**Instructions on how to use May 2009 MiniBooNE public data for muon-to-electron neutrino oscillation fits**

MiniBooNE Collaboration

(Dated: May 2009)

**The May 2009 dataset reflects the different fit method involved in the updated analysis, which simultaneously fits to $\nu_e$ and $\nu_\mu$ event distributions when looking for a $\nu_\mu \rightarrow \nu_e$ oscillation signal. Any signal prediction assumed during the fit does not affect the $\nu_\mu$ event distribution; the $\nu_\mu$ sample is only used in the fit in order to constrain the $\nu_e$ prediction in fitting for $\nu_\mu \rightarrow \nu_e$ oscillations, by taking advantage of $\nu_e$ and $\nu_\mu$ flux and cross section correlations. These correlations enter through the off-diagonal elements of the covariance matrix involved in the fit. See the following text for more details.**

## I MINIBOONE PUBLIC DATA AND NOTATION

The following MiniBooNE information is made publicly available for the updated electron neutrino appearance analysis:

- 1D array of electron neutrino bin boundaries in $E_\nu^{QE}$ (reconstructed neutrino energy). Numbers given assume 8 (11) bins for $E > 475$ MeV ($E > 300$ MeV) fits. This is therefore an array of 9 (12) numbers.

- 1D array of electron neutrino data events per $E_\nu^{QE}$ bin after $\nu_e$ selection cuts: $D_i^{\nu_e}$, with $i = 1, ..., 8(11)$. This is an array of 8 (11) integer numbers.

- 1D array of muon neutrino charged-current quasi-elastic (CCQE) data events per $E_\nu^{QE}$ bin after $\nu_\mu$ selection cuts: $D_i^{\nu_\mu}$, with $i = 1, ..., 8$. Note: The $E_\nu^{QE}$ bin boundaries for this array are different from the given electron neutrino $E_\nu^{QE}$ bin boundaries (the bin boundaries are (in MeV): 0 500 700 900 1100 1300 1500 1700 1900). This is an array of 8 integer numbers.

- 1D array of predicted electron neutrino background events per $E_\nu^{QE}$ bin after $\nu_e$ selection cuts: $B_i^{\nu_e}$, with $i = 1, ..., 8(11)$. This array contains 8 (11) real numbers.

- 1D array of predicted muon neutrino CCQE events per $E_\nu^{QE}$ bin after $\nu_\mu$ selection cuts: $P_i^{\nu_\mu}$, with $i = 1, ..., 8$. Note: The $E_\nu^{QE}$ bin boundaries for this array are different from the given electron neutrino $E_\nu^{QE}$ bin boundaries (the bin boundaries are (in MeV): 0 500 700 900 1100 1300 1500 1700 1900). This array contains 8 real numbers.

- 2D array of fractional covariance matrix in 24 (30) bins of $E_\nu^{QE}$ (8 (11) bins for predicted full $\nu_\mu \to \nu_e$ transmutation events + 8 (11) bins for predicted electron neutrino events + 8 bins for predicted muon neutrino events, side-by-side). Fractional refers to dividing the covariance matrix per pair of numbers of predicted events per $E_\nu^{QE}$ bin:

$$M_{ij}^{3\times3}/(P_i \cdot P_j), \tag{1}$$

where

  - $P_i = F_i$ for $i = 1, ..., 8(11)$, $F_i$ being the predicted full $\nu_\mu \to \nu_e$ transmutation events as a function of the same $E_\nu^{QE}$ binning used for the $\nu_e$ data sample (see below)
  - $P_i = B_{i-8(11)}^{\nu_e}$ for $i = 9(12), ..., 16(23)$
  - $P_i = P_{i-8(11)-8(11)}^{\nu_\mu}$ for $i = 17(24), ..., 24(30)$

This matrix contains systematic uncertainties and correlations between all three samples, and statistical uncertainties for the electron neutrino background prediction and the muon neutrino CCQE prediction. An additional statistical uncertainty must be added separately according to the signal prediction (see below). This is an array of 24×24 (30×30) real numbers.

- ntuple file of 17,037 MC full $\nu_\mu \to \nu_e$ transmutation events after $\nu_e$ selection cuts, with one row per event, and four columns with the following information:
  - $E_\nu^{QE}$: reconstructed neutrino energy (MeV)
  - $E_\nu^{true}$: true neutrino energy (MeV)
  - $L_\nu^{true}$: true neutrino baseline (distance between neutrino production and detection points) (cm)
  - $w$: weight for this full transmutation event, without including any oscillation probability in the weight, and with weight normalization such that the sum of all weights divided by the number of events in the file equals the number of expected full transmutation events across all $E_\nu^{QE}$ after $\nu_e$ selection cuts.

Note: full transmutation events are obtained from a $\nu_\mu$ flux prediction for which neutrino interactions are generated assuming all these events are $\nu_e$'s as opposed to $\nu_\mu$'s, and then reconstructed and selected according to $\nu_e$ selection cuts. The condition on the weights written above is:

$$\sum_{n=1}^{N} w_n/N = T \tag{2}$$

where N=17,037 is the number of events (rows) in the ntuple file, $w_n$ is the weight for that event, and T is the MC prediction for the number of full transmutation events passing $\nu_e$ cuts. From this file, one can also get the full transmutation prediction for each $E_\nu^{QE}$ bin: $T_i$, with $i = 1, ...8(11)$, and $\sum_i T_i \equiv T$. In total, this file contains 4×N real numbers.

