

A Markov Chain Monte Carlo oscillation data analysis using SK atmospheric and T2K beam neutrinos

Daniel Barrow¹, Aoi Eguchi³, Mo Jia², Junjie Jiang², Wei Shi², Michael
Wilking², and Clarence Wret¹

¹University of Oxford, UK

²Stony Brook University, USA

³University of Tokyo, Japan

September 22, 2023

Todo list

■ CW: (JJ)Syst constraint plots need slightly adjustment, will do 80

Abstract

This document describes the implementation of model updates in MaCh3 framework, MaCh3 results of the updated sensitivity (Asimov A fit), and fit to the real data of T2K near and far detector beam samples combined with the SK atmospheric samples. Comparisons to T2K OA2020 and joint-fit results from the P-Theta group are also documented. This analysis uses 1.1531×10^{21} POT in FHC and 8.336×10^{20} POT in RHC for the near-detector fit corresponding to runs 2-9, 1.9664×10^{21} POT in FHC, 1.6346×10^{21} POT in RHC corresponding to T2K far detector beam runs 1-10 and the full SK-IV livetime of 3244.4 days. Asimov and data fit results include smearing of Δm_{32}^2 from fake-data studies, and results with and without reactor constraint are studied. The impact of a flat prior on $\sin \delta_{\text{CP}}$ and a flat prior on δ_{CP} is also studied for the Jarlskog invariant.

The fit to data with the reactor constraint on $\sin^2 \theta_{13}$ applied finds a Bayes factor of normal ordering over inverted ordering $B(\text{NO}/\text{IO})$ of 7.33. This compares to the T2K

15 results of $B(\text{NO}/\text{IO}) = 4.2$ in the 2020 analysis. $\delta_{\text{CP}} = 0, \pi$ is excluded above 2σ with a
16 flat prior on δ_{CP} and the 2σ level is just included with a flat prior on $\sin \delta_{\text{CP}}$. A Jarlskog
17 invariant of $J = 0$ is excluded above 2σ with a flat prior on δ_{CP} , and is just excluded at
18 2σ with a flat prior on $\sin \delta_{\text{CP}}$. The Bayes factor $B[(\sin \delta_{\text{CP}} < 0)/(\sin \delta_{\text{CP}} \geq 0)]$ is 50.0
19 with the reactor constraint applied and a flat prior on δ_{CP} . The Bayes factor of upper
20 octant over lower octant $B(\text{UO}/\text{LO})$ is 1.78, which shows a weaker preference of the
21 upper octant compared to OA2020 [1] ($B(\text{UO}/\text{LO}) = 3.36$), due to SK's preference for
22 lower octant and T2K's preference for upper octant. The posterior predictive p-values
23 find that the samples are well-predicted by the model, with an overall $p = 0.42$ for the
24 rate-based p-value, and $p = 0.11$ for the shape-based p-value.

25 Contents

26	1 Introduction	5
27	2 Model updates implementation in MaCh3	6
28	2.1 Updates of atmospheric detector errors	6
29	2.2 Treatment of SubGeV 1Re1de excess	6
30	2.2.1 Updated Adler angle dials	6
31	2.2.2 New SK PID errors	6
32	2.2.3 MC with new decay electron cut	7
33	2.3 Correlated far detector systematics	7
34	3 Results from fitting Asimov data	8
35	3.1 Results applying the reactor constraint	10
36	3.1.1 Parameter constraints	10
37	3.1.2 Triangle plots	19
38	3.1.3 Bayes factors for atmospheric parameters	23
39	3.2 Results without applying the reactor constraint	24
40	3.2.1 Parameter constraints	24
41	3.2.2 Triangle plots	32
42	3.2.3 Bayes factors for atmospheric parameters	36
43	3.3 Bayes factors for CP violation	37
44	4 Results from fitting real data	38
45	4.1 Results applying the reactor constraint	40
46	4.1.1 Parameter constraints, prior flat in δ_{CP}	40
47	4.1.2 Parameter constraints, prior flat in $\sin \delta_{\text{CP}}$	45
48	4.1.3 Triangle plots	53
49	4.1.4 Bayes factors for atmospheric parameters	57
50	4.2 Results without applying the reactor constraint	58
51	4.2.1 Parameter constraints	58
52	4.2.2 Triangle plots	67
53	4.2.3 Bayes factors for atmospheric parameters	71
54	4.3 Impact of a flat prior on δ_{CP} or $\sin \delta_{\text{CP}}$	72
55	4.4 Impact of reactor constraint	74
56	4.5 Posterior Predictive Spectra and p -values	77

57	4.5.1	Posterior predictive spectra	77
58	4.5.2	Impact of uncertainty sources	77
59	4.5.3	Posterior predictive p -values	78
60	4.6	Systematic constraints	80
61	4.7	Comparison to OA2020 MaCh3 results	96
62	4.7.1	Parameter constraints	96
63	4.7.2	Bayes factors for atmospheric parameters	101
64	5	Results comparisons	102
65	5.1	Joint-fit results comparison between MaCh3 and PTheta	102
66	5.2	Joint-fit results and OA2020 from MaCh3 and PTheta	108
67	5.3	Potential explanation of results differences	111
68	5.4	Comparison of the systematic posterior distributions	116
69	6	Summary	120

1 Introduction

This note documents the results and discussion of the joint-fit of near detector samples (ND280 data from runs 2-9) and far detector samples (5 T2K beam samples from runs 1-10 and 18 SK atmospheric samples of 3224.4 days from SK-IV) performed in MaCh3 framework.

MaCh3 is an oscillation fitter which uses Bayesian inference and Markov chain Monte Carlo (MCMC) to sample the posterior probability in the high dimensional systematic phase space and using marginalisation to estimate the oscillation parameters posterior probability distributions in the T2K experiment. It is extended to simultaneously fit SK atmospheric and T2K near and far detector samples [2]. Running MCMC for smooth enough contours requires a large amount of computing resources. For results presented in this note, we utilized both the GPU and CPU resources from ComputeCanada clusters and Summit cluster.

The joint fit samples are discussed in further detail in TN428 [3]. The oscillation probability calculation method is documented in TN425 [4]. The joint-fit flux and cross-section model originally described in TN422 [5] and is updated in TN456 [6]. This work is based on the updated joint fit model and the implementation of the updates in MaCh3 framework is documented in Sec. 2. The analysis methods in this work are mostly the same as in the sensitivity studies documented in TN426 [2] and this note studies the sensitivity with the model updates, and the data fit.

Sec. 3 describes the Asimov A fit results for both with and without reactor constraint of $\sin^2 \theta_{13}$. Both one dimensional and two dimensional contours of oscillation parameters are presented. In addition, the Jarlskog invariant posterior probability distributions are also presented. Bayes factors of octant hypotheses, mass ordering and CP violation are calculated and discussed.

Sec. 4 describes the data fit results for both with and without reactor constraint of $\sin^2 \theta_{13}$. Both one dimensional and two dimensional contours of oscillation parameters are presented. In addition, the Jarlskog invariant posterior probability distributions are also presented. Bayes factors of octant hypotheses, mass ordering and CP violation are calculated and discussed. Effect of reactor constraint is discussed by overlaying the contours with and without the constraint. Finally, the joint-fit data fit results (with reactor constraint) are compared to OA2020 MaCh3 results (T2K-only fit, see TN393 [1]).

Sec. 5 describes the comparison of the joint-fit data fit results between MaCh3 and P-Theta as well as OA2020 results (TN393 [1] and TN397 [7]). Discussion of the results differences between fitters is documented as well.

2 Model updates implementation in MaCh3

This section describes the implementation of model updates in the MaCh3 framework. The motivation of the updates can be found in the joint-fit model updates TN456 [6] Sec. 2. All involved systematics can be found in the joint-fit fitter validation TN471 [8] (Tab. 6, 7, 8).

2.1 Updates of atmospheric detector errors

Joint-fit analysis decides to use the same fitQun version and CCQE model for SK and T2K samples as T2K official analysis. Hence, SK atmospheric detector errors need to be re-evaluated. The atmospheric detector errors are hard coded in MaCh3 and thus updating the values is straight-forward. The updated errors are documented in Sec. 3 of TN456 [6] and validated through fitter validation recorded in TN471 [8].

2.2 Treatment of SubGeV 1Re1de excess

To deal with the data/MC discrepancy observed in down-going SubGeV CC1pi dominated sample, three different ways are adopted. Details are documented in Sec. 4.4 from TN456 [6].

2.2.1 Updated Adler angle dials

The ad-hoc Adler angle uncertainty was introduced in the model for the sensitivity study last year. In the latest model, the previous Adler angle dial is split into two dials with different dependence on lepton momentum. The parameterization and prior constraints are also changed.

New splines are generated using the same weight files as P-Theta (see Sec. 4.4.4 in TN456 [6] v1.1). The Adler angle dials are implemented as splined cross-section systematics that apply to CC1pi and NC1pi true resonance interaction modes (including both charged and neutral pion) for all beam samples and atmospheric SubGeV samples. Validation is done and documented in fitter validation TN471 [8].

2.2.2 New SK PID errors

To account for the potentially poor performance of e/μ PID at low lepton momentum for CC1 π dominated samples, two 50% correlated systematics are developed (see TN456 [6]). Two migration dials are implemented; one moves events between the atmospheric single ring

130 one decay electron and muon-like samples, and the other moves events between the beam
131 FHC 1Re1de and FHC(RHC)1Rmu samples.

132 Technically, these two systematics are implemented in the same class as the correlated
133 detector systematics and the lack of correlation is coded between them and other detector
134 systematics.

135 **2.2.3 MC with new decay electron cut**

136 The T2K MC was produced with a new cut on the delayed hit clusters reconstructed by
137 fitQun to reduce the background from neutron capture signals in tagging decay electrons.
138 Similarly, the SK ATMPD group updated the atmospheric MC with this new cut. Thus, the
139 input MC for joint-fit is now updated with this new cut.

140 Technically, MaCh3 updates a few reconstructed fitQun variables in the reading in MC
141 inputs process accordingly. Same input is used by other fitters in joint-fit and event rates
142 validation between fitters is done.

143 **2.3 Correlated far detector systematics**

144 The motivation and validation of the correlated far-detector model are documented in Sec.
145 5 of TN456 [6]. In the current MaCh3 framework, systematics are divided into un-correlated
146 groups so that each group has its own class to store all the necessary methods. To implement
147 this correlated model, a new class of the correlated detector systematics has been developed
148 to implement the correlation, which inherit most of the methods for each class of systematics
149 from the previous un-correlated model. The validation of this implementation can be found
150 in the fitter validations, documented in TN471 [8].

151 MaCh3 uses different sets of T2K beam detector systematics compared to P-Theta and
152 Osc3++ due to the different fit-binning variables used in MaCh3 and P-Theta (documented
153 in TN456 [6]). MaCh3 uses the set of systematics binned in reconstructed energy for all T2K
154 beam samples (TN456 [6] Tab. 8) while P-Theta and Osc3++ bin in reconstructed lepton
155 momentum for beam e-like samples and reconstructed energy for beam muon-like samples
156 (TN456 [6] Tab. 9).

157 Another technical difference is that the covariance matrix of the SK detector systematics
158 used in MaCh3 and P-Theta are different in that MaCh3 does not absorb the effect of
159 secondary interaction and photo-nuclear systematics into the matrix while P-Theta does.
160 The SI and photo-nuclear systematics in MaCh3 are instead implemented in another class

161 dedicated to cross-section systematics. This has no impact on physics conclusions in the
 162 oscillation analysis, but should be kept in mind when comparing the impact due to difference
 163 classes of systematics.

164 3 Results from fitting Asimov data

165 The “Asimov data” refers to the far-detector predictions built by reweighting the simulated
 166 far-detector samples using the post-BANFF tune, and applying the set of known oscillation
 167 parameters in [Tab. 1](#). The set of oscillation parameters named “Asimov A” are applied
 168 to produce all results shown in this section. The parameter sets are the same as used in
 169 the sensitivity studies TN426 [\[2\]](#). The Asimov studies are present for two main purposes:
 170 validation of the analysis (since the parameter set that generated the data is known), and
 171 to provide an estimate of the sensitivity assuming no statistical fluctuations.

Parameter	Asimov A	Asimov B
$\sin^2 \theta_{12}$	0.307	0.307
$\sin^2 \theta_{23}$	0.528	0.45
$\sin^2 \theta_{13}$	0.0218	0.0218
Δm_{12}^2 (eV ²)	7.53×10^{-5}	7.53×10^{-5}
Δm_{23}^2 (eV ²)	2.509×10^{-3}	2.509×10^{-3}
δ_{CP}	-1.601	0.0

Table 1: Oscillation parameter sets Asimov A and B, set A is used throughout this analysis.

172 When running an Asimov fit, MaCh3 fits the far-detector Asimov data and the near-
 173 detector real data simultaneously. Hence, the results are comparable to the P-Theta Asimov
 174 studies using the post-BANFF data fit constraints, with the small caveat of there being
 175 slightly different approaches in MaCh3 and BANFF’s ND treatment. The full MaCh3 ND
 176 treatment can be found in the near-detector fit note, TN395 [\[9\]](#), and some details specific to
 177 the joint fit with SK are provided in the sensitivity studies note, TN426 [\[2\]](#).

178 Due to the model updates—especially the correlated detector systematics—work was
 179 done to ensure reasonable behaviour of the MCMC fitting technique of all systematics by
 180 looking at the traces of total negative log-likelihood (-LLH) and parameter values, and the
 181 autocorrelations for all parameters, from a series of test chains. The aim of the tuning was
 182 to achieve 0.2 autocorrelation after 20,000 lag, with no parameters behaving significantly
 183 differently to others whilst maintaining high acceptance rates. This aim was achieved.

184 The Markov Chains were computed such that one chain started from where a previous
 185 chain stopped due to the wall time of the computing clusters. The deepest chains were
 186 550,000 steps deep. For the Asimov fit, a total of 384 independent chains were run in
 187 parallel, producing 203,320,244 steps in total including the burn-in period of 80,000 steps
 188 per independent chain. This gave 172,599,860 steps after burn-in in total. For results shown
 189 in this section, only post burn-in steps are used. Good acceptance rates ($\sim 16 - 20\%$) and
 190 relatively short burn-in periods (80,000 steps) were achieved. The trace of the total negative
 191 log-likelihood of all chains for the Asimov A analysis are shown in Fig. 1.

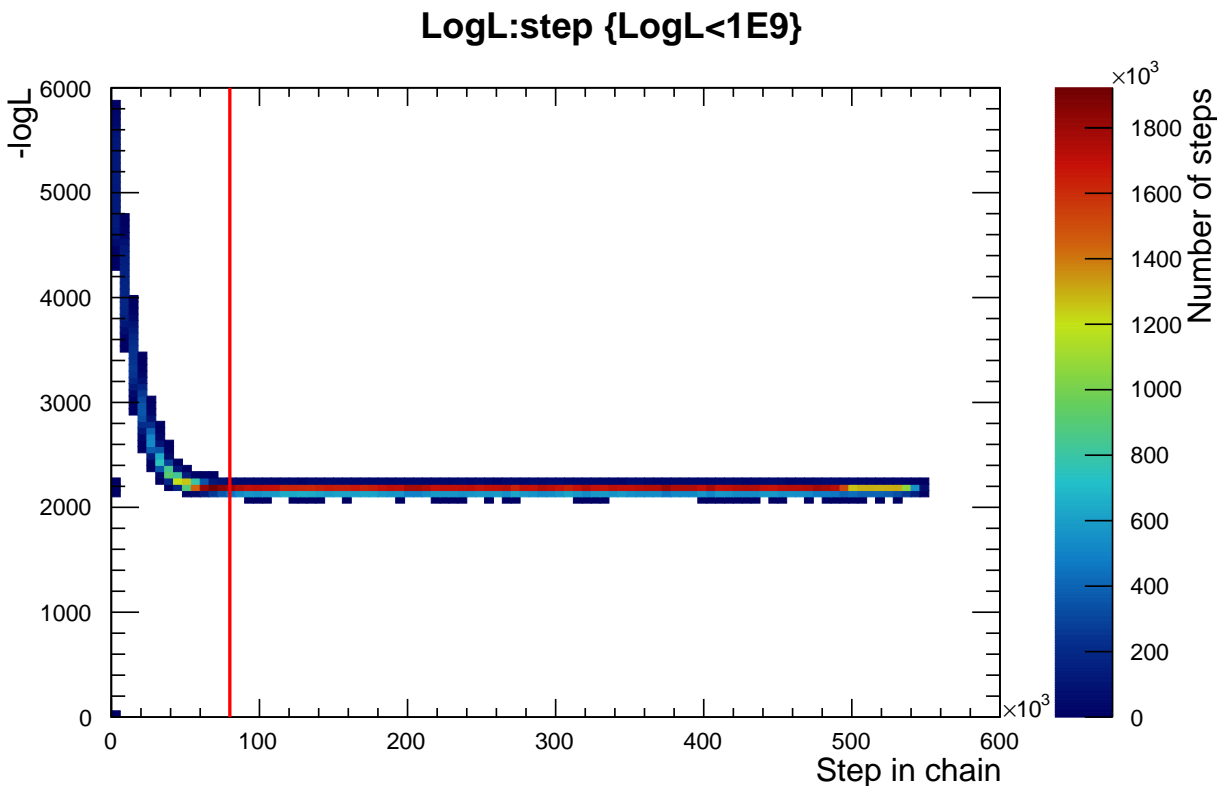


Figure 1: Collection of the traces of the negative log total likelihood from 384 independent chains in the Asimov A fit. The red vertical line shows the chosen burn-in cut at 80,000 steps.

192 This section first presents Asimov results with the reactor constraint applied, and then
 193 without it applied.

194 **3.1 Results applying the reactor constraint**

195 For these results, the reactor constraint on $\sin^2 \theta_{13}$ is used with the same value as the T2K
 196 OA2020 analysis TN393 [1] from the PDG [10]; $\sin^2 \theta_{13} = 0.0218 \pm 0.0007$.

197 **3.1.1 Parameter constraints**

198 Each parameter constraint in this section is shown in normal ordering ($\Delta m_{32}^2 > 0$), inverted
 199 ordering ($\Delta m_{32}^2 < 0$), and over both orderings.

200 **One-dimensional posteriors** Fig. 2, Fig. 3, Fig. 4 and Fig. 5 show the marginalised
 201 one-dimensional posterior probability for δ_{CP} , Δm_{32}^2 , $\sin^2 \theta_{13}$ and $\sin^2 \theta_{23}$, respectively.

