

Statistical Methods for the Neutrino Mass Hierarchy Determination

Emilio Ciuffoli*

IMP, CAS

E-mail: emilio@impcas.ac.cn

In the next decades several experiments will attempt to determine the neutrino mass hierarchy, i.e. the sign of Δm_{31}^2 . Since the two hierarchies are non-nested hypotheses, $\Delta\chi^2 = \chi_{IH}^2 - \chi_{NH}^2$ does not follow a one-degree-of-freedom chi-square distribution (for example, it can be negative): it is possible to prove that, under certain assumptions, it follows a Gaussian distribution with $\sigma = 2\sqrt{\Delta\chi^2}$; these assumptions seem to be verified for reactor neutrino experiments, at least in some simplified cases, but, for example, not for accelerator neutrino experiments. I will discuss the possible definitions of sensitivity, using both the Bayesian and the frequentist perspectives, how they should be modified if the conditions for Gaussianity are not fulfilled, and in which cases the two procedures used to treat the pull parameters (marginalization and minimization) yield the same $\Delta\chi^2$. I will focus on the advantages and disadvantages of the different approaches proposed, and in particular on their physical interpretations.

*The European Physical Society Conference on High Energy Physics
5-12 July, 2017
Venice*

*Speaker.

1. Statistical Distribution of $\Delta\chi^2$

One of the open problems in neutrino physics is the determination of the mass hierarchy (MH), *i.e.* the sign of $\Delta m_{31}^2 = m_3^2 - m_1^2$, where m_3, m_1 are two of the neutrino mass eigenvalues: if this difference is positive, the hierarchy is called normal (NH), if negative, it is called inverted (IH). In the next decade several experiments will try to perform this measurement, in particular we will discuss more in detail the MH determination using accelerator neutrino experiments (or long baseline neutrino experiments, LBNE) and reactor neutrino experiments (RNE).

In order to determine the MH it is convenient to define a test statistic, namely a function of the experimental data, whose value is related to the hierarchy (for example, it could be positive if the hierarchy is normal, negative if it is inverted). In this presentation, we will consider $\Delta\chi^2$ as test statistic, defined as

$$\Delta\chi^2 = \chi_{IH}^2 - \chi_{NH}^2 = -2\ln \frac{P(D|\hat{\theta}, IH)}{P(D|\hat{\theta}', NH)} \quad (1.1)$$

Where $\hat{\theta}'$ and $\hat{\theta}$ are the best fit values for the eventual pull parameters considered. $\Delta\chi^2$ is not the only possible test statistic: possible alternatives for LBNE and RNE are proposed, for example, in [1] and [2].

From the statistical point of view, the main problem is that the two hierarchies are disjoint hypotheses: this means that Wilks' theorem cannot be applied and $\Delta\chi^2$ does not follow a chi-square distribution; as a consequence, if for example we are using a frequentist approach, the confidence in the mass hierarchy determination (expressed as usual as the number of Gaussian standard deviations σ 's) is not necessarily equal to $\sqrt{\Delta\chi^2}$.

The non-nested problem arise from the fact that $\text{Sign}(\Delta m_{31}^2)$ can only assume two discrete values, +1 and -1; it is possible to avoid this problem introducing an additional pull parameter: for the MH determination this was suggested first in [3]. Indeed, in RNE only the absolute values of the Δm^2 's can be measured, however the mass hierarchy can still be determined from the relation $|\Delta m_{31}^2| = |\Delta m_{32}^2| \pm |\Delta m_{21}^2|$. We can introduce a new pull parameter η , writing $|\Delta m_{31}^2| = |\Delta m_{33}^2| + (1 - 2\eta)|\Delta m_{21}^2|$; for $\eta = 0$ the hierarchy is normal, for $\eta = 1$ the hierarchy is inverted; this method, however, cannot be used in LBNE. It is also possible to use the more general approach described in [4], where a linear combination of the spectra obtained using two different hypothesis is considered, *i.e.* $\eta f(E) + (1 - \eta)g(E)$. We can write the $\chi^2 = \chi^2(\eta)$ as a function of η (the generalization to the case with additional pull parameters is trivial); we can define now two $\Delta\chi^2$'s, $\Delta\chi_{NH}^2$ and $\Delta\chi_{IH}^2$:

$$\Delta\chi_{NH}^2 = \chi^2(0) - \chi^2(\hat{\eta}) \quad \Delta\chi_{IH}^2 = \chi^2(1) - \chi^2(\hat{\eta}) \quad (1.2)$$

