PHYSICAL REVIEW D **87,** 033004 (2013)

Systematic uncertainties in long-baseline neutrino oscillations for large θ_{13}

Pilar Coloma,[1] Patrick Huber,[1] Joachim Kopp,[2] and Walter Winter[3]

[1]*Center for Neutrino Physics, Virginia Tech, Blacksburg, Virginia 24061, USA*

[2]*Fermilab, P.O. Box 500, Batavia, Illinois 60510-0500, USA and Max-Planck-Institut für Kernphysik, P.O. Box 103980, 69029 Heidelberg, Germany*

[3]*Institut für Theoretische Physik und Astrophysik, Universität Würzburg, D-97074 Würzburg, Germany*

(Received 3 October 2012; published 11 February 2013)

We study the physics potential of future long-baseline neutrino oscillation experiments at large θ_{13}, focusing especially on systematic uncertainties. We discuss superbeams, beta beams, and neutrino factories, and for the first time compare these experiments on an equal footing with respect to systematic errors. We explicitly simulate near detectors for all experiments; we use the same implementation of systematic uncertainties for all experiments; and we fully correlate the uncertainties among detectors, oscillation channels, and beam polarizations as appropriate. As our primary performance indicator, we use the achievable precision in the measurement of the *CP*-violating phase δ. We find that a neutrino factory is the only instrument that can measure δ with a precision similar to that of its quark sector counterpart. All neutrino beams operating at peak energies $\gtrsim 2$ GeV are quite robust with respect to systematic uncertainties, whereas especially beta beams and T2HK suffer from large cross-section uncertainties in the quasielastic regime, combined with their inability to measure the appearance signal cross sections at the near detector. A noteworthy exception is the combination of a $\gamma = 100$ beta beam with an SPL-based superbeam, in which all relevant cross sections can be measured in a self-consistent way. This provides a performance second only to that of the neutrino factory. For other superbeam experiments such as LBNO and the setups studied in the context of the LBNE reconfiguration effort, statistics turns out to be the bottleneck. In almost all cases, the near detector is not critical to control systematics, since the combined fit of appearance and disappearance data already constrains the impact of systematics to be small, provided that the three-active-flavor oscillation framework is valid.

DOI: 10.1103/PhysRevD.87.033004 PACS numbers: 14.60.Pq, 14.60.Lm

I. INTRODUCTION

The story of large θ_{13} has unfolded in fast succession from first hints in global fits [1–4], via direct indications from T2K [5], MINOS [6] and Double Chooz [7], to a discovery by Daya Bay [8], which was soon confirmed by RENO [9]. A recent global fit yields $\sin^2\theta_{13} = 0.023 \pm 0.0023$ [10] (see also Refs. [11,12], which find very similar values), where the error bars are entirely dominated by the reactor measurements. The precision of reactor experiments on θ_{13} will continue to improve and is not expected to be exceeded by beam experiments any time soon (see, for instance, Ref. [13]).

The most important open questions in neutrino oscillations, within the context of three active flavors, are the determination of the neutrino mass hierarchy [$\mathrm{sgn}(\Delta m^2_{31})$] and the measurement of the *CP*-violating phase δ. While there might already be some weak evidence for $\delta \sim \pi$ from global fits [11,12], high-confidence-level *CP* violation (*CPV*) and mass hierarchy measurements cannot be performed with existing facilities, such as Daya Bay, RENO, Double Chooz, T2K, and NOνA, in spite of the relatively large value of θ_{13} [14]. In the most aggressive scenario, i.e., for upgraded proton drivers for both T2K and NOνA and mutually optimized neutrino-antineutrino running plans, *CPV* could be established at a 3σ confidence level for only 25% of all values of δ. Therefore, a next generation of experiments is mandatory, and a decision towards one of the proposed technologies—superbeam upgrades, a beta beam or a neutrino factory—will soon be needed.

The determination of the mass hierarchy need not necessarily be performed in long-baseline experiments, given the relatively large value of θ_{13}. An independent determination of the mass hierarchy may be provided from the combination of T2K, NOνA and INO [15], from new proposals such as PINGU [16], from a reactor experiment with a relatively long baseline [17–19],[1] or from the combination of reactor and long-baseline experiments with very high precision [21,22]. Almost all of the long-baseline experiments studied in this work would allow for a high-confidence-level mass hierarchy discovery because of the sufficient length of the baselines and the chosen neutrino energies (see, for instance, Refs. [23–25]). Those setups with shorter baselines ($\lesssim 500$ km), where it is not possible to determine the mass hierarchy from the long-baseline data alone, would have a very massive detector. In these cases, a large sample of atmospheric neutrino events will be available, which in combination with the beam data allows for an extraction of the mass hierarchy [26–29].

[1]This may be rather challenging from the experimental point of view; see Ref. [20], for instance.

Therefore, we will not focus on this observable in this study, and no sign degeneracies will be considered either.

Regarding δ, the main focus in the literature so far has been on the question of whether *CPV* can be detected; i.e., whether the *CP*-conserving cases $\delta = 0,\ \pi$ can be excluded. The discovery of leptonic *CPV* would support thermal leptogenesis [30], which could potentially lead to an explanation of the observed baryon-antibaryon asymmetry of the Universe—although a direct connection to the *CP*-violating phases in the high-energy theory can only be established in a model-dependent way. The *CP* asymmetry in vacuum is linearly proportional to $\sin\delta$, and great efforts have been made to optimize neutrino oscillation facilities for maximal sensitivity to this term. However, there are good reasons why $\cos\delta$ is also interesting. For example, if the neutrino mass matrix is determined by a symmetry to have the tribimaximal (bimaximal) form, corrections originating from the charged lepton mass matrix may lead to the sum rule [31–33], $\theta_{12} \simeq 35°(45°) + \theta_{13}\cos\delta$. It is obvious that establishing such sum rules, which usually depend on $\cos\delta$, requires the measurement of the $\cos\delta$ term. An ideal long-baseline experiment would therefore have a relatively "flat" performance independent of δ and would be able to measure both terms with similar precision. The ultimate goal will be the measurement of δ with a precision comparable to that achieved in the quark sector.

In order to capture the whole parameter space in δ for fixed θ_{13} (or a relatively small range of θ_{13}), so-called "*CP* patterns" were proposed in Refs. [34,35] to quantify the achievable precision as a function of the true δ. In Ref. [13], this dependence was studied in detail for different types of experiments, and the main factors that affect the achievable precision were identified. From the results presented in Ref. [13], it is clear that especially narrow-band beams and setups with short baselines are typically optimized for *CPV* measurement; i.e., there is good precision in the measurement of δ around the particular values 0 and π. On the other hand, more complicated (asymmetric) patterns arise in wide-band beams or in the presence of matter effects. Therefore, one of the challenges in comparing future experiments based on their achievable precision in δ, which we will call $\Delta\delta$, is that this precision strongly depends on δ itself. To facilitate the comparison between different facilities, we will use as a performance indicator the fraction of possible values of δ for which a certain precision can be obtained, similar to earlier figures showing the *CPV* performance. We thus treat the whole parameter space on equal footing, and our conclusions on the relative sensitivity of different experiments will not depend on any assumed "true" value of δ.[2] The disadvantage is that one cannot read off from the plot at which value of δ the performance of an experiment peaks. Note, however, that the discussed setups are typically optimized for *CPV*, which means that the optimal performance is in most cases achieved close to $\delta = 0$ or π [13]. A somewhat more subtle technical point, elaborated on in Refs. [34,35], is the fact that the absolute performance can depend in a nontrivial way on the chosen confidence level. For instance, $\Delta\chi^2$ may behave in a highly non-Gaussian way far from the best-fit point, in particular if the mass hierarchy degeneracy cannot be resolved. Here we assume that we are in the Gaussian limit, and sign degeneracies have not been considered. However, we note that the fraction of δ values for which a certain sensitivity is reached is also useful in the non-Gaussian case.