202 For the δ_{CP} result, $\sin \delta_{\text{CP}} = 0$ is excluded at exactly 2σ over both orderings. For $\sin^2 \theta_{23}$,
 203 the correct octant is found, although both octants are well within the 1σ credible interval in
 204 both orderings.

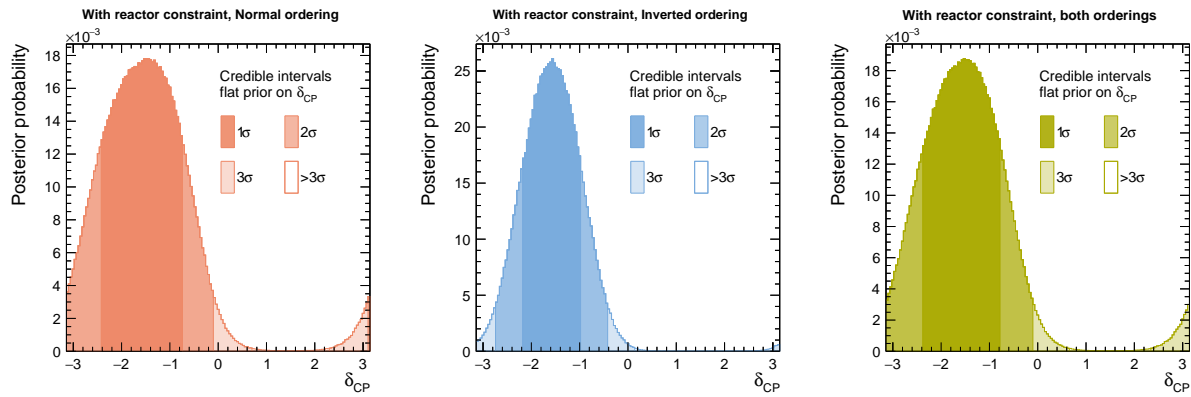


Figure 2: δ_{CP} for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

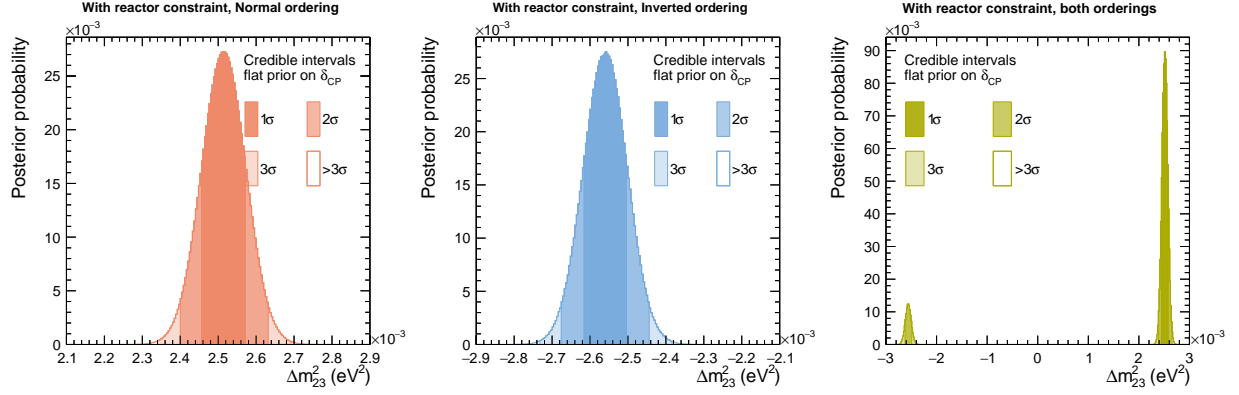


Figure 3: Δm_{32}^2 for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

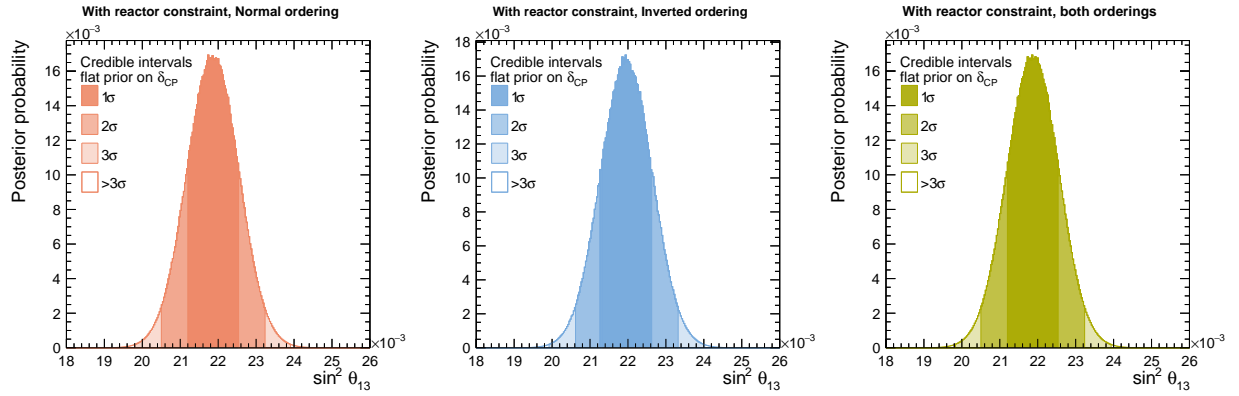


Figure 4: $\sin^2 \theta_{13}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

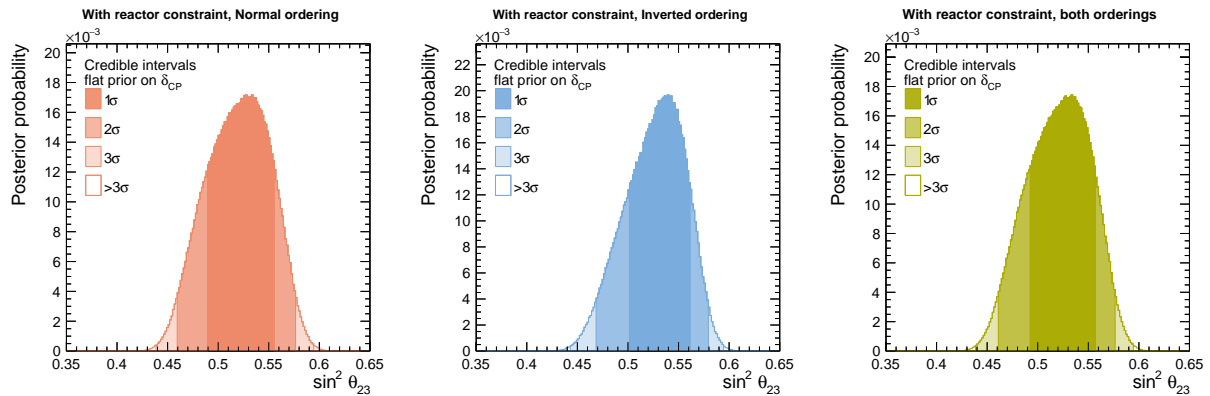


Figure 5: $\sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

The Jarlskog invariant The posterior distribution of Jarlskog invariant, which is defined as:

$$J = s_{13}c_{13}^2 s_{12}c_{12}s_{23}c_{23} \sin \delta_{CP} \quad (1)$$

where s_{ij} refers to $\sin \theta_{ij}$ and c_{ij} refers to $\cos \theta_{ij}$. More detailed description can be found in TN393 [1] (Appendix H in v9).

Fig. 6 and Fig. 7 show the marginalised Jarlskog invariant (J) posterior for normal ordering, inverted ordering and both ordering with flat prior on δ_{CP} and $\sin \delta_{CP}$, respectively. As in the mainline T2K analysis, the marginal likelihood for J is computed after the MCMC is run by computing J for each step in the chain after burn-in.

As was found in Fig. 2, $\sin \delta_{CP} = 0 (J = 0)$ is just excluded at 2σ credible interval over both orderings with a flat prior in δ_{CP} . For the prior flat in $\sin \delta_{CP}$, $J = 0$ is included in the 2σ credible interval.

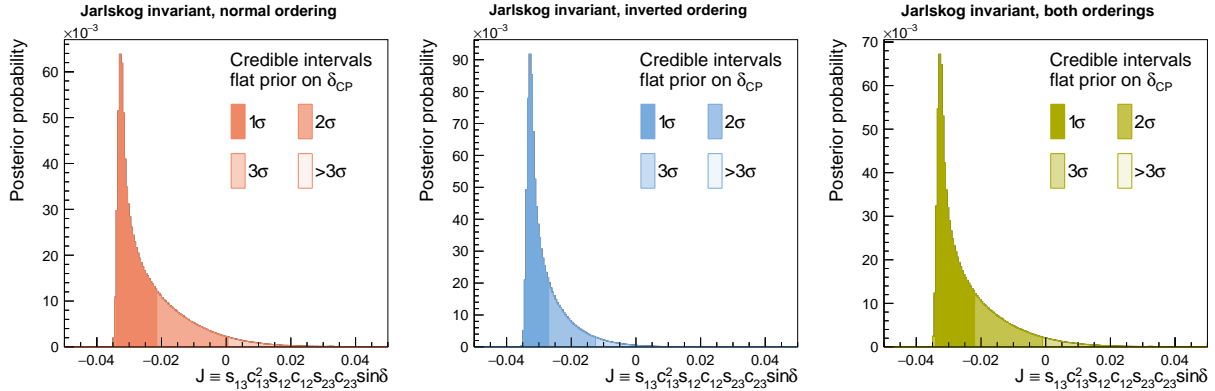


Figure 6: J for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

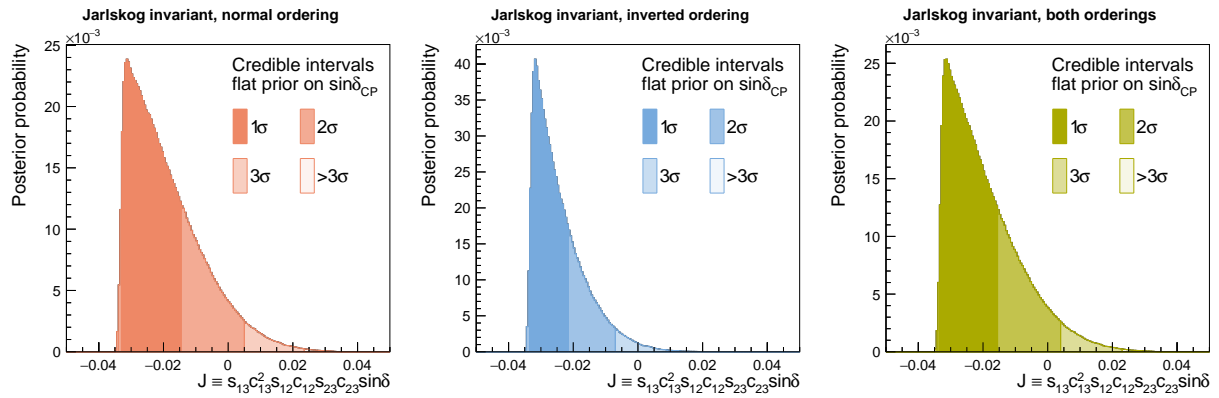


Figure 7: J for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with flat prior on $\sin \delta_{CP}$.

214 **Two-dimensional posteriors** Fig. 8, Fig. 9, Fig. 10, Fig. 11, Fig. 12, Fig. 13 and Fig. 14
 215 show the two-dimensional marginal posteriors for $\delta_{\text{CP}} - \sin^2 \theta_{13}$, $\delta_{\text{CP}} - \sin^2 \theta_{23}$, $\delta_{\text{CP}} - \Delta m_{32}^2$,
 216 $\Delta m_{32}^2 - \sin^2 \theta_{13}$, $\Delta m_{32}^2 - \sin^2 \theta_{23}$, $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ and $\sin^2 \theta_{23} - \Delta m_{32}^2$, respectively. These
 217 are shown separately for normal ordering, inverted ordering and both orderings, all applying
 218 the reactor constraint.

219 It is evident that the current MCMC statistics is not sufficient for 3σ credible intervals in
 220 2D with the current binning in the oscillation parameters. No statement is based on the 3σ
 221 confidence interval, so computing resources were not allocated to ensure stable 3σ contours.

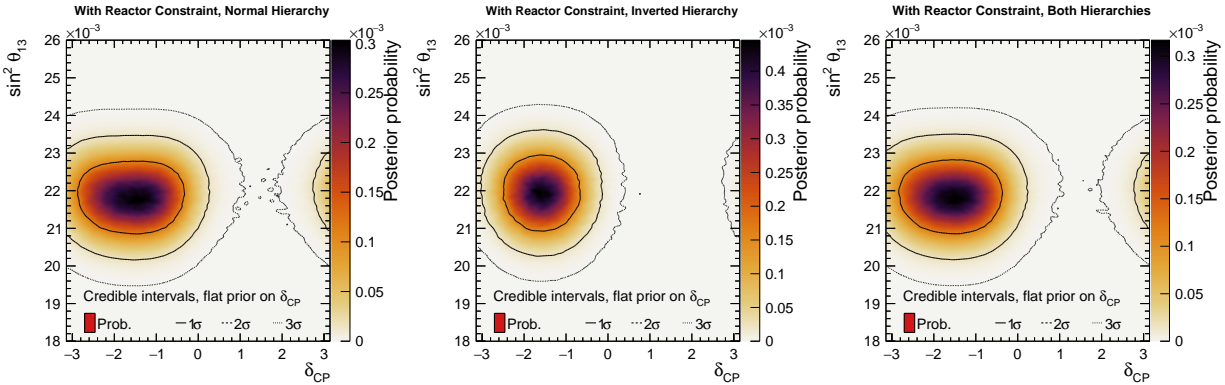


Figure 8: $\delta_{\text{CP}} - \sin^2 \theta_{13}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

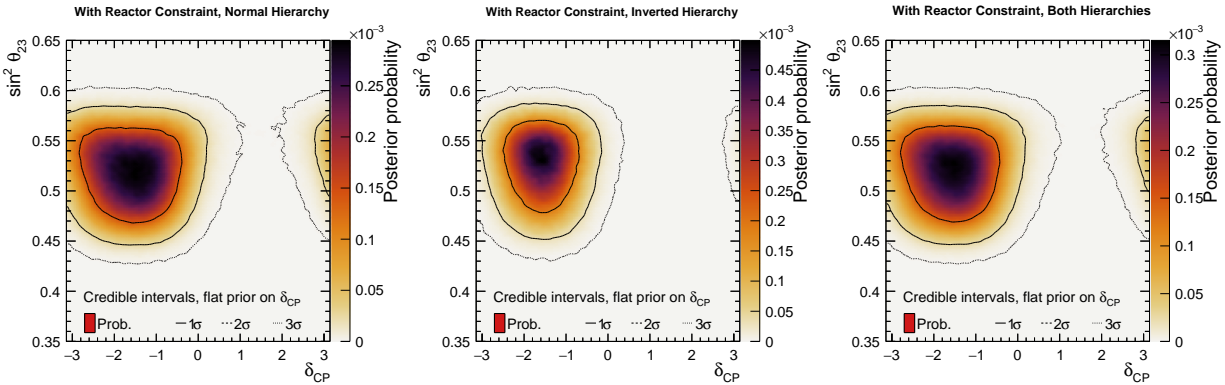


Figure 9: $\delta_{\text{CP}} - \sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

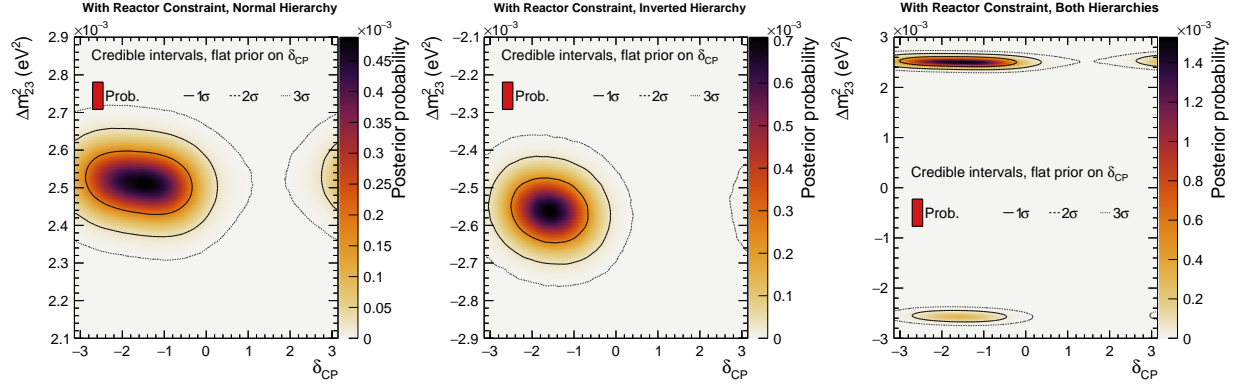


Figure 10: $\delta_{\text{CP}} - \Delta m_{32}^2$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

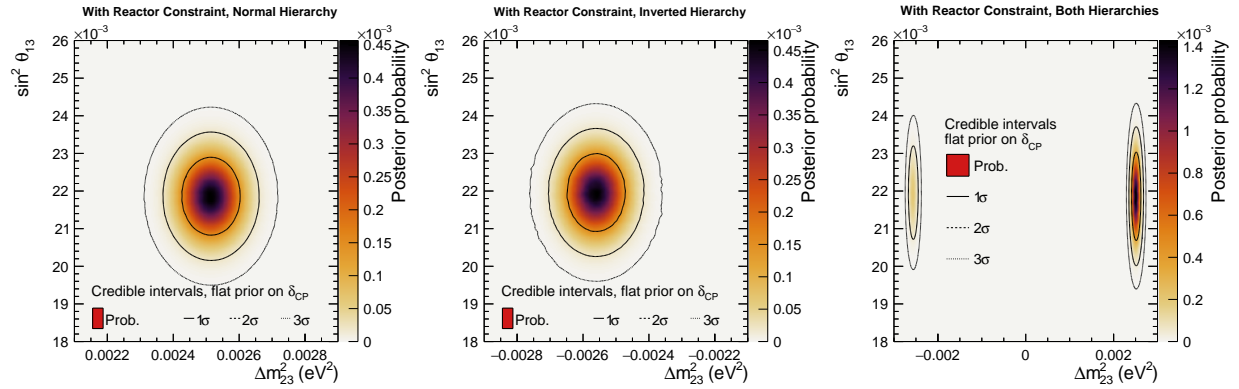


Figure 11: $\Delta m_{32}^2 - \sin^2 \theta_{13}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