where $\hat{\eta}$ is the best fit value for η . Both $\Delta\chi^2$'s now follow a chi-square distribution, since in both cases we are dealing with nested hypotheses: " $\eta = 0(1)$ " versus " $\eta \in \mathcal{R}$ ", however the $\Delta\chi^2$ defined in 1.1 is the difference between these two quantities, indeed

$$\Delta\chi_{IH}^2 - \Delta\chi_{NH}^2 = \chi^2(1) - \chi^2(\hat{\eta}) - (\chi^2(0) - \chi^2(\hat{\eta})) = \chi_{IH}^2 - \chi_{NH}^2 = \Delta\chi^2 \quad (1.3)$$

and hence does not follow a chi-square distribution.

It was proven that if there are no pull parameters ("*simple vs simple*" scenario), the $\Delta\chi^2$ defined in 1.1 follows a Gaussian distribution, with $\mu = \overline{\Delta\chi^2}$ and $\sigma = 2\sqrt{|\Delta\chi^2|}$ [5]; this is also true

when pull parameters are taken into account, if some conditions are satisfied [6, 7, 8]. In particular, it seems that these conditions are satisfied in RNE (at least in some simplified models), but not in LBNE, as can be seen from Fig. 1; more details on the models used can be found in [9]. We will discuss now how to quantify the sensitivity in the mass hierarchy determination, using a frequentist or a Bayesian approach.

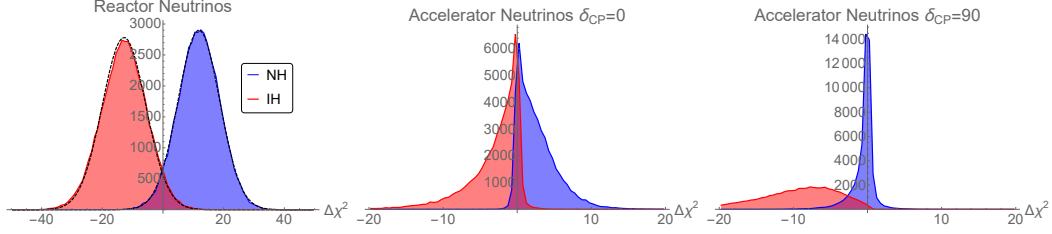


Figure 1: Statistical distribution of $\Delta\chi^2$ for RNE (left panel) and LBNE (central and right panels)

2. Sensitivity to Mass Hierarchy

2.1 Frequentist Approach

In a frequentist hypotheses test we are comparing a hypothesis H_0 , often called null hypothesis, versus an alternative hypothesis H_1 : we define a critical region, if the value of the test statistic T is inside that region, H_0 is excluded, if it is outside H_0 is accepted. The confidence level (CL) is given by $1 - \alpha$, where α is the probability of rejecting H_0 even if it is true, while the power is $1 - \beta$, where β is the probability of not rejecting H_0 if H_1 is true. In the case of the MH we can perform a hypothesis test by testing each hierarchy separately [8]: we define two thresholds, $T_{c,NH}$ and $T_{c,IH}$: if $T < T_{c,NH}$ the normal hierarchy is rejected, if $T > T_{c,IH}$ the inverted hierarchy is rejected. It is worth mentioning that, depending on the choice of $T_{c,NH}$ and $T_{c,IH}$, it is possible to accept or reject both hierarchies at the same time; moreover the CL depends on the statistical distribution of $\Delta\chi^2$. Since the CL depends only on the choice of $T_{c,NH}$ and $T_{c,IH}$, not on the actual result of the experiment (which tells us only if such a CL is achieved or not), it can be convenient to use the hypothesis test to quantify the sensitivity to the MH of future experiments. There are a couple of possible definition of sensitivity often used in literature [7, 8]; in the Gaussian, symmetric case (*i.e.* when $\mu_{NH} = \overline{\Delta\chi^2} = -\mu_{IH}$; μ_{NH} and μ_{IH} are the average value of $\Delta\chi^2$ for the NH and IH) we have

- **Median sensitivity:** $T_{c,NH(IH)} = -(+)\overline{\Delta\chi^2}$; CL (expressed as number of σ 's) is $\sqrt{\overline{\Delta\chi^2}}$, but $\beta = 0.5$
- **Mean Sensitivity:** $T_{c,NH} = T_{c,IH} = 0$, CL is $\sqrt{\overline{\Delta\chi^2}}/2$, but $\beta = \alpha$

Another common criterium for a discovery is the p-value, defined as a probability of finding a more extreme result than the observed one if H_0 is true. In the case of the MH, both the p-values for NH and IH must be reported: indeed, let's assume that an experiment was able to exclude the NH at 5σ 's; this however is not sufficient to conclude that the hierarchy is inverted, because the scenario would be quite different if the IH was compatible with the data at 1σ level or if it was also excluded at 5σ 's.

