Off-Axis-NOTE-SIM-12

A Study of an Off-axis Detector Performance as a Function of Sampling Frequency

L.Camilleri, A. Para

Abstract

This memo describes the studies made to determine the dependence of the sensitivity of an off-axis detector as a function of the sampling frequency (the thickness of the absorber placed between successive active planes). This involved optimizing the cuts used in the rejection of background from ν_μNC, ν_μCC and beam ν_e while maintaining a high efficiency for the oscillation signal: $\nu_\mu \rightarrow \nu_e$.

1 The generated event sample

The generation of Monte Carlo events and their reconstruction was described in a previous memo [1]. The basic geometry, Fig. 1 used consisted in the following sequence of material:

- **zair** cm of air,
- **zpla** cm of plastic,
- a glass RPC with X and Y readout, amounting to 5% of a radiation length..

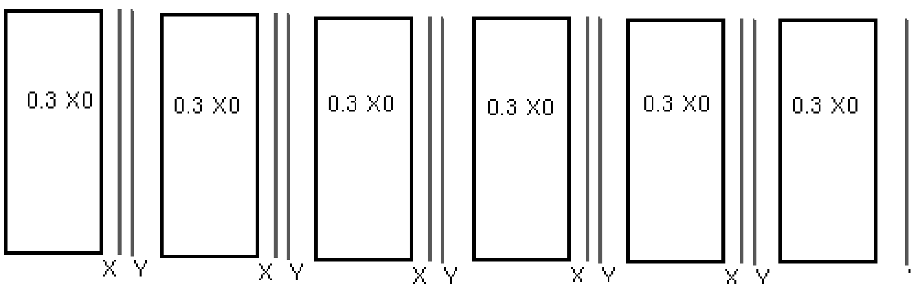

Figure 1: An example of the distribution of absorber and x,y readout planes

zair and **zpla** were chosen such as to yield a material with density equal to 0.7 gm.cm^{-3}, the density of particle board which will probably be used in the detector.

Event samples with two different geometries, described in Table 1, were generated for the analysis described in this note.

Table 1: Geometries generated for the analysis

zair (cm)	zpla (cm)	RPC (% rad.len)	Total sampling frequency (% rad.len)
5	12.5	5	30
2.5	6.25	5	17.5

By ignoring the X **AND** Y read out of every other RPC in the 30% radiation length sample of events, a third sampling frequency, namely 60% of a radiation length, could also be studied.

Furthermore, by ignoring the Y read out in the odd planes and the X readout in the even planes, a detector with alternate X **OR** Y readout in successive planes could also be studied.

The read out strip width was taken to be 3cm.

For each of the geometries studied charged current (CC) and neutral current (NC) interactions for both ν_μ and ν_e were generated. The ν_μ interactions were used to study the background. The ν_e interactions were used to study both the intrinsic ν_e background present in the beam and the signal from $\nu_\mu \to \nu_e$ oscillations.

For the 30% radiation length sampling frequency geometry $\bar{\nu}_\mu$ and $\bar{\nu}_e$ samples of events were also generated.

2 Ntuple generation

For each of the two geometries, the events were reconstructed and Ntuples were generated using the program described in [1]. These samples are referred to as 6CM and 12CM.

By ignoring the read out of every other RPC in the 30% radiation length sample of events, a third sampling frequency, namely 60% of a radiation length, could also be studied. This is referred to as DOUBLE.

Furthermore, by ignoring the Y read out in the odd planes and the X readout in the even planes, a detector with alternate X or Y readout in successive planes could also be studied. This is referred to as XORY.

The antineutrino run with 30% radiation length sampling is referred to as 12BAR.

These ntuples are stored in the directory:

/afs/fnal.gov/files/code/off-axis/off-axis_02/ntuples.

The ntuples for the five configurations, 6CM, 12CM, DOUBLE, XORY and 12BAR are stored in separate subdirectories as listed in Table 2.