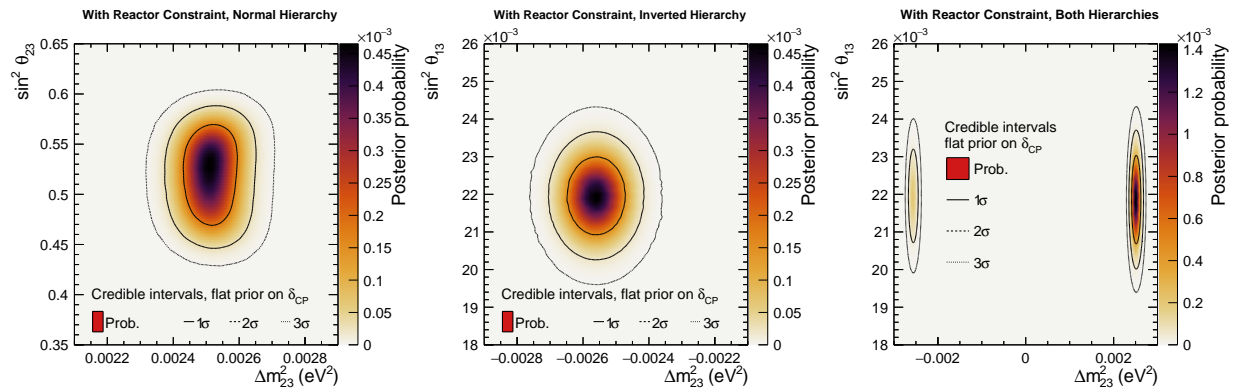


Figure 12: $\Delta m_{32}^2 - \sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

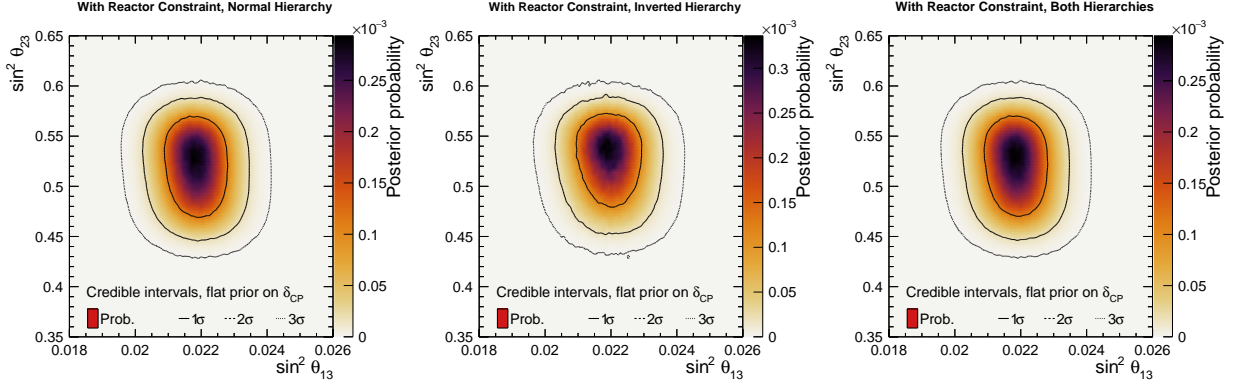


Figure 13: $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

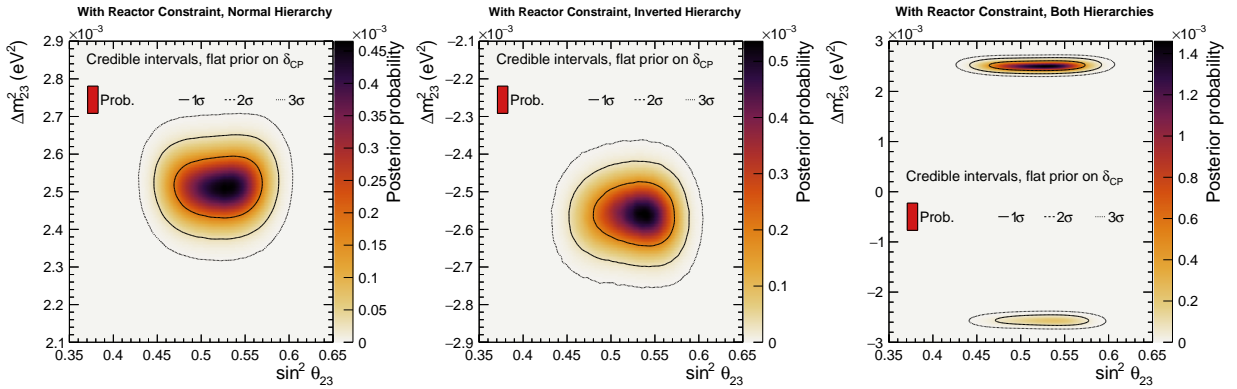


Figure 14: $\sin^2 \theta_{23} - \Delta m^2_{32}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

222 **The Jarlskog invariant** Fig. 15 and Fig. 16 show the two-dimensional $J - \sin^2 \theta_{23}$ poste-
 223 riors for normal ordering, inverted ordering and both orderings, with a flat prior on δ_{CP} and
 224 $\sin \delta_{\text{CP}}$ respectively. Fig. 17 shows the two-dimensional $J - \delta_{\text{CP}}$ posteriors for the orderings,
 225 with a flat prior on δ_{CP} .

226 As expected from the one-dimensional result in Fig. 6 and Fig. 7, the flat prior on $\sin \delta_{\text{CP}}$
 227 results in a weaker constraint at 1σ credible interval. However, at 2σ level the two are very
 228 similar, and at 3σ level the result with a flat prior on $\sin \delta_{\text{CP}}$ is slightly narrower than that
 229 with a flat prior on δ_{CP} . The correlation between J and $\sin^2 \theta_{23}$ remains similar to that seen
 230 in T2K's analyses.

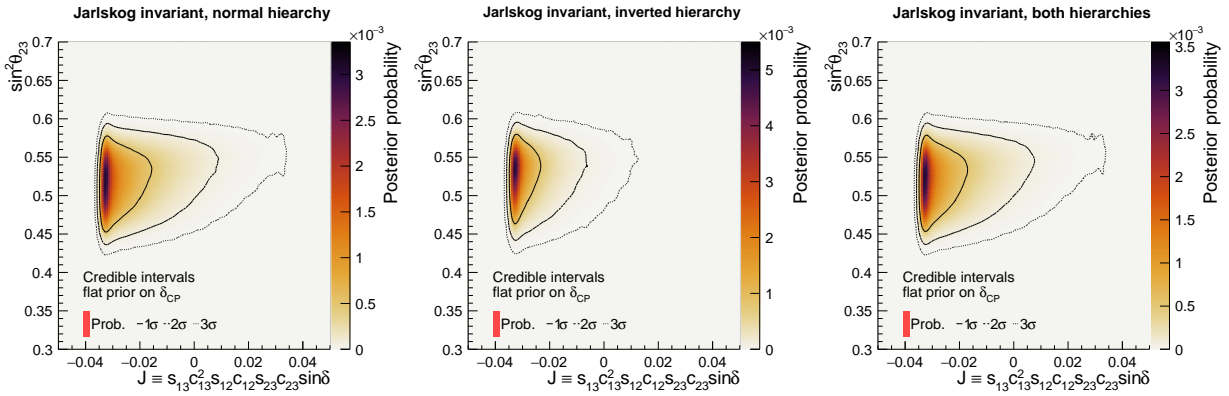


Figure 15: $J - \sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

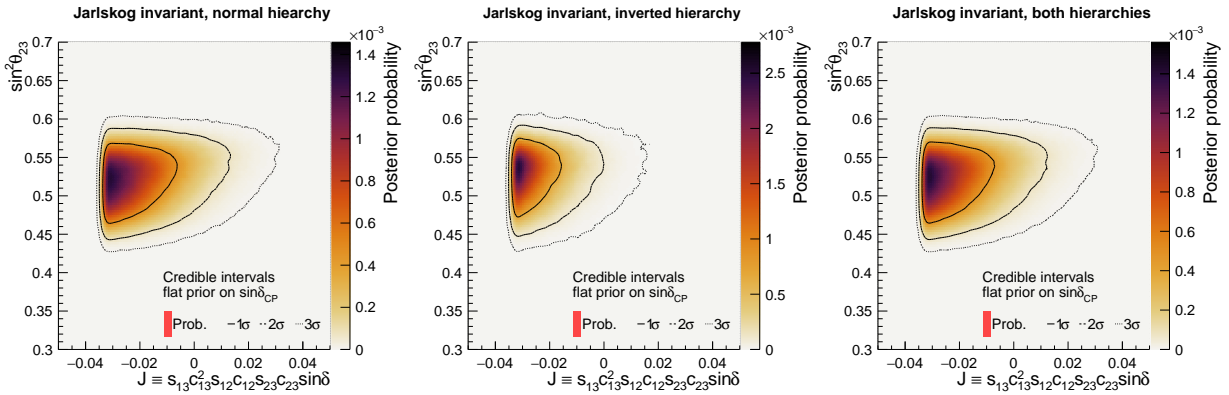


Figure 16: $J - \sin^2 \theta_{23}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

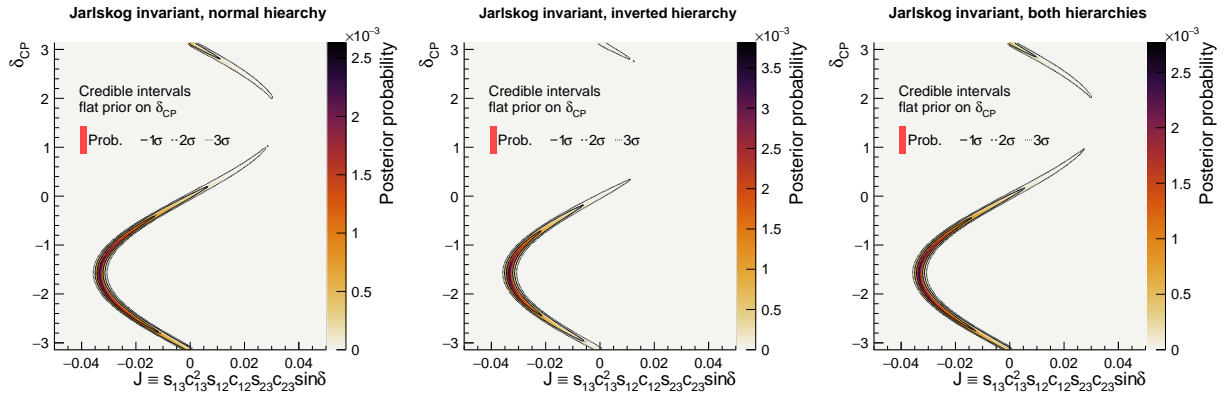


Figure 17: $J - \delta_{CP}$ for the Asimov A analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

231 **3.1.2 Triangle plots**

232 **Fig. 18**, **Fig. 19**, and **Fig. 20** show the one and two-dimensional oscillation parameters'
233 posteriors, displayed in triangular form. These all apply the reactor constraint and use a flat
234 prior on δ_{CP} .

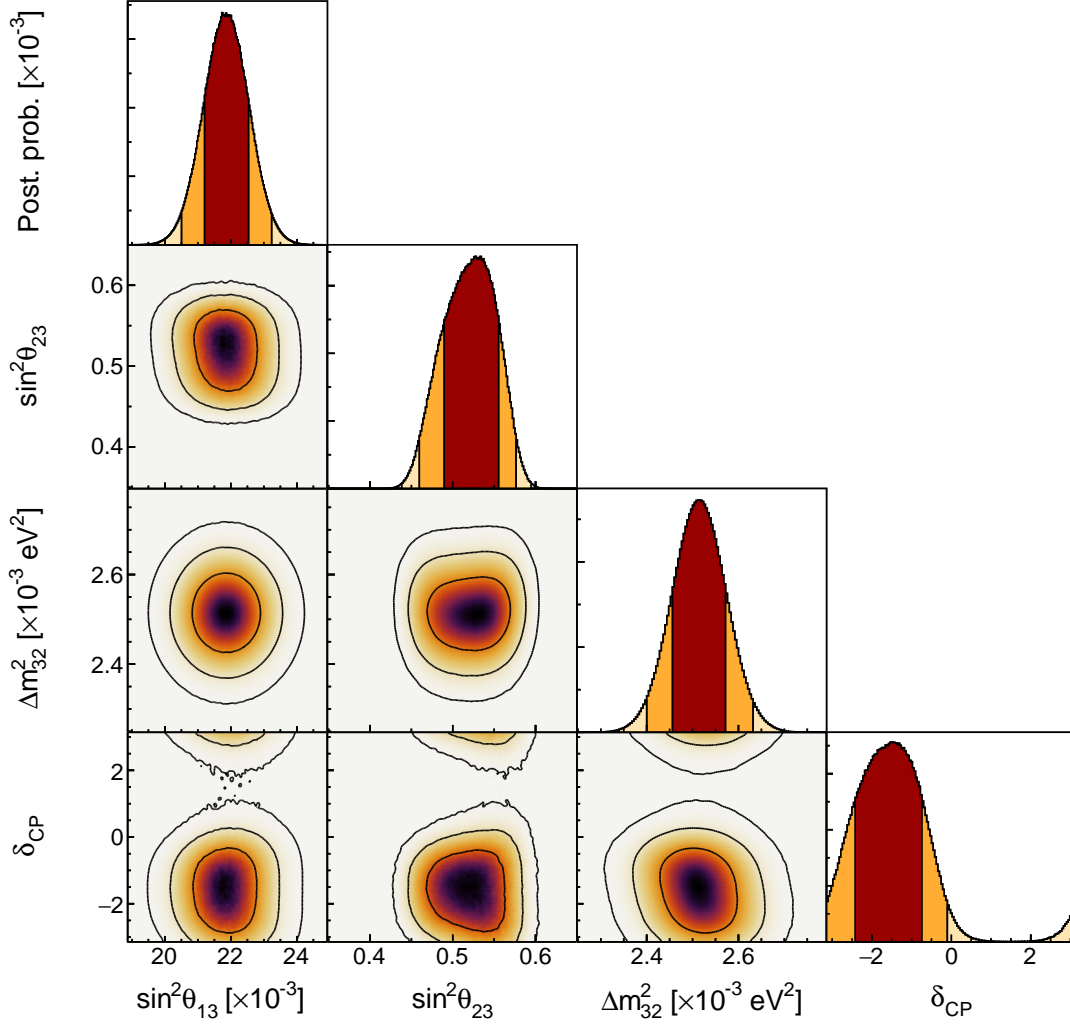


Figure 18: Triangle plots for the Asimov A analysis, with the reactor constraint applied, for the normal ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

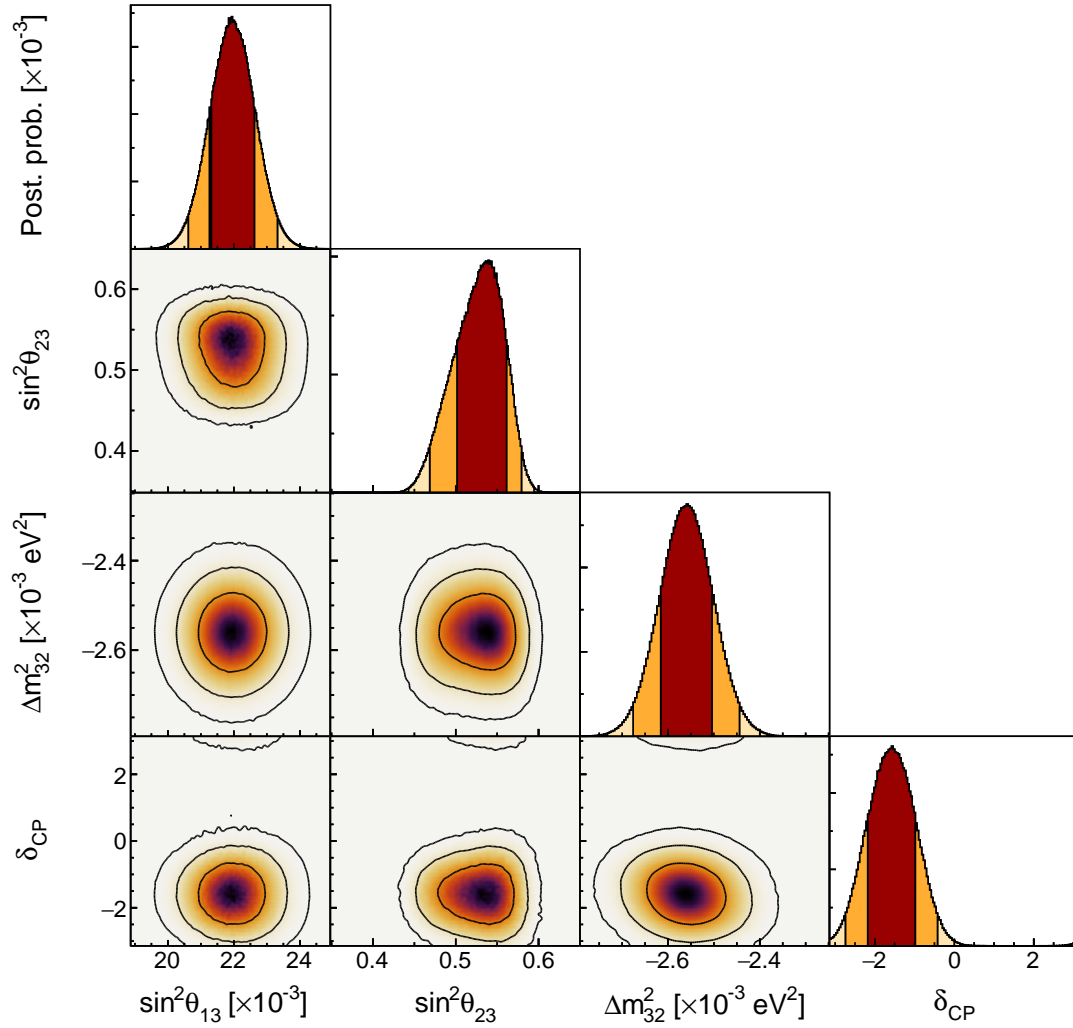


Figure 19: Triangle plots for the Asimov A analysis, with the reactor constraint applied, for the inverted ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

