Neutrino mass hierarchy from nuclear reactor experiments

Emilio Ciuffoli, Jarah Evslin, and Xinmin Zhang

TPCSF, IHEP, CAS, YuQuanLu 19B, Beijing 10049, China

Ten years from now reactor neutrino experiments will attempt to determine which neutrino mass eigenstate is the most massive. In this paper we present the results of more than seven million detailed simulations of such experiments, studying the dependence of the probability of successfully determining the mass hierarchy upon the analysis method, the neutrino mass matrix parameters, reactor flux models, geoneutrinos and, in particular, combinations of baselines. We show that a recently reported spurious dependence of the data analysis upon the high energy tail of the reactor spectrum can be removed by using a weighted Fourier transform. We determine the optimal baselines and corresponding detector locations. For most values of the CP-violating, leptonic Dirac phase δ, a degeneracy prevents NOνA and T2K from determining either δ or the hierarchy. We determine the confidence with which a reactor experiment can determine the hierarchy, breaking the degeneracy.

I. INTRODUCTION

Last year the Daya Bay [1,2] and RENO [3] experiments demonstrated that the neutrino mass matrix mixing angle θ_{13} is nonzero and several times larger than had been suspected just two years earlier. With the discovery that θ_{13} is nonzero, at least three qualitative questions remain to be answered in the standard model plus three massive neutrinos. First, it is not known whether the second or third neutrino flavor is the most massive. This choice is known as the neutrino mass hierarchy. Second, it is not known whether the leptonic sector has a nonzero CP-violating Dirac phase δ. Third, the most precise determination of the mixing angle, θ_{23}, by the MINOS Collaboration is [4]

$$\sin^2(2\theta_{23}) = 0.96 \pm 0.04. \tag{1}$$

This equation has eight distinct solutions, known as the octants of θ_{23}. Two of these solutions lead to inequivalent mass matrices, corresponding to θ_{23} modulo $90°$ greater than or less than $45°$.

The large value of θ_{13} means that an appearance experiment, searching for ν_e's in a beam of accelerator ν_μ's, may be sensitive to CP violation in the leptonic sector. The trouble is that, at the baselines of the T2K experiment in Japan and the NOνA experiment in the U.S., the CP phase δ is degenerate with the mass hierarchy, the octant of θ_{23} and the precise value of θ_{13}. Following the analysis in Ref. [5], by combining data from the appearance of ν_e's in a ν_μ beam with the appearance of $\bar\nu_e$'s in a $\bar\nu_\mu$ beam in its second three year run, given sufficient funding NOνA may be able to separate the octant and θ_{13}, which determine the total $\nu_e + \bar\nu_e$ appearance, from the hierarchy and δ, which roughly determine the difference $\nu_e - \bar\nu_e$. However for most values of δ even the combination of T2K and NOνA will not be able to break the degeneracy between δ and the hierarchy, and so will not be able to determine either.

Fortunately the large value of θ_{13} also implies that 1–3 neutrino oscillations, those with amplitudes proportional to $\sin^2(2\theta_{13})$, are large enough to be observed in reactor neutrino experiments at medium baselines (40–60 km). These oscillations are almost periodic in the inverse energy E, but their small aperiodicity may be used to determine the neutrino mass hierarchy [6] thus breaking the degeneracy at T2K and NOνA and so allowing a determination of the CP-violating Dirac phase. More precisely, combined with a hierarchy determination, the NOνA appearance mode can only determine $\sin(\delta)$ because the transformation $\delta \to \pi - \delta$ is degenerate with a slight shift in θ_{23} or θ_{13}. However the effective mass difference determined by a reactor experiment, which will be discussed below, differs from the δ-dependent atmospheric effective mass [7] determined by the disappearance channel of experiments such as MINOS and NOνA. This difference may yield a 1σ determination of $\cos(\delta)$. Such reactor experiments are not only possible but indeed *will* be performed within the next decade [8,9].

Furthermore, these reactor neutrino experiments may well have two detectors so as to reduce their sensitivity to systematic errors resulting from the detector's unknown energy response. This would offer a unique opportunity to measure δ using a strategy similar to that of the DAEδALUS experiment described in Ref. [10]. One would need a high intensity stationary pion source located at about 10 km from one detector and 20 km from the other. The fact that only one source is required makes such an experiment both cheaper than DAEδALUS would be at Long-Baseline Neutrino Experiment and also more precise, as it eliminates the uncertainties due to the relative strengths of the sources.

In this paper we present the main results of a series of simulations of medium baseline reactor experiments. We will present the reliability of the determinations of the hierarchy in such experiments and will also find the optimal baselines for their detectors.

II. SIMULATIONS

We consider reactor experiments in which $\bar{\nu}_e$'s are detected via inverse β decay by a detector with a 20 kton target mass consisting of 12% hydrogen. This is the mass of the proposed Daya Bay II detector, but it is 1.1 times the mass of the proposed detector for RENO 50, and so to interpret the results below for RENO 50 one needs to multiply all times by a factor of 1.1. We assume a detector resolution of $3\%/\sqrt{E(\text{MeV})}$, where E is the prompt energy which is 0.8 MeV less than the $\bar{\nu}_e$ energy. We assume a perfectly understood detector energy response and also ignore backgrounds except for the simulations reported in the last section of this paper, which include a simplified model of geoneutrinos and include various models of unknown energy response. As a result of these approximations our results are overly optimistic.

We have performed three and six year simulations using several reactor flux models with various cutoffs using fluxes arising from over 100 combinations of baselines. For each combination, we have simulated 5000 experimental runs with each hierarchy. The data analysis methods described below are then used to attempt to determine the hierarchy. We report the percentage p of experiments for which the hierarchy is determined correctly.

In most of our simulations we use the value of ΔM_{21}^2 from Ref. [11], $\sin^2(2\theta_{12})$ from Ref. [12], $|\Delta M_{32}^2|$ determined by combining ν and $\bar{\nu}$ mass differences from Ref. [4] and $\sin^2(2\theta_{13})$ from [2]

$$\Delta M_{21}^2 = 7.59 \times 10^{-5}~\text{eV}^2, \qquad \sin^2(2\theta_{12}) = 0.857$$

$$|\Delta M_{32}^2| = 2.41 \times 10^{-3}~\text{eV}^2, \qquad \sin^2(2\theta_{13}) = 0.089.$$

We have also systematically studied the effects of shifting these parameters, as will be described below. The fact that we hold these parameters fixed affects the chance of successfully determining the hierarchy. In the last section of our note we will report the results of another series of simulations in which we allowed the most relevant of these parameters to vary according to their current experimental uncertainties.