Table 2: Storage and number of files available(10000 events per file)

Configuration	Subdirectory for ν_e, $\bar{\nu}_e$	Subdirectory for ν_μ, $\bar{\nu}_\mu$	# ν_e, $\bar{\nu}_e$ files	# of ν_μ,$\bar{\nu}_\mu$ files
6CM	6cm	6cm	10	30
12CM	nue	numu	10	30
DOUBLE	double	double	10	50
XORY	xory	xory	10	30
12BAR	nuebar	numubar	10	9

3 The beam spectra and oscillation parameters used

In the subsequent analysis the detector was assumed to be located 10km off-axis at 732km. The beam spectra used to weight the events are in arrays of 1000 bins of 20 MeV between 0 and 20 GeV stored in files to be found in the directory:

/afs/fnal.gov/files/code/off-axis/off-axis_01/beam

The files used are;

- ν_μ spectrum: **me_numu_r10e5_z735e5.vec**

  This file was used for three different purposes:

  - It was used to weight the generated ν_μCC events in order to study the background from these events. However, since the ν_μ spectrum at long baselines is depleted though oscillations to ν_τ, for this source of background, the weight read from the file was multiplied by the survival probability for a ν_μ at that distance and energy:

    $$1.0 - \sin^2 2\theta_{\mu\tau} \sin^2 \frac{1.27\Delta \mathrm{m}^2 \mathrm{L}}{\mathrm{E}_\nu}$$

    with

    $$\sin^2 2\theta_{\mu\tau} = 1.0 \quad \text{and} \quad \Delta m^2 = 2.5 \times 10^{-3}\ \mathrm{eV}^2.$$

  - It was also used to weight the generated ν_μNC events to study the background from this source. No survival probability was needed here since ν_μNC and ν_τNC events are identical.
  - Finally it was used to weight the generated ν_eCC events in order to study the $\nu_\mu \to \nu_e$ signal events. Here the weight read from the file was multiplied by the $\nu_\mu \to \nu_e$ oscillation probability. This was taken to be:

$$\sin^2\theta_{23}\sin^2 2\theta_{13}\sin^2\frac{1.27\Delta m^2 L}{E_\nu}$$

which becomes for $\theta_{23} = 45$ degrees, and $sin^2 2\theta_{13} = 0.1$ (the CHOOZ' limit):

$$0.05\sin^2\frac{1.27\Delta m^2 L}{E_\nu}$$

- ν_e spectrum: **me_nue_r10e5_z735e5.vec**

  This was used to weight the generated ν_eCC events in order to study the background from the intrinsic ν_e in the beam.

When running at a yearly rate of 4.0×10^{20} protons on target the ν_μ file gave **107.594** ν_μCC events per kiloton.year and the ν_e file gave a rate of **2.177** ν_eCC events per kiloton.year.

Similarly for the antineutrino run the files used were:

- $\bar{\nu}_\mu$ spectrum: **me_anti_numubar_r10e5_z735e5.vec**
- $\bar{\nu}_e$ spectrum: **me_anti_nuebar_r10e5_z735e5.vec**

with the corresponding normalizations of **32.378** $\bar{\nu}_\mu$CC events and **0.707** $\bar{\nu}_e$CC events per kiloton.year.

It should be noted that the background from the minor components of the beam (anti neutrinos in neutrino running) was not included in these calculations except in the section dealing with the comparison of neutrino and antineutrino running. The contribution of ν_μ and ν_e is indeed important in antineutrino running and cannot be neglected.

4 The overall background rejection strategy

The strategy was to reduce the ν_μNC background to a level below that of the intrinsic ν_e in the beam, while keeping the signal ν_eCC efficiency as high as possible . This was done in two steps: loose cuts followed by the use of likelihood functions to distinguish between signal ν_eCC and background ν_μNC events. The purpose of the loose cuts was twofold. Some of them were used to reject ν_μCC and beam ν_eCC events. In addition they all served in defining a definite range of variables for the likelihood functions used in the second step.

For each of the configurations, half of the files available were used to study the background, define the loose cuts, generate the probability density functions used in the likelihood functions and decide on the cut on this likelihood that maximized the sensitivity of the experiment. The cuts and likelihood functions were then applied to the second half of the files, thus obtaining an unbiased measure of the sensitivity.

