Statistical Issues on the Neutrino Mass Hierarchy using the Standard Algorithm

L. Stanco[1], A. Lokhov[2], X. Ding[3,4] and F. Sawy[1,5]

[1]INFN, Sezione di Padova, 35131 Padova, Italy
[2] Institute for Nuclear Research RAS, 117312 Moscow, Russia
[3] INFN Sezione di Milano, Milan 20133, Italy
[4] Gran Sasso Science Insitute, F. Crispi 7, L'Aquila, AQ, Italy 67100
[5] Dipartimento di Fisica e Astronomia dell'Universita di Padova, 35131 Padova, Italy; Sawy@pd.infn.it

May 9, 2019

The Goal to achieve is ....

Advances in statistical methods may play a decisive role in the discovery reached at neutrino physics experiments. So that evaluating the used statistical methods and updating them is a necessary step in building a robust statistical analysis for answering the open questions in neutrino physics.

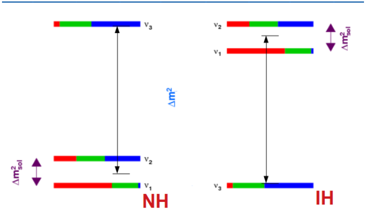
Does the neutrino spectrum follow IH model or NH model?

$$P_{\bar{\nu}_e \to \bar{\nu}_e} = 1 - \cos^4\theta_{13}\sin^2 2\theta_{12}\sin^2(\Delta_{21}) - \cos^2\theta_{12}\sin^2 2\theta_{13}\sin^2(\Delta_{31}) - \sin^2\theta_{12}\sin^2 2\theta_{13}\sin^2(\Delta_{32})$$

$$\Delta_{ij} \equiv \frac{L\Delta m^2_{ij}}{4E_\nu}$$

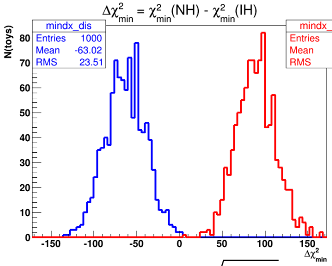
$$\Delta m^2_{31} = \Delta m^2 + \frac{\delta m^2_{sol}}{2}$$

$$\Delta m^2_{32} = \Delta m^2 - \frac{\delta m^2_{sol}}{2}$$



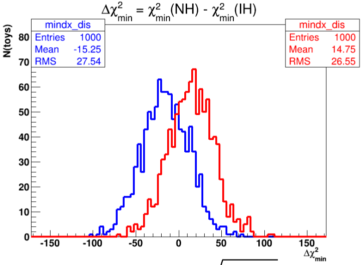
Contents

Issue I: The Limited Power of $\Delta\chi^2$



$$\sigma_{\Delta\chi^2} \equiv 2\sqrt{\overline{\Delta\chi^2}}$$

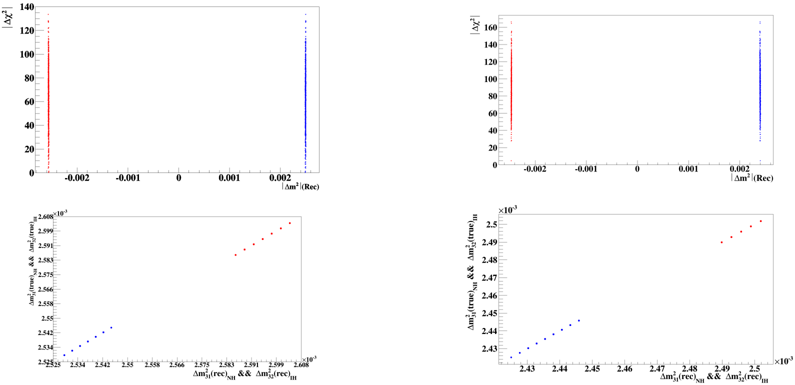
Infinity energy resolution		
μ_{NH}	-63.02	
σ_{NH}	23.51	
μ_{IH}	89.39	
σ_{IH}	22.83	
n"σ"(NH)	6.485(z-test)	7.94(approximation)
n"σ"(IH)	6.676(z-test)	9.45(approximation)