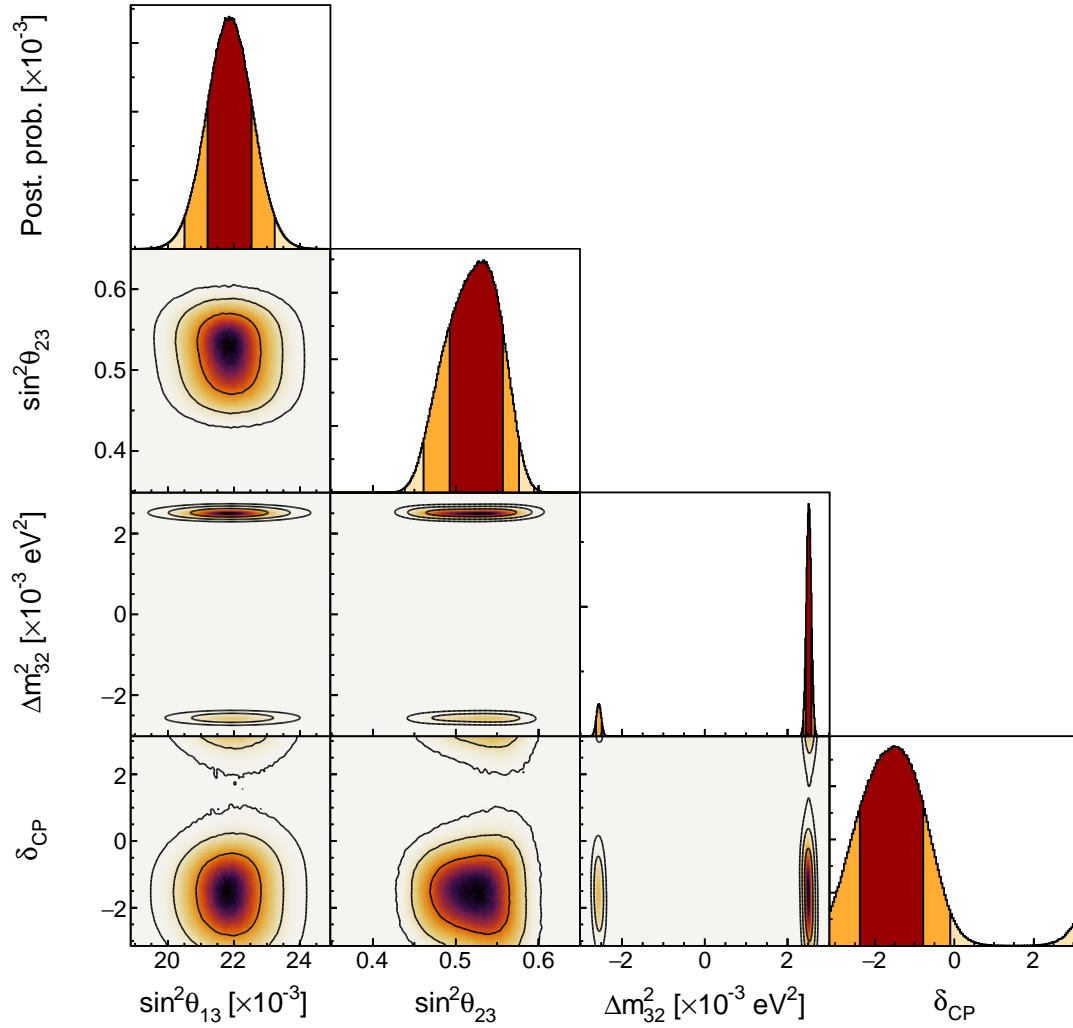


Figure 20: Triangle plots for the Asimov A analysis, with the reactor constraint applied, for both mass orderings, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

235 **3.1.3 Bayes factors for atmospheric parameters**

236 The posterior probabilities of the Asimov A analysis with applying the reactor constraint are
 237 summarized in **Tab. 2**. The Bayes factor of upper octant over lower octant $B(\text{UO}/\text{LO}) =$
 238 2.72, and the Bayes factor of normal ordering over inverted ordering $B(\text{NO}/\text{IO}) = 7.26$. The
 239 Bayes factor for the combination of upper octant and normal ordering is $B([\text{UO}+\text{NO}]/\text{Other}) =$
 1.73.

	$\sin^2 \theta_{23} < 0.5$	$\sin^2 \theta_{23} > 0.5$	Sum
NO ($\Delta m_{32}^2 > 0$)	0.25	0.63	0.88
IO ($\Delta m_{32}^2 < 0$)	0.02	0.10	0.12
Sum	0.27	0.73	1.00

Table 2: Posterior probabilities for the mass ordering and the octant in the Asimov A analysis, with the reactor constraint applied.

240

241 **3.2 Results without applying the reactor constraint**

242 This section shows the Asimov A analysis without applying the reactor constraint on $\sin^2 \theta_{13}$.
 243 The same Asimov A chains are used with a burn-in period of 80,000 steps per independent
 244 chain.

245 **3.2.1 Parameter constraints**

246 Each parameter constraint in this section is shown in normal ordering ($\Delta m_{32}^2 > 0$), inverted
 247 ordering ($\Delta m_{32}^2 < 0$), and over both orderings.

248 **One-dimensional posteriors** Fig. 21, Fig. 22, Fig. 23 and Fig. 24 show the marginalised
 249 one-dimensional posterior for δ_{CP} , Δm_{32}^2 , $\sin^2 \theta_{13}$ and $\sin^2 \theta_{23}$, respectively.

250 For the δ_{CP} result, a small region around $\delta_{CP} \sim \pi/2$ is excluded at 3σ confidence level.

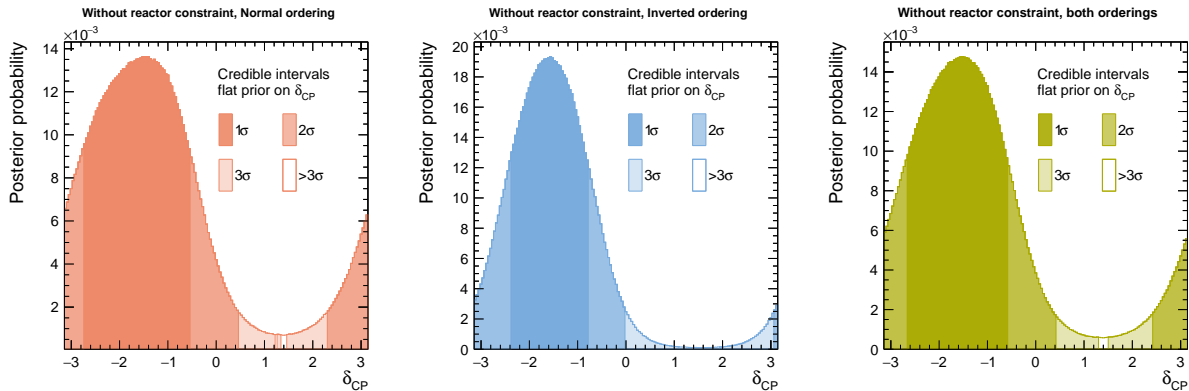


Figure 21: δ_{CP} for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

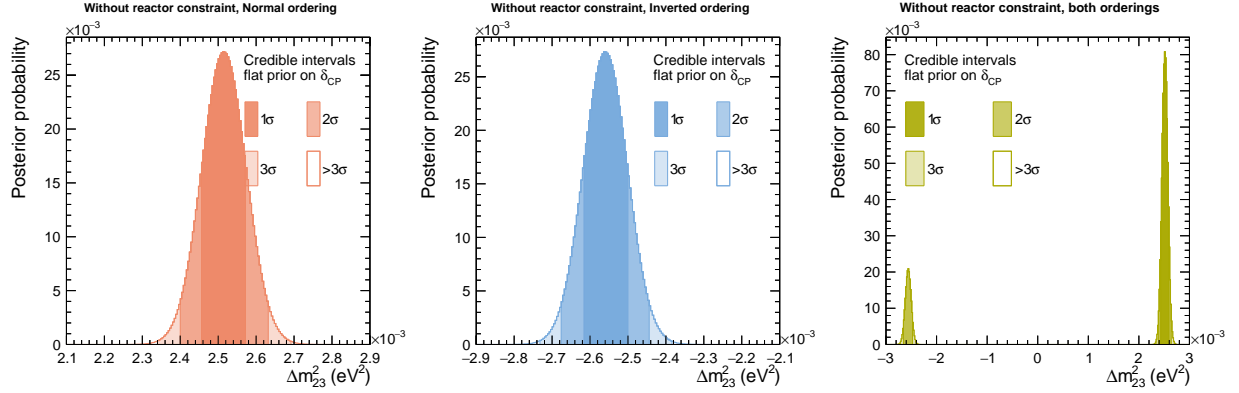


Figure 22: Δm_{32}^2 for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

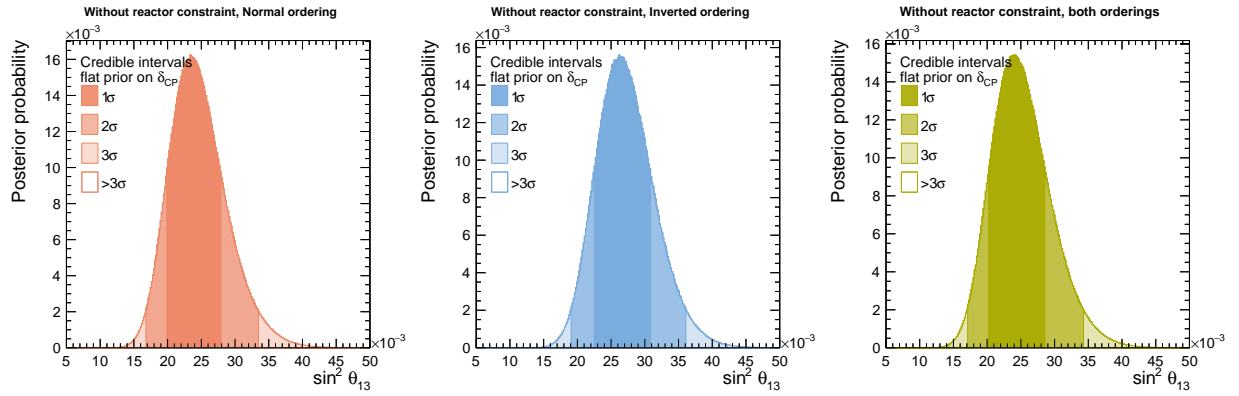


Figure 23: $\sin^2 \theta_{13}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

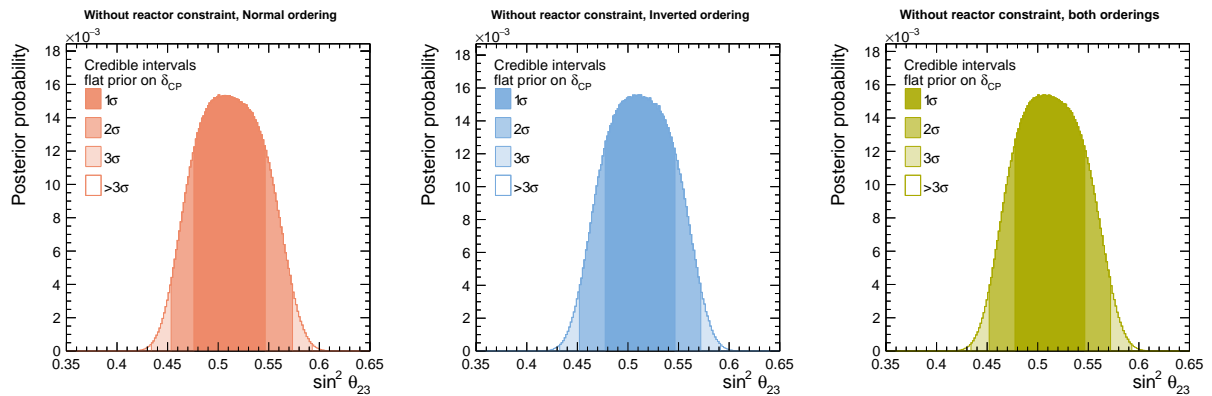


Figure 24: $\sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

251 **The Jarlskog invariant** Fig. 25 and Fig. 26 show the marginalised one-dimensional Jarl-
 252 skog invariant posteriors with flat prior on δ_{CP} and $\sin \delta_{\text{CP}}$ respectively. They are shown
 253 without reactor constraint for normal, inverted and both mass orderings. The positive J
 254 solution $J \sim 0.04$ is visible in the normal ordering and when marginalising over both or-
 255 derings, with a weakly bimodal posterior, and the strongest excluded region is between
 $J = [0.02, 0.035]$.

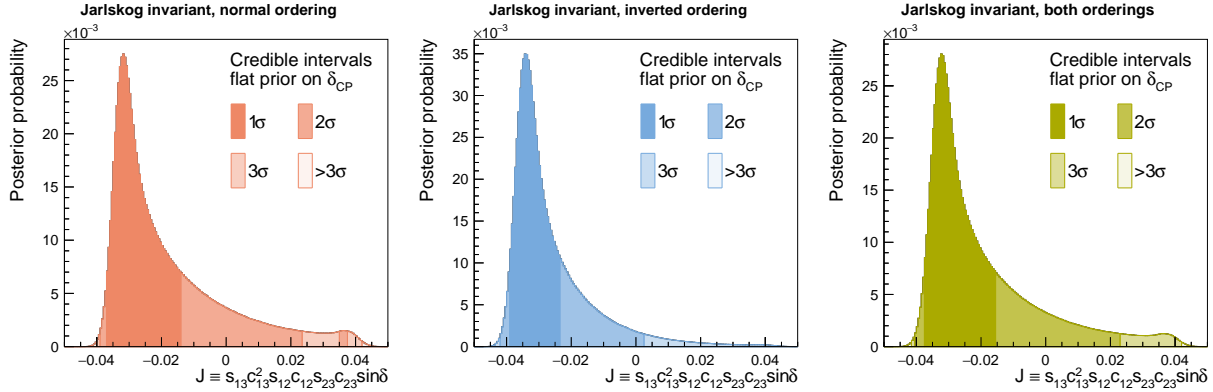


Figure 25: J for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

256

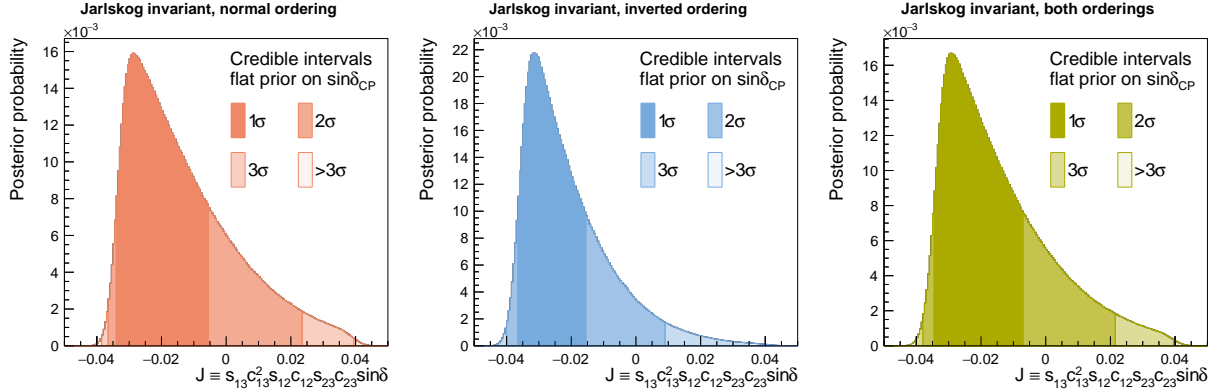


Figure 26: J for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

257 **Two-dimensional posteriors** Fig. 27, Fig. 28, Fig. 29, Fig. 30, Fig. 31, Fig. 32 and Fig. 33
 258 show the two-dimensional marginal posteriors for $\delta_{\text{CP}} - \sin^2 \theta_{13}$, $\delta_{\text{CP}} - \sin^2 \theta_{23}$, $\delta_{\text{CP}} - \Delta m_{32}^2$,
 259 $\Delta m_{32}^2 - \sin^2 \theta_{13}$, $\Delta m_{32}^2 - \sin^2 \theta_{23}$, $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ and $\sin^2 \theta_{23} - \Delta m_{32}^2$, respectively. These
 260 are shown separately for normal ordering, inverted ordering and both orderings, all applying
 261 the reactor constraint.

