Reassessing the sensitivity to leptonic CP violation

Mattias Blennow,1,* Pilar Coloma,2,† and Enrique Fernandez-Martinez3,4,‡

1Department of Theoretical Physics,
School of Engineering Sciences, KTH Royal Institute of Technology,
AlbaNova University Center, 106 91 Stockholm, Sweden
2Center for Neutrino Physics, Virginia Tech, Blacksburg, VA 24061, USA
3Departamento de Física Teórica, Universidad Autónoma de Madrid,
Cantoblanco E-28049 Madrid, Spain
4Instituto de Física Teórica UAM/CSIC,
Calle Nicolás Cabrera 13-15, Cantoblanco E-28049 Madrid, Spain

Abstract

We address the validity of the usual procedure to determine the sensitivity of neutrino oscillation experiments to CP violation. An explicit calibration of the test statistic is performed through Monte Carlo simulations for several experimental setups. We find that the usual “Δχ²” test statistic is close to being χ²-distributed under the null hypothesis of CP conservation for all facilities under study. Good agreement is also found between the results obtained using the Asimov data set and the median of the Monte Carlo distributions for CP violating values. The only exception is the NOνA experiment, for which large deviations are observed.

*Electronic address: emb@kth.se
†Electronic address: pcoloma@vt.edu
‡Electronic address: enrique.fernandez-martinez@uam.es

arXiv:1407.3274v2 [hep-ph] 15 Aug 2014
I. INTRODUCTION

Mixing in the lepton sector of the Standard Model is described by the unitary Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [1–5]. In the standard three family scenario, it can be parametrized by three mixing angles ($\theta_{12}$, $\theta_{23}$, and $\theta_{13}$), a Dirac CP violating phase ($\delta$) and, if neutrinos are Majorana particles, two additional Majorana CP phases. Neutrino oscillations are sensitive to the values of all mixing angles and the CP violating phase $\delta$, together with the two independent mass squared differences ($\Delta m_{21}^2$ and $\Delta m_{31}^2$, defined as $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$). Unlike for the CKM matrix, the mixing angles of the PMNS matrix have been experimentally found to be large, with $\theta_{13} \simeq 9^\circ$ being the smallest mixing angle [6–10] and the other two angles being much larger ($\theta_{12} \sim 33^\circ$ and $\theta_{23} \sim 45^\circ$ [11]). Thus, the Jarlskog invariant $J = \cos \theta_{13} \sin 2\theta_{13} \sin \theta_{12} \sin \theta_{23} \sin \delta$ can be potentially as large as $\sim 0.035$ for maximally CP violating values of $\delta$, four orders of magnitude higher than the currently measured value for its counterpart in the quark sector: $J_{\text{CKM}} = (2.96^{+0.20}_{-0.16}) \cdot 10^{-5}$ [12]. Since it has been shown that, within the context of Standard Model Electroweak Baryogenesis, $J_{\text{CKM}}$ is not large enough to account for the observed Baryon asymmetry of the Universe [13, 14], the discovery of an additional source of CP violation (such as $\delta$ in the PMNS matrix) could open new possibilities for alternative generation mechanisms.

Given the current knowledge on the neutrino oscillation parameters, the focus for the next generation of neutrino oscillation experiments will be to determine the neutrino mass ordering, the CP violating phase $\delta$, and the octant of the mixing angle $\theta_{23}$. For the determination of the mass ordering, there has been a lively discussion in the literature on whether or not the common way of assessing the sensitivity of a future experiment is applicable [15–18]. As shown in Ref. [17], the common approach does give a reasonable approximation of the median sensitivity, although not for the reasons typically used in the argumentation for it. The small correction was mainly based on the one-sided nature of the hypothesis test and only to a minor degree on the non-gaussianity of the statistical distributions. Similarly, deviations from a $\chi^2$ distribution of the test statistic should be tested for in the search for $\delta$. Indeed, one of the requirements for the validity of the common approach when making sensitivity analyses is that the change in number of events forms a linear space upon variations of $\delta$. However, this requirement will necessarily be violated at some level since $\delta$ is periodic and a change by $2\pi$ will leave the number of events unchanged. In addition, there is no guarantee that the predicted data without statistical fluctuations, as used in the common approach, will be representative. In the present work, these assumptions are tested by explicit Monte Carlo simulations in order to find out exactly how much the sensitivity analyses are affected by these assumptions. To do so, we start from the basic frequentist definitions and apply the Feldman–Cousins approach [19] in order to determine the sensitivity of several experimental setups.

