An application of the Strong Confidence to the CHOOZ experiment with frequentist inclusion of systematics

D. Nicolò[a,b] G. Punzi[a,c] G. Signorelli[a,c*]

[a)] *INFN Sez. di Pisa (Italy)*

[b)] *University of Pisa (Italy)*

[c)] *Scuola Normale Superiore di Pisa (Italy)*

March 11, 2002

Abstract

We apply a new prescription in confidence interval estimation, based on the frequentist method of strong Confidence Level (sCL), to make an inference on neutrino oscillation parameters based on the result of the CHOOZ experiment. Limits are obtained at 90% sCL, which include also exact frequentist treatment of systematics. Neutrino oscillations $\overline{\nu}_{\rm e} \to \overline{\nu}_{\rm x}$ are excluded for $\Delta m^2 \geq 8 \cdot 10^{-4}\,{\rm eV}^2$ at maximum mixing and $\sin^2(2\theta) \geq 0.17$ at large Δm^2 values. These limits should be regarded as safer than previous quoted limits and do not need to be supplemented by any additional sensitivity information. Resulting limits are slightly looser than the previously published values that used an approximate treatment of systematics.

*Corresponding author: Giovanni.Signorelli@pi.infn.it

1 Introduction

We applied the concept of Strong Confidence and Strong Confidence Limit [1] to the analysis of the data of the CHOOZ experiment, the first long baseline search for neutrino oscillations, operated at 1 km from two nuclear reactors. Preliminary [2] and final results of this experiment [3] have already been published. The reader is referred to those articles for an introduction to the problem of neutrino oscillations, for a general description of the experiment and for a discussion of its data analysis. Previous limits to neutrino oscillations were derived following the unified approach. New limits are found to be slightly looser than previous ones because of two effects, as will be discussed later: one, inherent in the strong confidence definition; the other associated with a rigorous classical treatment of systematics.

Let us briefly introduce the notations used throughout this paper: we will call the observable x (neutrino flux or content of the various energy bins) and X the observable space; μ are the oscillation parameters $(\sin^2 2\theta, \Delta m^2)$ to be determined and $B(x)$ the confidence region for a particular observed value x.

2 The strong confidence method

The standard definition of COnfidence Level of a belt can be written as follows:

$$\forall \mu \quad p(x : B(x) \not\ni \mu | \mu) \leq 1 - CL. \tag{1}$$

A confidence belt B is said to have a strong Confidence Level[1] equal to sCL if

$$\forall \mu \, \forall \chi \quad \frac{p(x : x \in \chi, B(x) \not\ni \mu | \mu)}{\sup\limits_{\mu} p(x : x \in \chi | \mu)} \leq 1 - sCL, \tag{2}$$

where χ is a subset of X, the set of all possible experimental outcomes. Compare to the definition of "standard" CL, which is obtained putting $\chi = X$ in (2).

We don't want to enter here into the details and motivations which led to formula (2) and refer the reader to references [1] and [4], but we want just to stress some facts:

- the presence of the denominator in formula (2), which descends from the requirement that the set of possible limits must depend only on the shape of the likelihood function, assures that even unlikely experimental results are correctly weighted by taking into account their maximum occurrence probability;

- $sCL \leq CL$ for a given confidence belt (put $\chi = X$ in (2)). In other words sCL belts are "good" confidence belts, yet overcovering for every value of μ;

- despite the freedom in the construction of the Strong Band we can define a "core" region of μ-values which must belong to every band. This is obtained putting $\chi = x$ in (2), implying that all values of μ such that

$$\frac{p(x|\mu)}{\sup_{\mu} p(x|\mu)} \geq 1 - sCL \tag{3}$$

must belong to every $B(x)$; as a consequence of this, it is impossible to exclude any value of the parameter with a likelihood close to its maximum.

3 Construction of the Strong Belt

The construction of the Strong Belt is a generalization of the construction of the Neyman belt [5, 6]. We divide the $x - \mu$ space in cells and flag all the cells which according to (3) belong to the core. After that for each μ we must include some more cells[1] to satisfy the "coverage requirement" (2). This can be done in a number of ways and there is the need of an additional prescription to supplement the definition (2) in order to get rid of this arbitrariness. The most natural choice is to include at fixed μ the $x-$cells by decreasing likelihood ratio (LR), for the core region is determined by a cut on that same quantity (see eq. (3)). For a given value of the parameter, every observable x outside the band has a LR lower than every x in the band. The

[1] In general the core is not itself a strong band.

same property holds also for Feldman–Cousins construction [6], but here the LR threshold is determined by eq. (2).