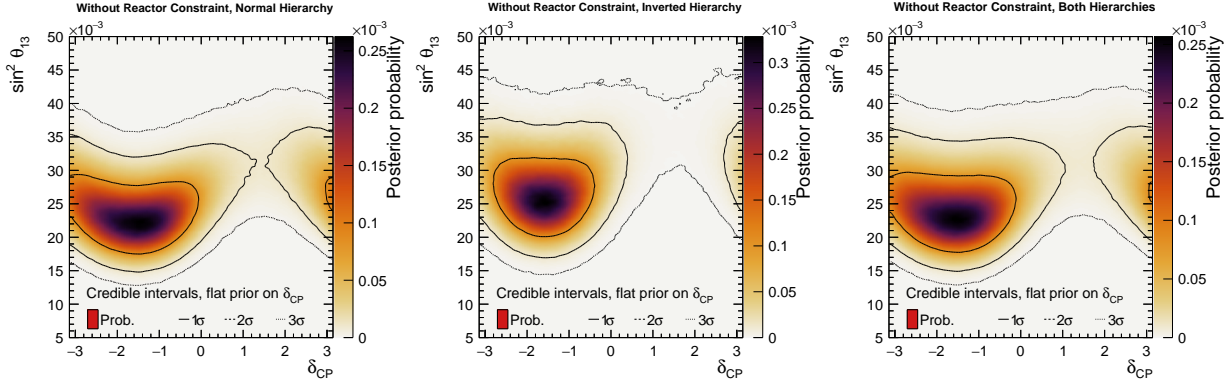


Figure 27: $\delta_{\text{CP}} - \sin^2 \theta_{13}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

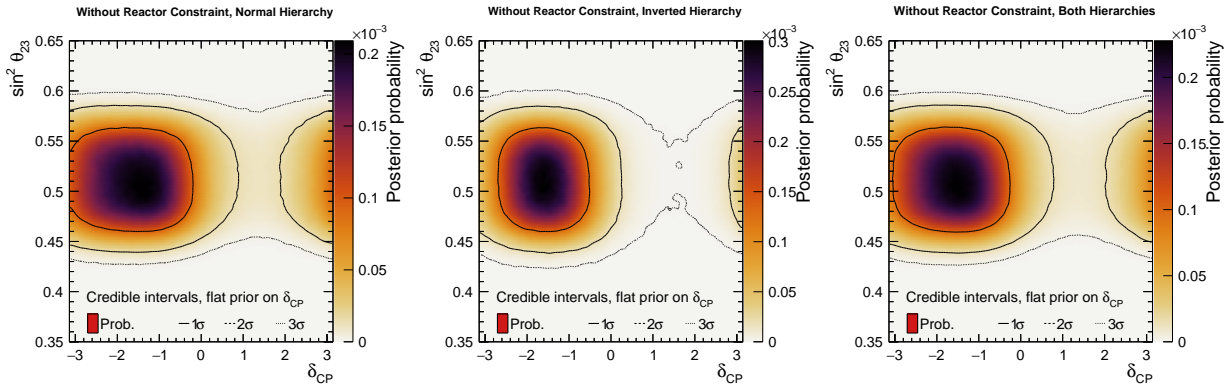


Figure 28: $\delta_{\text{CP}} - \sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

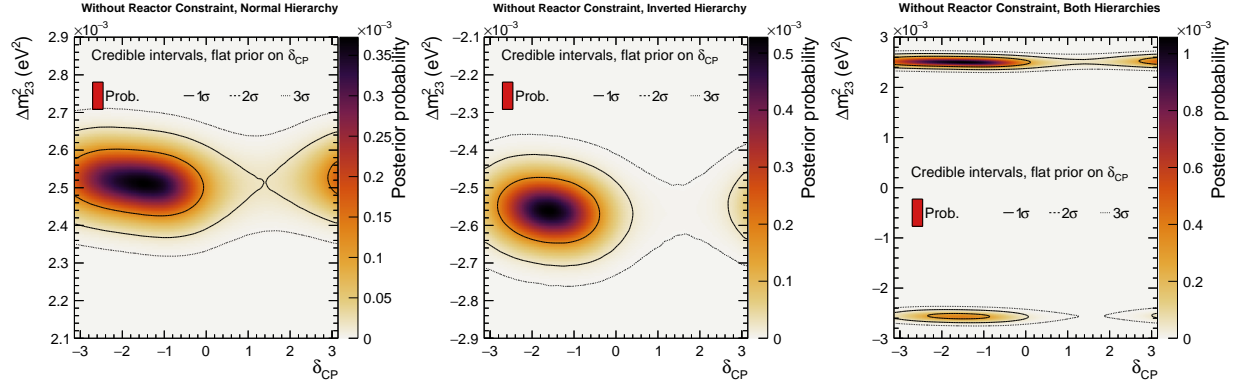


Figure 29: $\delta_{CP} - \Delta m_{32}^2$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

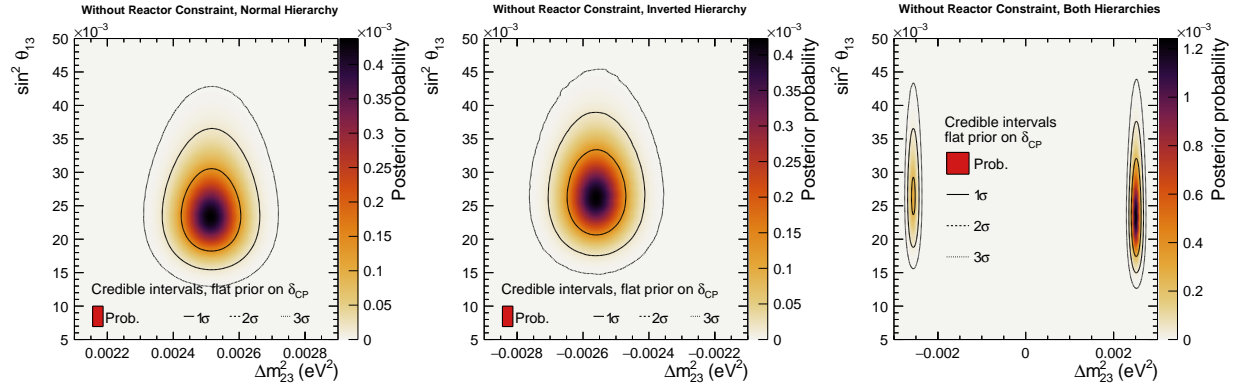


Figure 30: $\Delta m_{32}^2 - \sin^2 \theta_{13}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

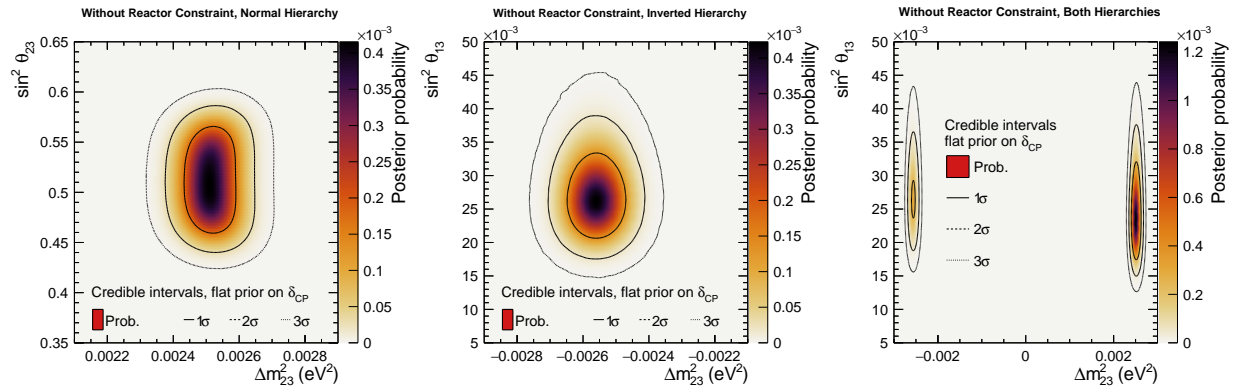


Figure 31: $\Delta m_{32}^2 - \sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

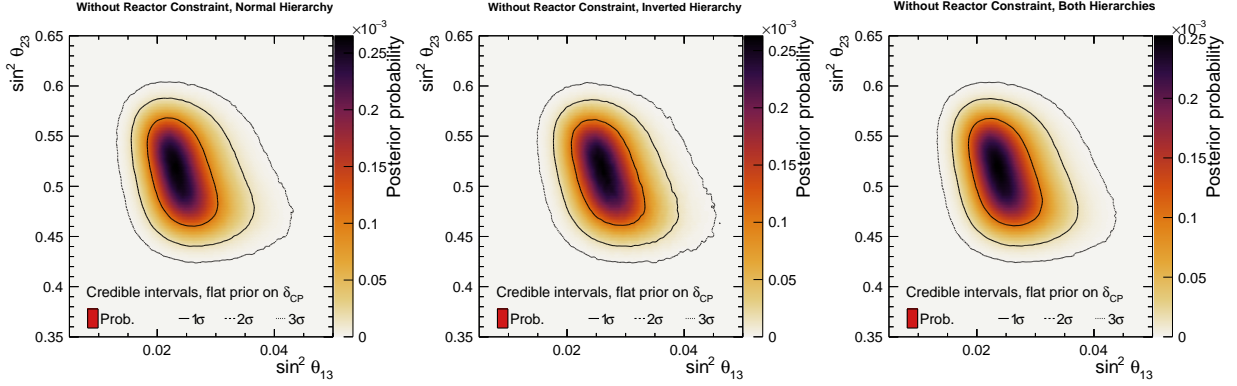


Figure 32: $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

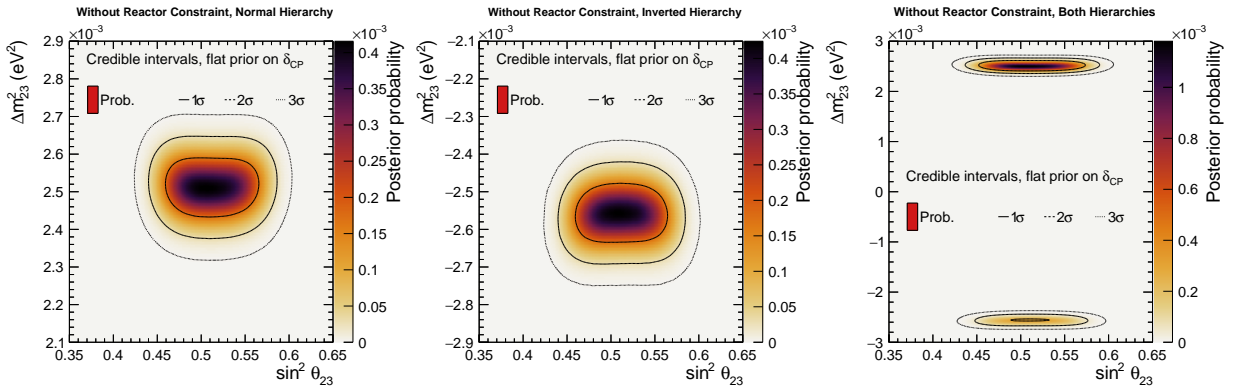


Figure 33: $\sin^2 \theta_{23} - \Delta m_{32}^2$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

262 **The Jarlskog invariant** Fig. 34 and Fig. 35 show the two-dimensional $J - \sin^2 \theta_{23}$ poste-
 263 riors for normal ordering, inverted ordering and both orderings, with a flat prior on δ_{CP} and
 264 $\sin \delta_{\text{CP}}$ respectively. Fig. 36 shows the two-dimensional $J - \delta_{\text{CP}}$ posteriors without reactor
 265 constraint for the orderings, with a flat prior on δ_{CP} . As with the one-dimensional case,
 266 the $J \sim 0.04$ solution is evident for the result with a flat prior on δ_{CP} , and the posterior
 267 is relatively flat in $\sin^2 \theta_{23}$ in this region. Interestingly, the $J = 0.04$ is not present for the
 posterior with a flat prior on $\sin \delta_{\text{CP}}$.

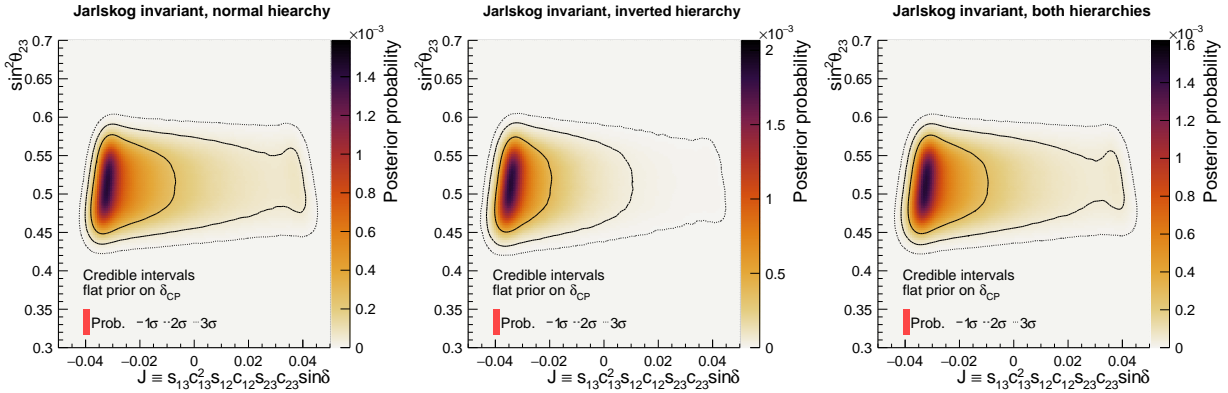


Figure 34: $J - \sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

268

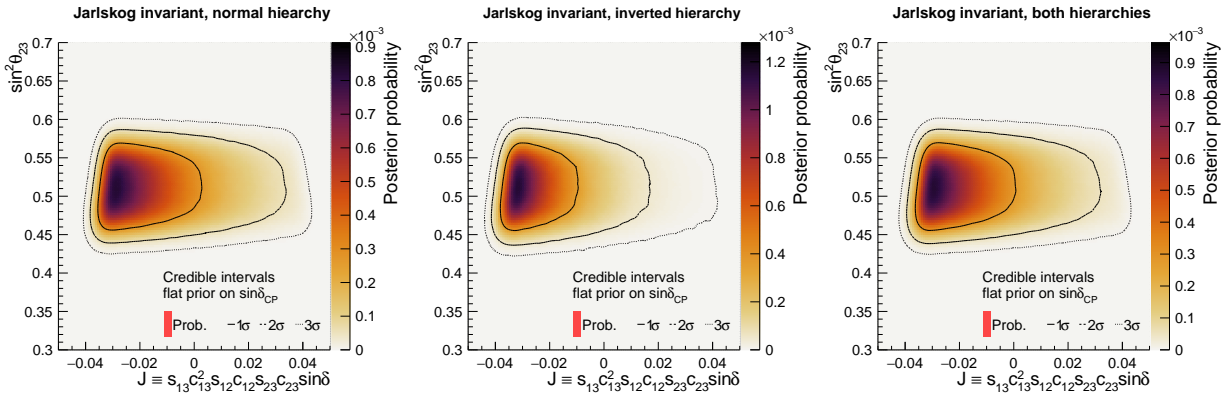


Figure 35: $J - \sin^2 \theta_{23}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

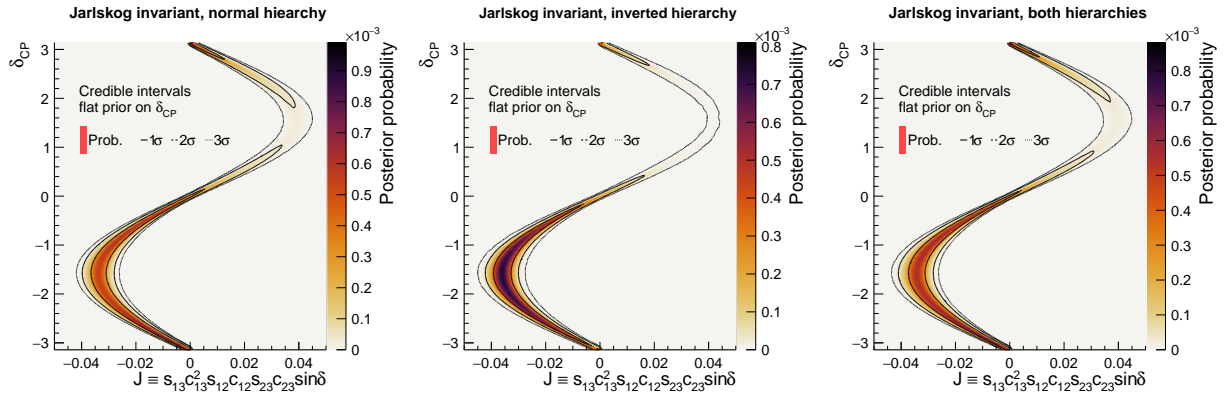


Figure 36: $J - \delta_{CP}$ for the Asimov A analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

269 **3.2.2 Triangle plots**

270 **Fig. 37**, **Fig. 38**, and **Fig. 39** show the one and two-dimensional oscillation parameters'
271 posteriors, displayed in triangular form. These do not apply the reactor constraint, and use
272 a flat prior on δ_{CP} .

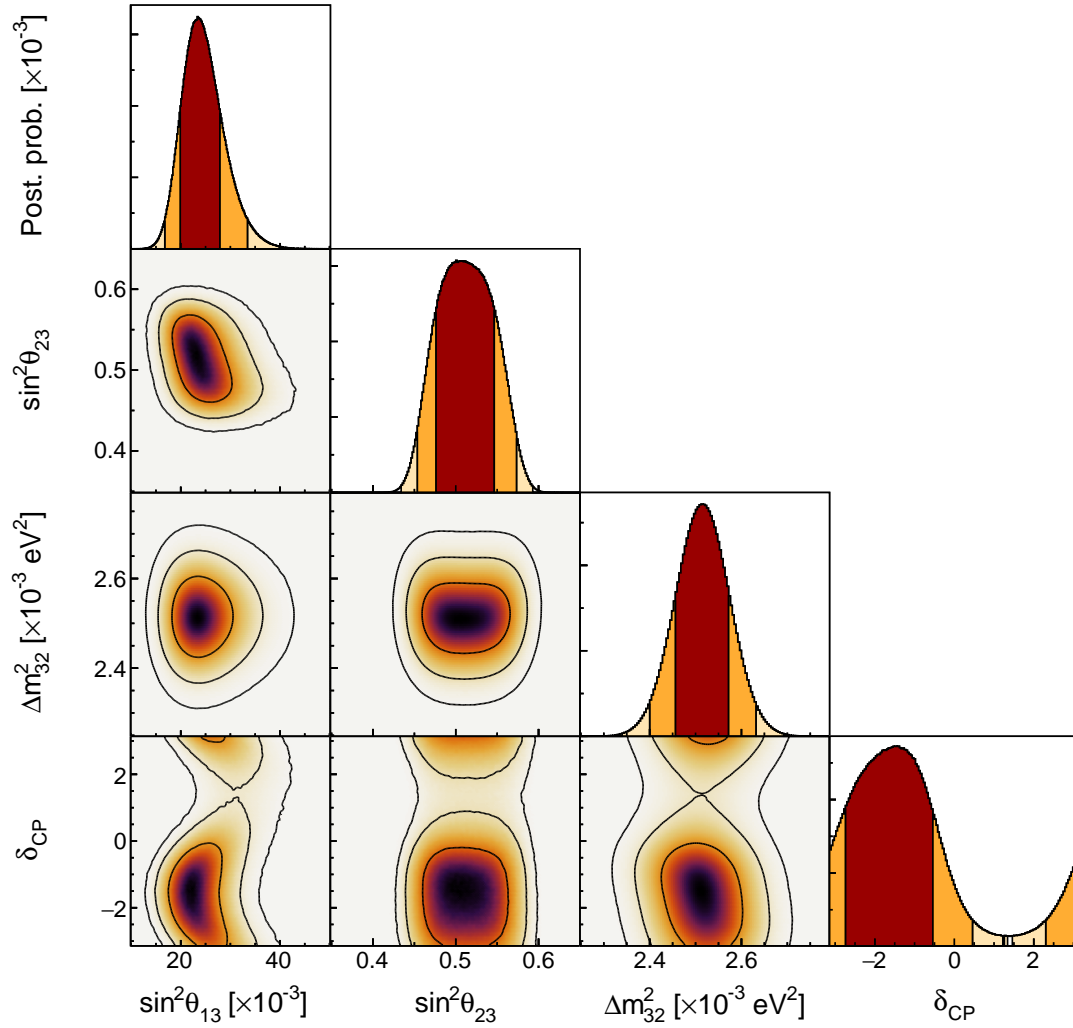


Figure 37: Triangle plots for the Asimov A analysis, without the reactor constraint applied, for the normal ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