The following variables were defined based on the variables available in the ntuples and described in [1]:

ang= $\sqrt{(}\text{angx}^2 + \text{angy}^2)$, the three-dimensional angle of the candidate electron with respect to the beam.

tothit = nhit_x + nhit_y, total number of hit strips in the event.

ensumht = h_roa_x + h_roa_y, total number of hit strips associated to the electron candidate (the track with the largest number of hit strips).

why = ensumht/tothit, the fraction of the event hits associated to the electron candidate.

chimax, the maximum of chix and chiy, the χ^2 of the fits to the electron candidate track in the x and y views.

peet = $\sqrt{(}\text{peet\_x}^2 + \text{peet\_y}^2)$, the overall net transverse momentum of the event.

emult = ensumht/(nplx+nply), the mean number of hit strips per plane associated to the candidate electron.

plgapmin, the lowest of plgapx and plgapy, provided they were not zero. plgapx and plgapy are the last hit gap before the first gap encountered along the track.

5 The loose cuts

Clearly the value of some of the above variables depended on the sampling frequency. Therefore some of the cuts used were configuration dependent.

5.1 Configuration-independent cuts

.

The following cuts were the same for all configurations:

1.3 < **emult** < **3.0** The lower cut is tighter than necessary for NC rejection but is necessary for rejecting muons in ν_μCC events. This can be seen in Fig. 2.

ang < **2.0**

chimax < **100.0**

plgapmin $\neq$ **1.0**

5.2 Configuration-dependent cuts

The next set was configuration dependent:

ensumht

The range of ensumht for accepted events is given in the second column of Table 3. For events with **why** > **0.95** a tighter cut was applied: ensumht had to be larger than the value given in the third column of the table.

tothit

This is equivalent to a total energy cut.

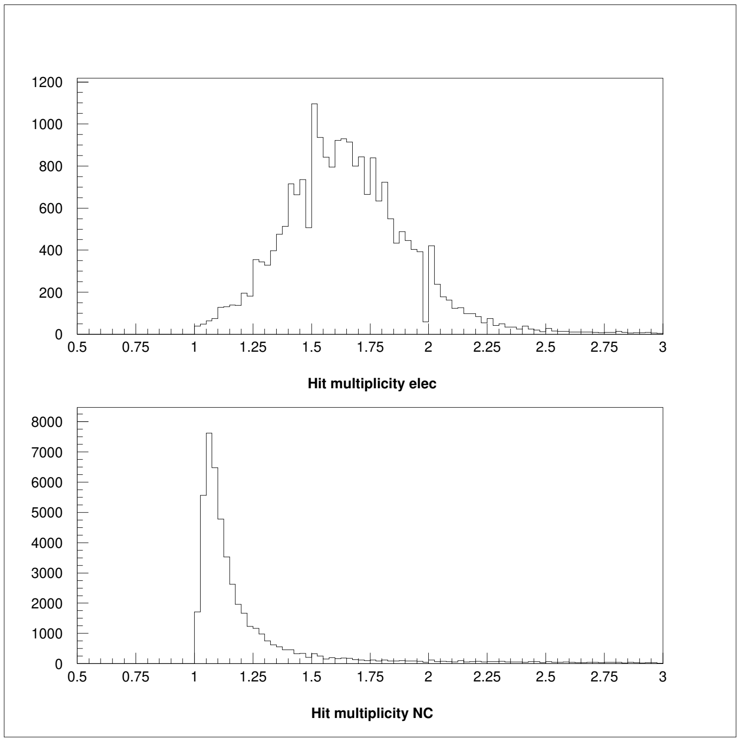

Figure 2: The mean number of hit strips per plane associated to the track. Top plot for signal electrons and bottom plot for CC background
.

The upper cut on this variable is tighter than necessary for NC rejection but is necessary for rejecting intrinsic ν_eCC events which are mostly of higher energy than the signal ν_eCC. This can be seen in Fig. 3.

peet

6 The likelihood functions

Having applied these cuts, three two-dimensional probability density functions (pdf) were generated for signal events (PS) and for ν_μNC events (PB). The pdf's used were:

- **ensumht vs why** referred to as **PS(sy)** and **PB(sy)**, Fig. 4.
- **ang vs tothit** referred to as **PS(at)** and **PB(at)**, Fig. 5.
- **chimax vs plgapmin** referred to as **PS(cp)** and **PB(cp)**, Fig. 6.