The determination of the neutrino mass hierarchy from 1–3 oscillations in the $\bar{\nu}_e$ spectrum proceeds as follows. At energies E/MeV greater than $L/12$ km, deviations from periodicity in $1/E$ are too small to be measured and the wave number k determines [7]

$$\Delta M_{\text{eff}}^2 = \cos^2(\theta_{12})|\Delta M_{31}^2| + \sin^2(\theta_{12})|\Delta M_{32}^2|. \quad (2)$$

At low energies the deviation from periodicity is large and it determines various combinations of $|\Delta M_{31}^2|$ and $|\Delta M_{32}^2|$ given in Ref. [13]; for example the energy of the 16th oscillation peak is proportional to $|\Delta M_{31}^2|$. To determine the mass hierarchy one needs to combine two distinct combinations of $|\Delta M_{31}^2|$ and $|\Delta M_{32}^2|$; thus one must combine the high and low energy parts of the spectrum. For

example, the mass hierarchy is normal (inverted) if $|\Delta M_{31}^2|$ is greater (less) than ΔM_{eff}^2.

The most studied algorithm which determines the hierarchy given a reactor $\bar{\nu}_e$ spectrum is that of Ref. [14]. One first finds the Fourier transform of the measured spectrum, as suggested in Ref. [15]. 1–3 oscillations have a wave number k of about $|\Delta M_{32}^2|$ and so the peak structure of the transform at $k \sim |\Delta M_{32}^2|$ is sensitive to these oscillations and thus the hierarchy. In Refs. [14,16] it was shown that the heights of the peaks can be combined into two real numbers RL and PV such that $RL + PV$ is positive if and only if the hierarchy is normal. In Ref. [13] two more observables were added, one mixing information from the Fourier sine and cosine transforms and one using a nonlinear Fourier transform with the same aperiodicity as the 1–3 oscillations. We will report analyses of our simulations using both $RL + PV$ and also using a neural network which finds the combination of all four observables which best determines the hierarchy.

III. CHALLENGES

An obstruction to this analysis has been described in Ref. [17], which observed that $RL + PV$ is very sensitive to the choice of model of the reactor neutrino flux and to variations of ΔM_{eff}^2 smaller than the precision of its determination by MINOS [18]. Following Ref. [19] we have reduced this spurious dependence, by employing a weighted Fourier transform in which higher energies are weighted less heavily, providing a soft cutoff. Figure 1 shows the average values of $RL + PV$ in simulated data using a normal and a weighted Fourier transform. The normal Fourier transform leads to the fluctuating dependence found in Ref. [17], but these fluctuations are much smaller with a weighted transform. A steeper weight would
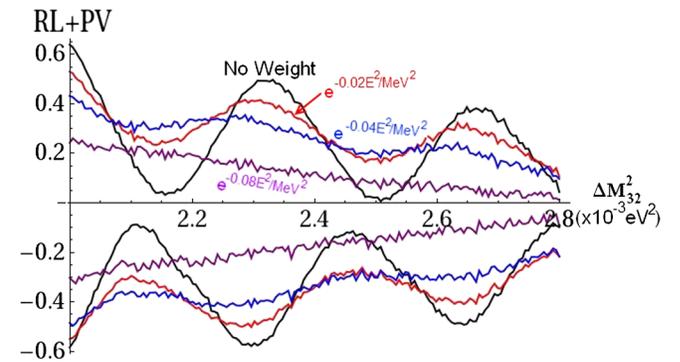

FIG. 1 (color online). Simulated average values of $RL + PV$ obtained from 100,000 neutrinos observed at a baseline of 58 km assuming the numerically interpolated reactor spectra from the 1980s [20,33,34]. The black curve uses an unweighted Fourier transform while the purple, blue and red curves respectively use weights of $\exp(-0.02E^2/\text{MeV}^2)$, $\exp(-0.04E^2/\text{MeV}^2)$ and $\exp(-0.08E^2/\text{MeV}^2)$. Notice that the first weight is so flat that the oscillations remain quite large, while the last is so strong that the $RL + PV$ signal is suppressed.
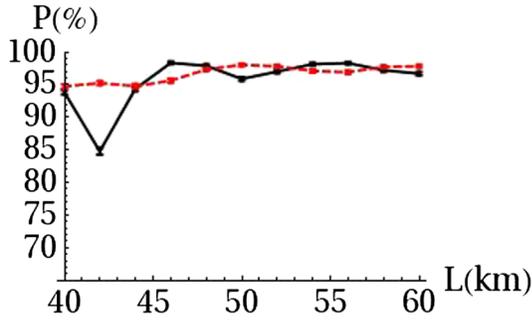
FIG. 2 (color online). The probability of success, calculated using $RL + PV$, with simulations of six years of flux arising from a reactor complex at a distance L and another complex 500 meters farther away. The error bars are the statistical errors arising from the number of simulations. The black solid and red dashed curves correspond to simulations using the quadratic fit to the reactor flux models of the 1980s [20] and the quintic fit to the recent calculation of Ref. [35].

further reduce the amplitude of the fluctuations, but would also reduce the average value of $RL + PV$, weakening the hierarchy dependence. Thus we find that the spurious dependence observed in Ref. [17] can be eliminated via a simple modification of the analysis.

Does our choice of flux model matter? Different flux models differ primarily in the high energy tail. Therefore the choice of flux models does have a large effect on the unweighted $RL + PV$, as can be seen in Fig. 2. However, once weights are included, the choice of flux model at most baselines is important at about the same level as the statistical errors arising from our finite number of experiments, as can be seen in Fig. 3. Therefore we will choose the quadratic flux model of Ref. [20] for easier comparison with previous studies, such as that of Ref. [16].

We will now discuss a more serious challenge which afflicts otherwise promising sites for such an experiment.

