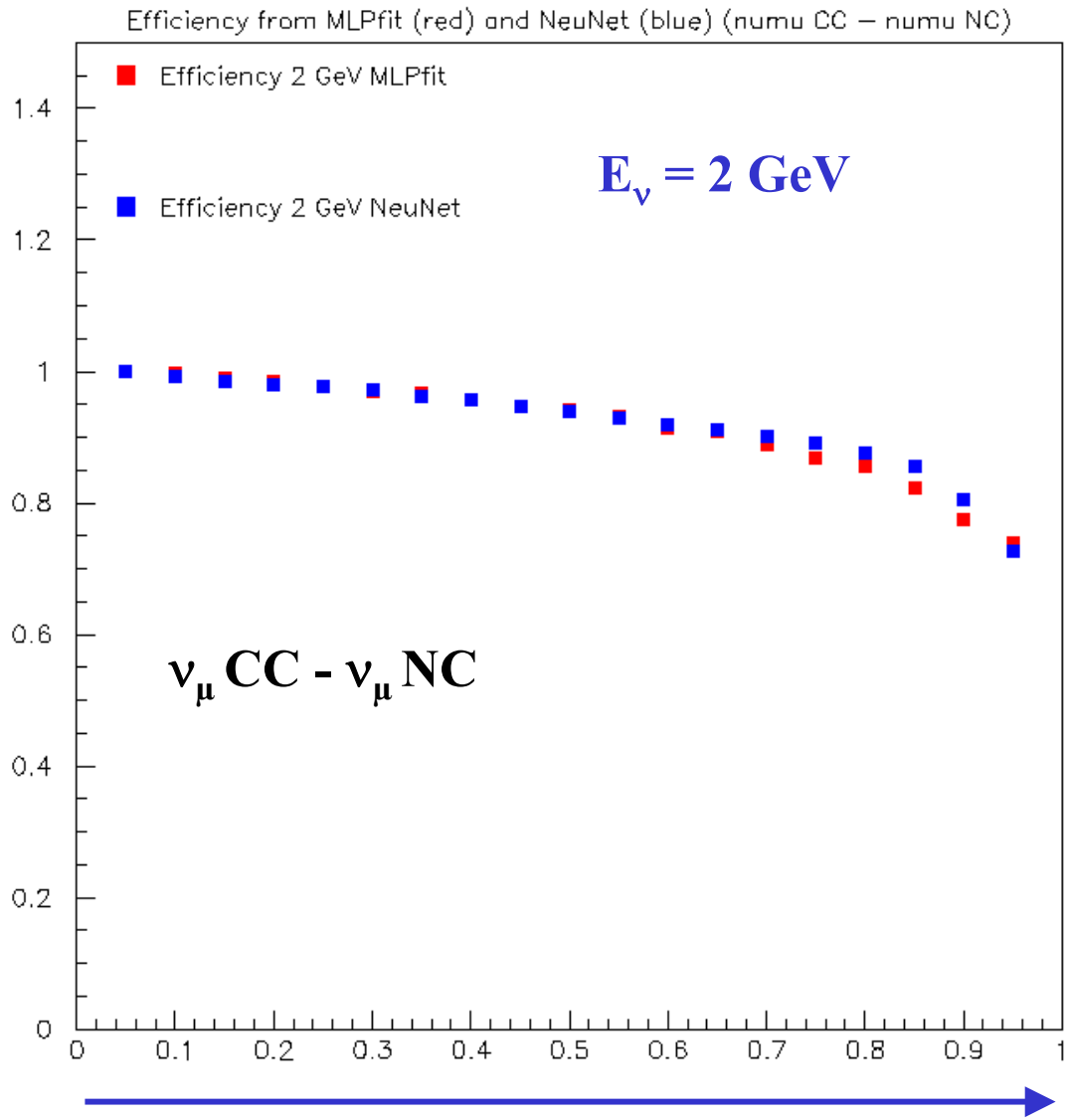

Efficiency from MLPfit (red) and NeuNet (blue) (nue CC – NC)
Efficiency 2 GeV MLPfit
Efficiency 2 GeV
SNNS
$E_\nu = 2$ GeV
ν_e CC -NC
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{cut}


Efficiency from MLPfit (red) and NeuNet (blue) (nue CC – NC)
Efficiency 0.5 GeV MLPfit
Efficiency 0.5 GeV SNNS
$E_\nu = 0.5$ GeV
ν_e CC -NC