All possible subsets χ of non-flagged cells including the highest LR cell are considered and validity of eq. (2) is checked. In case of failure this highest-rank cell is added to the band and the entire procedure is repeated with the next-higher LR cell, until eq. (2) holds definitely.

3.1 Inclusion of systematic parameters

In real life, the *pdf* often depends on additional parameters (*e. g.* the normalization of absolute neutrino flux from reactors, detector efficiencies and calibrations), which we will call α, which are sources of systematic errors.

A correct way of taking into account systematics in a purely frequentist framework would correspond to building a confidence region in the $x-(\mu,\alpha)$ space and then *projecting* the region on the $x-\mu$ space; this would assure the proper coverage for every value of α. This approach (which is flagged by "correct systematics" in what follows) is seldom (almost never) followed in practice, due to the heavy computational resources needed. The two most commonly used approximate methods are:

1. Bayesian integration over the systematic parameters;

2. replacement of the likelihood with the one obtained by maximizing with respect to α for every μ (profile likelihood) [9, 10, 11].

In the first case one must assume a prior distribution for α, which can then be integrated out. However, a classical treatment does not allow such a Bayesian contamination: in fact, this procedure usually leads to under-coverage. The second method does not assure proper coverage for every value of α (this is assured only for that particular value that maximizes the likelihood for the given value of μ). In many cases this does not lead to significant under-coverage, so it can be considered a reasonable approximation. All published limits on neutrino oscillations derived by the unified approach were obtained in this way [3, 12, 13].

Conversely, the definition of strong limits is easily generalized to include nuisance parameters consistently and without approximations. When taking

systematics into account, eq. (2) is rewritten in a straightforward manner by the substitution $\mu \to (\mu, \alpha)$. Further manipulations allow us to write it in a more useful and clear form which eliminates the dependence on α, since we are interested in quoting limits on μ:

- condition $(\mu, \alpha) \notin B(x), \forall \alpha$ implies that $\mu \notin B_\mu(x)$, B_μ being the projection of the band on the μ space;
- since the inequality (2) must hold $\forall \alpha$, it is also true that:

$$\forall \mu \, \forall \chi : \quad \frac{\sup\limits_{\alpha} p(x : x \in \chi, B_\mu(x) \not\ni \mu | \mu, \alpha)}{\sup\limits_{\mu} \sup\limits_{\alpha} p(x : x \in \chi | \mu, \alpha)} \leq 1 - sCL. \tag{4}$$

In using this formula, it is natural to order x-cells by decreasing profile likelihood ratio LR_{prof}, given by

$$LR_{\text{prof}} = \frac{\sup\limits_{\alpha} p(x|\mu, \alpha)}{\sup\limits_{\mu} \sup\limits_{\alpha} p(x|\mu, \alpha)} ; \tag{5}$$

as in eq. (3), a cut on this variable provides the core of the strong confidence interval. Therefore, strong belts with inclusion of systematics are obtained by applying the procedure described in the previous subsection with minor modifications.[2] The important point is that the *pdf* is replaced by a new *pdf'* which is its maximization with respect to α thus preserving the dimensionality of the problem, unlike the "correct systematics" case in which we add new dimensions for α, build the belt in a $\dim(\mu) + \dim(\alpha) + \dim(x)$ space and only in the end project everything back on the μ-space. This simplification, which is strictly related to the Strong Confidence requirement, yields a much simpler exact treatment than other ordering prescriptions.

[2] Another point worth noting is that the above ordering somehow minimizes the width of the confidence region in μ, by forcing the band to expand in α-direction rather than in μ-direction.

4 Strong limits for the CHOOZ experiment: total counting rate

As an application of the described procedure we calculated the Strong Limits for the CHOOZ experiment. Since CHOOZ is a disappearance search for neutrino oscillations, a positive evidence would result both in a deficit on the $\overline{\nu}_{\rm e}$ counting rate expected. and in a deformation of the $\overline{\nu}_{\rm e}$ energy spectrum. In this paper we deal only with the total counting rate ("integral" analysis). The procedure has been generalized in [4] to exploit the information contained in the whole spectrum.