II. STATISTICAL APPROACH

The most common way of quantifying the experimental sensitivity to leptonic CP violation is to quote the confidence level at which the CP conserving values of $\delta$ (0 and $\pi$) can
be rejected. In the literature, this is typically computed by constructing the test statistic

\[ S = \min_{\delta=0,\pi} \chi^2 - \min_{\text{global}} \chi^2, \]  

(1)

where \( \chi^2 = -2 \log \mathcal{L} \), and \( \mathcal{L} \) is the likelihood of observing the data given a particular set of oscillation parameters. It is then assumed that \( S \) is \( \chi^2 \)-distributed with one degree of freedom, based on the implications of Wilks’ theorem \[20\]. In addition, the Asimov data set\(^1\) \[21\], i.e., the event rates without statistical fluctuations, is assumed to be representative for the experimental outcome and is thus used to determine the expected confidence level (CL) at which the CP conservation hypothesis would be rejected. However, as for the case of the neutrino mass ordering \[15, 17\], it is not clear to what degree the assumptions underlying Wilks’ theorem are violated when testing CP conservation in this fashion, resulting in a need to explicitly test this hypothesis.

The procedure to test the CP conservation hypothesis at a given CL can be summarized as follows: First, the distribution of the test statistic \( S \) is found by simulating a large number of realizations of the experiments based on the predicted event rates under the assumption of CP conservation. The value of \( S \) is then computed for each realization, which provides the distribution of \( S \). CP conservation will be rejected at CL \( x \) if the measured value of \( S \) is among the \( 1-x \) fraction of largest values in the distribution. This automatically defines a critical value, \( S_c(x) \), such that CP conservation is rejected at CL \( x \) if \( S > S_c(x) \). By construction, \( S_c(x) \) is the inverse of the cumulative distribution function (CDF) of \( S \) under the CP conserving hypothesis.

The above construction is only concerned with the test of CP conservation for a given data set, i.e., once the experiment has already taken data. The expected performance of future facilities will depend on the true value of \( \delta \). Thus, in a frequentist approach the performance of the facility must be quantified for each value of \( \delta \) separately. In addition, due to statistical fluctuations, different realizations of a given facility will lead to a different significance at which CP conservation can be rejected. Therefore, the convention is to define the expected sensitivity of a given experiment as the CL obtained for the median of the distribution, and is typically shown as a function of the value of \( \delta \) itself. This is usually referred to as the median sensitivity and it will not necessarily coincide with the significance computed with the Asimov data set.

A calibration of the \( \chi^2 \), following the procedure described above, was performed in Ref. \[22\]. The T2HK experiment \[23\] was used as an example, and the critical values were found to be significantly smaller than for the \( \chi^2 \) distribution with 1 degree of freedom in the region \( \sin^2 2\theta_{13} \gtrsim 10^{-2} \). However, the study in Ref. \[22\] was restricted to values of \( \sin^2 2\theta_{13} \lesssim 5 \times 10^{-2} \), and was done using a different test statistic than the one considered in the present work.

\(^1\) So named after the Franchise short story by Isaac Asimov, where an entire electorate was replaced by one single representative.
A. External constraints

The common approach when dealing with external constraints on the parameters (such as previous determinations of oscillation parameters or systematic errors) is to include an additional term in the $\chi^2$ of the form

$$\chi^2 = \chi^2_0(\xi) + \frac{(\xi - \bar{\xi})^2}{\sigma^2_\xi},$$

(2)

where $\xi$ is the constrained parameter, $\chi^2_0(\xi)$ is the $\chi^2$ provided by the experiment itself for a fixed value of $\xi$, and $\sigma_\xi$ is the error in the determination of $\xi$ which has a measured central value of $\bar{\xi}$. In order to calibrate the $\chi^2$ to do a proper hypothesis test, statistical fluctuations must also be considered for the experiment determining $\xi$. This can be approximated by considering the experiment to be measuring $\xi$ with a Gaussian error of $\sigma_\xi$. From this follows that, when calibrating the critical value $S_c$ for a fixed $\xi$, the values of $\xi$ should be chosen according to a normal distribution with mean $\xi$ and standard deviation $\sigma_\xi$. The test statistic is then again defined through Eq. (1) and the critical value is computed accordingly.