The key issue for long-baseline experiments at large θ_{13} is systematics. It is well known that signal normalization uncertainties especially affect neutrino oscillation measurements for large θ_{13}; see, e.g., Refs. [23,36,37]. While in phenomenological studies, near detectors are only in rare cases explicitly included or discussed (see, e.g., Ref. [36] for T2HK and Ref. [38] for the neutrino factory), it is usually assumed that these can be described by an effective systematic error in the far detector in the range from 1% to 10%. The chosen values are "educated guesses" in the absence of explicit near-detector simulations. This is unsatisfactory given the large impact of systematic uncertainties at large θ_{13}. Indeed, it is not even sufficient to use realistic numbers for the systematic errors, but it is equally important to implement them in an appropriate way, in particular taking into account correlations between the errors affecting different oscillation channels, different parts of the energy spectrum, etc. For instance, most conventional simulations assume that systematics are uncorrelated among all oscillation channels, but fully correlated among all energy bins and backgrounds. In the real world, cross sections are correlated among all channels measuring the same final flavor, fluxes among all channels in the same beam, etc. Furthermore, it is known that the matter density uncertainty affects the measurements for large θ_{13} for experiments with long baselines and high energies; see, e.g., Refs. [39,40].

In this study, we will explore the effect of systematic errors on the achievable precision in different experiments, and we will provide a detailed comparison between different setups under the same assumptions for the systematics. Our systematics treatment is an extension of the one used for multidetector reactor experiment simulations [41–43]. In particular,

(1) We use a detailed, physics-based and self-consistent systematics implementation including correlations, which is comparable for all experiments.
(2) We explicitly simulate the near detectors, with comparable assumptions regarding statistics and geometry for all experiments.

[2]It should be kept in mind, however, that the experiment which reaches the best overall precision in δ may not yield the best *CPV* discovery potential (and vice versa), since the latter depends on the achievable precision around the specific values $\delta = 0,\ \pi$.

(3) We do not only choose particular values for the systematic errors; we also study ranges which span the gamut from conservative to optimistic.
(4) We use exactly the same assumptions for cross-section and matter density uncertainties for all experiments. For the systematic uncertainties that depend on the particular type of neutrino beam (for instance, flux uncertainties or intrinsic beam backgrounds) we consider the same values for all experiments of the same type.

For the experiment definitions and simulations, we have modified the AEDL (Abstract Experiment Definition Language) of the GLoBES software [44,45], which allows now for a flexible systematics implementation entirely in AEDL (without the need to write C code).[3]

The paper is organized as follows: In Sec. II, the experiment simulation is summarized, while details, including a comparison between old (using an effective description of the errors in the far detector) and new systematics implementation, can be found in the supplementary material available for this paper [46]. A comparison of the performance of all setups is presented in Sec. III, where the dependence on exposure is also discussed. In Sec. IV, we identify for each experiment the relevant performance bottlenecks and provide guidance on which quantities should be optimized in each case. Finally, we summarize and conclude in Sec. V.

II. SIMULATION TECHNIQUES AND SYSTEMATICS TREATMENT

A. Experimental setups

Table I summarizes the main features of the setups studied in this work. More details are given in the supplementary material [46]. We have chosen four representative *benchmark* setups for long-baseline neutrino oscillation experiments:

Beta beam: A high-γ ($\gamma = 350$) beta beam [47,48] has been chosen, since it provides a very good *CPV* discovery potential, even comparable to the one obtained at a neutrino factory; see, e.g., Ref. [49]. The relatively long baseline ($L = 650$ km) is enough to guarantee a measurement of the mass hierarchy given the large value of θ_{13} (see, for instance, Ref. [24]). The beam is aimed at a 500 kt Water Čerenkov (WC) detector. This setup will be referred to as BB350.

Neutrino factory: We consider a low-energy version of the neutrino factory, with a parent muon energy of 10 GeV and a Magnetized Iron Neutrino Detector (MIND) detector placed at a baseline of 2000 km [50]. This is the setup currently under consideration within the International Design Study for a Neutrino Factory (IDS-NF) [51]. It will be referred to as NF10 hereafter.

Off-axis conventional neutrino beam: Here we follow the T2HK proposal, given its high relevance in the literature and the fact that a Letter of Intent (LoI) has already been submitted [28]. The experiment uses a WC detector with a fiducial mass of 560 kt, placed at a distance of 295 km from the source.

On-axis conventional neutrino beam: We study a setup with a relatively high-energy flux (taken from Ref. [52]) and a 100 kt Liquid Argon (LAr) detector at 2300 km from the source. This corresponds to one of the configurations under consideration within LAGUNA [53] and LAGUNA-LBNO [54]. We have checked that the Fermilab-to-DUSEL Long-Baseline Neutrino Experiment (LBNE), with a 34 kt LAr detector at a baseline of 1290 km [55], has a very similar performance. We will therefore refer to this setup as WBB in the rest of this paper, since the conclusions extracted from its performance would be generally applicable to both LBNO and LBNE.[4]

In addition to these setups, we will also consider four alternative setups with high relevance in the literature. In particular, we will discuss two out of the three options considered during the LBNE reconfiguration process: a new Fermilab-based beam line aimed at a 10 kt LAr surface detector placed at Homestake ($\mathsf{LBNE}_{\mathsf{mini}}$), and the existing NuMI beam with a new 30 kt LAr surface detector placed at the Ash River site ($\mathsf{NO}\nu\mathsf{A}^{+}$).[5] Moreover, we consider a lower-energy version of the neutrino factory (NF5), with a muon energy of 5 GeV and a baseline of 1300 km. (The detector technology in this case is still a MIND in order to make a direct comparison to the NF10 setup easier.) The fourth alternative setup discussed in this paper is a combination of a low-γ ($\gamma = 100$) beta beam with the SPL [27], labeled as BB + SPL. We use this setup to study whether a combination of different channels (in this case, *CPT* conjugates) can reduce the impact of systematic errors. Further details on the experiment implementation as well as the analysis techniques can be found in Secs. 1 and 2 of the supplementary material [46].

Finally, sometimes we will compare the results for the setups listed in Table I to the results that would be obtained by 2020 from the combination of the present facilities; that is, T2K, $\mathsf{NO}\nu\mathsf{A}$ and reactors. In order to do so, we assume that T2K and $\mathsf{NO}\nu\mathsf{A}$ will have run for five and four years per polarity by that date, respectively, and that the precision on the measurement of θ_{13} will be

[3]This is one of the key modifications of the software which is expected to be included in the GLoBES 4.0 release.

[4]We find a slightly worse performance for the LBNE setup, though.

[5]The third option considered within the LBNE reconfiguration process consists of a 15 kt LAr underground detector placed at the MINOS site in Soudan. However, we have checked that the performance of this setup is much inferior to that of all other setups discussed here, and therefore we do not consider it any further.

TABLE I. Main features of the setups considered in this work. From left to right, the columns list the names of the setups, the approximate peak energy of the neutrino flux, the baseline, off-axis angle, detector technology, fiducial detector mass, beam power (for the conventional and superbeam setups), the useful parent decays per year (ion decays for the beta beams and muon decays for the neutrino factories), and the running time in years for each polarity. Note that the neutrino and antineutrino running is simultaneous for the neutrino factory setups NF10 and NF5. (The μ^+ and μ^- circulate in different directions within the ring.) For beta beams, the number of useful ion decays is different for the two polarities, so we quote the number of useful ^{18}Ne (^{6}He) decays per year separately.