2.2 Bayesian Approach

Using the frequentist approach it is only possible to calculate the probability $P(D|MH)$ of obtaining the data D given a certain hypothesis, *i.e.* the compatibility of each hierarchy with the data; $P(MH|D)$, instead, *i.e.* the probability for the hierarchy to be normal or inverted given the experimental data, can be calculated only using the Bayesian theorem, for example

$$P(NH|D) = \frac{\pi(NH)}{\pi(NH) + \pi(IH)K^{-1}} \quad K = \frac{P(D|NH)}{P(D|IH)} = e^{-\Delta\chi^2/2} \quad (2.1)$$

Where $\pi(MH)$ are the priors for the normal or inverted hierarchy; the choice of priors can influence significantly the final result, however in the case of the MH there is a very natural choice, *i.e.* symmetric priors where $\pi(NH) = \pi(IH) = 0.5$. K is called the Bayes factor, in the last step we used the definition of $\chi^2 = -2\ln P(D|H)$, however it is important to underline that, if pull parameters are considered, the $\Delta\chi^2$ used in Eq. 2.1 can be quite different from the one defined in 1.1; indeed, using the Bayesian approach the eventual pull parameters must be marginalized, *i.e.* integrated over, weighted by their priors, while in the frequentist approach they must be minimized. In general the two procedures yield different results, however using Laplace method it is possible to verify that, if certain conditions are fulfilled (namely, if χ^2 is strongly peaked around the minimum and if the determinant of the covariance matrix calculated at the minimum is the same for the two hierarchies), the two procedures are equivalent: this method cannot be applied to LBNE, but it works with reasonable precision for RNE [9].

Finally, it is important to notice that the Bayesian and the frequentist approaches give different and complementary information: the first one gives us the probability for the hierarchy to be normal or inverted, while the second one tells us the compatibility of each hierarchy with the data: for a more complete analysis both should be used.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grants No. 11605247 and 11375201) and by the Chinese Academy of Sciences President's International Fellowship Initiative Grant No. 2015PM063.

References

- [1] L. Stanco, S. Dusini and M. Tenti, "Determination of the neutrino mass hierarchy with a new statistical method", Phys.Rev. D95 (2017) no.5, 053002, doi:10.1103/PhysRevD.95.053002;
- [2] L. Stanco (INFN, Padua) *et al.*, "A new way to determine the neutrino mass hierarchy at reactors" arXiv:1707.07651 [hep-ph]
- [3] Capozzi, Lisi and Marrone, "Neutrino mass hierarchy and electron neutrino oscillation parameters with one hundred thousand reactor events", Phys.Rev. D89 (2014), 013001, doi:10.1103/PhysRevD.89.013001
- [4] S. Algeri, J. Conrad and D.A. van Dyk, "A method for comparing non-nested models with application to astrophysical searches for new physics", MNRAS: Letters, 2016, doi:10.1093/mnrasl/slw025

- [5] X. Qian, et al., “Statistical Evaluation of Experimental Determinations of Neutrino Mass Hierarchy”, *Phys. Rev. D* 86 (2012) 113011, doi:10.1103/PhysRevD.86.113011
- [6] S.F. Ge et al., “Determination of mass hierarchy with medium baseline reactor neutrino experiments”, *JHEP* 1305 (2013) 131, doi:10.1007/JHEP05(2013)131
- [7] E. Ciuffoli, J. Evans and X.M. Zhang, “Confidence in a neutrino mass hierarchy determination”, *JHEP* 1401 (2014) 095, doi:10.1007/JHEP01(2014)095
- [8] Blennow et al., “Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering”, *JHEP* 1403 (2014) 028, doi:10.1007/JHEP03(2014)028
- [9] E. Ciuffoli, “Statistical Methods for the Neutrino Mass Hierarchy”, Proceedings of Prospects in Neutrino Physics (NuPhys2016), London, UK