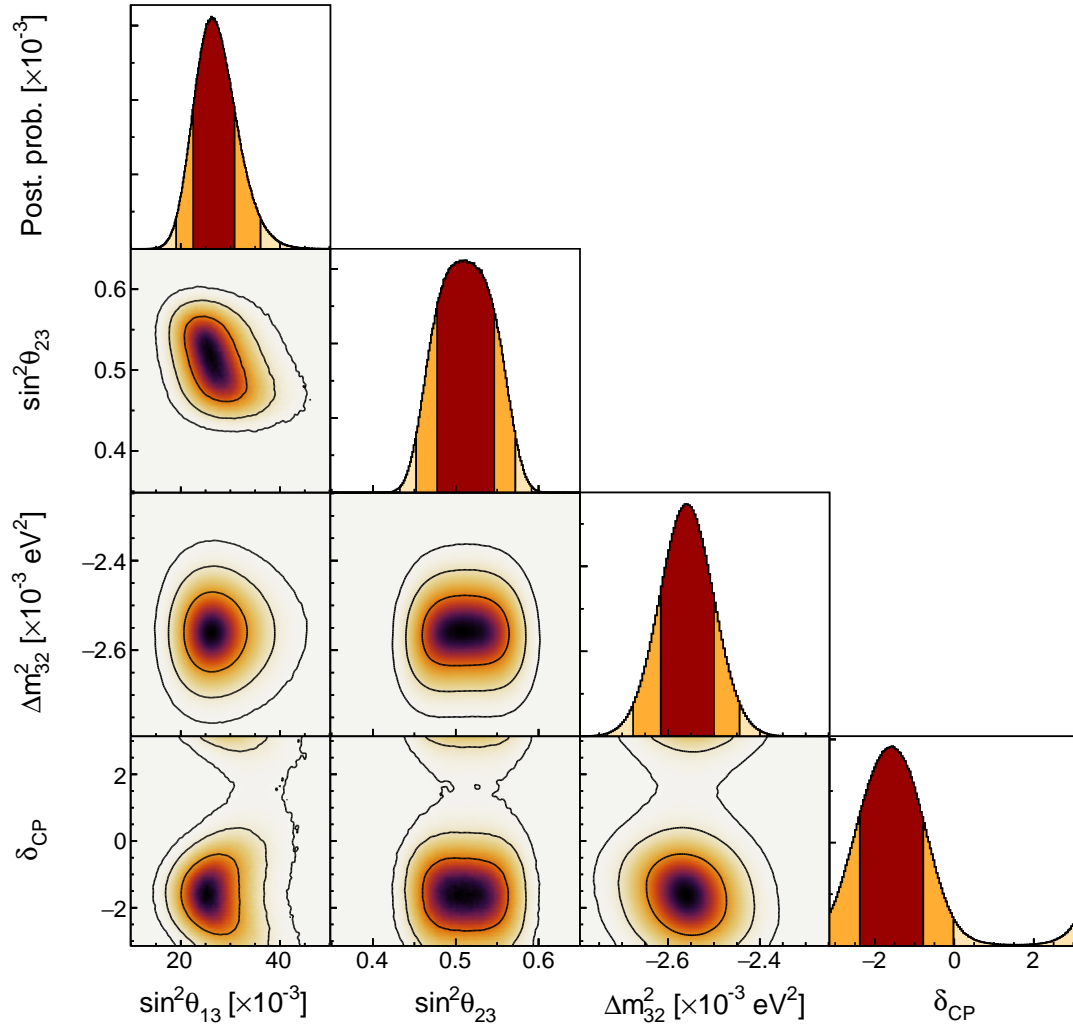


Figure 38: Triangle plots for the Asimov A analysis, without the reactor constraint applied, for the inverted ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

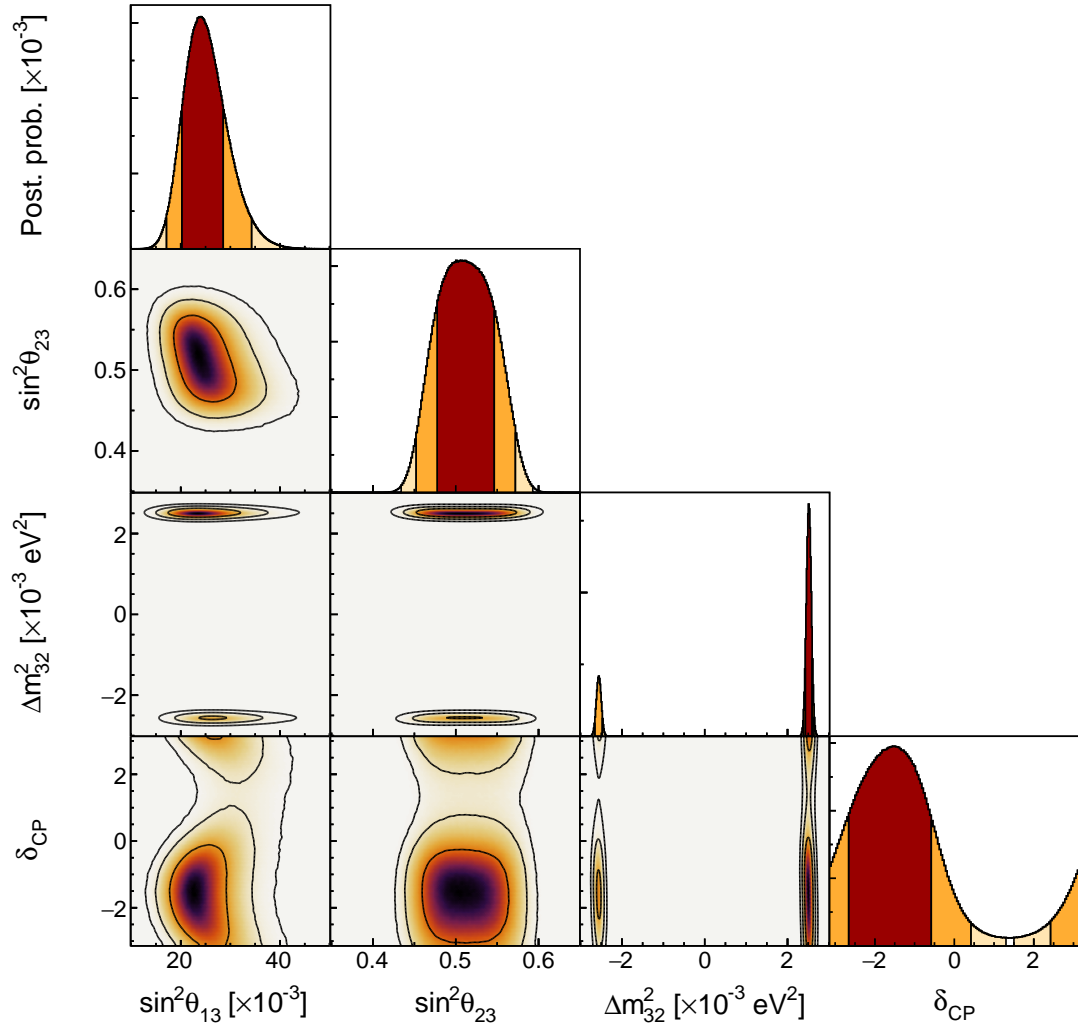


Figure 39: Triangle plots for the Asimov A analysis, without the reactor constraint applied, for both orderings, with a flat prior on δ_{CP} . The 1, 2, and 3σ credible intervals are overlaid on the posterior.

273 **3.2.3 Bayes factors for atmospheric parameters**

274 The posterior probabilities of the Asimov A analysis without applying the reactor constraint
 275 are summarized in **Tab. 3**. The Bayes factor of upper octant over lower octant $B(\text{UO}/\text{LO}) =$
 276 1.67, and the Bayes factor of normal ordering over inverted ordering $B(\text{NO}/\text{IO}) = 3.98$. The
 277 Bayes factor for the combination of upper octant and normal ordering is very weak, with
 278 $B([\text{UO}+\text{NO}]/\text{Other}) = 0.98$.

	$\sin^2 \theta_{23} < 0.5$	$\sin^2 \theta_{23} > 0.5$	Sum
NO	0.30	0.50	0.80
IO	0.08	0.13	0.20
Sum	0.38	0.62	1.00

Table 3: Posterior probabilities for the mass ordering and the octant in the Asimov A analysis, without the reactor constraint applied.

279 **3.3 Bayes factors for CP violation**

280 **Tab. 4** summarises the posterior probabilities of $\sin \delta_{\text{CP}} \geq 0$ and $\sin \delta_{\text{CP}} < 0$ with and without
 281 the reactor constraint. The Bayes factor $B[(\sin \delta_{\text{CP}} < 0)/(\sin \delta_{\text{CP}} \geq 0)]$ is 24.5(6.3) with(out)
 282 the reactor constraint applied.

	$\sin \delta_{\text{CP}} \geq 0$	$\sin \delta_{\text{CP}} < 0$	Sum
No RC	0.137	0.863	1.000
RC	0.039	0.961	1.000

Table 4: Posterior probabilities for $\sin \delta_{\text{CP}}$ in the Asimov A analysis with and without reactor constraint

283 4 Results from fitting real data

284 In this section we move from Asimov studies to analysing the real data at all detectors’
285 selections; T2K ND280, T2K SK, and SK atmospheric. After inspecting traces and auto-
286 correlations, the same MCMC configuration is used for the data analysis as with the analysis
287 of Asimov A “data”. For the data analysis, part of the steps are collected from ORNL
288 Summit and part are from Alliance Canada. 277,661,111 steps were collected in total, with
289 good acceptance rates ($\sim 15 - 19\%$) and a burn-in of 80,000 steps, with the number of steps
290 after burn-in being 230,940,527. Similar to the Asimov A analysis, the auto-correlation
291 goal was reaching 0.2 by lag 20,000. The trace of the total negative log likelihood for all
292 independent chains are shown in [Fig. 40](#). Results shown in this section are produced using
293 steps collected after burn-in.

294 As with the Asimov A analysis, smearing from the fake-data studies TN457 [\[11\]](#) have
295 been added, and are applied by randomly smearing Δm_{32}^2 by $3.6 \times 10^{-5} \text{ eV}^2$.

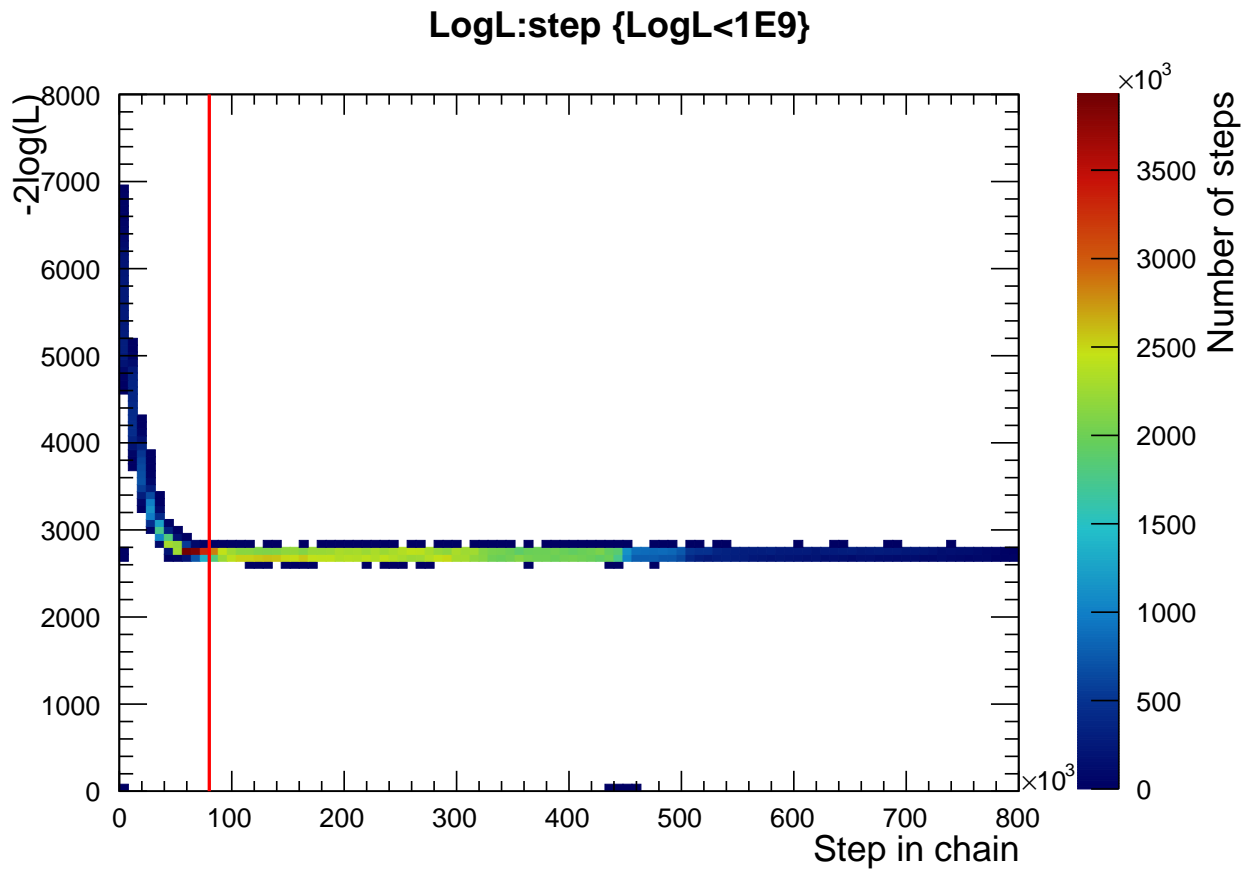


Figure 40: Collection of the traces of the negative log total likelihood from data fit chains. The red line is the cut for post burn-in at 80,000 steps. Note that a few steps post burn-in (step $\sim 450 \times 10^3$) have $-\text{LogL} = 0$ near zero. They appeared after processing the raw chain output files, and were not present in the original files. These steps are excluded in the analysis.

296 4.1 Results applying the reactor constraint

297 For these results, the reactor constraint on $\sin^2 \theta_{13}$ is used with the same value as the T2K
 298 OA2020 analysis TN393 [1] from the PDG [10]; $\sin^2 \theta_{13} = 0.0218 \pm 0.0007$.

299 4.1.1 Parameter constraints, prior flat in δ_{CP}

300 Each parameter constraint in this section is shown in normal ordering ($\Delta m_{32}^2 > 0$), inverted
 301 ordering ($\Delta m_{32}^2 < 0$), and over both orderings.

302 **One-dimensional posteriors** Fig. 41, Fig. 42, Fig. 43 and Fig. 44 show the marginalised
 303 one-dimensional posterior for δ_{CP} , Δm_{32}^2 , $\sin^2 \theta_{13}$ and $\sin^2 \theta_{23}$, respectively.

304 The δ_{CP} result in Fig. 41 show $\delta_{\text{CP}} = 0$ is excluded at greater than 2σ , and $\sin \delta_{\text{CP}} = 0$
 305 excluded at more than 2σ for all choices of ordering (normal, inverted and both). As further
 306 investigated in Tab. 5, after applying the interval shifts from the fake data study (Tab. 4
 307 in TN457 [11]), the 2σ intervals for both normal and inverted ordering exclude $\delta_{\text{CP}} = 0$.
 308 These results are studied further with the Jarlskog invariant, including the impact of the
 309 prior being flat in δ_{CP} or $\sin \delta_{\text{CP}}$.

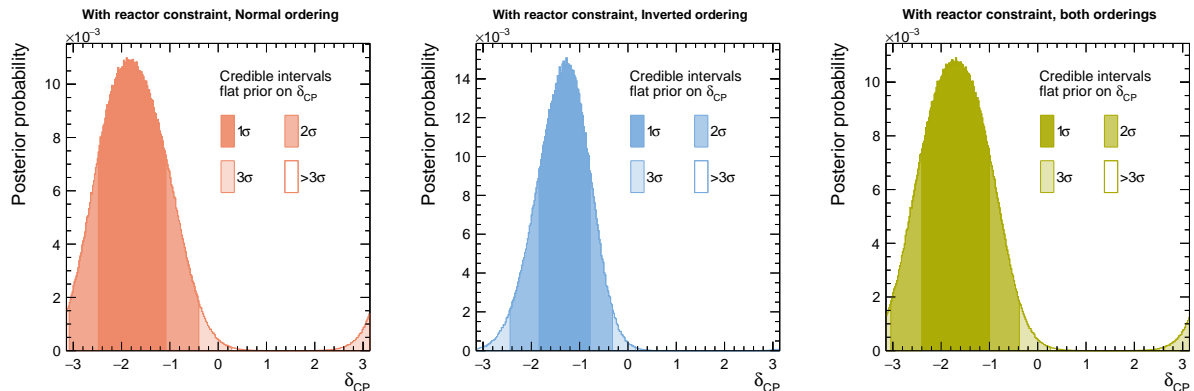


Figure 41: δ_{CP} for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

310 The Δm_{32}^2 result in Fig. 42 shows a preference for the normal ordering with a Bayes
 311 factor of 7.33 calculated in Sec. 4.1.4.