4.1 Likelihood function

Our observable x is chosen to be the ratio of measured to expected $\overline{\nu}_{\rm e}$-flux, which gives a direct estimate of the oscillation probability. The physics parameters to be inferred are the usual oscillation parameters $\mu = (\sin^2(2\theta), \Delta m^2)$. The expected flux in absence of oscillations is known at a $\sigma_{\rm syst} = 2.7\%$ level [3] which is of the same order of magnitude as the statistical uncertainty ($\sigma_{\rm stat} = 2.8\%$). We are then forced to introduce an additional "nuisance" parameter α to account for our lack of knowledge of the absolute flux. Other systematic effects related to the spectral shape (energy calibration for instance) are negligible. The likelihood is then written, apart from multiplicative factors, as follows:

$$p(x|\mu,\alpha) \propto \exp\left[-\frac{1}{2}\left(\frac{x-\alpha\mathcal{P}(\mu)}{\sigma_{\rm stat}}\right)^2\right] \times \exp\left[-\frac{1}{2}\left(\frac{\alpha-1}{\sigma_{\rm syst}}\right)^2\right], \tag{6}$$

$\mathcal{P}(\mu)$ being the survival probability averaged over the energy spectrum.

We considered an observable range from 0.75 to 1.25 divided into 50 cells. The domain of oscillation parameters ($0 < \sin^2(2\theta) < 1$, $10^{-4} < \Delta m^2 < 1\,{\rm eV}^2$) was divided into 100×100 cells (with a constant binning in $\log \Delta m^2$); the range $0.95 < \alpha < 1.05$ was divided into 10 cells[3]. We computed $B_\mu(\overline{x})$,

[3] We verified that values of α outside the considered range give no further contribution

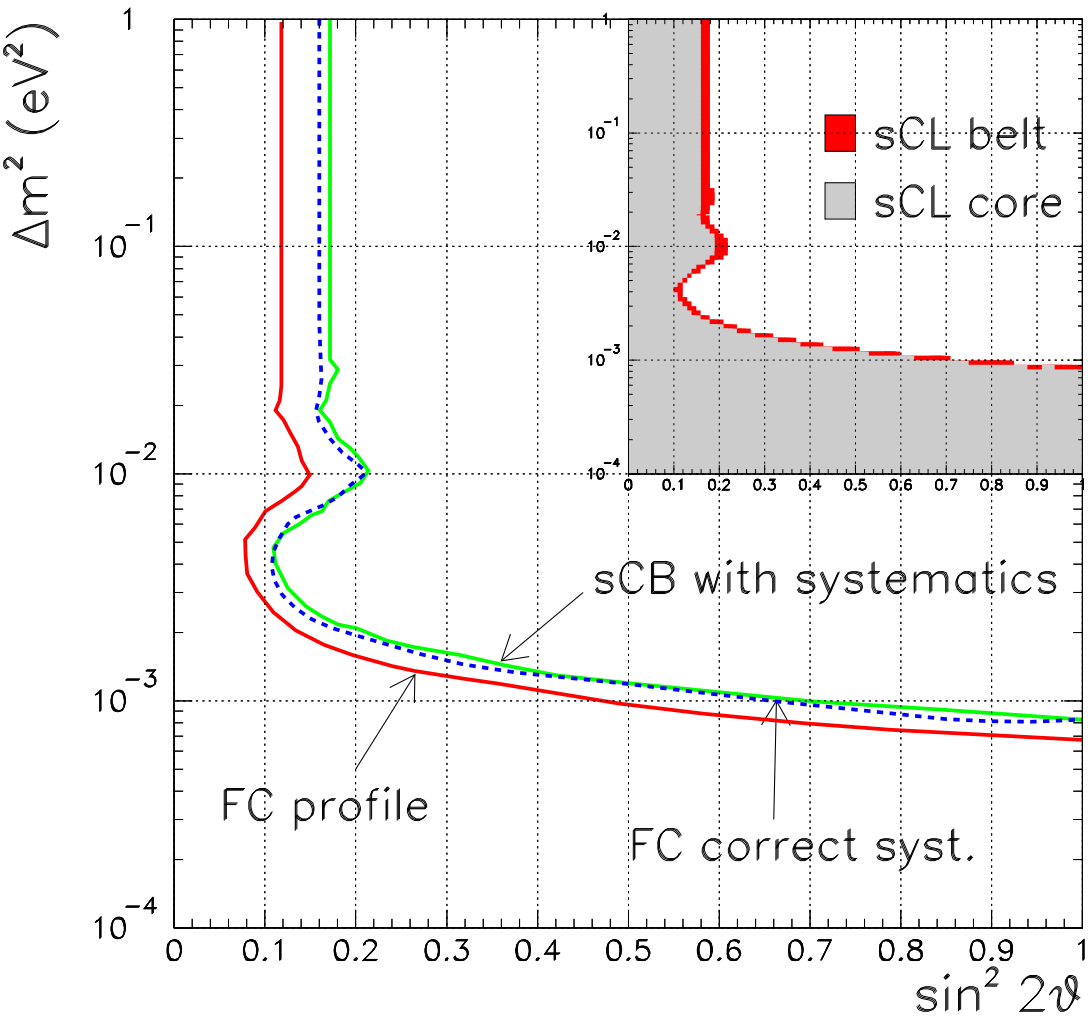

Figure 1: 90% confidence belts calculated using the integral of the spectrum. Three belts are shown: strong belt (light solid line), Feldman–Cousins belt with "correct systematics" treatment (dashed line) and Feldman–Cousins belt derived by profile likelihood ratio ordering (dark solid line). In the inset the difference between the core and the whole strong band is shown.