	Setup	E_ν^{peak}	L	OA	Detector	kt	MW	Decays/yr	$(t_\nu, t_{\bar{\nu}})$
Benchmark	BB350	1.2	650	···	WC	500	···	$1.1(2.8) \times 10^{18}$	(5,5)
	NF10	5.0	2000	···	MIND	100	···	7×10^{20}	(10,10)
	WBB	4.5	2300	···	LAr	100	0.8	···	(5,5)
	T2HK	0.6	295	2.5°	WC	560	1.66	···	(1.5,3.5)
Alternative	BB100	0.3	130	···	WC	500	···	$1.1(2.8) \times 10^{18}$	(5,5)
	+SPL			···			4	···	(2,8)
	NF5	2.5	1290	···	MIND	100	···	7×10^{20}	(10,10)
	$\text{LBNE}_{\text{mini}}$	4.0	1290	···	LAr	10	0.7	···	(5,5)
	NOνA$^+$	2.0	810	0.8°	LAr	30	0.7	···	(5,5)
2020	T2K	0.6	295	2.5°	WC	22.5	0.75	···	(5,5)
	NOνA	2.0	810	0.8°	TASD	15	0.7	···	(4,4)

dominated by the systematic error reachable at Daya Bay. In order to simulate this combination, we have followed Ref. [14].

B. Systematics

In this study, we treat systematic uncertainties for all experiments in the same framework. We implement beam flux uncertainties, fiducial mass uncertainties, cross-section uncertainties, and the matter density uncertainty in the same way for all setups, whereas background uncertainties can only be the same within each class of experiments—superbeams, beta beams, and neutrino factories. For each systematic error, we consider default, optimistic and conservative values; see Table II. Note that the error estimates given in this table do *not* include the impact of the near detectors, which instead are explicitly simulated in our study. For details on how systematics is implemented and these values are obtained, see Secs. 2 and 3 of the supplementary material [46]. For example, cross-section (times efficiency) errors will affect all channels in near and far detectors sensitive to the same cross sections; i.e., they are fully correlated among these. As we shall demonstrate, these systematics correlations, as part of the implementation, are more important than the actually used values. In the supplementary material [46] (see Sec. 4), we also compare the new systematics treatment including near detectors with the older effective treatment in terms of signal and background normalization errors. We have found a relatively good agreement, except for T2HK and BB350, for which the near-far extrapolation depends strongly on the poorly known ratio of ν_e and ν_μ cross sections in the quasielastic scattering (QE) regime. Therefore, the performance of these experiments strongly depends on the systematics assumptions.

C. Oscillation parameters and observables

The following input values for the oscillation parameters, in agreement with the allowed ranges at 1σ from global fits [10–12], are used for all simulations in this paper:

$$\Delta m^2_{21} = 7.64 \times 10^{-5}\ \text{eV}^2, \qquad \Delta m^2_{31} = 2.45 \times 10^{-3}\ \text{eV}^2;$$
$$\theta_{21} = 34.2°, \qquad \theta_{23} = 45°, \qquad \theta_{13} = 9.2°.$$

Unless otherwise stated, a normal mass hierarchy is assumed. In our fits, we include Gaussian priors with a 1σ width of 3% for the solar parameters, 8% for θ_{23} and 4% for Δm^2_{31}. No external priors were used for θ_{13} and δ. We have checked that adding a prior on θ_{13} corresponding to the expected precision of the final Daya Bay measurement [8] does not affect our results, except for a mild improvement in the measurement of δ around $\pm 90°$ for some facilities where the intrinsic degeneracy is still present.

Our results will be presented in the following terms:

Achievable precision on δ, $\Delta\delta$*:* We will show the fraction of possible values of δ for which a given precision at 1σ (1 d.o.f.) is achieved.

CPV discovery potential: This is defined as the ability of a given facility to rule out the *CP* conservation hypothesis (i.e., $\delta = 0, \pi$) at a given confidence level. Our results for this observable will always be shown at 3σ (1 d.o.f.).

III. PERFORMANCE COMPARISON

The nominal performance of all setups listed in Table I is compared in Fig. 1, using the default values for the systematic errors according to Table II. In the left panel, the fraction of δ values for which a given precision in δ can be achieved is shown. Considering only the benchmark setups BB350, NF10, WBB, and T2HK, it can be seen that

TABLE II. The systematic errors considered in our analysis for superbeams (SB), beta beams (BB), and neutrino factories (NF). Numerical values are shown for optimistic, default, and conservative assumptions. All numbers are based on external input and do not yet include any information from the near detector. Different fiducial volume uncertainties are implemented for the far detector (FD) and near detector (ND), as indicated. Note that the background uncertainties listed here affect only detector-related backgrounds due to neutral current events, charge or flavor misidentification backgrounds, whereas the uncertainties related to the intrinsic beam background for superbeams are treated as flux errors. Cross section uncertainties are treated independently in the quasielastic scattering (QE), resonance production (RES) and deep-inelastic scattering (DIS) regimes. The uncertainty due to different detection efficiencies for the different lepton flavors has also been added in quadrature, and therefore only the error on the effective ratio is shown. Blank spaces indicate the cases when ν_e and ν_μ cross sections are allowed to vary in a completely independent way.

	SB			BB			NF		
Systematics	Optimistic	Default	Conservative	Optimistic	Default	Conservative	Optimistic	Default	Conservative
Fiducial volume ND	0.2%	0.5%	1%	0.2%	0.5%	1%	0.2%	0.5%	1%
Fiducial volume FD (including near-far extrap.)	1%	2.5%	5%	1%	2.5%	5%	1%	2.5%	5%
Flux error signal ν	5%	7.5%	10%	1%	2%	2.5%	0.1%	0.5%	1%
Flux error background ν	10%	15%	20%		correlated			correlated	
Flux error signal $\bar{\nu}$	10%	15%	20%	1%	2%	2.5%	0.1%	0.5%	1%
Flux error background $\bar{\nu}$	20%	30%	40%		correlated			correlated	
Background uncertainty	5%	7.5%	10%	5%	7.5%	10%	10%	15%	20%
Cross secs. $\times$ eff. QE	10%	15%	20%	10%	15%	20%	10%	15%	20%
Cross secs. $\times$ eff. RES	10%	15%	20%	10%	15%	20%	10%	15%	20%
Cross secs. $\times$ eff. DIS	5%	7.5%	10%	5%	7.5%	10%	5%	7.5%	10%
Effective ratio ν_e/ν_μ QE	3.5%	11%	...	3.5%	11%	...	...	...	...
Effective ratio ν_e/ν_μ RES	2.7%	5.4%	...	2.7%	5.4%	...	...	...	...
Effective ratio ν_e/ν_μ DIS	2.5%	5.1%	...	2.5%	5.1%	...	...	...	...
Matter density	1%	2%	5%	1%	2%	5%	1%	2%	5%

the neutrino factory outperforms all other options by a factor of 2. It is the only experiment which can achieve a precision comparable to that obtained in the quark sector, where the CP phase is determined to be $\gamma = 70.4^{+4.3}_{-4.4}{}^\circ$ [56], depicted by the vertical gray band in the left panel. We also show in this figure, in addition to the setups listed in Table I, the results that would be obtained by the year 2020 from the combination of T2K, NOνA and reactors.

