

Appendix B

A Log-Likelihood Multi-Variable Fit

This Appendix describes the mechanism of the log-likelihood fit that is used in the multi-variable fit in chapter 7 of this thesis. A Physics model is assumed, which is described by a set of parameters, denoted by α . The method is not restrictive in the Physics model used: in the case of 2-generation mixing, α corresponds to $\{\Delta m^2, \sin^2 2\theta\}$; for 3-generation mixing $\alpha \equiv \{\Delta m^2\}$; no oscillations is also an option, in which case α can be thought of as the special case of 2-generation mixing where $\sin^2 2\theta = 0$. The aim of the fit is to describe the **GOLD** data in terms of a sample comprised by **MC**(α) and **ROCK**.

B.1 Mechanism of log-likelihood fit

We will use the formalism of the binned log-likelihood which, with simple arguments, can be shown to be equivalent to the extended log-likelihood [32][70]. Let the **GOLD** distribution be G_i , the **MC** distribution be $M_i(\alpha)$, where α represents the neutrino mixing model and parameters, and the **ROCK** distribution be R_i , which is not a function of α , where $i = 1 \dots n$ and n is the number of bins in each distribution. The normalised to unity **MC** and **ROCK** distributions are represented by $m_i(\alpha)$ and r_i respectively. Then, the function describing the data is:

$$F_i(\alpha) = A(\alpha)m_i(\alpha) + X(\alpha)r_i, \quad (\text{B.1})$$

where $A(\alpha)$ and $X(\alpha)$ are the numbers of **MC** and **ROCK** events respectively to be determined by the fit at each set of mixing parameters.

In each bin, the Poisson probability for obtaining G_i GOLD events when $F_i(\alpha)$ are expected is given by:

$$P(G_i; F_i(\alpha)) = \frac{1}{G_i!} e^{-F_i(\alpha)} F_i(\alpha)^{G_i} \quad (\text{B.2})$$

The binned likelihood is then defined by the product of the probabilities for each bin:

$$\mathcal{L}(\alpha) = \prod_{i=1}^n P(G_i; F_i(\alpha)). \quad (\text{B.3})$$

However, the MC normalisation, $A(\alpha)$ in Eq. **B.1** is known with an error of 20% [43]. The likelihood $\mathcal{L}(\alpha)$ will be multiplied by an additional factor that contains this information and constrains the range of $A(\alpha)$:

$$\mathcal{L}(\alpha) = \prod_{i=1}^n P(G_i; F_i(\alpha)) \times \exp - \frac{1}{2} \left(\frac{1 - A(\alpha)/A^0(\alpha)}{\sigma_{A^0}} \right)^2, \quad (\text{B.4})$$

where $A^0(\alpha)$ is the expected number of MC events when absolutely normalised to the experiment's exposure and $\sigma_{A^0} = 0.2$ is the fractional error of this normalisation estimation [19][71]. The binned *log*-likelihood is then given by:

$$\ell(\alpha) = \ln \mathcal{L}(\alpha) = \sum_{i=1}^n \ln [P(G_i; F_i(\alpha))] - \frac{1}{2} \left(\frac{1 - A(\alpha)/A^0(\alpha)}{\sigma_{A^0}} \right)^2. \quad (\text{B.5})$$

The log-likelihood will be used instead of the likelihood: it is a mathematically better-behaved function while retaining the stationary points. From Eq. **B.2** we obtain the final expression for the log-likelihood as a function of the oscillation parameters:

$$\ell(\alpha) = \sum_{i=1}^n [G_i \ln F_i(\alpha) - F_i(\alpha) - \ln G_i!] - \frac{1}{2} \left(\frac{1 - A(\alpha)/A^0(\alpha)}{\sigma_{A^0}} \right)^2. \quad (\text{B.6})$$

The MC(α) plus ROCK populations, $A(\alpha)$ and $X(\alpha)$, are not constrained to sum up to the event population of the GOLD sample. We expect $A(\alpha) + X(\alpha)$ to be close but not equal to

$\sum_{i=1,n} G_i$. In fact, $A(\alpha) + X(\alpha) = \sum_{i=1,n} G_i$ would have been true if the additional term that constrains the $MC(\alpha)$ normalisation about the absolute expectation had not been used. This is a property of the binned log-likelihood [16].