$$\sigma_{\Delta\chi^2} \neq 2\sqrt{\overline{\Delta\chi^2}}$$

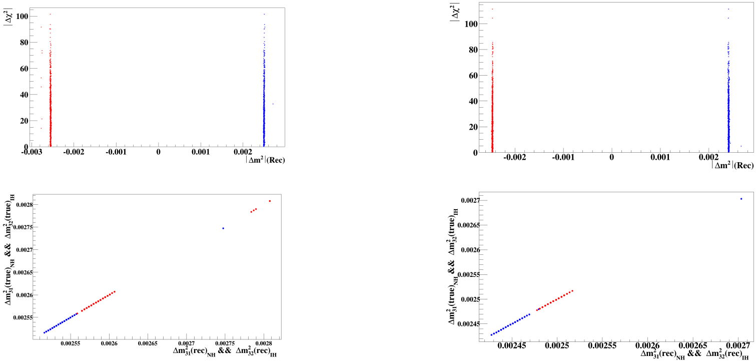
relative energy resolution $\frac{3\%}{\sqrt{E}}$		
μ_{NH}	-15.21	
σ_{NH}	27.52	
μ_{IH}	14.69	
σ_{IH}	26.55	
n"σ"(NH)	1.086(z-test)	3.9(approximation)
n"σ"(IH)	1.120(z-test)	3.8(approximation)

The precision of the fit assuming infinity energy resolution



$|\Delta\chi^2|$ vs $|\Delta m^2|_{(Rec)}$ for 1000 (NH) + 1000 (IH) toy JUNO simulations generated at $\Delta m^2 = 2.500 \times 10^{-3}$ for **NH hypothesis (left panel)** and $\Delta m^2 = -2.460 \times 10^{-3}$ for **IH hypothesis (right panel)** with six years of exposure and the ten near reactor cores assuming infinity energy resolution.

The precision of the fit assuming 3% relativity energy resolution



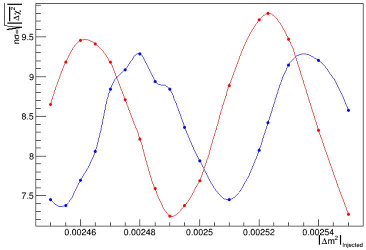
$|\Delta\chi^2|$ vs $|\Delta m^2|_{(Rec)}$ for 1000 (NH) + 1000 (IH) toy JUNO simulations generated at $\Delta m^2 = 2.500 \times 10^{-3}$ for **NH hypothesis (left panel)** and $\Delta m^2 = -2.460 \times 10^{-3}$ for **IH hypothesis (right panel)** with six years of exposure and the ten near reactor cores assuming 3% relativity energy resolution.

To conclude this point,

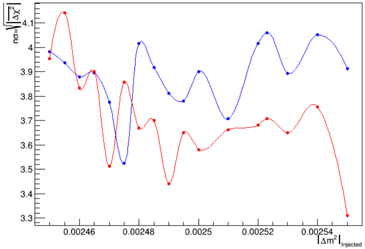
The $\Delta\chi^2$ estimator provides us with different results due to different simulation procedures. When the simulation is performed on a single event basis and not on a semi-analytical basis, it does not take into account the correlation between the side-bins due to systematic uncertainties, the significance drastically drops. The systematic uncertainties due to the $\frac{3\%}{\sqrt{E}}$ relatively energy resolution causes unbalanced immigration effect between bins that consequently create side-bin correlations leading to significant reduction in the experiment sensitivity . That invalids the use of the standard approximation at $\frac{3\%}{\sqrt{E}}$ relatively energy resolution.

1. The Draw-backs of the Standard Method Using $1D - \Delta\chi^2$
   - Issue I: The Limited Power of $\Delta\chi^2$
   - Issue II: The oscillation of significance with Δm^2
2. The Draw-backs of the Standard Method Using $2D - \chi^2$
   - Issue III: Non-bright Results using χ^2 as a Bi-Dimensional
   - Issue II: The oscillation of significance with Δm^2

Assuming that the approximation, $n\sigma = \sqrt{|\overline{\Delta\chi^2}|}$, is valid



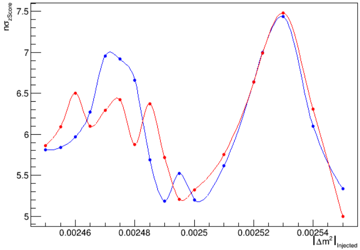
The oscillation of significance using the standard method with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming an **infinite energy resolution** where blue line for NH sample and red line for IH sample.