Table 3: Cut values for ensumht

Configuration	Cut range	why >0.95 events
6CM	35 < ensumht < 180	60
12CM	24 < ensumht < 100	40
DOUBLE	12 < ensumht < 50	20
XORY	12 < ensumht < 50	20

Table 4: Cut values for tothit

Configuration	Cut range
6CM	60 < tothit < 200
12CM	32 < tothit < 100
DOUBLE	16 < tothit < 50
XORY	16 < tothit < 50

The procedure to generate each pdf was to generate a two-dimensional histogram of the relevant variables, smooth it and normalize it to unity.

Given the values of the 6 variables, ensumht, why, ang, tothit, chimax and plgapmin, of an event, the probability that the event was of signal origin was computed as:

$$\mathrm{PS(sy)} \times \mathrm{PS(at)} \times \mathrm{PS(cp)}$$

Similarly the probability of the event being of background origin was:

$$\mathrm{PB(sy)} \times \mathrm{PB(at)} \times \mathrm{PB(cp)}$$

and the likelihood was:

$$\mathrm{L} = \log(\frac{\mathrm{PS(sy)} \times \mathrm{PS(at)} \times \mathrm{PS(cp)}}{\mathrm{PB(sy)} \times \mathrm{PB(at)} \times \mathrm{PB(cp)}})$$

Note that different pdf's were calculated and used for each of the five configurations. The kumacs to generate the likelihoods for the signal ν_eCC and the

Table 5: Cut values for peet

Configuration	Cut range
6CM	peet < 80
12CM	peet < 40
DOUBLE	peet < 20
XORY	peet < 20

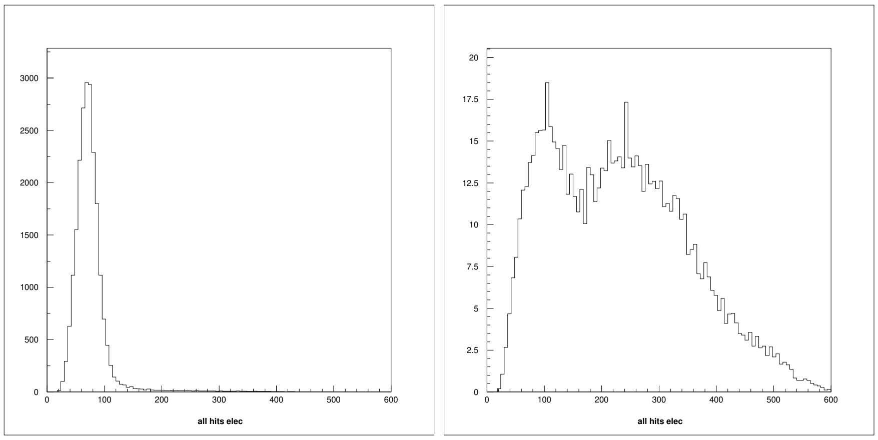

Figure 3: The distribution of total hits in the event for signal ν_e (left plot) and intrinsic beam ν_e (right plot)

fortran routines they call are named:
likelsyxxxxx.kumac, **likelatxxxxx.kumac** and **likelcpxxxxx.kumac**.
likelsyxxxxx.for, **likelatxxxxx.for** and **likelcpxxxxx.for**.
where xxxxx stands for 6cm, 12cm, double, 12xory and 12bar.

They produce pdf arrays called:
likelsy.xxxxx, **likelat.xxxxx** and **likelcp.xxxxx**.

The corresponding names for the NC background are:
likncsyxxxxx.kumac, **likncatxxxxx.kumac** and **liknccpxxxxx.kumac**.
likncsyxxxxx.for, **likncatxxxxx.for** and **liknccpxxxxx.for**.
likncsy.xxxxx, **likncat.xxxxx** and **liknccp.xxxxx**.

The pdf functions were then used on the unbiased sample of events by running the following kumac and fortran routines:
chainxxxxx.kumac and **chainxxxxx.for** for signal ν_eCC and NC background.
, and
chainnuexxxxx.kumac and **chainnuexxxxx.for** for intrinsic beamν_eCC.