We would like to point out the remarkable performance of BB + SPL, which outperforms any of the other superbeam or beta-beam options. As we will discuss in Sec. IV, the reason for this is the reduction in systematic errors related to the cross sections. For the other alternative setups, which can be regarded as smaller versions of the respective original proposals, the precision varies strongly as a function of the true δ. This is due to the fact that intrinsic degeneracies appear around $\delta = \pm 90^\circ$ superimposed to the true solutions, effectively worsening the observable precision on δ (see also Ref. [13]). In the particular case of the NF5, this is due to the coarse energy binning that we have used for our simulations, which is fixed by the available migration matrices.[6]

It is interesting to compare the precision on δ in the left panel of Fig. 1 with the CPV discovery potential in the right panel, which mainly corresponds to the precision at specific values of δ (0 and π). Again, neutrino factories emerge as the optimal setups, being able to observe CPV at 3σ for $\sim$85% of the values of δ. Compared to the left panel, NF10 and NF5 perform very similarly, which is expected from optimization studies [50]. However, this optimization does not include the full parameter space, which is much better covered by NF10 (left panel).[7] The performance of almost all setups is reduced for larger $\sin^2 2\theta_{13}$. The exceptions are NOνA$^+$ and WBB, which benefit from the increased statistics for larger values of θ_{13}. Note that the 3σ CPV discovery potential of T2HK is comparable to that of NOνA$^+$, whereas T2HK is clearly a better precision instrument (left panel). As far as the comparison between LBNE$_{\text{mini}}$ and NOνA$^+$ is concerned, both perform similarly in the left panel, with NOνA$^+$ exhibiting a stronger dependence on δ, as explained above. For CPV discovery (right panel), however, LBNE$_{\text{mini}}$ cannot compete at all, because it does not reach 3σ for any value of δ. Finally, it should be noticed that the combination of T2K, NOνA and reactors (the 2020 line) would indeed be able to observe CPV at 3σ for some values of δ, although the fraction of δ values for which this is possible remains well below 10% (see also Ref. [14]).

[6]We have checked that, if the bin size is reduced by a factor of 2, the dependence on δ is largely reduced, since the intrinsic degeneracies are better resolved.

[7]In fact, the meeting point of the two curves in the left panel roughly corresponds to the values of δ which are relevant for a good CPV discovery potential, and therefore the CPV sensitivity is the same.

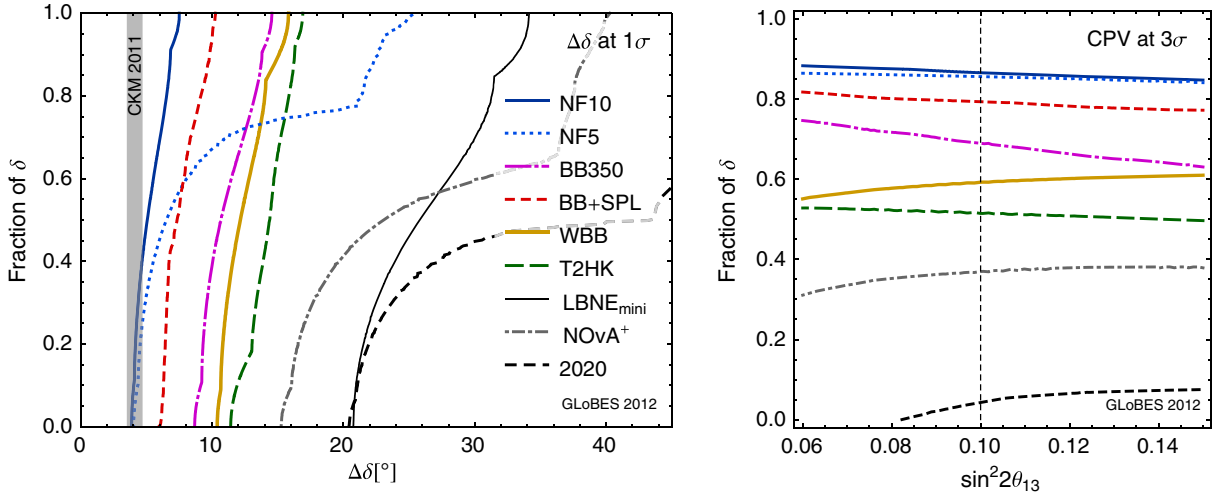

FIG. 1 (color online). Comparison between the different setups from Table I for the default systematic errors listed in Table II (including near detectors). We have also included in the comparison the results that would be obtained by 2020 through the combination of T2K, NOνA and reactors. Left panel: Fraction of δ as a function of the precision at 1σ for $\sin^2 2\theta_{13} = 0.1$. Right panel: Fraction of δ for which *CPV* can be established at 3σ as a function of $\sin^2 2\theta_{13}$ in the currently allowed range. A true normal hierarchy has been assumed, and no sign degeneracies have been accounted for. In the right panel, $\text{LBNE}_{\text{mini}}$ is not shown, because it does not reach 3σ sensitivity to *CPV*. The vertical dotted line in the right panel corresponds to $\sin^2 2\theta_{13} = 0.1$, which is the true value chosen for θ_{13} in the left panel. In the left panel, the vertical gray band depicts the current precision for the *CPV* phase in the quark sector, taken from Ref. [56].

The comparison between different setups depends not only on systematic uncertainties but also on exposure. We therefore show in Fig. 2 the exposure dependence of the precision on δ for all setups in Table I. Here, exposure ($\mathcal{E}$) is defined as the beam intensity (protons on target or useful parent decays) $\times$ running time $\times$ detector mass, and only the relative exposure compared to the nominal values $\mathcal{E}_0$ from Table I is shown. The left panel of Fig. 2 shows results for the benchmark setups, while the right panel displays results for the alternative setups. The bands reflect the variation of performance between the optimistic and conservative choices for the systematics uncertainties (see Table II).

Figure 2 reveals several interesting features. First, for some experiments the gradient at the nominal exposure is significantly larger than for others. In particular, WBB, $\text{LBNE}_{\text{mini}}$, and NO$\nu$A$^+$ operate in the statistics-limited regime ($\Delta\delta \propto 1/\sqrt{\mathcal{E}}$), where the systematics contribution is small. This makes the exposure the most relevant performance bottleneck for them (see also Sec. IV). Comparing $\text{LBNE}_{\text{mini}}$ with NOνA$^+$, NOνA$^+$ clearly exhibits a larger dependence on systematics. This dependence increases with exposure. In most other cases, when optimistic values are chosen for the systematics (lower edges of the bands), the scaling with exposure seems to be dominated by statistics ($\Delta\delta \propto 1/\sqrt{\mathcal{E}}$); while for conservative values (upper edges), the setups start to be more dominated by systematics, and the curves are less steep. In these cases, the difference between optimistic and conservative systematics increases significantly with exposure. An interesting exception is BB + SPL, for which systematics are equally important regardless of the exposure. In this case, the dominant systematics (cross-section ratios) are reduced by the combination of the two beams, so that even under conservative assumptions for the systematic uncertainties, the performance of BB + SPL is still dominated by statistics.

IV. PERFORMANCE BOTTLENECKS AND THE ROLE OF THE NEAR DETECTORS

Here we discuss the most important limiting factors for the performance of each individual experiment; i.e., the key factors to be watched in the design and optimization of the experiment. As we will demonstrate, there is typically one dominant performance bottleneck, which is, however, different from experiment to experiment. We study the impact of

(1) Systematic errors, including possible correlations.
(2) Exposure.
(3) The near detector.

In order to identify the key systematic errors, we start by taking all of them at their default values and then switching off each group of systematic errors (flux errors, cross-section uncertainties, etc.) independently. This method reveals the systematic uncertainties that have the greatest impact, and which uncertainties are irrelevant for the measurement of δ.[8]

[8]We have also checked that the inverse procedure, i.e., starting from the statistics-only limit and switching on each group of systematic errors independently, leads to similar conclusions. However, this second procedure is less intuitive, and therefore we will not show its results in the following.