## II HOW TO USE MINIBOONE DATA IN $\nu_\mu \to \nu_e$ OSCILLATION FITS

In the following, we describe only how to form a $\chi^2$ from MiniBooNE data and predictions, for any given $\nu_\mu \to \nu_e$ oscillation model. How to use these $\chi^2$ numbers depends on the specific analysis the user wishes to make.

In the MiniBooNE fit for oscillation parameters, the $\nu_e$ event distribution (as a function of $E_\nu^{QE}$ is fit simultaneously with the $\nu_\mu$ event distribution (as a different function of $E_\nu^{QE}$). The $\chi^2$ is calculated using a matrix that contains the systematic correlations between the $\nu_\mu$, $\nu_e$ signal, and $\nu_e$ background samples.

The $\chi^2$ is defined as follows:

$$\chi^2 = \sum_{i,j=1}^{N_e+N_\mu} (D_i - P_i) M_{ij}^{-1} (D_j - P_j) \tag{3}$$

where:

- $N_e = 8(11)$ is the number of $E_\nu^{QE}$ bins for observed electron neutrino events for $E > 200$ MeV ($E > 475$ MeV) fits

- $N_\mu = 8$ is the number of $E_\nu^{QE}$ bins for observed muon neutrino events

- $D_i = (D_j^{\nu_e}(j = 1, ..., N_e), D_j^{\nu_\mu}(j = 1, ..., N_\mu))$ is a 1D side-by-side array of observed electron neutrino events and observed muon neutrino events as functions of $E_\nu^{QE}$

- $P_i = ((B_j^{\nu_e} + S_j)(j = 1, ..., N_e), P_j^{\nu_\mu}(j = 1, ..., N_\mu))$ is a 1D side-by-side array of predicted electron neutrino events from background, $B^{\nu_e}$ plus any possible signal, $S_i$, from $\nu_\mu \to \nu_e$ oscillations, and predicted muon neutrino events, $B^{\nu_\mu}$.

  The number of signal events, $S_j$ predicted per $E_\nu^{QE}$ bin $j = 1, ..., N_e$ is obtained as follows:

$$S_j = P_{\mu\to e}(E_\nu^{true}, L_\nu^{true}; osc.parameters) \cdot T_j \tag{4}$$

where $P_{\mu\to e}$ is the oscillation probability, and the oscillation parameters are, for example $\Delta m^2$ and $\sin^2 2\theta_{\mu e}$. In order to compute the number of signal events, $S_i$, one can proceed as follows:

- loop through all events in the full transmutation file, $n = 1, ..., N$
- if event $n$ falls into $E_\nu^{QE}$ bin $i$ (based on the number on the first column in the full transmutation ntuple file), then increment the signal prediction in bin $i$ by:

$$S_i = S_i + P_{\mu\to e}(E_\nu^{true}, L_\nu^{true}; osc.parameters) \cdot w_n/N \tag{5}$$

  where $E_\nu^{true}, L_\nu^{true}, w_n$ are the numbers in the second, third, and fourth column in the full transmutation ntuple file for that event.

- $M_{ij}^{-1}$ is the inverse of the total $2 \times 2$ covariance matrix $M_{ij}$ obtained from the pieces above as follows:

  - First, the full $3\times$ three covariance matrix for the signal prediction, $S_i$, $\nu_e$ background prediction, and $\nu_\mu$ CCQE prediction is formed from the given fractional covariance matrix:

$$M_{ij}^{3\times 3,total} = M_{ij}^{3\times 3} \cdot (P_i \cdot P_j) \tag{6}$$

    where $i, j = 1, ..., N_e + N_e + N_\mu$.

  - A statistical error contribution from the signal prediction, $S_i$ is added to the diagonal elements of $M_{ij}^{3\times 3,total}$ for $i = 1, ..., N_e$:

$$M_{ij}^{3\times 3,total} = M_{ij}^{3\times 3,total} + \delta_{ij} \cdot S_i; \qquad i = 1, ..., N_e \tag{7}$$

  - Then, $M_{ij}^{3\times 3,total}$, which has the form ($\nu_e$ signal, $\nu_e$ background, $\nu_\mu$ CCQE), is collapsed to the total $2\times 2$ covariance matrix $M_{ij}$, which has the form ($\nu_e$

background + $\nu_e$ signal, $\nu_\mu$ CCQE). This is done by superimposing the $N_e \times N_e$-dimensional blocks of the covariance matrix, bin-by-bin, in the following manner: the top left ($\nu_e$ signal) block of the covariance matrix with the central ($\nu_e$ background) block; the top middle block with the central ($\nu_e$ background) block; the middle left block with the central ($\nu_e$ background) block; the bottom left block with the bottom middle block; and the top right block with the middle right block.

– The collapsed 2×2 error matrix is then inverted to give $M_{ij}^{-1}$.

The number of degrees of freedom (*ndf*) to be used in the fit in order to obtain $\chi^2$-probabilities should be the number of $\nu_e$ bins, plus the number of $\nu_\mu$ bins, minus the number of oscillation parameters (in the case of the example oscillation fit program, two, one for $\Delta m^2$ and one for $\sin^2 2\theta$), minus 1, to account for an effective normalization correction we have introduced in the fit, using a comparison of the $\nu_\mu$ CCQE prediction to the $\nu_\mu$ CCQE data:

$$ndf = N_e + N_\mu - N_{fit\ params} - 1 = 8(11) + 8 - 2 - 1 = 13(16) \tag{8}$$