As a determination of the hierarchy requires a medium baseline, the neutrino flux arriving from each reactor will necessarily be quite low. This means both that the detector must be very large and also that flux from multiple reactors must be used. Multiple reactors are available, especially in places like Japan, Korea and China's Guangdong province where such experiments may be built. However multiple reactors imply multiple baselines, and so $\bar{\nu}$'s will arrive at the detector with their 1–3 oscillations out of phase. Neutrinos from different reactors are not coherent with each other, wave functions are not added, but probabilities are added and this destroys the fine structure of the spectrum whose precise measurement is essential to a determination of the hierarchy.

IV. RESULTS

We now provide the most systematic analysis of this interference effect to date. Our goal is to illustrate the effect of multiple baselines on the chance of success of the experiment and on the optimal location. We will
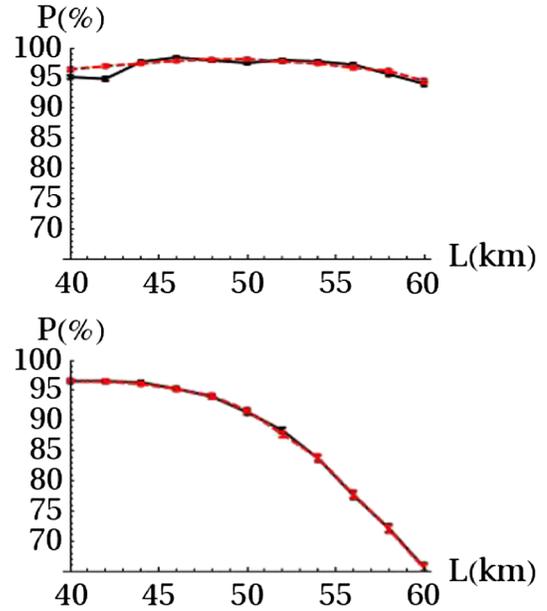

FIG. 3 (color online). As in Fig. 2, but now including a shallow weight of $e^{-0.02E^2}$ in the upper panel and a steeper weight of $e^{-0.08E^2}$ in the lower panel. Note that in both cases the model dependence of Fig. 2 reduced, suggesting that this dependence is caused by neutrinos in the high energy tail. The later spectrum reduces the dependence by more, but at a cost of hierarchy-dependent information at moderately high energies and as a result increases statistical errors arising from the limited number of events in each bin. As a result the first weight outperforms the second at the long baselines where statistical errors are important.

consider two 18 GW thermal capacity idealized point reactor neutrino sources. By defining an effective baseline difference, our results can be applied to configurations with many reactors. The reactor flux is normalized such that, including ν oscillations, at 58 km each 18 GW complex yields 25,000 $\bar{\nu}_e$ in three years.

In Figs. 4–6, we display the probability p of successfully determining the hierarchy with three and six years of live time for various combinations of baselines. The solid and dashed curves are analyses using respectively $RL + PV$ and a neural network optimizing 24 coefficients corresponding to the 4 hierarchy indicators of Ref. [13] and to 6 weights: $e^{-0.08E^2}$, $e^{-0.02E^2}$, $e^{-0.04(E-3.6)}$, $e^{-E/8}$, $e^{-0.1(E-5.25)^2}$ and $e^{-(E-3.6)^3/100}$. In all, the weighted cosine and sine Fourier transforms used for the ith indicator are

$$F_c^i(k) = \sum_j w^i(E_j) N(E_j) \cos\left(\frac{kL}{E_j}\right)$$

$$F_s^i(k) = \sum_j w^i(E_j) N(E_j) \sin\left(\frac{kL}{E_j}\right)$$

$$w^i(E) = a_1^i e^{-0.08\left(\frac{E}{\text{MeV}}\right)^2} + a_2^i e^{-0.02\left(\frac{E}{\text{MeV}}\right)^2} + a_3^i e^{-0.04\left(\frac{E}{\text{MeV}}-3.6\right)^2}$$
$$+ a_4^i e^{-\frac{E}{8\text{MeV}}} + a_5^i e^{-0.1\left(\frac{E}{\text{MeV}}-5.25\right)^2} + a_6^i e^{-0.01\left(\frac{E}{\text{MeV}}-3.6\right)^3}.$$
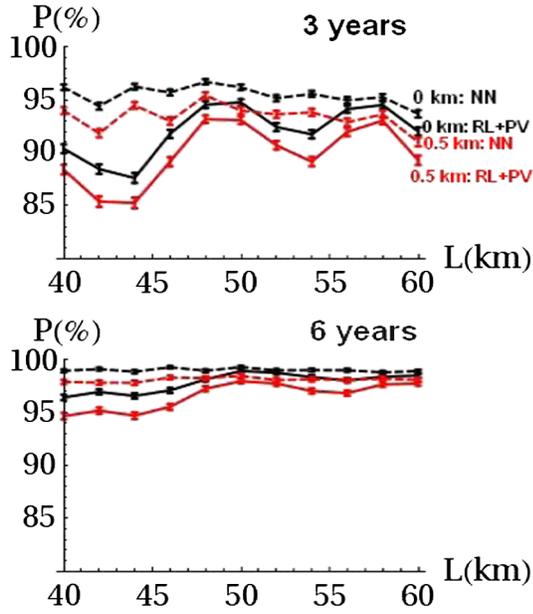
$$(3)$$

FIG. 4 (color online). Chance of success, p, after three years (first panel) and six years (second panel) of live time, with neutrinos from two 18 GW reactor complexes at distinct baselines. The horizontal axis is the shortest baseline; the color is the difference between baselines: 0 m (black) and 500 m (red). The solid curves use $RL + PV$ and the dashed curves a neural network. In each case smaller baseline differences lead to greater probabilities of success.

FIG. 6 (color online). As in Fig. 4 but now the baseline differences are 5 km (black), 10 km (red), 15 km (blue) and 20 km (purple). In each case smaller baseline differences lead to smaller probabilities of success.