All these routines and likelihood functions as well as links to the event files directories can be found in the directory:

/afs/fnal/files/code/off-axis/off-axis_01/code/analysis

An example of the signal and NC background likelihoods for the 12CM configuration is shown in Fig 7. In this case a cut on the likelihood of 3.25 was chosen (see next section).

7 Results

A Figure of Merit (FOM) was defined as:

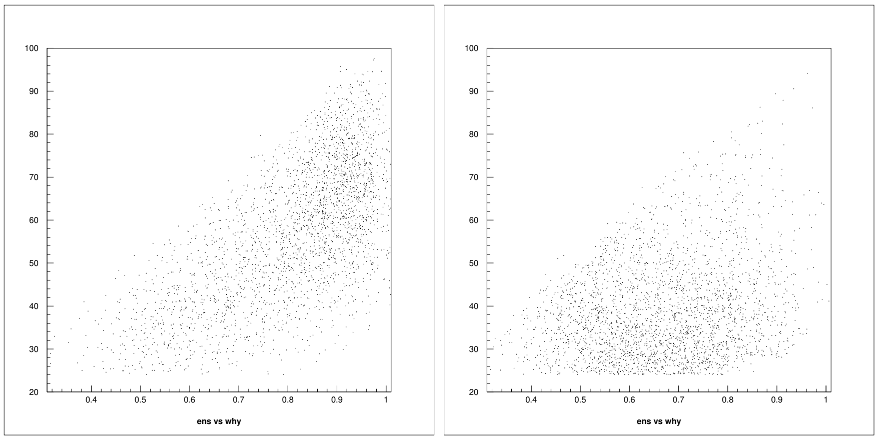

Figure 4: The ensumht vs why pdf two-dimensional distribution for signal(left plot) and NC background (right plot)

$$\text{FOM} = \frac{\text{Number of signal events}}{\sqrt{\text{Total number of background events}}}$$

The FOM was calculated assuming:

- 4.0×10^{20} protons on target per year
- a 5 year run
- a 50 kiloton detector
- an 85% fiducial mass.

For each configuration, the numbers of signal events and of background (ν_μNC, ν_μCC and beam ν_e), and, from these, the FOM were calculated as a function of likelihood cut. The FOM is a slowly varying function of the likelihood cut as is shown in the example of Fig. 8. However the background changes by a factor of three over this range of likelihood cut and the signal by a factor of two as can be seen in Fig. 8. The cut that gave the maximum FOM was selected and used in the analysis of the unbiased event files

The signal efficiency and background rejection obtained are listed in Table 6 for the 12CM configuration as an example.

With the detector and running scenario listed above the expected number of produced events are as follows:

Events = Flux x Years x Mass x Fiducial

Number of ν_μCC events before oscillations:

107.594 x 5 x 50 x 0.85 = **22863.7**

Number of intrinsic beam ν_eCC events:

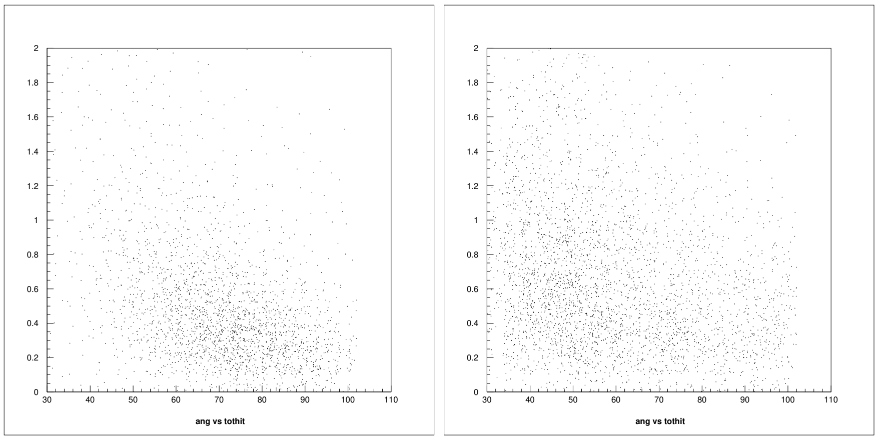

Figure 5: The ang vs tothit pdf two-dimensional distribution for signal(left plot) and NC background (right plot)