312 The posterior for $\sin^2 \theta_{23}$ in Fig. 44 shows a very weak preference for the upper octant
 313 when marginalising over both mass orderings, with the lower octant included in the 1σ
 314 credible interval. For the normal ordering, there is almost no octant preference at all. For
 315 the inverted ordering the upper octant is more strongly preferred, although the 1σ credible

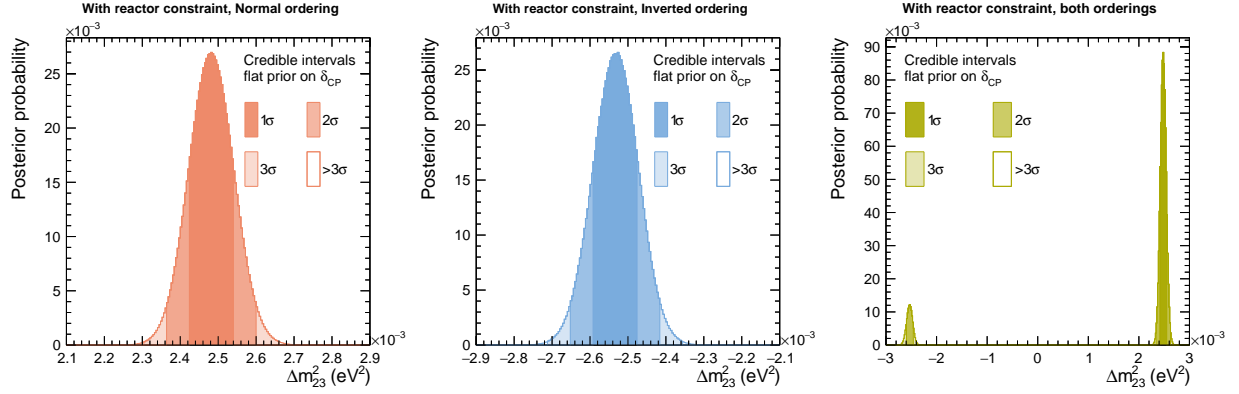


Figure 42: Δm_{32}^2 for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

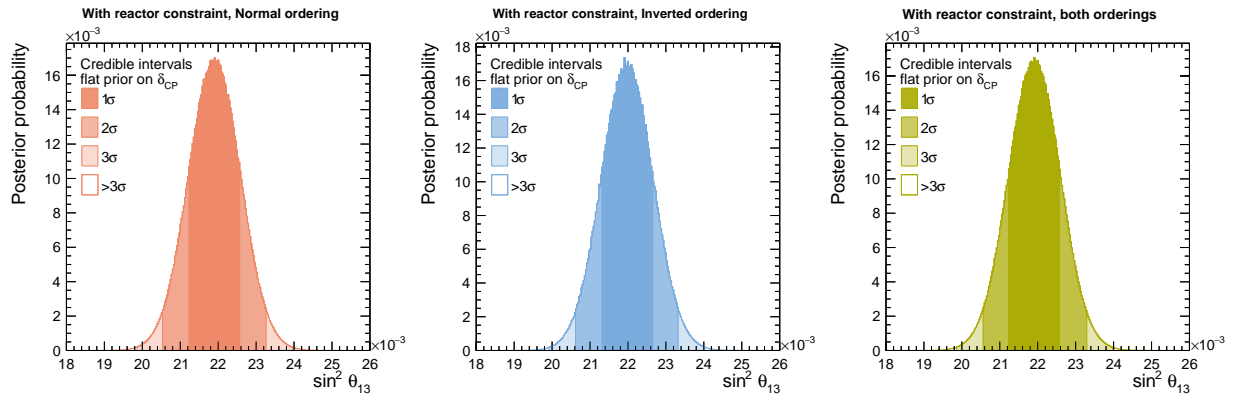


Figure 43: $\sin^2 \theta_{13}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

316 interval just excludes maximal mixing around $\sin^2 \theta_{23} \sim 0.513$. This is confirmed by the
 317 Bayes factor studies, which finds $B(UO/LO) = 1.78$ for both orderings, $B(UO/LO)|_{NO} =$
 318 1.63 , and $B(UO/LO)|_{IO} = 3.33$.

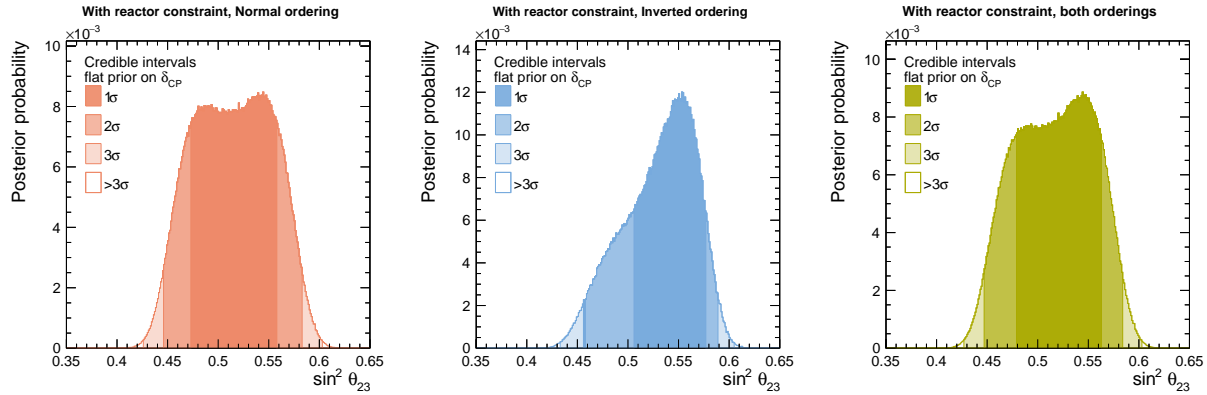


Figure 44: $\sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

Normal Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-2.50127	-1.10056	-3.06155	-0.380193	2.58131	0.260132
Increment	-0.034	0.122	-0.025	0.081	-0.024	0.059
Fit+Increment	-2.53527	-0.97856	-3.08655	-0.299193	2.55731	0.319132
Pass	TRUE		TRUE		FALSE	

Inverted Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-1.82092	-0.780396	-2.42123	-0.300152	-3.02153	0.140071
Increment	-0.034	0.122	-0.025	0.081	-0.024	0.059
Fit+Increment	-1.85492	-0.658396	-2.44623	-0.219152	-3.04553	0.199071
Pass	TRUE		TRUE		FALSE	

Both Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-2.42123	-1.02052	-3.02153	-0.380193	2.62133	0.260132
Increment	-0.034	0.122	-0.025	0.081	-0.024	0.059
Fit+Increment	-2.45523	-0.89852	-3.04653	-0.299193	2.59733	0.319132
Pass	TRUE		TRUE		FALSE	

Table 5: Summary table of δ_{CP} credible interval edges for flat prior of δ_{CP} from the data fit (with reactor constraint) and the edges after applying the shifts from the fake data study TN457 [11] for normal ordering (upper table), inverted ordering (middle table) and both orderings (lower table) respectively. 1σ , 2σ and 3σ intervals are investigated. “Fit” refers to the credible interval edges from MaCh3 data fit results with reactor constraint; “Increment” refers to the extreme shift values from the fake data study results (Table 4 of TN457 v1.3 [11]); “Fit+Increment” refers to the credible interval edges after applying the shift values from fake data studies. The criteria of “Pass” is to check whether the credible interval after applying the shift contains any one of the δ_{CP} values $-\pi, 0, \pi$. “LE” refers to the lower edge of the interval and “UE” refers to the upper edge. Note that “LE” is not necessarily smaller than “UE” because of wrapping through $-\pi = \delta_{\text{CP}} = \pi$.

Normal Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-2.88146	-0.700355	3.02153	-0.180091	2.52128	0.380193
Increment	-0.047	0.168	-0.02	0.086	-0.02	0.054
Fit+Increment	-2.92846	-0.532355	3.00153	-0.094091	2.50128	0.434193
Pass	TRUE		FALSE		FALSE	

Inverted Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-2.08106	-0.400203	-2.66135	-0.140071	-3.14159	0.300152
Increment	-0.047	0.168	-0.02	0.086	-0.02	0.054
Fit+Increment	-2.12806	-0.232203	-2.68135	-0.054071	3.12159	0.354152
Pass	TRUE		TRUE		FALSE	

Both Ordering	1sig		2sig		3sig	
	LE	UE	LE	UE	LE	UE
Fit	-2.82143	-0.660335	3.04154	-0.180091	2.54129	0.380193
Increment	-0.047	0.168	-0.02	0.086	-0.02	0.054
Fit+Increment	-2.86843	-0.492335	3.02154	-0.094091	2.52129	0.434193
Pass	TRUE		FALSE		FALSE	

Table 6: Summary table of δ_{CP} credible interval edges for flat prior of $\sin \delta_{\text{CP}}$ from the data fit (with reactor constraint) and the edges after applying the shifts from the fake data study TN457 [11] for normal ordering (upper table), inverted ordering (middle table) and both orderings (lower table) respectively. 1σ , 2σ and 3σ intervals are investigated. “Fit” refers to the credible interval edges from MaCh3 data fit results with reactor constraint; “Increment” refers to the extreme shift values from the fake data study results (Table 4 of TN457 v1.3 [11]); “Fit+Increment” refers to the credible interval edges after applying the shift values from fake data studies. The criteria of “Pass” is to check whether the credible interval after applying the shift contains any one of the δ_{CP} values $-\pi, 0, \pi$. “LE” refers to the lower edge of the interval and “UE” refers to the upper edge. Note that “LE” is not necessarily smaller than “UE” because of wrapping through $-\pi = \delta_{\text{CP}} = \pi$. For simplification, the interval edges for a specific sigma are the most extreme values among all connected intervals.

319 **4.1.2 Parameter constraints, prior flat in $\sin \delta_{CP}$**

320 The results until now have applied a flat prior on δ_{CP} , as is commonplace in T2K and SK
 321 analyses. However, the variable which the flat prior is being applied has the ability to shape
 322 the posterior, especially if the constraint from the data on the posterior is weak. For example,
 323 a prior flat in δ_{CP} is not the same as a prior flat in $\sin \delta_{CP}$, since the former down-weights
 324 probabilities around $\pm\pi/2$. The $P(\nu_\mu \rightarrow \nu_e) - P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ has dependence on $\sin \delta_{CP}$, not
 325 δ_{CP} . Hence, studying the impact of the shape of the prior on δ_{CP} is important, especially
 326 when intending to make definitive statements on particular credible intervals. All results in
 327 this section applies the reactor constraint on $\sin^2 \theta_{13}$.

328 **Fig. 45** shows the one-dimensional posteriors for each mass ordering for δ_{CP} with a prior
 329 flat in $\sin \delta_{CP}$. Compared to the prior flat in δ_{CP} , this prior choice enforces a strong zero prior
 330 at the maximal CP violation points, $\cos \delta_{CP} = 0$, i.e. when $\delta_{CP} = \pm\pi/2$. Hence, the credible
 331 intervals are discontinuous around $\delta_{CP} = \pi/2$, which is near the maximal posterior density
 when a prior flat in δ_{CP} was used.

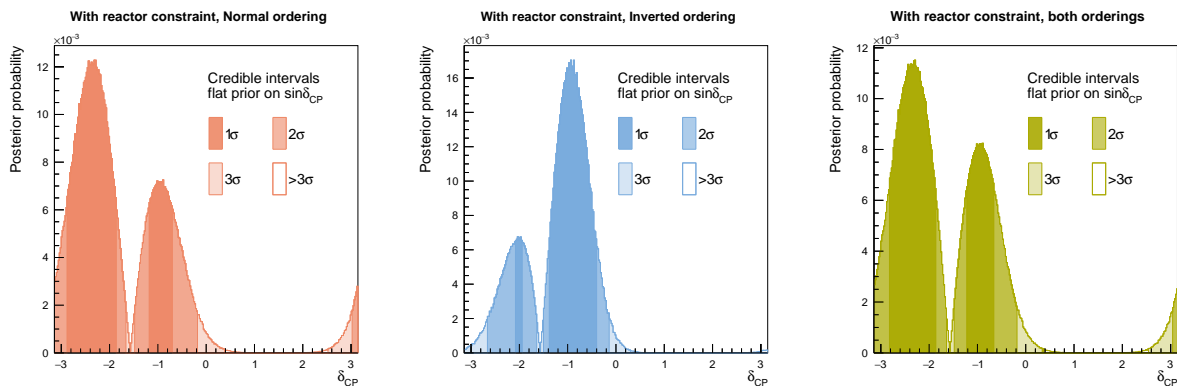


Figure 45: δ_{CP} for the real data analysis, with the reactor constraint applied and a flat prior on $\sin \delta_{CP}$, for normal (left), inverted (center), and both (right) orderings.

332

333 **Fig. 46** and **Fig. 47** show the Δm_{32}^2 and $\sin^2 \theta_{13}$ posteriors, respectively, which show little
 334 dependence on this choice of priors, owing largely to the lack of correlation with δ_{CP} .

335 **Fig. 48** shows the results for $\sin^2 \theta_{23}$, where we see a slightly more peaked distribution
 336 for the normal ordering, but otherwise looks the same.

337 The one-dimensional posteriors are compared in detail for the two choices of priors in
 338 **Sec. 4.3**.

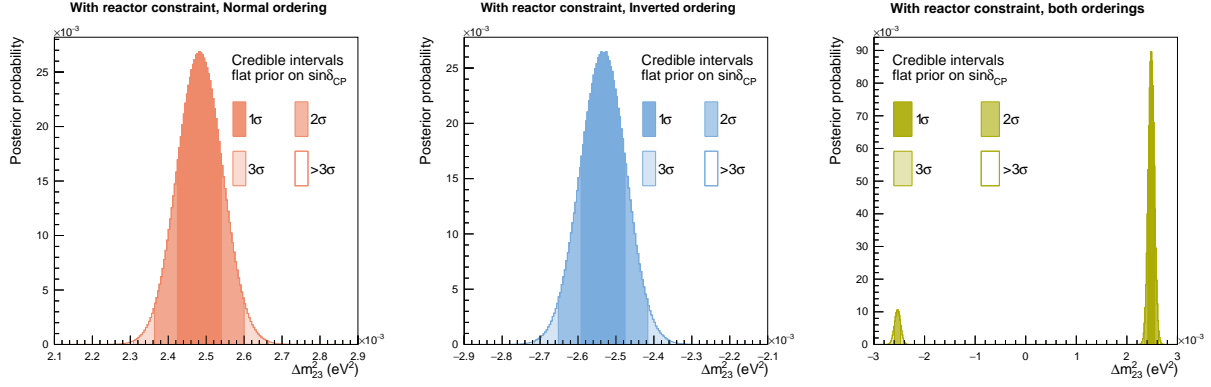


Figure 46: Δm_{32}^2 for the real data analysis, with the reactor constraint applied and a flat prior on $\sin \delta_{CP}$, for normal (left), inverted (center), and both (right) orderings.

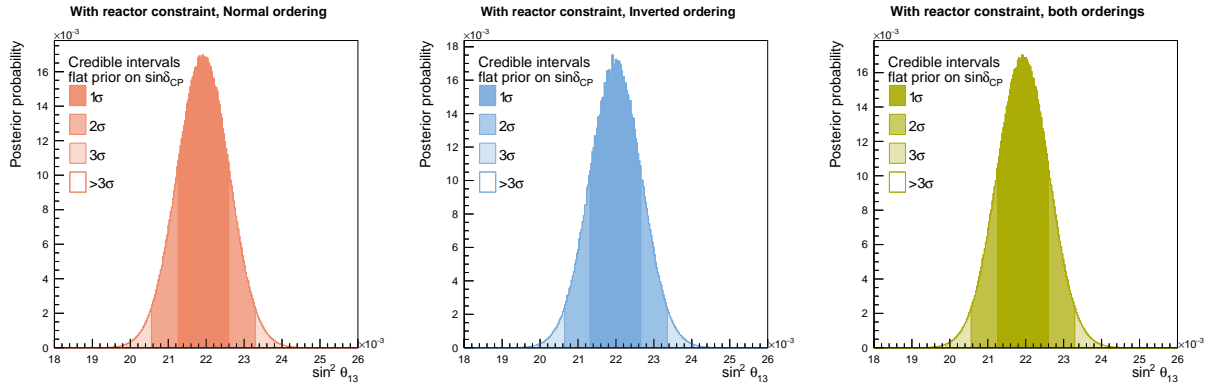


Figure 47: $\sin^2 \theta_{13}$ for the real data analysis, with the reactor constraint applied and a flat prior on $\sin \delta_{CP}$, for normal (left), inverted (center), and both (right) orderings.

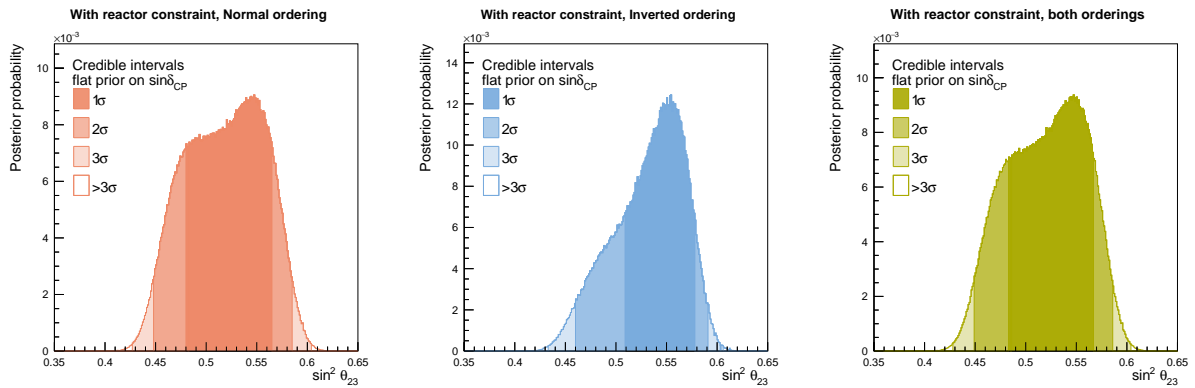


Figure 48: $\sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied and a flat prior on $\sin \delta_{CP}$, for normal (left), inverted (center), and both (right) orderings.

339 **The Jarlskog invariant** Fig. 49 and Fig. 50 show the marginalised Jarlskog invariant (J)
 340 posterior for normal ordering, inverted ordering and both ordering, shown separately with a
 341 flat prior on δ_{CP} and flat prior on $\sin \delta_{\text{CP}}$.

342 For the J with flat prior on δ_{CP} , $J = 0$ is comfortably excluded at more than 2σ credible
 343 level for all choices of mass ordering. With a prior flat in $\sin \delta_{\text{CP}}$, the constraint on J lessens,
 and $J = 0$ is just excluded at 2σ credible interval, driven by the normal ordering.