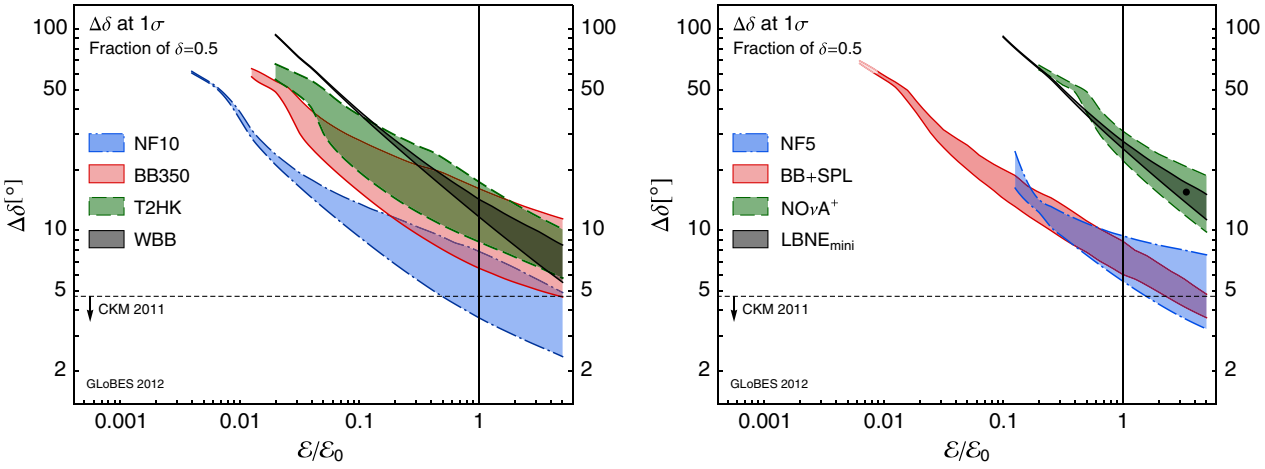

FIG. 2 (color online). Error on δ (at 1σ, for $\sin^2 2\theta_{13} = 0.1$) as a function of exposure, where the bands reflect the variation in the results due to different assumptions for the systematic errors between the optimistic (lower edges) and conservative (upper edges) values in Table II. In the left panel, the results for the benchmark setups from Table I are shown, while the right panel shows the results for the alternative setups. The nominal exposure $\mathcal{E}_0$, to which the exposure $\mathcal{E}$ on the horizontal axis is normalized, is the one given in Table I. Here, near detectors are included, and the results are shown for the median values of δ (where the fraction of δ is 50%). The current precision on the CP phase in the CKM matrix V_{CKM} is also indicated. In the right panel, the black dot indicates the luminosity for the original LBNE configuration (34 kt LAr detector [55]).

The impact of switching off groups of systematic uncertainties for the different experiments is shown in Fig.s 3 and 4 for the benchmark and alternative setups, respectively, and for the default values of the systematics listed in Table II. In both figures, the upper colored bars show how much the precision in δ would improve for each experiment if a given systematic error were switched off. (Only those groups of systematic errors that actually have a sizeable impact for each facility are shown.) For each experiment, the precision that would be reached in the statistics-only limit is also shown (''all off''). As mentioned above, the different bars do not typically add up to the ''all off'' one, nor does the dominant systematics alone account for it. The reason is that correlations among systematic uncertainties and between systematics and oscillation parameters, i.e., the difference between the ''all off'' case and the other bars, can be interpreted as the importance of these correlations. The impact of doubling the exposure (see also Fig. 2), as well as the performance loss each experiment would suffer if no near detector were available, is also shown for each experiment.

Note that the edges of the bars shown in Figs. 3 and 4 correspond to the medians of the corresponding $\Delta\delta$ distributions; i.e., for 50% of all δ values, the precision will be better than the $\Delta\delta$ value shown in the figures, and for the other 50% it will be worse. The δ values corresponding to the left and right edges of any given bar need not be the same, since the median may change from one edge of the bar to the other. Small differences in the results would appear if instead we chose a fixed value of δ for all bars corresponding to the same experiment, or if we sampled the $\Delta\delta$ distributions not at their median, but at a fraction of δ other than 50%. Nevertheless, we expect our general conclusions to remain unchanged.

Let us first discuss the impact that different systematic uncertainties have on the benchmark setups (Fig. 3). For NF10, the most important systematic uncertainty is the one on the matter density, as has been established earlier [39]. Improving the flux error or the understanding of the deep-inelastic scattering (DIS) cross sections marginally helps, with the relative importance of these two depending slightly on the value of δ. For BB350, the sensitivity is limited mostly by the errors on the QE and resonance production (RES) effective cross-section ratios (the event rate is substantial in both regimes). Correlations among these turn out to be important because they lead to an effective shape error. For WBB, the impact of systematics is generally small, while the sensitivity is mainly limited by the exposure (see also Fig. 2). It should be noted here that WBB has been simulated with a LAr detector, which is the least studied among the detector technologies considered here. For instance, it is the only detector for which no tabulated detector response functions (''migration matrices'') from detailed Monte Carlo simulations are available for the signal reconstruction up to now. For T2HK, the impact of systematic uncertainties (in particular on the QE cross-section ratio and the intrinsic beam background) is generally large. Exposure is also important, but it is not the dominant limitation.

An interesting question is how much the near detectors actually help in the precision measurement of δ. We therefore show in Fig. 3 how the predicted δ precision changes when the near detector is not included in the analysis (''no ND''). A somewhat surprising result is that omitting the near detector affects the achievable precision of none of

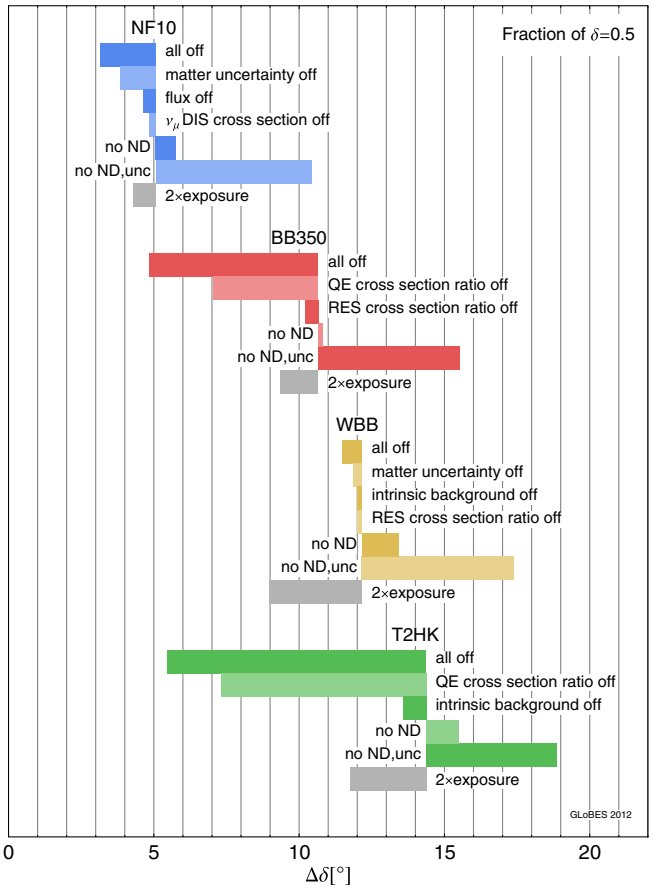

FIG. 3 (color online). Dependence of the achievable precision in δ (at 1σ, for $\sin^2 2\theta_{13} = 0.1$) for the benchmark setups in Table I on systematic uncertainties, exposure, and near detectors. The bars show the improvement in the precision of δ compared to the default scenario if the dominant systematic errors are switched off separately. Here "all off" refers to the statistics-only limit; "matter uncertainty off" to no matter density uncertainty; "flux off" to no flux errors; "DIS ν_μ cross section off" to no DIS effective cross-section errors for neutrinos and antineutrinos; "cross section ratio off" to fully correlated effective cross-section errors for ν_e and ν_μ, and for $\bar{\nu}_e$ and $\bar{\nu}_\mu$; and "intrinsic background off" to no uncertainty on the intrinsic beam backgrounds. The effect of doubling the exposure is also shown, as well as two sets of results without a near detector: for "no ND," systematic uncertainties are still correlated between oscillation channels at the far detector; while for "no ND, unc," correlations between appearance and disappearance channels are not included. The $\Delta\delta$ values shown here correspond to the median value of δ (i.e., for 50% of δ values, the precision would be better; for the other 50% it would be worse).