Table 6: Signal efficiency and background rejection for the 12CM configuration

	Background Rejection		
Signal efficiency	ν_μNC	ν_μCC	Beam ν_e
0.35	1.9 x 10^{-3}	5.6 x 10^{-4}	5.6 x 10^{-2}

2.177 x 5 x 50 x 0.85 = **462.6**
Number of $\nu_\mu \to \nu_e$ signal events:
Events = ν_μCC events x oscillation prob.
22863.7 x 0.0293 = **669.7**

The numbers of signal events and of background events from the three sources at the optimum FOM are shown in Table 7 for the four configurations, assuming the running conditions listed above.

8 The sensitivity as a function of absorber thickness

Three configurations were used in this analysis; 6CM,12CM and DOUBLE.

The results are shown in Table 8.

The uncertainty on these numbers is less than 1.0. Whereas reducing the absorber thickness by a factor of two does not improve the sensitivity significantly, doubling the thickness does worsen it by **21%**.

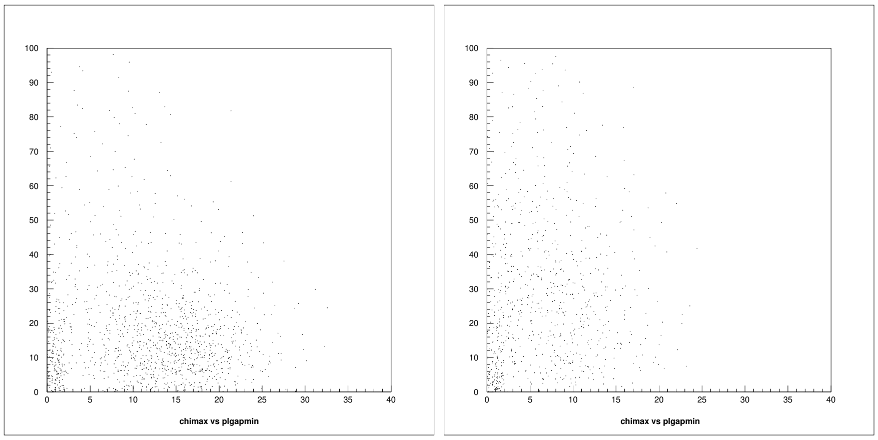

Figure 6: The chimax vs plgapmin pdf two-dimensional distribution for signal(left plot) and NC background (right plot)

Table 7: Numbers of signal and background events at the optimum FOM

		Background				
Configuration	Signal	ν_μNC	ν_μCC	Beam ν_e	Sum backg.	FOM
6CM	208.9	7.6	2.1	24.3	34.0	35.8
12CM	232.2	12.9	5.3	25.8	44.0	35.0
DOUBLE	217.7	20.6	8.7	27.3	56.6	28.9
XORY	213.6	13.6	4.9	24.3	42.8	32.6

9 The sensitivity as a function of active plane positions

Given a certain number of active readout planes, is it better to have both an x and a y plane together every D(cm) of aborber, Fig 9, or to alternate x and y every 0.5D(cm) of absorber, Fig 10?

Separating the x and y planes rather than lumping them together improves the FOM by **10%**.

The origin of this improvement was investigated. The effect of the loose cuts was the same on the two configurations. It was only after the likelihoods that the improvement was seen. This was due equally to the **ensumht vs why** and the **ang vs tothit** likelihoods. The reason for this improvement was traced to an improvement in the energy resolution of the detector when using a finer sampling, even though at each sampling only an x or a y are available. It goes from 17.2% in the DOUBLE configuration that samples every 0.6 rad. length to 15.0% in the XORY configuration that samples every 0.3 rad. length.