While the above is true for the calibration of the critical value for the test statistic $S_c$, the known outcome of the external experiments should be taken into account when computing the median expected outcome by fixing $\bar{\xi}$ to the measured value. This procedure will then reflect the probability of reaching a given confidence level based on the already known outcome of the external constraints. Note that, if external constraints from a hypothetical future experiment are implemented, the outcome of such an experiment is unknown and should therefore still be chosen according to the outcome distribution.

In our simulations, we have followed this prescription for the existing constraints on the neutrino oscillation parameters as well as for including constraints on the systematic errors. The calibration of the test statistic has been performed for the best fit values of the parameters, but is expected not to be crucially dependent on this.

III. SIMULATION DETAILS

We simulate the long-baseline experiments LBNE [24], T2HK [23], ESSνSB [25], and NOνA [26] using the GLoBES software [27, 28] to find their respective Asimov data sets as functions of the parameter values. The simulation details for the LBNE, T2HK, ESSνSB, and NOνA experiments were implemented as in Refs. [17], [29], [25], and [17], respectively. For the ESSνSB, we use the configuration with a 540 km baseline and 2.5 GeV protons. Once the Asimov data sets were computed, realizations of the experiments were constructed by applying Poisson statistics individually to the number of events in each bin based on the expected rates. This was implemented using the MonteCUBES software [30]. The value of $S$ was then computed for each realization in order to find the distribution of $S$ under

\footnote{Note that this is \textit{not} equivalent to picking $\xi$ from a random distribution. The test statistic must still be separately calibrated for each simple hypothesis.}

\footnote{The only modification is that in the present work the beam power of the LBNE experiment has been increased to 1.2 MW, and the detector mass has been fixed to 34 kt.}
FIG. 1: Distribution of the test statistic $S$ for the different simulated experiments. We show the value of $1 - \text{CDF}(S)$, where CDF is the cumulative distribution function. The values of $1 - \text{CDF}$ corresponding to $1\sigma$, $2\sigma$, and $3\sigma$ are shown as horizontal lines. For comparison, the red line shows the result for a $\chi^2$ distribution with one degree of freedom.

the CP conserving and CP violating hypotheses. In all cases, 5 % (10 %) systematic errors were used for the signal (background) event rates. These are bin-to-bin correlated, but uncorrelated between signal and backgrounds and between different oscillation channels. The true values of the oscillation parameters have been set according to the best fits in Ref. [11], with the exception of $\theta_{23}$ which has been set to 45°. Marginalization is performed on $\theta_{12}$, $\sin^2 2\theta_{13}$, $\sin^2 2\theta_{23}$, $\Delta m^2_{21}$ and $\Delta m^2_{31}$, using gaussian priors in agreement with present experimental uncertainties from Ref. [11]. Sign degeneracies are fully taken into account in the minimizations.

In order to compute the critical values of $S$, we simulate $10^5$ realizations for $\delta = 0$ and $\pi$ and for both neutrino mass orderings. This is sufficient to reliably find the value of $S_c$ up to $\sim 3.5\sigma$ level, corresponding to a CL of $\sim 99.95$ %. On the other hand, for the computation of the expected outcomes we are only interested in obtaining the value of $S$ for the median of the distributions. Since this is much less sensitive to statistical fluctuations than the sampling of the tails, only $10^3$ realizations of the experiments are simulated in order to obtain the median of $S$ for these values.

IV. RESULTS

Figure 1 shows our results for the distribution of the test statistic $S$ for the experimental setups considered in this work. As can be seen from the figure, the CDFs are generally close to a $\chi^2$ distribution with one degree of freedom for almost all experiments under consideration. The only exception is NOvA, for which large deviations are observed and the critical values corresponding to a given CL are considerably reduced in comparison to the
values obtained under the assumption of a $\chi^2$-distributed test statistic. On the other hand, for the other experiments the CDFs fall slightly slower than a $\chi^2$ distribution and therefore somewhat larger values are obtained. This outcome agrees with the naive expectation that the largest deviations from Wilks’ theorem should be more relevant in the experiments with the poorest sensitivity to the value of $\delta$. Indeed, one of the requirements for the applicability of Wilks’ theorem is that, when varying the parameter that is being tested for, in this case $\delta$, the subsequent change in the observables used to determine it should constitute a linear space. This requirement will obviously be violated at some level for $\delta$ since a change by $2\pi$ will render no change in the number of events. However, next generation facilities will be able to constrain very precisely the value of $\delta$, and within a small range of values of $\delta$ the linearity condition for the number of events is expected to be approximately satisfied. Since the precision in the determination of $\delta$ by NO$\nu$A will be poorer than for any of the next generation facilities considered here, the larger deviation from the $\chi^2$ distribution observed for NO$\nu$A could be attributed to a stronger violation of this requirement.