the setups considered here by more than 1–2 degrees. Also, in none of the cases is the near detector the most critical factor. The main reason is that, even without a near detector, most systematic errors are correlated among the different oscillation channels, so that the nuisance parameters are constrained by the requirement of self-consistency among the different far-detector channels (in particular, appearance and disappearance). Note that this self-consistency requirement relies entirely on the validity of the three-active-flavor oscillation framework. Thus, in the absence of a near detector, it is doubtful that meaningful bounds on physics beyond the Standard Model (such as sterile neutrinos or nonstandard interactions [57,58]) could be obtained.

In order to illustrate the importance of correlations between appearance and disappearance data, we also show in Fig. 3 results for the case where the near detector is omitted and, in addition, correlations between appearance and disappearance channels in the far detector are not included; i.e., the appearance and disappearance data sets are assigned independent systematic errors ("no ND, unc"). In this case, as expected, the near detector plays a crucial role for all setups. The difference between these bars and the bars labeled "no ND" shows explicitly the importance of correlations and how disappearance data can be used to constrain systematic errors in the appearance sample in a very efficient way. The effect of the disappearance channel is particularly relevant for the BB350 setup, for instance. In this case, since the ν_e disappearance is small, the far detector is particularly useful in constraining systematic effects related to flux uncertainties and ν_e cross-section measurements. Therefore, the near detector does not provide any additional information and can be removed from the analysis with practically no impact on the precision. In addition, for BB350 (T2HK), the near detectors do not provide a measurement of the $\overset{(-)}{\nu}_\mu$ ($\overset{(-)}{\nu}_e$) cross sections needed for the far-detector appearance measurement, but only of the flavor-conjugate cross section. (In T2HK, the near detector is in principle sensitive also to the $\overset{(-)}{\nu}_e$ cross sections due to the intrinsic beam backgrounds, but the statistics in these channels is too small to allow for a precise cross-section measurement.) The experiment which benefits most from the near detector is WBB, where having the near detector is in principle more important than improving any of the systematic errors. The reason for this is that this setup is statistically limited, and is therefore not able to constrain both the nuisance parameters and the atmospheric oscillation parameters independently in the analysis. Thus, increasing the nominal exposure would be much more critical in this case.

Figure 4 shows how the performance of the alternative setups in Table I depends on systematics, exposure, and the near detectors. We notice that for NF5, the relative impact of the systematic errors (including the matter density uncertainty) is smaller, whereas exposure is somewhat more important than for NF10. In addition, NF5 is the only experiment for which the near detector is more important than systematics or exposure. For $\text{LBNE}_{\text{mini}}$ and $\text{NO}\nu\text{A}^+$, the near detector also has a larger effect than the systematic uncertainties, but the main limitation for these statistics-dominated experiments is exposure. In fact, an increase in statistics may also render the near detector unnecessary because, similar to WBB discussed above, $\text{LBNE}_{\text{mini}}$ and $\text{NO}\nu\text{A}^+$ need the near detector mainly because they are

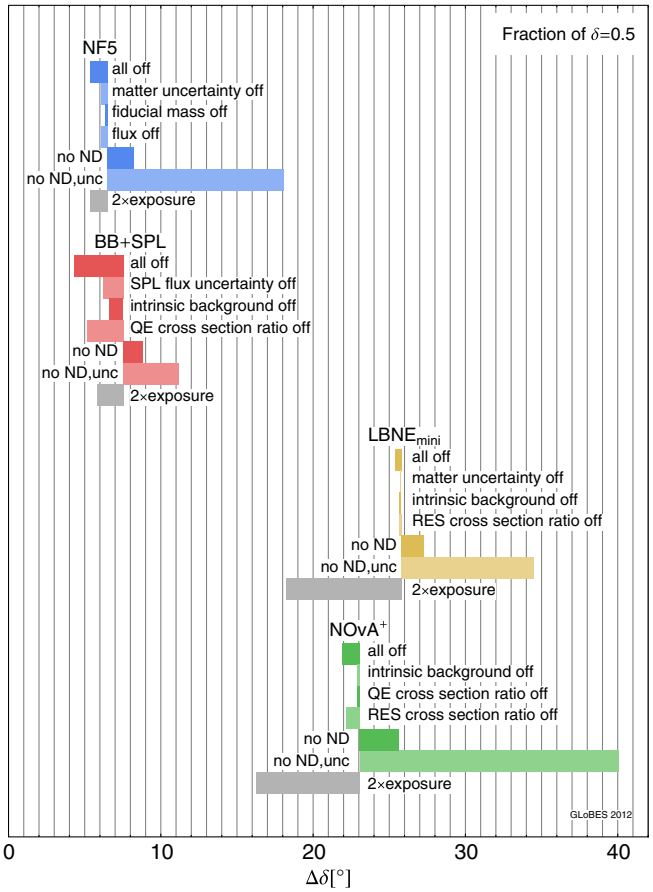

FIG. 4 (color online). Dependence of the performance of the alternative setups in Table I on systematic uncertainties, exposure, and near detectors. The meaning of the labels and abbreviations is the same as in Fig. 3.

unable to constrain both the systematic nuisance parameters and the atmospheric oscillation parameters independently using disappearance data. Thus, in the optimization of experiments of this type, the benefits of a near detector and of increased statistics have to be carefully weighed against each other.

An interesting question one could ask at this point is whether there are feasible ways of reducing systematic uncertainties, especially the ones on the QE cross sections. An interesting example for this is BB + SPL. In this setup, both ν_e and ν_μ cross sections can be measured precisely in the same detector, which reduces the impact of systematics and increases the absolute performance. This can be seen from the reduced length of the "all off" bar for BB + SPL compared to the BB350 or T2HK cases, which all operate in the low-energy regime. Some further improvement would be achieved if the SPL flux uncertainty were reduced, though. Note that BB + SPL could in principle even compete with NF5 if the exposure could be significantly increased, the cross-section ratios could be better constrained, or the SPL flux could be better understood. This shows how a combination of facilities can be of great help in reducing the impact of systematics on their performance. A similar effect would be obtained if an independent measurement of the ν_e/ν_μ cross-section ratio were performed for both neutrinos and antineutrinos. The proposed low-energy muon storage ring experiment ν-STORM [59] would be ideal for this measurement.

An additional method to reduce the impact of systematics could be a facility optimized for the second oscillation peak; see, for instance, Ref. [37], where an SPL-based experiment with a detector at 650 km instead of 130 km is proposed. This would be useful to increase the *CPV* discovery potential of the facility as well as to reduce the impact of systematic errors. Note that in Ref. [37], correlations between systematic uncertainties were not taken into account, and near detectors were not simulated explicitly. We have checked that the conclusions still hold in the case where full correlations are taken into account. Indeed, the setup proposed in Ref. [37] exhibits the least dependence on systematic errors among all the experiments compared here.