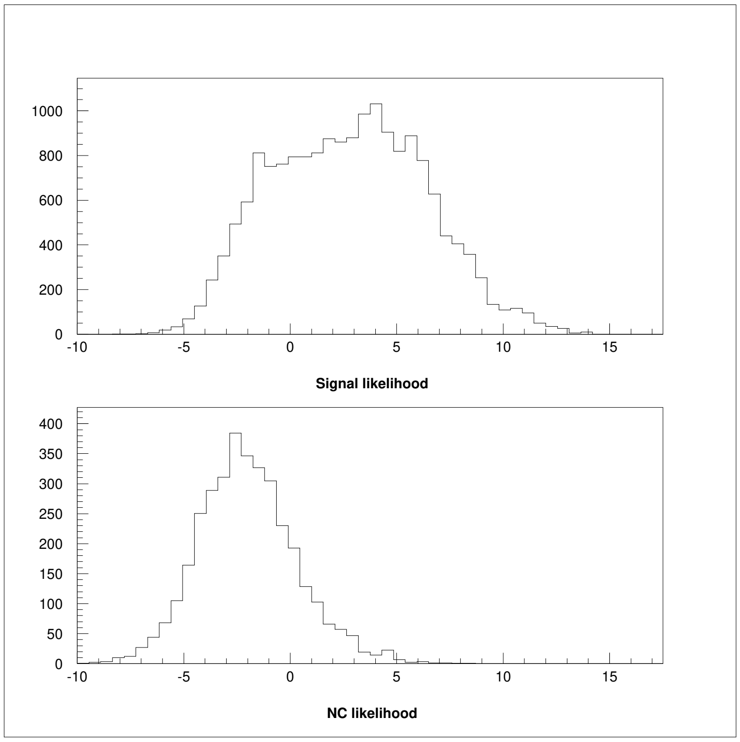

Figure 7: The signal and NC likelihoods for the 12CM configuration

10 Differences between $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$

Since determining θ_{13} and δ_{CP} will require antineutrino running as well as neutrinos, the performance of the detector for antineutrinos was also investigated. Antineutrino events were generated, ntuples were produced and likelihoods were computed in the same way as described for neutrinos using the 30% radiation length geometry. This sample is referred to as 12BAR and is to be compared to the neutrino 12CM configuration.

As explained earlier the background arising from the ν_μ and ν_e components of the antineutrino beam had to be taken into account. For completeness the background from the $\bar{\nu}_\mu$ and $\bar{\nu}_e$ components of the neutrino beam were also computed.

The corresponding normalization beam flux files were:

- ν_μ in antineutrino running:
  me_anti_numu_r10e5_z735e5.vec
- ν_e in antineutrino running:
  me_anti_nue_r10e5_z735e5.vec

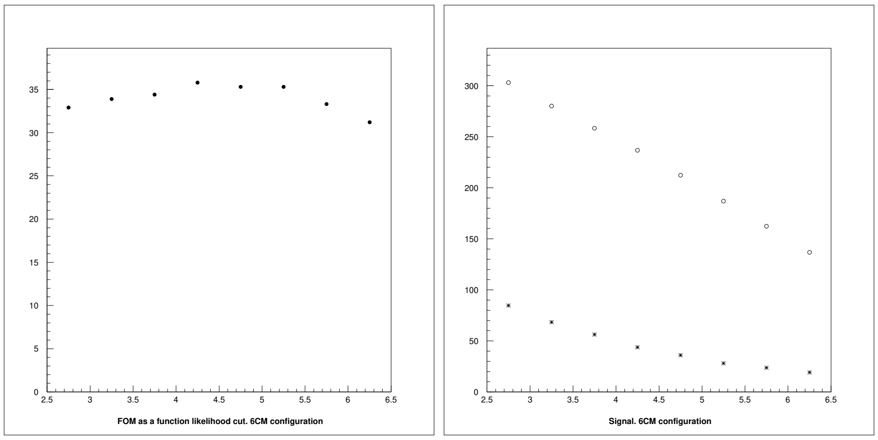

Figure 8: Left plot: The Figure of Merit as a function of likelihood cut. Right plot: The numbers of signal (open circles) and background (asterisk) events as a function of likelihood cut

Table 8: Dependence of the Figure of Merit on the Sampling Frequency

Configuration	Radiation length(%)	Likelihood Cut	FOM
6CM	17.5	4.50	35.8
12CM	30	3.25	35.0
DOUBLE	60	2.50	28.9

- $\bar{\nu}_\mu$ in neutrino running:
  me_numubar_r10e5_z735e5.vec
- $\bar{\nu}_e$ in neutrino running:
  me_nuebar_r10e5_z735e5.vec

The signal efficiency and background rejection at the best FOM are summarized in Table 10.