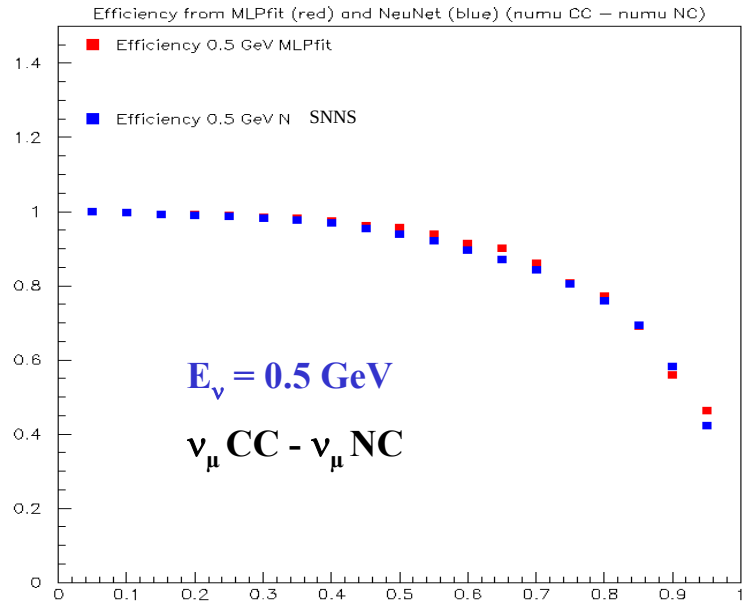

Efficiency from MLPfit (red) and NeuNet (blue) (nue CC – NC)
Efficiency 3 GeV MLPfit
Efficiency 3 GeV SNNS
$E_\nu = 3$ GeV
ν_e CC -NC

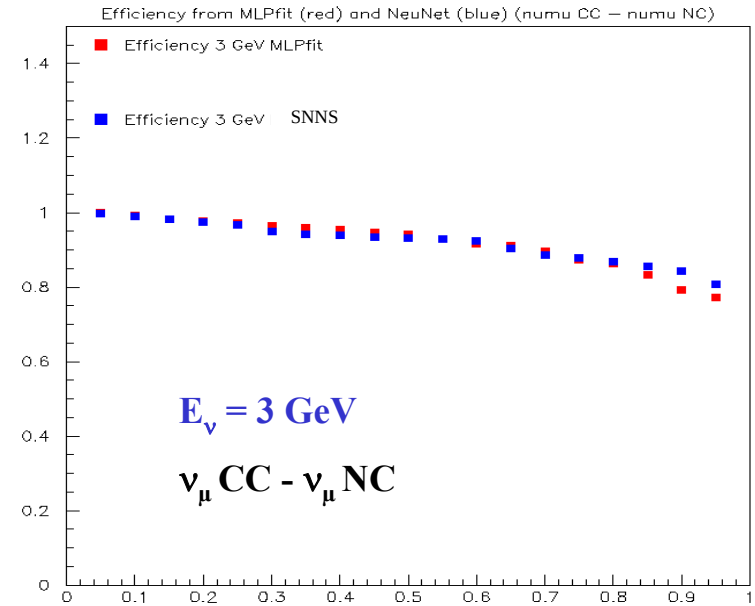

Efficiency from MLPfit (red) and NeuNet (blue) (nue CC – NC)
Efficiency 1 GeV MLPfit
Efficiency 1 GeV SNNS
$E_\nu = 1$ GeV
ν_e CC -NC