For each set of neutrino mixing parameters, α , the solution for the parameters $A(\alpha)$ and $X(\alpha)$ is found by maximising the log likelihood of Eq. **B.6**.

B.2 Determination of mixing parameters

Once the mechanism for maximising the likelihood for a particular mixing model and mixing parameters has been set up, one can loop over all parameter space and maximise the log-likelihood as a function of the mixing parameters as well as $A(\alpha)$ and $X(\alpha)$. The maximum in $\ell(\alpha)$ then defines the preferred value of mixing parameters, α^* . The **MC** and **ROCK** populations at that point are denoted by A^* and X^* respectively. The exclusion regions in neutrino mixing parameters around α^* can then be drawn for the desired confidence levels.

B.3 Determination of the Quality of Fit

The log-likelihood fit described above will provide us with the confidence levels that the solution be within a region of the neutrino mixing parameters given that the neutrino oscillation hypothesis is correct. However, the quality of the fit is not answered for. Two methods will be used and are highlighted below.

B.3.1 A χ^2 -test

A simple way to address the issue will be to perform a χ^2 -test of the **GOLD** distribution, G_i , against the best-fit model (**MC** plus **ROCK**) distribution, $F_i(\alpha^*)$. Care will be needed in determining the degrees of freedom: these will be equal to $n - m - 2$, where n is the number

of bins in the distributions ($n = 24$ for the multi-variable analysis in the present thesis), m is the number of neutrino mixing parameters that are fitted ($m = 2$ for $\nu_\mu \leftrightarrow \nu_e$, $\nu_\mu \leftrightarrow \nu_\tau$ 2-generation mixing and $m = 1$ for 3-generation maximal mixing and $m = 0$ for no mixing) and 2 degrees of freedom are lost for the determination of the MC and ROCK populations.

B.3.2 The Likelihood of the log-Likelihood

A more interesting approach to the quality of fit is to ask the following question: given the solution (maximum log-likelihood) found by the data fit for that particular neutrino mixing hypothesis, what is the probability for obtaining the maximum value of log-likelihood that we did for the experiment? In other words, what is the likelihood of the log-likelihood? The following procedure will be adopted to address the question:

- A large number of MC-simulated experiments will be produced with the true neutrino mixing parameters corresponding to the solution α^* obtained by fitting the GOLD data. Each of these experiments is normalised to A^* events and a number of X^* ROCK events is added on top – see below for details.
- These simulated experiments will be fitted for neutrino oscillations the same way the GOLD data was.
- For each simulated experiment, the maximum point of the log-likelihood function will be found and the log-likelihood value itself is recorded.
- Let $L(\ell)$ be the normalised to unity distribution of log-likelihood values obtained from all simulated experiments. Then, the probability of an experiment giving a log-likelihood of ℓ_a is simply the integral of $L(\ell)$ from $-\infty$ to ℓ_a :

$$P(\ell_a) = \int_{-\infty}^{\ell_a} L(\ell) d\ell \quad (\text{B.7})$$

B.3.3 Generation of Pseudo-Simulated Experiments

In practice, the generation of a large number of simulated experiments, albeit of moderate statistics, is a lengthy and very time consuming job. The following procedure is adopted in simulating a background-contaminated experiment with little effort and no compromise:

- The very large **MCAF** sample is used in order to extract a very large statistics experiment at α^* , the neutrino mixing parameters favoured by the data fit. This MC experiment is normalised to a number of A^* neutrino events. This results to a distribution M_i^* that corresponds to the average experiment at the solution α^* .
- The **ROCK** sample is normalised down to X^* , the number of background events that maximise the log-likelihood at α^* . This is taken as the average background sample, which is described by the distribution R_i^* .
- The population of each bin of the average MC and **ROCK** distributions at α^* is allowed to fluctuate about its mean, M_i^* and R_i^* , according to Poisson statistics. The resulting MC and **ROCK** distributions are added together to form a simulated neutrino mixing experiment with background contamination.

The above can be repeated to generate a large number of simulated experiments. These can then be used in order to extract the distribution of maximum log-likelihood, as described above. In practice 1000 experiments are simulated.