As can be seen in Fig. 1, if the weights suppress the high energy tail too weakly, such as $e^{-0.02(\frac{E}{\text{MeV}})^2}$, then the spurious fluctuations persist. On the other hand if they suppress it too strongly, such as $e^{-0.08(\frac{E}{\text{MeV}})^2}$, then statistical errors are increased and so the $RL + PV$ hierarchy signal
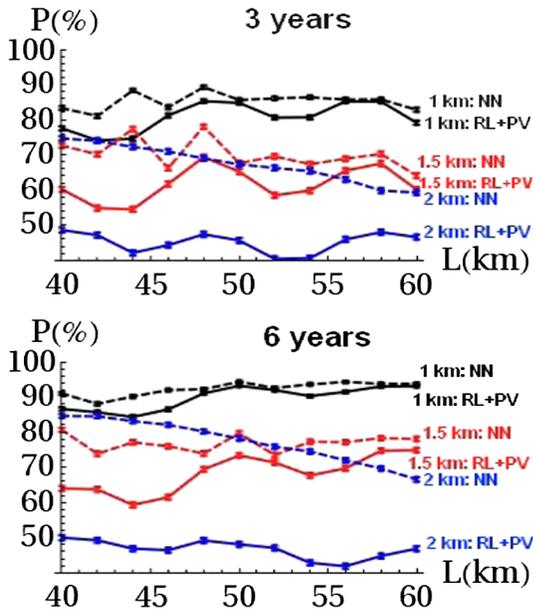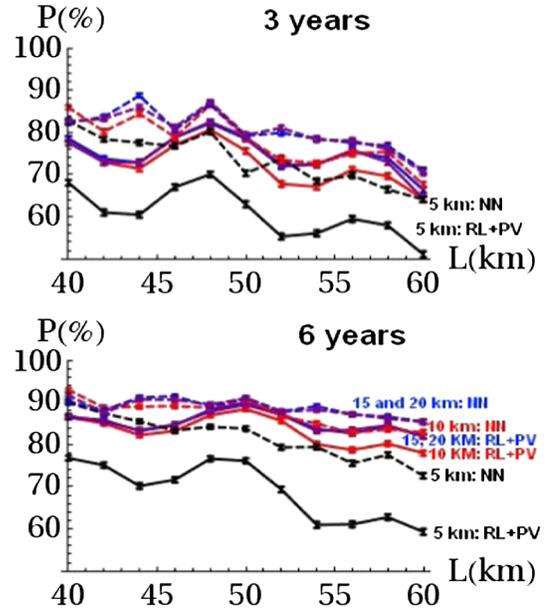
is weakened. The weights used here have been chosen to strike a balance between these two effects.

To determine the coefficient matrix a, each run of 5000 simulations/hierarchies was divided into groups. For each group, we determined the chance of success independently by using coefficients obtained by training the neural network on different groups. In order to avoid overfitting, coefficients obtained by training the neural network on a given group are never used to evaluate its chance of success. The probability of success that is reported below is the average probability of success of all of the groups. Note in particular that the optimal coefficients are distinct for each experimental run; for example the nonlinear Fourier transform is much more heavily weighted at baselines below about 54 km, where it significantly outperforms the older indicators.

For example, in Fig. 7 we illustrate the four optimal weights $w^i(E)$ in the case most relevant to Daya Bay II, corresponding to six years of exposure to one reactor complex at 52 km and another at 52.5 km. The corresponding coefficient matrix is

$$
a = \begin{pmatrix} 0.07 & 2.53 & 1.71 & -3.58 & 0.19 & 0.61 \\ -0.35 & 0.93 & 1.73 & -3.35 & -0.13 & 2.13 \\ -1.61 & 0.91 & 1.83 & -1.51 & 3.71 & 0.86 \\ 0.05 & 0.66 & -0.48 & -1.79 & 3.66 & -0.51 \end{pmatrix}.
$$

$$(4)$$

With the number of simulations that we have run, these optimal coefficient matrices are fairly stable against

FIG. 5 (color online). As in Fig. 4 but now the baseline differences are 1 km (black), 1.5 km (red) and 2 km (blue).
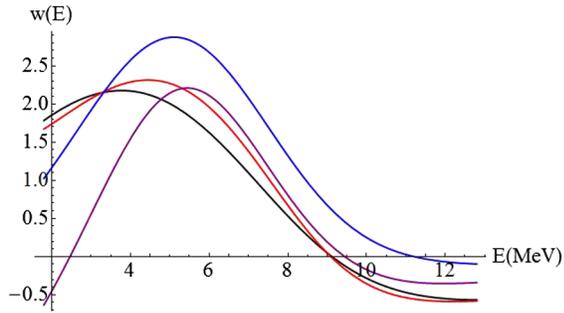
FIG. 7 (color online). In order of decreasing y-intercept we plot the weight functions $w^i(E)$ for (1) RL (black), (2) PV (red), (3) the mixed sine cosine indicator of Ref. [13] (blue) and (4) the weighted Fourier transform of Ref. [13] (purple) which optimize the chance of successfully determining the neutrino mass hierarchy in six years with one reactor complex at 52 km and another at 52.5 km.

statistical fluctuations. However they do exhibit some dependence upon the baselines and the running time of the experiment.

Figures 4 and 5 present the three and six year probabilities of success for baseline differences between 0 and 2 km. A 2 to 5 km baseline difference causes a lower chance of success. With these analysis methods, a 95% chance of a successful hierarchy determination in six years is only possible if the baseline difference is appreciably less than 1 km, about the effective difference for the BaiMianShi and Mudeungsan sites [21] for Daya Bay II and RENO 50. As the preferred Guemseong site for RENO 50 receives neutrinos from only 18 GW of reactors, the times reported here must be doubled, in addition to the factor of 1.1 discussed earlier. This is not the case for the Munmyeong or Munsusan sites.

Figure 6 illustrates that, for baseline differences of 5 to 20 km, the far detector is a background. As a result a larger baseline difference and a shorter baseline to the near reactor both increase the signal/background ratio and so p. These strong backgrounds reduce the optimal baseline.

Interference and flux from more distant reactors is a big problem for differences above 500 meters, arising when a detector is not perpendicular to a long linear array like the Daya Bay—Ling Ao complex. Sites such as BaiYunZhang are perpendicular and so enjoy identical baselines, but use flux from only a single array. In particular Daya Bay II sites such as BaiMianShi are not competitive. The old sites for Reno 50 will face similar problems, as the largest mountains are at a 45 degree angle to the reactor array. New sites such as Guemseong and Munmyeong have negligible interference, although the latter must endure reasonably large backgrounds from reactors 130 km away. Munsusan on the other hand will suffer from a baseline difference of order 700 meters.