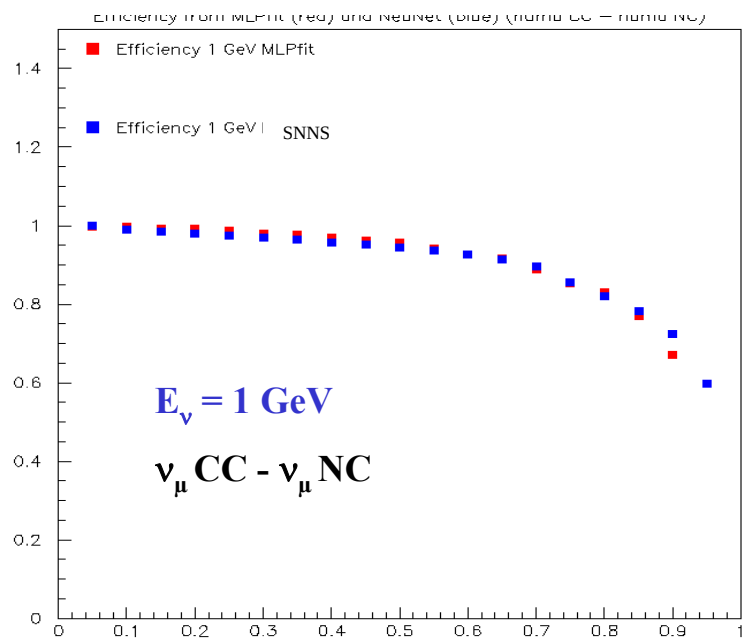

Efficiency from MLPfit (red) and NeuNet (blue) (nue CC – NC)
Efficiency 5 GeV MLPfit
Efficiency 5 GeV SNNS
$E_\nu = 5$ GeV
ν_e CC -NC

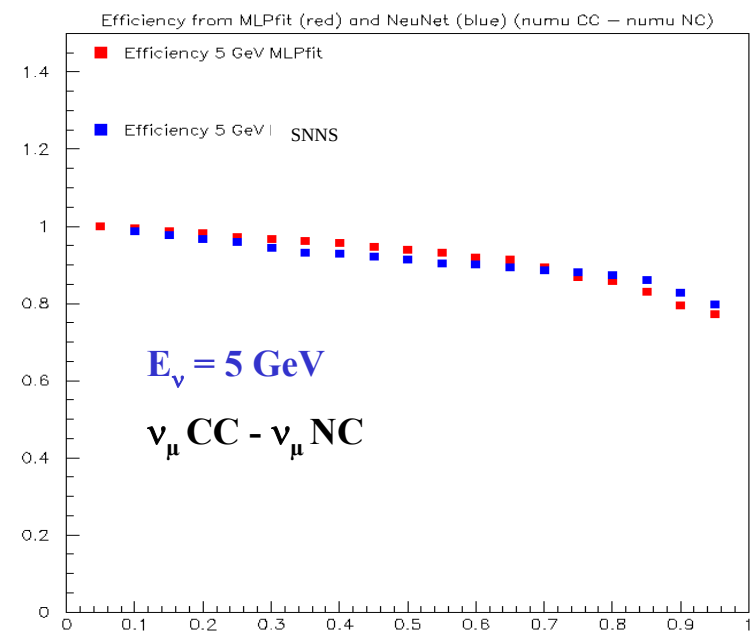

Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – numu NC)
Efficiency 2 GeV MLPfit
Efficiency 2 GeV NeuNet
$E_\nu = 2$ GeV
ν_μ CC - ν_μ NC
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

F_{cut}


Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – numu NC)
Efficiency 0.5 GeV MLPfit
Efficiency 0.5 GeV N
SNNS
$E_\nu = 0.5$ GeV
ν_μ CC - ν_μ NC


Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – numu NC)
Efficiency 3 GeV MLPfit
Efficiency 3 GeV
SNNS
$E_\nu = 3$ GeV
ν_μ CC - ν_μ NC


Efficiency 1 GeV MLPfit
Efficiency 1 GeV
SNNS
$E_\nu = 1$ GeV
ν_μ CC - ν_μ NC


Efficiency from MLPfit (red) and NeuNet (blue) (numu CC – numu NC)
Efficiency 5 GeV MLPfit
Efficiency 5 GeV
SNNS
$E_\nu = 5$ GeV
ν_μ CC - ν_μ NC

Summary/ Comments

- ν_μ CC / ν_e CC / NC separation studied for various energies with MC in the FAR detector, using only event hit topology.
- Separation in all channels looks easy above 2 GeV.
- No attempt was made yet to use the correct a priori probabilities.
- Good agreement of results from the comparison of the MLPfit and SNNS packages .
- Work in progress

Future Work Plan

- **Expand** the set of input variables
- Use **tracking/shower** information
- Study ANN performance versus E_{vis} instead of E_{ν} (which is unknown!)
- Learn how to **interpret** the classification results in order to obtain **"signal"** and **"background"** at a given **CL.**
- Attempt to apply for the estimation of oscillation limits and parameters.