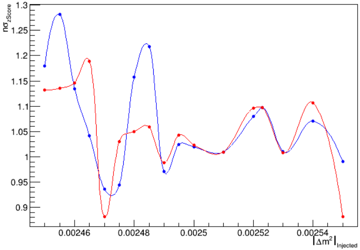
The variation of significance using the standard method with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming 3% **relativity energy resolution** where blue line for NH sample and red line for IH sample.

Issue II: The oscillation of significance with Δm^2

$$\text{Using z-test for 1D, } n\sigma = \frac{|\overline{\Delta\chi^2_{NH}}| - |\overline{\Delta\chi^2_{IH}}|}{\sigma(dispertion)}$$



The oscillation of significance using the standard method with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming an **infinite energy resolution** where blue line for NH sample and red line for IH sample.



The variation of significance using the standard method with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming 3% **relativity energy resolution** where blue line for NH sample and red line for IH sample.

To conclude this point,

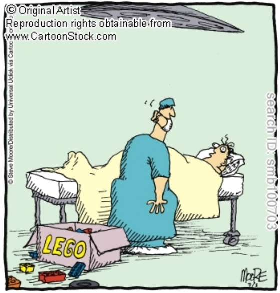
The oscillation of the experimental sensitivity with the assumed value of the neutrino atmospheric mass difference ($|\Delta m^2|$) implies that the standard method results have strong dependency on the input parameter value.

1 The Draw-backs of the Standard Method Using $1D - \Delta\chi^2$
- Issue I: The Limited Power of $\Delta\chi^2$
- Issue II: The oscillation of significance with Δm^2

2 The Draw-backs of the Standard Method Using $2D - \chi^2$
- Issue III: Non-bright Results using χ^2 as a Bi-Dimensional
- Issue II: The oscillation of significance with Δm^2

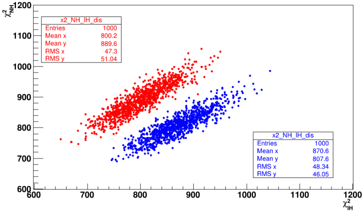
It is important to remember this

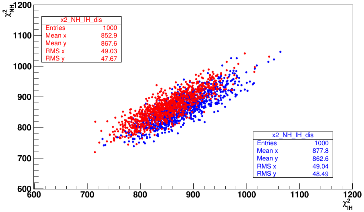
Evaluating the used statistical methods and updating them is a necessary step to build up a robust statistical analysis



"Surgery went well, Mr. Moore. I had a lot of fun rebuilding your knee joint."

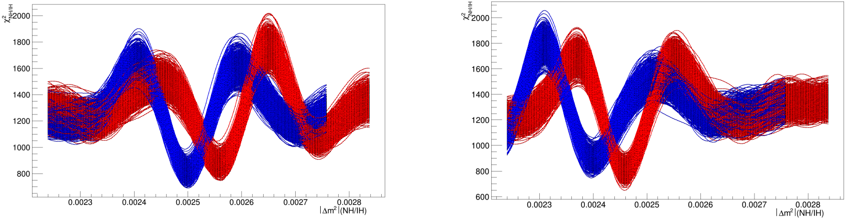
The Sensitivity Results using χ^2 as a Bi-Dimensional



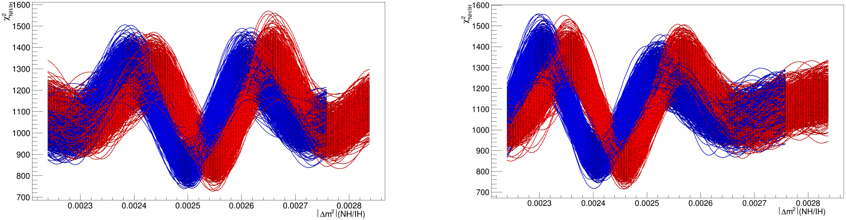


	0%		3%	
	NH	**IH**	**NH**	**IH**
μ_{NH}	807.6	889.6	862.6	867.6
σ_{NH}	46.05	51.04	48.49	47.67
μ_{IH}	870.6	800.2	877.8	852.9
σ_{IH}	48.34	47.3	49.04	49.03
nσ(NH)	1.145σ		0.603σ	
nσ(IH)	1.099σ		0.608σ	

Issue III: Non-bright Results using χ^2 as a Bi-Dimensional



The precision of the fit assuming infinity relativity energy resolution



The precision of the fit assuming 3% relativity energy resolution

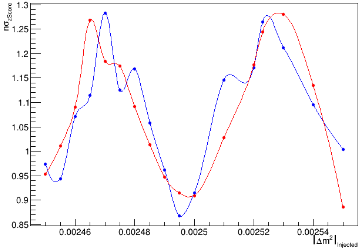
To conclude this point,

When $\chi^2_{min(NH)}$ and $\chi^2_{min(IH)}$ are drawn in two dimensional map, their strong positive correlation manifests χ^2 as a bi-dimensional estimator. The overlapping between the χ^2 distributions of the two hypotheses leading to reduction of the experimental sensitivity.