The other relatively significant deviation from a $\chi^2$ distribution can be seen for T2HK. We have checked that this extra deviation with respect to the better agreement showed by LBNE and the ESS can be attributed to the sign degeneracies, that play a very important role in the determination of $\delta$ from T2HK. A parametric degeneracy will again lead to situations in which the change in the number of events does not span a linear space, implying the non-applicability of Wilks’ theorem. When the mass hierarchy is assumed to be known, the distribution obtained for T2HK lies on top of those obtained for LBNE and the ESS as expected.

A priori, a variation on the critical values $S_c$ with respect to the ones obtained for a $\chi^2$ distribution does not necessarily imply a change in the sensitivity of an experiment. This will depend on whether the Asimov data can be taken as a good approximation of the median result for values of $\delta \notin \{0, \pi\}$. In Fig. 2, the median value of the distribution of $S$ is shown as a function of $\delta$ for all experiments under consideration, and the results are compared to the values obtained from the Asimov data. We find that the Asimov data set is very close to the median value of $S$ for all future experiments, while significant deviations are found for NO$\nu$A. It thus follows that the Asimov data may be taken as a good approximation of the median test statistic when considering the more sensitive experiments, while for NO$\nu$A Monte Carlo simulation would be advisable.

Finally, in Fig. 3 we compare the computed sensitivity to CP violation using a full analysis based on Monte Carlo simulation with the results obtained in the common approach (i.e., assuming the cut values of a $\chi^2$ distribution with one degree of freedom and the Asimov data set). From this figure, a reasonable agreement in the sensitivity can be observed with respect to the results obtained by simply taking the $\sqrt{S}$ for the Asimov data. The only exception is for NO$\nu$A in the region around $\delta \sim 90^\circ$ where, even if the critical value $S_c$ from Fig. 1 is significantly lower than for a $\chi^2$ distribution, the Asimov data set considerably overestimates the median result from the Monte Carlo (see Fig. 2). We also note that the Monte Carlo results do not go to zero around the CP conserving values of $\delta$. Instead the median sensitivity is around 0.67$\sigma$, corresponding to a confidence level of 50 %. This should be expected as the median confidence level obtained if CP is conserved should be 50 %. The Asimov data cannot reflect this since any fluctuations around it will increase the value of the test statistic. It is therefore not a good approximation of the median in a region close
FIG. 2: The test statistic $S$ for the simulated experiments, as a function of the true value of $\delta$ (in degrees). The median values obtained from Monte Carlo simulations are shown as solid lines, whereas the values corresponding to the predicted Asimov data set are shown as dashed lines.

FIG. 3: Predicted median sensitivity for rejecting CP conservation, as a function of the true value of $\delta$ (in degrees). Solid lines show the results using the true median and Monte Carlo calibrated distributions of $S$, while dashed lines show the results using the common approach of taking $\sqrt{S}$ for the Asimov data set.
to the CP conserving values of $\delta$.

V. SUMMARY AND OUTLOOK

In this letter, we have studied the validity of the common approach used to compute the sensitivity of future long baseline experiments to leptonic CP violation. By explicit Monte Carlo simulation we have found that the test statistic (defined in Eq. (1)) is close to being $\chi^2$-distributed for almost all the experiments under consideration. We have also found that the median value of the distribution is well approximated by the Asimov data set in most cases. This results in a sensitivity which is very similar, although slightly worse, than what is typically quoted in the literature.

An exception to the above rule is the NO$\nu$A experiment. While we find that the test statistic is still not $\chi^2$-distributed and is peaked towards significantly smaller values, we also find that the Asimov data is not representative for the distribution of the test statistic. Thus, when the sensitivity was computed using the median values from the Monte Carlo simulations instead of the values coming from the Asimov data, the results were found to be similar to the ones obtained in the common approach.

