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

MiniBooNE Collaboration

(Dated: May 2009)

The May 2009 dataset reflects the different fit method involved in the antineutrino analysis, which mirrors the updated neutrino analysis. This method simultaneously fits to $\bar{\nu}_e$ and $\bar{\nu}_\mu$ event distributions when looking for a $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation signal. Any signal prediction assumed during the fit does not affect the $\bar{\nu}_\mu$ event distribution; the $\bar{\nu}_\mu$ sample is only used in the fit in order to constrain the $\bar{\nu}_e$ prediction in fitting for $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations, by taking advantage of $\bar{\nu}_e$ and $\bar{\nu}_\mu$ 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 antineutrino appearance analysis:

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

- 1D array of muon antineutrino charged-current quasi-elastic (CCQE) data events per E_ν^{QE} bin after $\bar{\nu}_\mu$ selection cuts: $D_i^{\bar{\nu}_\mu}$, with $i = 1, ..., 8$. Note: The E_ν^{QE} bin boundaries for this array are different from the given electron antineutrino E_ν^{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 antineutrino background events per E_ν^{QE} bin after $\bar{\nu}_e$ selection cuts: $B_i^{\bar{\nu}_e}$, with $i = 1, ..., 8(11)$. This array contains 8 (11) real numbers.

- 1D array of predicted muon antineutrino CCQE events per E_ν^{QE} bin after $\bar{\nu}_\mu$ selection cuts: $P_i^{\bar{\nu}_\mu}$, with $i = 1, ..., 8$. Note: The E_ν^{QE} bin boundaries for this array are different from the given electron antineutrino E_ν^{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_ν^{QE} (8 (11) bins for predicted full $\bar{\nu}_\mu \to \bar{\nu}_e$ transmutation events + 8 (11) bins for predicted electron antineutrino events + 8 bins for predicted muon antineutrino events, side-by-side). Fractional refers to dividing the covariance matrix per pair of numbers of predicted events per E_ν^{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 $\bar{\nu}_\mu \to \bar{\nu}_e$ transmutation events as a function of the same E_ν^{QE} binning used for the $\bar{\nu}_e$ data sample (see below)
  - $P_i = B_{i-8(11)}^{\bar{\nu}_e}$ for $i = 9(12), ..., 16(23)$
  - $P_i = P_{i-8(11)-8(11)}^{\bar{\nu}_\mu}$ for $i = 17(24), ..., 24(30)$

This matrix contains systematic uncertainties and correlations between all three samples, and statistical uncertainties for the electron antineutrino background prediction and the muon antineutrino 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 28,685 MC full $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ transmutation events after $\bar{\nu}_e$ selection cuts, with one row per event, and four columns with the following information:
  - E_ν^{QE}: reconstructed neutrino energy (MeV)
  - E_ν^{true}: true neutrino energy (MeV)
  - L_ν^{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_ν^{QE} after $\bar{\nu}_e$ selection cuts.

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

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

where N=28,685 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 $\bar{\nu}_e$ cuts. From this file, one can also get the full transmutation prediction for each E_ν^{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 $\bar{\nu}_\mu \to \bar{\nu}_e$ OSCILLATION FITS

In the following, we describe only how to form a χ^2 from MiniBooNE data and predictions, for any given $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillation model. How to use these χ^2 numbers depends on the specific analysis the user wishes to make.

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

The χ^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_ν^{QE} bins for observed electron antineutrino events for $E > 200$ MeV ($E > 475$ MeV) fits
- $N_\mu = 8$ is the number of E_ν^{QE} bins for observed muon antineutrino events
- $D_i = (D_j^{\bar{\nu}_e}(j = 1, ..., N_e), D_j^{\bar{\nu}_\mu}(j = 1, ..., N_\mu))$ is a 1D side-by-side array of observed electron antineutrino events and observed muon antineutrino events as functions of E_ν^{QE}
- $P_i = ((B_j^{\bar{\nu}_e} + S_j)(j = 1, ..., N_e), P_j^{\bar{\nu}_\mu}(j = 1, ..., N_\mu))$ is a 1D side-by-side array of predicted electron antineutrino events from background, $B^{\bar{\nu}_e}$ plus any possible signal, S_i, from $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations, and predicted muon antineutrino events, $B^{\bar{\nu}_\mu}$.

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

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

where $P_{\mu\rightarrow e}$ is the oscillation probability, and the oscillation parameters are, for example Δm^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_ν^{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:

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

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 2×2 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, $\bar{\nu}_e$ background prediction, and $\bar{\nu}_\mu$ 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, $M_{ij}^{3\times3,total}$, which has the form (ν_e signal, ν_e background, ν_μ CCQE), is collapsed to the total 2×2 covariance matrix M_{ij}, which has the form (ν_e background + ν_e signal, ν_μ 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 (ν_e signal) block of the covariance matrix with the central (ν_e background) block; the top middle block with the central (ν_e background) block; the middle left block with the central (ν_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 inverted to give M_{ij}^{-1}.

The number of degrees of freedom (*ndf*) to be used in the fit in order to obtain χ^2-probabilities should be the number of ν_e bins, plus the number of ν_μ bins, minus the number of oscillation parameters (in the case of the example oscillation fit program, two, one for Δm^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 ν_μ CCQE prediction to the ν_μ CCQE data:

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