V. SUMMARY AND CONCLUSIONS

Systematic uncertainties in neutrino oscillation experiments are especially important for large θ_{13}. Hence, a dedicated comparison with a careful treatment of these uncertainties is needed to determine the optimal next-generation experiment, given the large value of θ_{13}. Also, the degree to which this optimization depends on the assumptions regarding systematic uncertainties should be carefully assessed.

In this study, we have analyzed and compared superbeams, beta beams, and neutrino factories on an equal footing, paying special attention to systematic uncertainties. In particular, a realistic implementation of systematic uncertainties in the simulations used to predict the sensitivity of future experiments depends not only on individual numbers for certain systematic errors, but also on how these errors are correlated among different detectors, oscillation channels, etc. In most previous studies, only a few types of systematic uncertainties were considered, and the respective error margins were chosen in order to account in an effective way for the real error menu. In particular, near detectors were typically not simulated explicitly, and correlations were neglected. In this paper, instead, we have used explicit near- and far-detector simulations with comparable assumptions and an improved systematics implementation which takes into account all possible correlations. Moreover, to allow for a simple and fair comparison of different facilities, we have used identical assumptions on external input (in particular, cross sections) wherever possible. Besides our default set of systematic errors, we also consider more conservative and more optimistic scenarios (see Table II), which should encompass the performance of a real experiment.

Table I summarizes the setups studied in this work. Since we expect that the mass hierarchy can de determined by all of the discussed setups (with the possible exception of NOνA$^+$), we have used the 3σ discovery potential for leptonic *CP* violation (*CPV*) and the achievable precision at 1σ in the measurement of δ ($\Delta\delta$) as our main performance indicators. While the first indicator depends on the performance of the experiment around the specific values $\delta = 0, \pi$, the second one treats all values of δ as equally important. Since the dependence of the experimental sensitivity on δ is in general complicated, we present our results in terms of the fraction of δ values for which a certain precision (or better) is achieved.

We have compared our new systematics implementation with the previous effective treatment and have found good agreement except for T2HK and BB350, for which the near–far extrapolation depends strongly on the poorly known ratio of ν_e and ν_μ QE cross sections. Therefore, the performance of these experiments strongly depends on the systematics assumptions, and it is difficult to make self-consistent predictions. We have also discussed the impact of the true value of δ on the measurement. While the performance is relatively uniform for most experiments, the precision attainable from NOνA$^+$ and NF5 in particular depends on δ. NF5, for instance, was clearly optimized for *CPV*, whereas a precise measurement independent of δ would require higher muon energies and longer baselines, as realized in the NF10 setup. Interestingly, T2HK does not exhibit such a strong dependence on δ, in spite of the narrow-beam spectrum.

For each experiment under consideration here, we have also identified the main limitations to a further increase in sensitivity. We have considered systematic uncertainties, exposure, and the impact of the near detector as possible bottlenecks. The results can be summarized as follows:

Superbeams can be divided into two classes: low and high energy. For the low-energy experiment T2HK, the fact that the ν_e cross sections needed for appearance measurements cannot be easily obtained from the near detector is clearly the most important limitation. This is especially relevant since it operates in the QE regime, where cross-section uncertainties are large and it is very difficult to relate the measured $\overset{(-)}{\nu}_\mu$ cross sections to the $\overset{(-)}{\nu}_e$ cross sections needed for the appearance measurement. Although the intrinsic beam backgrounds were included in our near-detector simulations, we could not identify a simple way of measuring the $\overset{(-)}{\nu}_e$ cross sections directly with the required precision. Uncertainties in the intrinsic beam background and the limited exposure are also limiting factors for T2HK, and we find that the availability of a near detector is of some importance. For WBB, LBNE$_{\text{mini}}$, and NOνA$^+$, which operate at higher energies, systematic errors can be controlled to the level needed, which in turn implies that, from the systematics point of view, very robust predictions can be made for these experiments. The critical issue for them is instead exposure. For instance, for LBNE$_{\text{mini}}$, investing in the far-detector mass may be more important than constructing a near detector. We conclude that for superbeams, the impact of systematic uncertainties depends mainly on the beam energy, especially because cross-section uncertainties are much smaller in the high-energy DIS region than in the low-energy QE region. The separation into narrow-band beams (T2HK, NOνA$^+$) and wide-band beams (WBB, LBNE$_{\text{mini}}$), on the other hand, has turned out not to be the primary issue.

Beta beams using ^{6}He and ^{18}Ne for the neutrino production also suffer from the fact that the ratio between the $\overset{(-)}{\nu}_e$ and $\overset{(-)}{\nu}_\mu$ QE cross sections is needed as an external input. If a low-γ beta beam is, however, combined with an SPL-based superbeam (BB + SPL), the performance is much better and much more robust than that of a high-γ BB350. In fact, BB + SPL is the only experiment that could compete with a neutrino factory. The reason is that BB + SPL uses both the $\overset{(-)}{\nu}_e \to \overset{(-)}{\nu}_\mu$ and $\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_e$ channels, so that both $\overset{(-)}{\nu}_e$ and $\overset{(-)}{\nu}_\mu$ cross sections can be measured.

Neutrino factories achieve the best absolute precision (comparable to that in the quark sector) and are very robust with respect to systematic errors. This is due to two main factors: firstly, the energy of the beam lies in the DIS regime where cross-section uncertainties are small; secondly, this is the only experiment where the final flavor cross section can be determined in a self-consistent way from the disappearance data. Depending on baseline and muon energy, the most relevant factors affecting their performance are the matter density uncertainty (for setups with longer baselines and high E_μ, such as NF10) or exposure and near detector (for setups with shorter baselines and low E_μ, such as NF5).

We also find, remarkably, that near detectors have a relatively small impact and help to improve the precision in δ by only about 1°–2° or less. The reason is that most systematic uncertainties are correlated between appearance and disappearance channels and can therefore be constrained by the far detector alone, provided that the statistics in the disappearance channel is good enough to break correlations between systematic effects and the atmospheric oscillation parameters. The near detector turns out to be practically useless at beta beam facilities: since the ν_e disappearance channel does not depend very much on the atmospheric parameters, the ν_e data from the far detector are even more useful for constraining flux and cross-section uncertainties than in other experiments. It should be kept in mind, however, that near detectors will still be required to constrain effects of new physics in neutrino oscillations. In addition, if a combined analysis

of appearance and disappearance data is not possible, a near detector proves to be critical in order to constrain cross-section and flux uncertainties, as expected. Moreover, a well-designed near-detector facility is an excellent safeguard against "unknown unknowns."

The most attractive superbeam option, as far as the impact of systematic uncertainties is concerned, is a high-energy, wide-band beam like LBNO or LBNE operating in the DIS regime. The $\text{LBNE}_{\text{mini}}$ experiment may be the first step towards such an experiment. However, both LBNO and LBNE have a limited discovery potential for *CPV*, and would suffer strongly from a reduction in statistics (notice that the WBB setup studied here had a LAr detector with a fiducial mass of 100 kt). The ultimate precision can be reached with a neutrino factory, which is the only experiment with a precision competitive with the one achieved in the quark sector. The BB + SPL combination of a $\gamma = 100$ beta beam and a superbeam is a very interesting option that is very robust with respect to systematic errors and has a performance closer to neutrino factory than any other superbeam or beta beam. Previous studies have underestimated the performance of BB + SPL because they did not take into account correlations between systematic uncertainties. Predictions for T2HK and BB350 heavily rely on external input on the flavor dependence of the QE cross sections. Here, an independent ν_e cross-section measurement—for instance, by a facility like ν-STORM—is necessary.