The minimal baseline difference at potential Daya Bay II sites that use flux from multiple reactor complexes is about 500 meters, corresponding to the Dongkeng site of

Ref. [21] which uses flux from the Taishan and Yangjiang complexes. Thus we find that the best case probability of determining the hierarchy is about 98%. This result depends upon the values of the neutrino mixing parameters; in general we have found that a 1σ increase in $\sin^2(\theta_{13})$ or $|\Delta M_{21}^2|$ or a 1σ decrease in $|\Delta M_{32}^2|$ can improve the hierarchy determination by 0.1σ–0.3σ.

The *disappearance* channel at NOνA may provide a 1%–2% determination of the atmospheric effective mass of Refs. [7,22]. For $\delta \sim \pi$, this mass differs from the high energy reactor effective mass (2) by about 1.5% and from the low energy reactor effective mass $|\Delta M_{31}^2|$ by nearly 3%. As the sign of these differences depends upon the hierarchy, NOνA disappearance data can improve the hierarchy determination at Daya Bay II. However there is no such advantage if $\delta \sim 0$, as the atmospheric effective mass would be nearly equal to that of Eq. (2), which will be measured more precisely at Daya Bay II.

V. SOURCES OF ERROR

In this paper we have determined the probability that, using various Fourier transform based analyses, a medium baseline reactor neutrino experiment can successfully determine the hierarchy. This chance of success and the associated number of σ's is a somewhat pessimistic indicator of the sensitivity of the experiment because those experiments that fail typically have low values of $|RL + PV|$ due to statistical fluctuations. As we have not considered systematic errors, this situation can be cured by simply running the experiment longer. More to the point, the median experiment, defined as an experiment yielding the *median* value of $|RL + PV|$, provides the correct hierarchy with significantly more confidence than the *mean* chance of success reported here. The results in Ref. [23] suggest for example that a 98% mean chance of success corresponds to a 3.5σ confidence for the median experiment. While that study used a χ^2 analysis and not a Fourier analysis, whenever it was possible to cross-check the two approaches we have found mutually consistent results. A similar consistency was found in Ref. [17].

At the same time, the results of this study are in many ways overoptimistic. A number of sources of error have been ignored which are certain to degrade the sensitivity to the hierarchy. The most prominent among these is the unknown nonlinear energy response of the detector. The energy of a reactor neutrino is determined by the number of photoelectrons measured by the various photomultiplier tubes (PMTs). As Daya Bay II and RENO 50 will have detectors whose diameters are of order the mean free path of light inside of their liquid scintillators, the number of photoelectrons also strongly depends on the position of the event. The position will be determined roughly by the distribution of photoelectrons among the PMTs and also by the timing at which the photons are detected in various locations.

However a determination of the number of photoelectrons expected for each position of the inverse β decay event is very challenging. It will require an extensive calibration campaign, and even then it is unclear whether it will reach the subpercent precision required for a determination of the hierarchy. In addition it is complicated by other factors, such as energy lost by the geometry-dependent Cherenkov radiation from the positrons and dead time in the electronics, as well as the instability of the scintillator itself over the long time scales over which such experiments will run.

The unknown energy response of the detector may well prove to be too great of an obstacle for the hierarchy to be determined. However, even if the target 1% knowledge of the energy response is achieved, it will still introduce a large systematic error into the spectrum which will reduce the confidence in a hierarchy determination. This error can be significantly reduced in a setup with two identical detectors [24–26]. However it not only significantly impacts the chance of successfully determining the hierarchy, but also the magnitude of its effect depends on its unknown energy and position dependence. Below we will consider its affect on the chance of success in the case of a simple model of the detector's energy response.

Much of the hierarchy determining power [23,27] of these experiments will come from a comparison of ΔM_{eff}^2 with the atmospheric splitting [7] determined from the disappearance channel at accelerator experiments like NOνA. This method relies upon a 1% difference between the effective splittings. The sign of this difference determines the hierarchy. However the atmospheric splitting itself also suffers from systematic effects; for example the nuclear effects of Ref. [28] can increase the measured atmospheric splitting by 2%, enough for the normal hierarchy to mimic the inverted hierarchy.

The good news is that some of the flux uncertainties which are important for the determination of θ_{13} and the reactor anomaly [29,30] are irrelevant for the determination of the hierarchy. The determination of the hierarchy relies entirely upon the locations of small, short oscillations in the spectrum and not upon the broad features caused by the uncertainties in flux models and energy-dependent acceptance. This insensitivity is built into the Fourier approach, as broad features affect the Fourier transform only at low wave numbers, far below those where the peaks and valleys determining RL and PV are located. Indeed, this is one of the main motivations for the Fourier approach.

The main remaining sources of error are the uncertainties in the neutrino mass matrix and various backgrounds. Both of these are capable of affecting the small scale structure of the spectrum and so the determination of the hierarchy. The mass matrix elements whose uncertainties have the largest affect on the determination of the

hierarchy are θ_{13} and $|\Delta M_{32}^2|$. There are several relevant backgrounds, the most important of which are of comparable size for these experiments. We will consider just one, a simplified model of geoneutrinos employed just to see what effect it has.

For the analysis of errors we consider six year runs with two neutrino sources with baselines which differ by 0, 500 meters, 1 kilometer, 1.5 kilometers and 2 kilometers. The results of our study, displayed in Figs. 8 and 9 for the $RL + PV$ and neutral network analyses respectively, are compared with the results of Figs. 4 and 5 which correspond to the black curves in those figures. In each new simulation $\sin^2(2\theta_{13})$ and $|\Delta M_{32}^2|$ are chosen from a Gaussian distribution centered at 0.089 and 2.4 × 10^{-3} eV respectively. These Gaussians have width σ equal to 0.013 and 10^{-4} eV. The blue curves correspond to simulations in which no backgrounds are included. The red curves also include a simple model of geoneutrinos. The total number of detected geoneutrinos in each simulation follows a Gaussian distribution centered at 2520 with a width of 20%. The energy distribution of the geoneutrinos is centered at 1.8 MeV with a width of 500 keV. The finite resolution of the detector is not convolved with the geoneutrino energy distribution; instead the width is taken to model the sum of the true width and the effect of finite resolution.