$\overline{x} = 1.01$ being the experimental result. It is not necessary to compute all the multi-dimensional band and to slice it with the $x = \overline{x}$ layer but we directly built $B_\mu(\overline{x})$ proceeding as follows: for each μ ($\sin^2(2\theta), \Delta m^2$ cell) we

to the projection of the band on μ-space.

consider only the x-cells with $LR_{\rm prof}(x) \leq LR_{\rm prof}(\overline{x})$ and use only them to form subsets χ on which we verify condition (4). If it fails for some χ, then $\mu \in B_\mu(\overline{x})$. In fact we should take the highest $LR_{\rm prof}$ cell out of χ and put it in the band, but the highest $LR_{\rm prof}(x)$ is just $\overline{x}$ which therefore belongs to the belt.

The results of our computation are shown in Fig. 1. The limits quoted are slightly looser, as one expects from the definition of strong confidence, than those obtained by using the Feldman–Cousins prescription with the correct inclusion of systematics: at large Δm^2 values, the strong upper limit on mixing is $\sin^2(2\theta) \leq 0.17$ (0.16 for FC). It is also remarkable that these limits are significantly higher than the one obtained with the FC+profile procedure (asymptotically $\sin^2(2\theta) \leq 0.11$) used in [3].

5 Computational requirements

CPU time required to perform such computations is not so heavy as one might think, although one has to handle a large amount of memory. Most of the job lies in numerical computation of the *pdf*, when it is not known analytically. Of course, dimensionality and graining of both observable and parameter space heavily affects both time and memory. Since calculation of both *pdf* and strong band can be performed independently for each value of the parameter, it is possible to spread the whole computational effort over different CPU's.

6 Conclusions

We derived 90% strong confidence limits on neutrino oscillations in the disappearance mode $\overline{\nu}_{\rm e} \rightarrow \overline{\nu}_{\rm x}$ based on the complete CHOOZ data sample. No evidence is found in the oscillation parameter region with $\Delta m^2 \geq 8 \cdot 10^{-4}\,{\rm eV}^2$ at maximum mixing and $\sin^2(2\theta) \geq 0.17$ at large Δm^2 values. These limits are looser than previously published ones because of two effects: one, inherent in the definition of strong confidence; the other associated with a classically correct treatment of systematics.

Acknowledgements

We wish to thank the CHOOZ collaboration for allowing us to use its data, in particular the Pisa group for useful discussions. We are also grateful to Flavia Donno for her support in computing.

References

[1] G. Punzi, *Stronger classical confidence limit*, Contribution to the "Workshop on Confidence Limits", CERN, Jan 17-18 2000, to pe published in the proceedings of the workshop.
G. Punzi, *A stronger classical definition of confidence limit*, INFN Pisa preprint INFN PI/AE 99/08 (also available on the INFN Pisa web site *http://www.pi.infn.it/library/preprints*).

[2] M. Apollonio *et al.*, Phys. Lett. B **420**, 397 (1998).

[3] M. Apollonio *et al.*, Phys. Lett. B **466**, 415 (1999).

[4] D. Nicolò, G. Signorelli, G. Punzi, in preparation.

[5] J. Neyman, Philos. Trans. R. Soc. London, **A236**, 333 (1937).

[6] G. J. Feldman, R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).

[7] C. Caso *et al.*, Eur. Phys. J. C**3** 1 (1998).

[8] B. P. Roe, M. B. Woodroofe, Phys. Rev. D **60**, 053009-1 (1999).

[9] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, Vol. 2, 5th ed. (Oxford University Press, New York, 1991).

[10] B. Achkar *et al.* [Bugey coll.] Nucl. Phys. B**434** (1995) 503-532

[11] G. Zacek *et al.* [Gösgen coll.] Phys. Rev. D **34**, (1986) 2621

[12] F. Boehm *et al.*, "Results from the Palo Verde neutrino oscillation experiment," Phys. Rev. D **62** (2000) 072002 [arXiv:hep-ex/0003022].

[13] P. Astier *et al.*, Phys. Lett. B **453**, 169 (1999).