ACKNOWLEDGMENTS

We would like to thank E. Fernandez-Martinez for illuminating discussions, and M. Mezzetto, who contributed during the early stages of this work. We thank M. Vagins for providing the T2HK fluxes in machine-readable format. W. W. would like to acknowledge support from DFG Grants No. WI 2639/3-1 and No. WI 2639/4-1. This work has been supported by the U.S. Department of Energy under Award No. DE-SC0003915. P. C., P. H., and W. W. would like to thank GGI Florence for hospitality during their stay within the "What's ν?" program. P. C. would also like to thank CERN and Fermilab for their hospitality during completion of this work. This work has been also supported by the EU FP7 Projects EURONU (No. CE212372) and INVISIBLES (Marie Curie Actions, No. PITN-GA-2011-289442). Fermilab is operated by the Fermi Research Alliance under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy.

[1] G. L. Fogli, E. Lisi, A. Marrone, A. Palazzo, and A. M. Rotunno, Phys. Rev. Lett. **101**, 141801 (2008).
[2] T. Schwetz, M. Tortola, and J. W. F. Valle, New J. Phys. **10**, 113011 (2008).
[3] M. C. Gonzalez-Garcia, M. Maltoni, and J. Salvado, J. High Energy Phys. 04 (2010) 056.
[4] T. Schwetz, M. Tortola, and J. Valle, New J. Phys. **13**, 063004 (2011).
[5] K. Abe *et al.* (T2K Collaboration), Phys. Rev. Lett. **107**, 041801 (2011).
[6] P. Adamson *et al.* (MINOS Collaboration), Phys. Rev. Lett. **107**, 181802 (2011).
[7] Y. Abe *et al.* (DOUBLE-CHOOZ Collaboration), Phys. Rev. Lett. **108**, 131801 (2012).
[8] F. An *et al.* (DAYA-BAY Collaboration), Phys. Rev. Lett. **108**, 171803 (2012).
[9] J. Ahn *et al.* (RENO Collaboration), Phys. Rev. Lett. **108**, 191802 (2012).
[10] M. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz, J. High Energy Phys. 12 (2012) 123.
[11] G. Fogli, E. Lisi, A. Marrone, D. Montanino, A. Palazzo, and A. Rotunno, Phys. Rev. D **86**, 013012 (2012).
[12] D. Forero, M. Tortola, and J. Valle, Phys. Rev. D **86**, 073012 (2012).
[13] P. Coloma, A. Donini, E. Fernandez-Martinez, and P. Hernandez, J. High Energy Phys. 06 (2012) 073.
[14] P. Huber, M. Lindner, T. Schwetz, and W. Winter, J. High Energy Phys. 11 (2009) 044.
[15] M. Blennow and T. Schwetz, J. High Energy Phys. 08 (2012) 058; 11 (2012) 098(E).
[16] E. K. Akhmedov, S. Razzaque, and A. Y. Smirnov, arXiv:1205.7071.
[17] S. T. Petcov and M. Piai, Phys. Lett. B **533**, 94 (2002).
[18] S. Choubey, S. Petcov, and M. Piai, Phys. Rev. D **68**, 113006 (2003).
[19] L. Zhan, Y. Wang, J. Cao, and L. Wen, Phys. Rev. D **79**, 073007 (2009).
[20] X. Qian *et al.*, arXiv:1208.1551.
[21] H. Nunokawa, S. J. Parke, and R. Z. Funchal, Phys. Rev. D **72**, 013009 (2005).
[22] A. de Gouvea, J. Jenkins, and B. Kayser, Phys. Rev. D **71**, 113009 (2005).
[23] V. Barger, P. Huber, D. Marfatia, and W. Winter, Phys. Rev. D **76**, 053005 (2007).
[24] W. Winter, Phys. Rev. D **78**, 037101 (2008).
[25] J. Tang and W. Winter, Phys. Rev. D **81**, 033005 (2010).
[26] P. Huber, M. Maltoni, and T. Schwetz, Phys. Rev. D **71**, 053006 (2005).
[27] J.-E. Campagne, M. Maltoni, M. Mezzetto, and T. Schwetz, J. High Energy Phys. 04 (2007) 003.
[28] K. Abe *et al.*, arXiv:1109.3262.
[29] V. Barger, R. Gandhi, P. Ghoshal, S. Goswami, D. Marfatia, S. Prakash, S. Raut, and S. Sankar, Phys. Rev. Lett. **109**, 091801 (2012).

[30] M. Fukugita and T. Yanagida, Phys. Lett. B **174**, 45 (1986).
[31] S. King, J. High Energy Phys. 08 (2005) 105.
[32] I. Masina, Phys. Lett. B **633**, 134 (2006).
[33] S. Antusch and S. F. King, Phys. Lett. B **631**, 42 (2005).
[34] W. Winter, Phys. Rev. D **70**, 033006 (2004).
[35] P. Huber, M. Lindner, and W. Winter, J. High Energy Phys. 05 (2005) 020.
[36] P. Huber, M. Mezzetto, and T. Schwetz, J. High Energy Phys. 03 (2008) 021.
[37] P. Coloma and E. Fernandez-Martinez, J. High Energy Phys. 04 (2012) 089.
[38] J. Tang and W. Winter, Phys. Rev. D **80**, 053001 (2009).
[39] P. Huber, M. Lindner, and W. Winter, Nucl. Phys. **B645**, 3 (2002).
[40] T. Ohlsson and W. Winter, Phys. Rev. D **68**, 073007 (2003).
[41] P. Huber, M. Lindner, T. Schwetz, and W. Winter, Nucl. Phys. **B665**, 487 (2003).
[42] P. Huber, M. Lindner, M. Rolinec, T. Schwetz, and W. Winter, Phys. Rev. D **70**, 073014 (2004).
[43] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter, J. High Energy Phys. 05 (2006) 072.
[44] P. Huber, M. Lindner, and W. Winter, Comput. Phys. Commun. **167**, 195 (2005); http://www.mpi-hd.mpg.de/lin/globes/.
[45] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter, Comput. Phys. Commun. **177**, 432 (2007).
[46] P. Coloma, P. Huber, J. Kopp, and W. Winter, arXiv:1209.5973; See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.87.033004 for [brief description].
[47] J. Burguet-Castell, D. Casper, J. Gomez-Cadenas, P. Hernandez, and F. Sanchez, Nucl. Phys. **B695**, 217 (2004).
[48] J. Burguet-Castell, D. Casper, E. Couce, J. Gomez-Cadenas, and P. Hernandez, Nucl. Phys. **B725**, 306 (2005).
[49] J. Bernabeu *et al.*, arXiv:1005.3146.
[50] S. K. Agarwalla, P. Huber, J. Tang, and W. Winter, J. High Energy Phys. 01 (2011) 120.
[51] International design study of the neutrino factory, http://www.ids-nf.org.
[52] A. Longhin, Proc. Sci., ICHEP2010 (2010) 325.
[53] D. Angus *et al.* (LAGUNA Collaboration), arXiv:1001.0077.
[54] A. Rubbia *et al.* (LAGUNA-LBNO Collaboration), CERN Report No. CERN-SPSC-2012-021 (unpublished); CERN Report No. SPSC-EOI-007 (unpublished).
[55] T. Akiri *et al.* (LBNE Collaboration), arXiv:1110.6249.
[56] A. Lenz, U. Nierste, J. Charles, S. Descotes-Genon, H. Lacker, S. Monteil, V. Niess, and S. T'Jampens, Phys. Rev. D **86**, 033008 (2012).
[57] J. Kopp, M. Lindner, T. Ota, and J. Sato, Phys. Rev. D **77**, 013007 (2008).
[58] S. Antusch, M. Blennow, E. Fernandez-Martinez, and J. Lopez-Pavon, Phys. Rev. D **80**, 033002 (2009).
[59] P. Kyberd *et al.*, arXiv:1206.0294.