The signal efficiency for a comparable level of major component background rejection is much better in antineutrino running. This is because of the hadronic

Table 9: Dependence of the Figure of Merit on the active plane position

Configuration	Likelihood Cut	FOM
DOUBLE	2.50	28.9
XORY	3.00	32.6

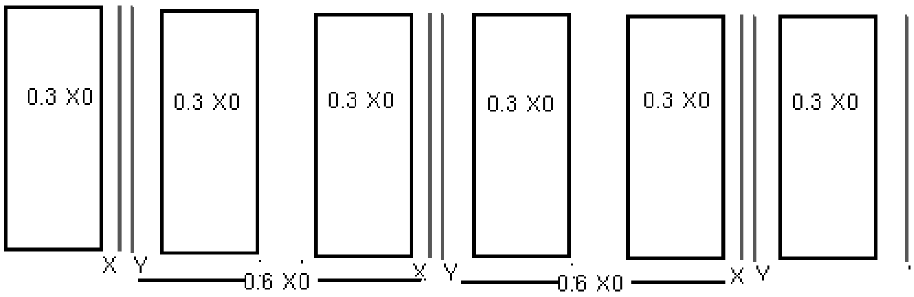

Figure 9: x and y plane together

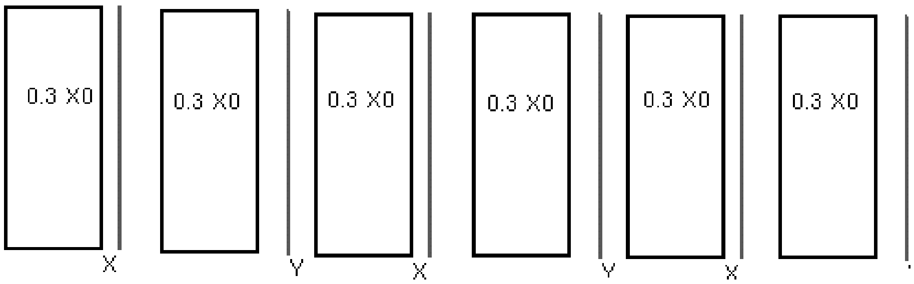

Figure 10: x or y plane alternating

energy in antineutrino events being lower, on average, than in neutrino events. The probability to simulate a signal electron of appreciable energy is therefore reduced in antineutrino events. However the background from the minor component (ν_μ and beam ν_e) is of comparable magnitude to the major component background.

The expected numbers of produced events for a 5 year antineutrino run are: **6880.3** $\bar{\nu}_\mu$CC events, **150.24** intrinsic beam $\bar{\nu}_e$CC events and **192.2** signal $\bar{\nu}_e$CC events.

Folding efficiencies and rejections results in the numbers of expected events shown in Table 11.

Thus, to have comparable significance in neutrino and antineutrino running, the antineutrino run will have to be $(33.8/19.7)^2$ times as long as the neutrino run, i.e. $\sim$3.0 times longer.

Table 10: Comparison of the neutrino and antineutrino signal efficiencies and background rejections for the 12CM configuration

		Background Rejection		
Configuration	Signal efficiency	ν_μNC	ν_μCC	Beam ν_e
12CM	0.35	1.9 x 10^{-3}	5.6 x 10^{-4}	5.6 x 10^{-2}
12BAR	0.50	2.0 x 10^{-3}	3.0 x 10^{-4}	6.7 x 10^{-2}

Table 11: A comparison of the numbers of events expected in neutrino and antineutrino running

		Background							
		Major component			Minor component				
		ν_μ	ν_μ	ν_e	ν_μ	ν_μ	ν_e	Sum	
Conf.	Signal	NC	CC	Beam	NC	CC	Beam	backg.	FOM
12CM	232.2	12.9	5.3	25.8	1.6	0.0	1.5	47.1	33.8
12BAR	96.1	4.6	0.8	10.1	4.3	0.6	3.3	23.7	19.7

References

[1] L. Camilleri, A. Para. The Reconstruction of ν_μ and ν_e Monte Carlo events. Off-Axis-NOTE-SIM-11.