Acknowledgments

We are very grateful to Thomas Schwetz for valuable discussions and encouragement. We warmly thank Peter Ballett for pointing out an inconsistency in the treatment of the systematic errors in the first version of the manuscript, and Pilar Hernández for illuminating discussions. This work was supported by the Göran Gustafsson Foundation [MB] and by the U.S. Department of Energy under award number DE-SC0003915 [PC]. EFM acknowledges financial support by the European Union through the FP7 Marie Curie Actions CIG NeuProbes (PCIG11-GA-2012-321582) and the ITN INVISIBLES (PITN-GA-2011-289442), and the Spanish MINECO through the “Ramón y Cajal” programme (RYC2011-07710) and through the project FPA2009-09017. We also thank the Spanish MINECO (Centro de excelencia Severo Ochoa Program) under grant SEV-2012-0249 as well as the Nordita Scientific program “News in Neutrino Physics”, where parts of this work was performed.

[1] B. Pontecorvo, Sov.Phys.JETP 6, 429 (1957).
[2] B. Pontecorvo, Sov.Phys.JETP 7, 172 (1958).
[3] Z. Maki, M. Nakagawa, Y. Ohmuki, and S. Sakata, Prog.Theor.Phys. 23, 1174 (1960).
[4] Z. Maki, M. Nakagawa, and S. Sakata, Prog.Theor.Phys. 28, 870 (1962).
[5] B. Pontecorvo, Sov.Phys.JETP 26, 984 (1968).
[6] F. An et al. (DAYA-BAY Collaboration), Phys.Rev.Lett. 108, 171803 (2012), 1203.1669.
[7] J. Ahn et al. (RENO collaboration), Phys.Rev.Lett. 108, 191802 (2012), 1204.0626.
[8] Y. Abe et al. (Double Chooz Collaboration), Phys.Rev. D86, 052008 (2012), 1207.6632.
[9] P. Adamson et al. (MINOS Collaboration), Phys.Rev.Lett. 107, 181802 (2011), 1108.0015.
[10] K. Abe et al. (T2K Collaboration), Phys.Rev.Lett. 107, 041801 (2011), 1106.2822.
[11] M. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz, JHEP 1212, 123 (2012), 1209.3023, URL http://www.nu-fit.org/.
[12] J. Beringer et al. (Particle Data Group), Phys.Rev. D86, 010001 (2012).
[13] M. Gavela, P. Hernandez, J. Orloff, and O. Pene, Mod.Phys.Lett. A9, 795 (1994), hep-ph/9312215.
[14] M. Gavela, P. Hernandez, J. Orloff, O. Pene, and C. Quimbay, Nucl.Phys. B430, 382 (1994), hep-ph/9406289.
[15] X. Qian, A. Tan, W. Wang, J. Ling, R. McKeown, et al., Phys.Rev. D86, 113011 (2012), 1210.3651.
[16] F. Capozzi, E. Lisi, and A. Marrone, Phys.Rev. D89, 013001 (2014), 1309.1638.
[17] M. Blennow, P. Coloma, P. Huber, and T. Schwetz, JHEP 1403, 028 (2014), 1311.1822.
[18] O. Vitells and A. Read (2013), 1311.4076.
[19] G. J. Feldman and R. D. Cousins, Phys.Rev. D57, 3873 (1998), physics/9711021.
[20] S. Wilks, Annals Math.Statist. 9, 60 (1938).
[21] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur.Phys.J. C71, 1554 (2011), 1007.1727.
[22] T. Schwetz, Phys.Lett. B648, 54 (2007), hep-ph/0612223.
[23] K. Abe, T. Abe, H. Aihara, Y. Fukuda, Y. Hayato, et al. (2011), 1109.3262.
[24] C. Adams et al. (LBNE Collaboration) (2013), 1307.7335.
[25] E. Baussan et al. (ESSnuSB Collaboration), Nuclear Physics B 885, 127 (2014), 1309.7022.
[26] D. Ayres et al. (NOvA Collaboration) (2004), hep-ex/0503053.
[27] P. Huber, M. Lindner, and W. Winter, Comput.Phys.Commun. 167, 195 (2005), hep-ph/0407333.
[28] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter, Comput.Phys.Commun. 177, 432 (2007), hep-ph/0701187.
[29] P. Coloma, P. Huber, J. Kopp, and W. Winter, Phys.Rev. D87, 033004 (2013), 1209.5973.
[30] M. Blennow and E. Fernandez-Martinez, Comput.Phys.Commun. 181, 227 (2010), 0903.3985.