The error bars shown reflect only statistical errors in the simulations, caused by the fact that for each setup we only ran 5000 simulations with each hierarchy. The blue and red curves agree to within these error bars, and so geoneutrinos contribute no noticeable effect to within the precision of our simulations. On the other hand, the effect of the uncertainty in θ_{13} and $|\Delta M_{32}^2|$ is significant. First of all, the fluctuating baseline dependence observed in Figs. 4 and 5 and then reported as the black curve in Figs. 8 and 9 is eliminated once uncertainties in the oscillation parameters are considered. In the case of the $RL + PV$ analysis in Fig. 8, the effect of this parameter variation on the central values of the probability of success is quite small. Considering the uncertainty in θ_{13} and $|\Delta M_{32}^2|$ one finds that the chance of success is reduced by of order 1% when the baseline difference is 0 or 500 meters, but is effectively unchanged when the difference is 2 km.

On the other hand, for all baseline differences one can see in Fig. 9 that the parameter uncertainty reduces the effectiveness of the neural network, although it continues to enjoy a significantly higher chance of success than $RL + PV$. This is because if the neutrino mass matrix is well known, the neural network optimizes itself to the corresponding parameters.

Once Daya Bay has completed its three year run, θ_{13} will be measured more precisely than has been considered here. Similarly, before Daya Bay II and RENO 50 begin, NOνA and MINOS+ will have measured $|\Delta M_{32}^2|$ more precisely than has been considered here. These more precise values
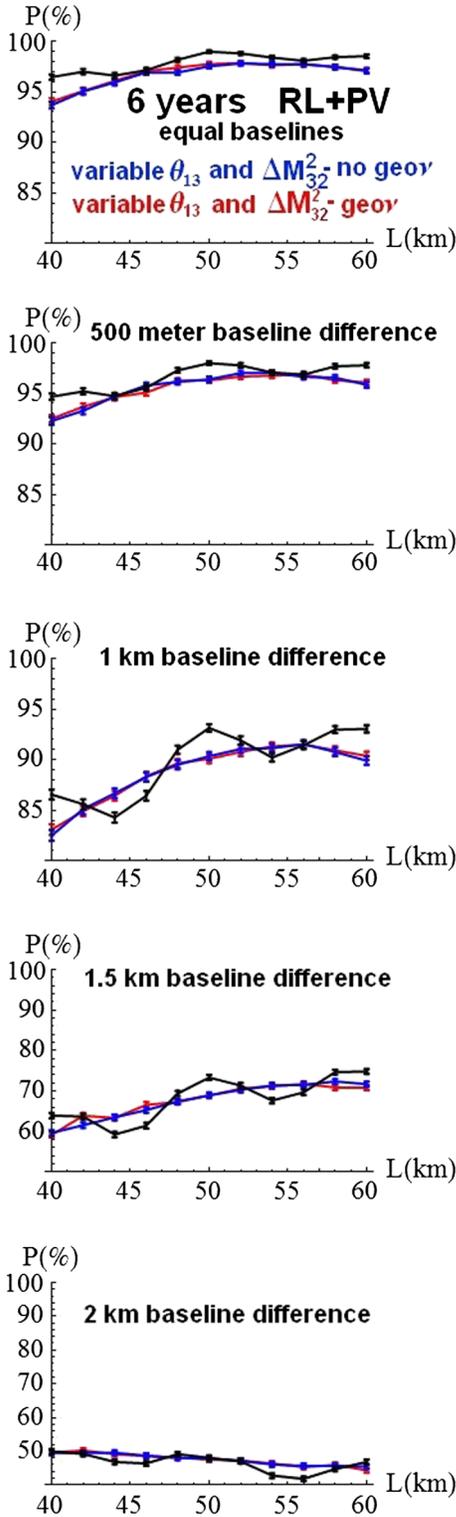
FIG. 8 (color online). The chance of success in six years as determined using $RL + PV$. The neutrinos arise from two sources whose baseline differences vary from 0 to 2 km. The black curves are the results of Figs. 4 and 5. In the blue and red curves the parameters θ_{13} and $|\Delta M_{32}^2|$ are Gaussian distributed with errors corresponding to the uncertainties to which they have been determined experimentally. The red curve also contains a simple model of a geoneutrino background.
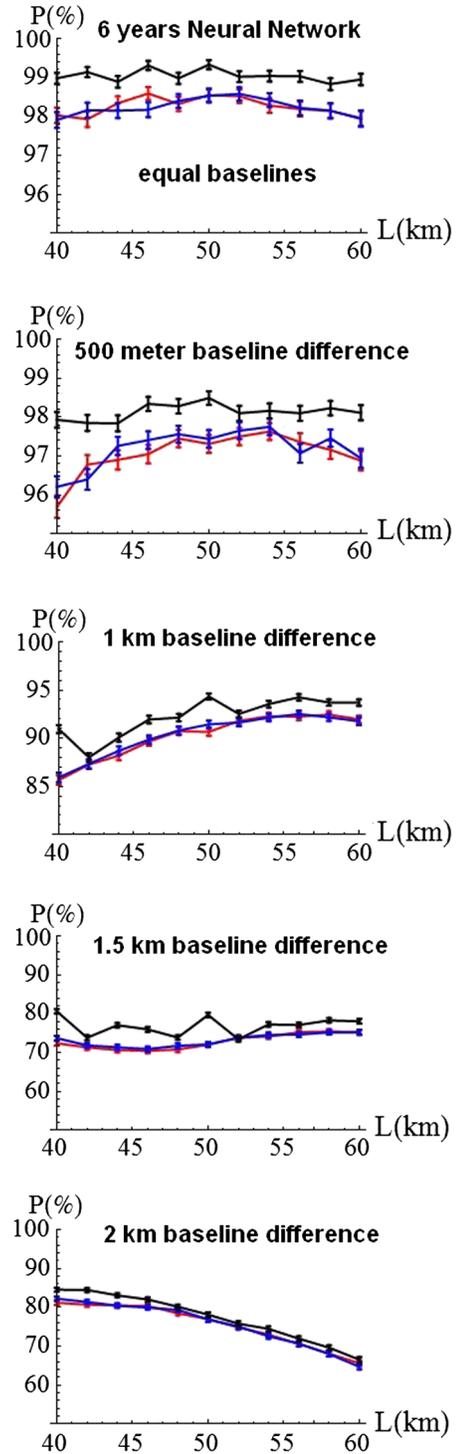

FIG. 9 (color online). As in Fig. 8 but the analysis is done using a neural network.

can be used to train the neural network, and so in this sense the final effect of the uncertainty of θ_{13} and $|\Delta M_{32}^2|$ on the chance of success will be smaller than the 0.1–0.2σ indicated in Fig. 9. Nonetheless, there are other albeit perhaps smaller sources of error which have not been considered.