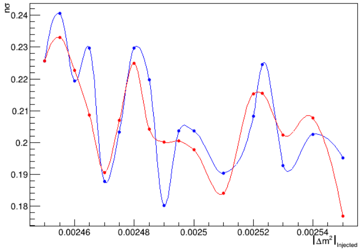
1. The Draw-backs of the Standard Method Using $1D - \Delta\chi^2$
   - Issue I: The Limited Power of $\Delta\chi^2$
   - Issue II: The oscillation of significance with Δm^2
2. The Draw-backs of the Standard Method Using $2D - \chi^2$
   - Issue III: Non-bright Results using χ^2 as a Bi-Dimensional
   - Issue II: The oscillation of significance with Δm^2

Using z-test for 2D,

$$n\sigma = \frac{\sqrt{(\mu_{IH}^{NH} - \mu_{IH}^{IH})^2 + (\mu_{NH}^{NH} - \mu_{NH}^{IH})^2}}{\sigma_{IH} + \sigma_{NH}}$$



The oscillation of significance using χ^2 as **bi-dimensional distribution** with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming an **infinite energy resolution** where blue line for NH sample and red line for IH sample.



The oscillation of significance using χ^2 as **bi-dimensional distribution** with $|\Delta m^2|_{inj}$ for 200(NH) + 200 (IH) JUNO-toy like simulations for 1 banchmark assuming an 3% **relativity energy resolution** where blue line for NH sample and red line for IH sample.

Finally,

To summarize, I will run through the three main the draw-backs of the standard method. Firstly, when the side-bins correlations are taking into account, the statistical assumptions are not valid any more and the limited power of the $\Delta\chi^2$ manifests itself. Secondly, the experimental sensitivity strongly depends on the value of the neutrino atmospheric mass difference. Thirdly, the overlapping between the χ^2 distributions of the two hypotheses leads to reduction of the experimental sensitivity.

In conclusion, it is up to you to improve the statistical analysis by realizing that you can efficiently deal with these draw-backs by simply evaluating them.

THANK YOU

It is often said that the language of science is mathematics. It could well be said that the language of experimental science is statistics.

The BackUP

1. How are the toy simulations done?
2. How are the fitting procedures?

The Toy Simulations

A toy simulations were based on a single event basis and the expected systematic errors via a Gaussian distribution centered at the expected mean and with the standard deviation of the estimated uncertainty can be added. For JUNO, a global $3\%/E(\text{MeV})$ resolution on the energy reconstruction is expected. The oscillation parameters have been taken from the most recent global fits.

	best-fit	3σ region
Sin^2_{12}	0.2970	0.2500 - 0.3540
$Sin^2_{13}(NH)$	0.02140	0.0185 - 0.0246
$Sin^2_{13}(IH)$	0.02180	0.0186 - 0.0248
δm^2_{sol}	7.37×10^{-5}	$6.93 \times 10^{-5} - 7.97 \times 10^{-5}$
$\Delta m^2(NH)$	2.500×10^{-3}	2.37×10^{-3} - 2.63×10^{-3}
$\Delta m^2(IH)$	2.460×10^{-3}	-2.60×10^{-3} to -2.33×10^{-3}

The Poisson statistical fluctuation is automatically included. Version "J17v1r1" of official JUNO Software is used for date simulations.

The Fitting Procedures

The fitting and minimization of χ^2 has required to use directly the ROOT minimization libraries, in particular the TMinuit algorithm. In the minimization procedure all the parameters were fixed to the best values that are indicated in assuming a very small error on it. One benchmark is referring to 6 years running at a distance ~ 52.5 km with a total power 36 GW and relative energy resolution $\frac{3\%}{\sqrt{E}}$. A total of 108357 signal events has been used in our simulations with a 10 keV bin energy width. All the oscillation parameters are in their best fitting values.