There are several potentially catastrophic sources of error at such experiments. For example, if after some time the scintillator's optical properties change as occurred at the Palo Verde or Chooz experiments, due to the large size of the detector the number of photoelectrons would drop so as to make the resolution too poor to observe 1–3 oscillations. As a result a determination of the hierarchy would be impossible. Even a change in optical properties similar to that observed in the first year of running at Daya Bay may have serious consequences for the hierarchy determination.

Another concern is the coherence of the first and second neutrino mass eigenstates with respect to the third eigenstate. At a fixed energy E, during a 50 km voyage from the reactor to the detector, their relative velocities will lead to a relative separation of roughly

$$\Delta L = \frac{|\Delta M_{32}^2|}{2E^2}(50\ \text{km}) = \frac{6 \times 10^{-11} m}{(E/\text{MeV})^2}. \quad (5)$$

If ΔL is of order the length of the wave packet, the 1–3 oscillation amplitude will be reduced and if the wave packets are shorter then there will be no 1–3 oscillations [31,32].

This threshold is an order of magnitude larger than the lower bound on the coherence length that may be obtained from the observation of 1–2 oscillations at KamLAND [12], which is marginally stronger than bounds that may be obtained from the observation of 1–3 oscillations at Daya Bay [2] and RENO [3]. Therefore it appears as though current experimental bounds do not preclude a coherence length of the neutrino wave packets shorter than ΔL, leaving open the possibility that decoherence of the neutrino mass eigenstates will be observed at medium baseline reactor experiments and as a result the hierarchy will not be determined. Indeed the coherence length is unlikely to be more than an order of magnitude longer than ΔL as a result of the phase shift caused by the nucleon recoil energy in the radioactive decay in which these neutrinos are created.

Of the sources of error which are certain to be present, perhaps the most serious is the detector's unknown energy response. Define E_o to be the energy which would on average be deduced for a neutrino of energy E_t using the results of all of the calibrations and simulations available. In other words E_o is the best guess that the experimenter will be able to make for the energy of a given neutrino when analyzing the data, ignoring statistical fluctuations. Below we have considered a simple model of this response

$$E_o = \frac{2|\Delta M_{32}^2|(\text{eV}^2) + 4\epsilon \cos^2(\theta_{12})\Delta M_{21}^2(\text{eV}^2)L(\text{m})}{2|\Delta M_{32}^2|(\text{eV}^2) + \frac{\phi}{1.27}\frac{E_t(\text{MeV})}{L(\text{m})}}E_t$$

$$+ \frac{(1-2\epsilon)\frac{\phi}{1.27}\frac{E_t(\text{MeV})}{L(\text{m})}}{2|\Delta M_{32}^2|(\text{eV}^2) + \frac{\phi}{1.27}\frac{E(\text{MeV})}{L(\text{m})}}E_t \quad (6)$$

where

$$\sin(\phi) = \frac{\cos^2(\theta_{12})\sin\left(\frac{2.54\Delta M_{21}^2(\text{eV}^2)L(\text{m})}{E_t(\text{MeV})}\right)}{\sqrt{1 - \sin^2(2\theta_{12})\sin^2\left(\frac{2.54\Delta M_{21}^2(\text{eV}^2)L(\text{m})}{E_t(\text{MeV})}\right)}}. \quad (7)$$

Here ϵ is a parameter such that $\epsilon = 0$ corresponds to a perfectly understood energy response and $\epsilon = 1$ corresponds to the model which was proposed in Ref. [17] as the worst case for the determination of the hierarchy. More precisely, the case $\epsilon = 1$ corresponds to a systematic error in the energy response which causes the normal hierarchy spectrum to resemble the inverted hierarchy spectrum. The ratio of the true to the observed energy in the case $\epsilon = 1$ is drawn in Fig. 10.

We consider 120 kton years of exposure to 36 GW of thermal capacity of reactors at a single baseline of 54 km, using an $RL + PV$ analysis. While the nonlinear response considered here poses a more serious problem in the case of the normal hierarchy than the inverted hierarchy, there is a similar model which is more dangerous in the case of the inverted hierarchy. Therefore we will report the average chance of success, corresponding
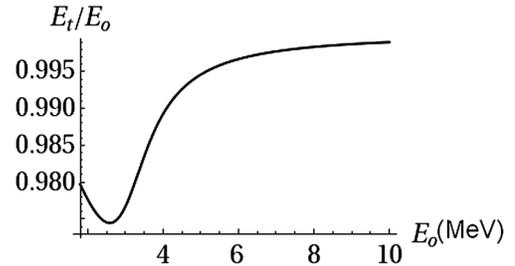

FIG. 10. The ratio of the observed to true energy corresponding to the case $\epsilon = 1$ of the unknown energy response model (6). This model corresponds to an unknown nonlinear energy response of about 2%.
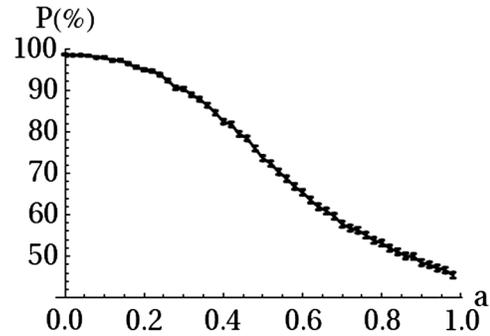

FIG. 11. The probability of correctly determining the hierarchy with 120 kton years of exposure to 36 GW of reactors at 54 km given the unknown energy response model of Eq. (6) for various values of the parameter ϵ.

to the average of the chance of correctly determining the hierarchy in the case in which the true hierarchy is normal and the chance in the case in which the true hierarchy is inverted. The results of 5000 simulations with each hierarchy and 50 values of ϵ are reported in Fig. 11.

One may observe that for ϵ greater than about $1/3$, corresponding to an unknown energy response of about 0.7%, the hierarchy can only be determined correctly in 89% of simulations using the $RL + PV$ Fourier transform based analysis. As has been anticipated in Ref. [13], this partially reflects a weakness in the Fourier analysis. Indeed, in the absence of nonlinearity the χ^2 approach of Ref. [26] obtained $\Delta\chi^2 = 18$ which according to Eq. (9) of Ref. [23] corresponds to a probability of successfully determining the hierarchy of 98.8%, similar to that obtained using the Fourier transform approach with a neural network. In the case $\epsilon = 1/3$ the χ^2 approach of Ref. [26] yielded $\Delta\chi^2 = 12$ corresponding to a probability of success of 97.6%, appreciably better than the probability of success obtained here using the Fourier transform.

However, the apparent superiority of the χ^2 approach in the presence of an unknown energy response results from the use of pull parameters to parametrize this response. In Ref. [26] three such parameters were introduced and chosen so as to minimize χ^2. We have repeated this analysis with no pull parameters and obtained $\Delta\chi^2 = 2$ for the case $\epsilon = 1/3$, which is much worse than the confidence obtained above using the Fourier transform analysis. The reason for this is that the Fourier analysis is insensitive to the overall mass scale, and so is only adversely affected by a part of the unknown energy response. One may apply an analogous pull parameter procedure to a Fourier method, to obtain a probability of success of order 97%. However by using a particular functional form of the spectrum to determine the pull parameters it would sacrifice much of the model independence which is the strength of the Fourier approach.

ACKNOWLEDGMENTS

[1] F. P. An *et al.* (DAYA-BAY Collaboration), Phys. Rev. Lett. **108**, 171803 (2012).

[2] D. Dwyer, in *Proceedings of the 25th International Conference on Neutrino Physics and Astrophysics, Kyoto, Japan, 2012*, http://neu2012.kek.jp/neu2012/programme.html.

[3] J. K. Ahn *et al.* (RENO Collaboration), Phys. Rev. Lett. **108**, 191802 (2012).

[4] R. Nichol, in *Proceedings of the 25th International Conference on Neutrino Physics and Astrophysics, Kyoto, Japan, 2012*, http://neu2012.kek.jp/neu2012/programme.html.

[5] S. K. Agarwalla, S. Prakash, and S. U. Sankar, J. High Energy Phys. 07 (2013) 131.

[6] S. T. Petcov and M. Piai, Phys. Lett. B **533**, 94 (2002); S. Choubey, S. T. Petcov, and M. Piai, Phys. Rev. D **68**, 113006 (2003).

[7] H. Nunokawa, S. J. Parke, and R. Zukanovich Funchal, Phys. Rev. D **72**, 013009 (2005).

[8] Y. Kim, in *Proceedings of the Laboratori Nazionali del Gran Sasso, Assergi, Italy, 2012*.

[9] Y. Wang, in *Proceedongs of the International Workshop on Neutrino Factories, Super Beams and Beta Beams, Williamsburg, Virginia, USA, 2012*.

[10] C. Aberle, A. Adelmann, J. Alonso, W. A. Barletta, R. Barlow, L. Bartoszek, A. Bungau, A. Calanna *et al.*, arXiv:1307.2949.

[11] K. Nakamura *et al.* (Particle Data Group Collaboration), J. Phys. G **37**, 075021 (2010).

[12] A. Gando *et al.* (KamLAND Collaboration), Phys. Rev. D **83**, 052002 (2011).

[13] E. Ciuffoli, J. Evslin, and X. Zhang, J. High Energy Phys. 03 (2013) 016.

[14] L. Zhan, Y. Wang, J. Cao, and L. Wen, Phys. Rev. D **78**, 111103 (2008).

[15] J. Learned, S. T. Dye, S. Pakvasa, and R. C. Svoboda, Phys. Rev. D **78**, 071302 (2008).

[16] L. Zhan, Y. Wang, J. Cao, and L. Wen, Phys. Rev. D **79**, 073007 (2009).

[17] X. Qian, D. A. Dwyer, R. D. McKeown, P. Vogel, W. Wang, and C. Zhang, Phys. Rev. D **87**, 033005 (2013).

[18] R. Nichol, in *Proceedings of the 25th International Conference on Neutrino Physics and Astrophysics, Kyoto, Japan, 2012*, http://neu2012.kek.jp/neu2012/programme.html; P. Adamson *et al.* (MINOS Collaboration), Phys. Rev. D **86**, 052007 (2012).

[19] E. Ciuffoli, J. Evslin, and X. Zhang, arXiv:1208.4800.

[20] P. Vogel and J. Engel, Phys. Rev. D **39**, 3378 (1989).

[21] E. Ciuffoli, J. Evslin, and X. Zhang, J. High Energy Phys. 12 (2012) 004.

[22] N. Okamura, Prog. Theor. Phys. **114**, 1045 (2005).

[23] E. Ciuffoli, J. Evslin, and X. Zhang, arXiv:1305.5150.

[24] E. Ciuffoli, J. Evslin, Z. Wang, C. Yang, X. Zhang, and W. Zhong, arXiv:1211.6818.

[25] S. Kettell, J. Ling, X. Qian, M. Yeh, C. Zhang, C.-J. Lin, K.-B. Luk, R. Johnson *et al.*, arXiv:1307.7419.

[26] E. Ciuffoli, J. Evslin, Z. Wang, C. Yang, X. Zhang, and W. Zhong, arXiv:1308.0591.

[27] Y.-F. Li, J. Cao, Y. Wang, and L. Zhan, Phys. Rev. D **88,** 013008 (2013).

[28] P. Coloma and P. Huber, arXiv:1307.1243.

[29] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, Phys. Rev. D **83,** 073006 (2011).

[30] E. Ciuffoli, J. Evslin, and H. Li, J. High Energy Phys. 12 (2012) 110.

[31] B. Kayser, Phys. Rev. D **24,** 110 (1981).

[32] E. Akhmedov, D. Hernandez, and A. Smirnov, J. High Energy Phys. 04 (2012) 052.

[33] K. Schreckenbach, G. Colvin, W. Gelletly, and F. Von Feilitzsch, Phys. Lett. **160B,** 325 (1985).

[34] A. A. Hahn, K. Schreckenbach, G. Colvin, B. Krusche, W. Gelletly, and F. Von Feilitzsch, Phys. Lett. B **218,** 365 (1989).

[35] P. Huber, Phys. Rev. C **84,** 024617 (2011); **85,** 029901(E) (2012).