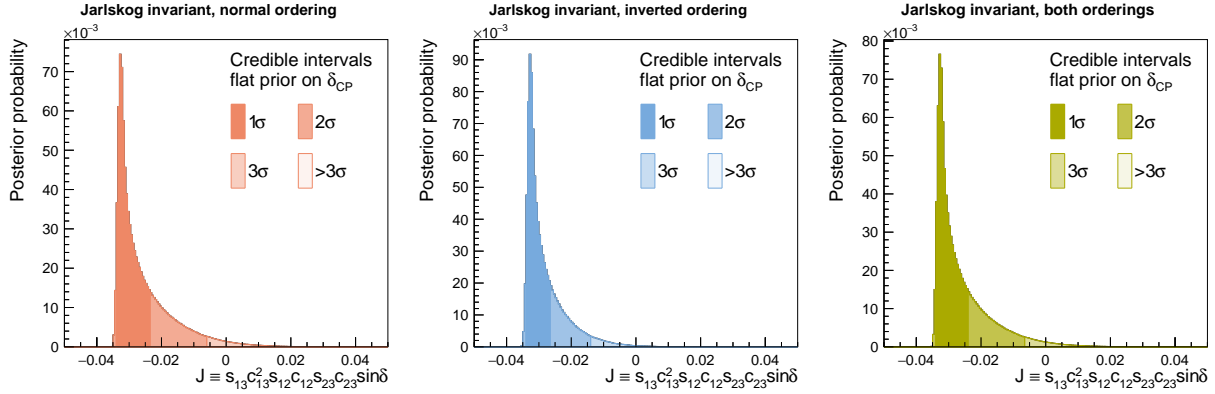


Figure 49: J for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

344

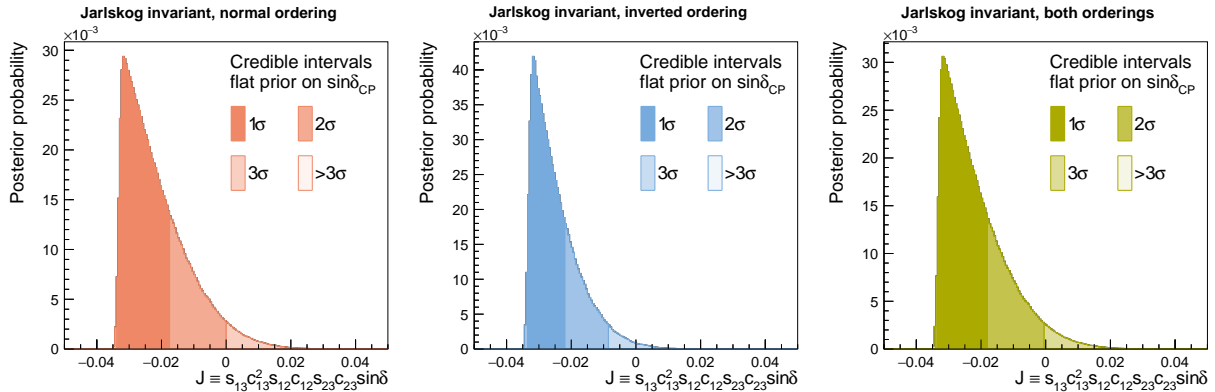


Figure 50: J for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

345 **Two-dimensional posteriors** Fig. 51, Fig. 52, Fig. 53, Fig. 54, Fig. 55, Fig. 56 and Fig. 57
 346 show the two-dimensional marginal posteriors for $\delta_{\text{CP}} - \sin^2 \theta_{13}$, $\delta_{\text{CP}} - \sin^2 \theta_{23}$, $\delta_{\text{CP}} - \Delta m_{32}^2$,
 347 $\Delta m_{32}^2 - \sin^2 \theta_{13}$, $\Delta m_{32}^2 - \sin^2 \theta_{23}$, $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ and $\sin^2 \theta_{23} - \Delta m_{32}^2$, respectively. These
 348 are shown separately for normal ordering, inverted ordering and both orderings, all applying
 349 the reactor constraint.

350 As with the Asimov A studies, the number of MCMC steps and the chosen binning is not
 351 sufficient to make reliable 3σ statements, and the 3σ credible interval should only be used
 352 to guide the reader for the approximate exclusion, and no accurate statistical statements
 353 should be made about 3σ exclusion.

354 In general, the results are roughly compatible with the expected sensitivity from the
 355 Asimov A analysis.

356 Compared to the main T2K-only results (see Fig 80 in TN393 [1]), the correlation between
 357 $\sin^2 \theta_{23}$ and $\sin^2 \theta_{13}$ in Fig. 56, and $\sin^2 \theta_{23}$ and δ_{CP} in Fig. 52, are relaxed in the joint analysis,
 358 and they are almost independent. Similar can be said about $\sin^2 \theta_{23}$ and Δm_{32}^2 , likely due
 to the maximum likelihood being around maximal mixing.

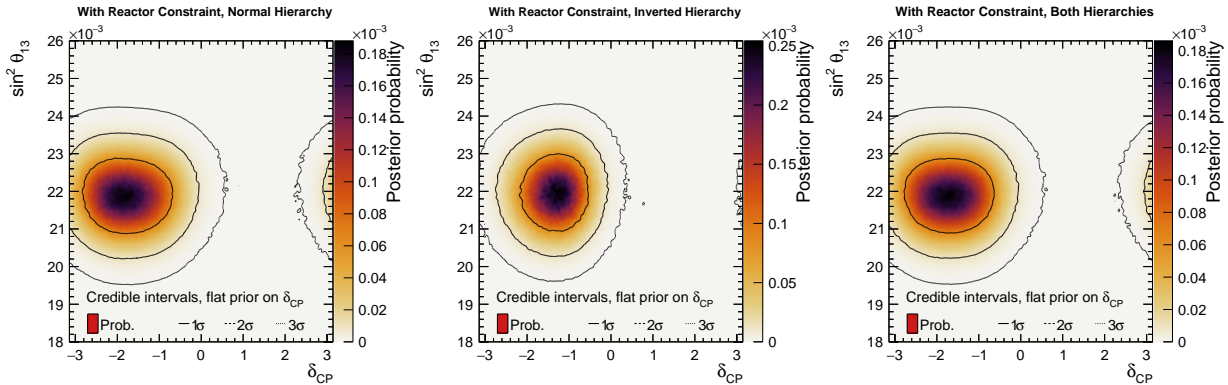


Figure 51: $\delta_{\text{CP}} - \sin^2 \theta_{13}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

359

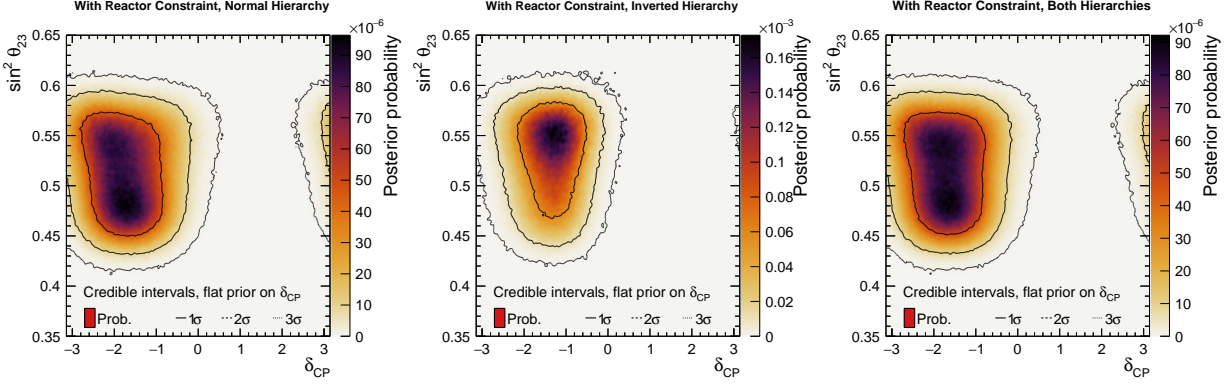


Figure 52: $\delta_{\text{CP}} - \sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

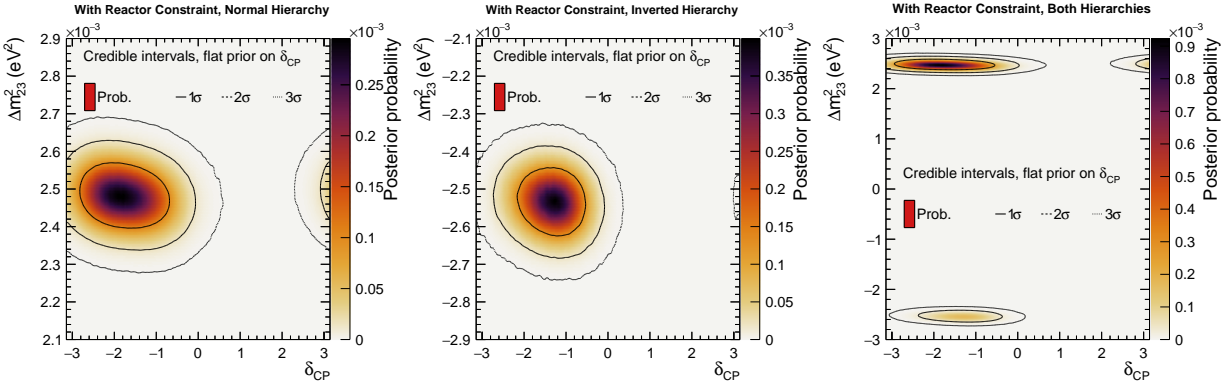


Figure 53: $\delta_{\text{CP}} - \Delta m_{32}^2$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

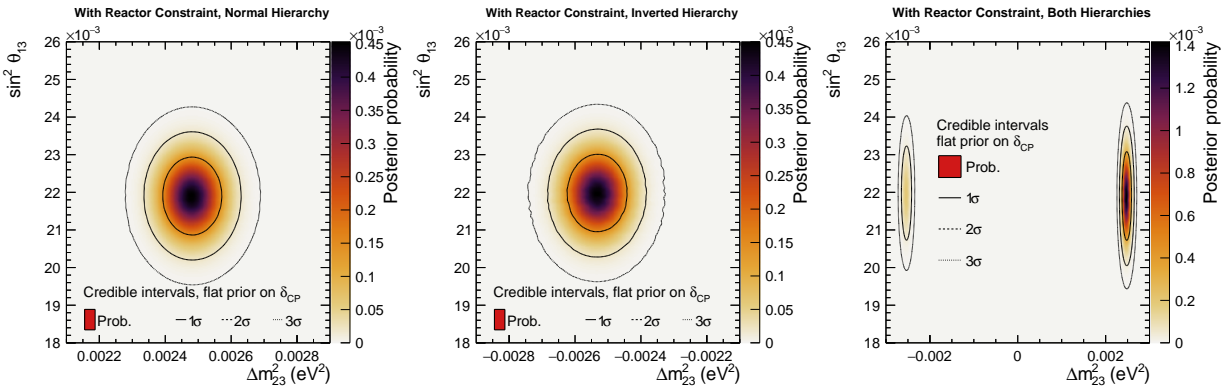


Figure 54: $\Delta m_{32}^2 - \sin^2 \theta_{13}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

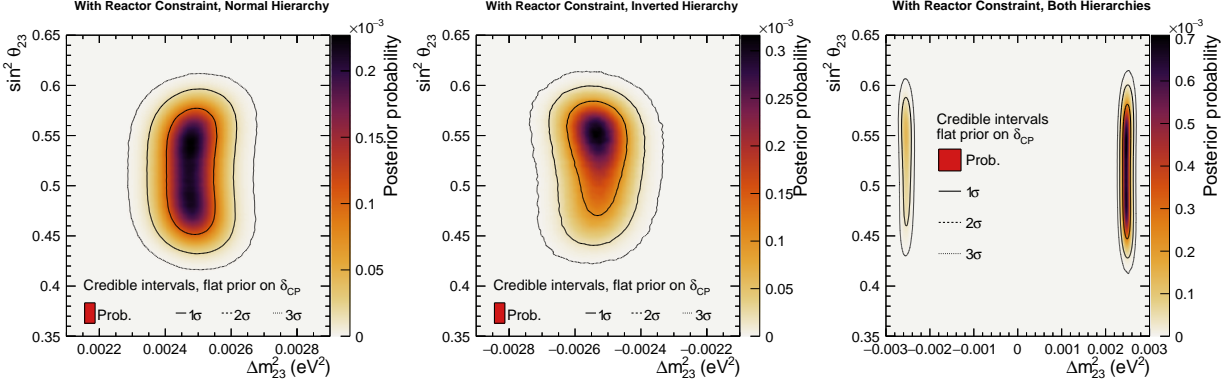


Figure 55: $\Delta m_{32}^2 - \sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

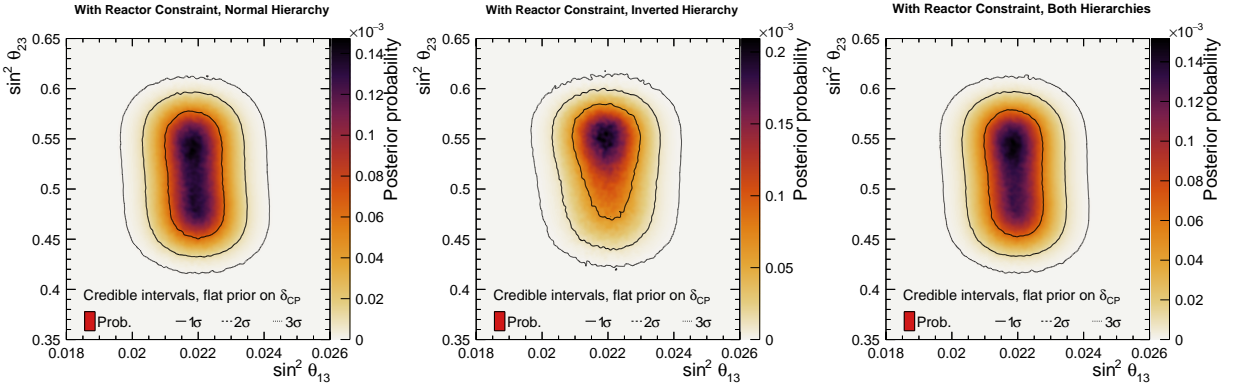


Figure 56: $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

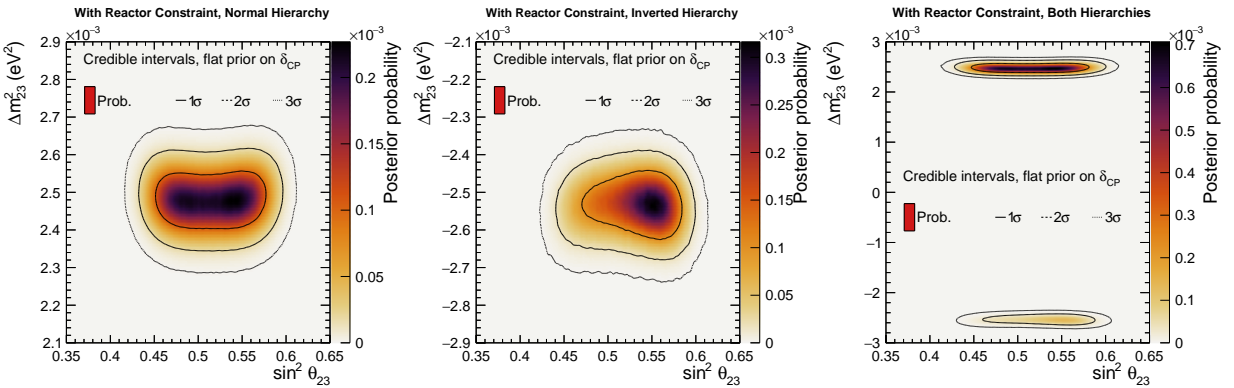


Figure 57: $\sin^2 \theta_{23} - \Delta m_{32}^2$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

360 **The Jarlskog invariant** Fig. 58 and Fig. 59 show the two-dimensional $J - \sin^2 \theta_{23}$ poste-
 361 riors for normal ordering, inverted ordering and both orderings, with a flat prior on δ_{CP} and
 362 $\sin \delta_{\text{CP}}$ respectively. Fig. 60 shows the two-dimensional $J - \delta_{\text{CP}}$ posteriors for the orderings,
 363 with a flat prior on δ_{CP} .

364 As in the Asimov A studies and the one-dimensional posterior from the data analysis,
 365 the posterior for the analysis using a prior flat in $\sin \delta_{\text{CP}}$ has a larger 1σ credible interval.
 366 The 2σ and 3σ intervals are much more similar than the 1σ interval for the two choices of
 367 flat prior.

368 We see the strongest constraint on J in the inverted hierarchy and the lower octant, as
 369 expected from the analysis of the one-dimensional posteriors from the data. This agrees
 370 with the expectation of the impact of the prior; a prior flat in δ_{CP} emphasises the extreme
 solutions at $|J| \sim 0.03$.

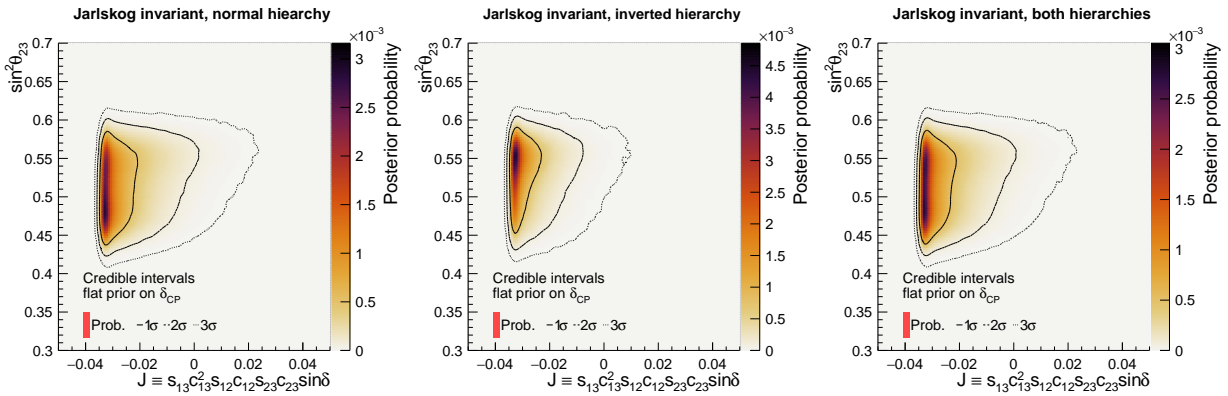


Figure 58: $J - \sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

371

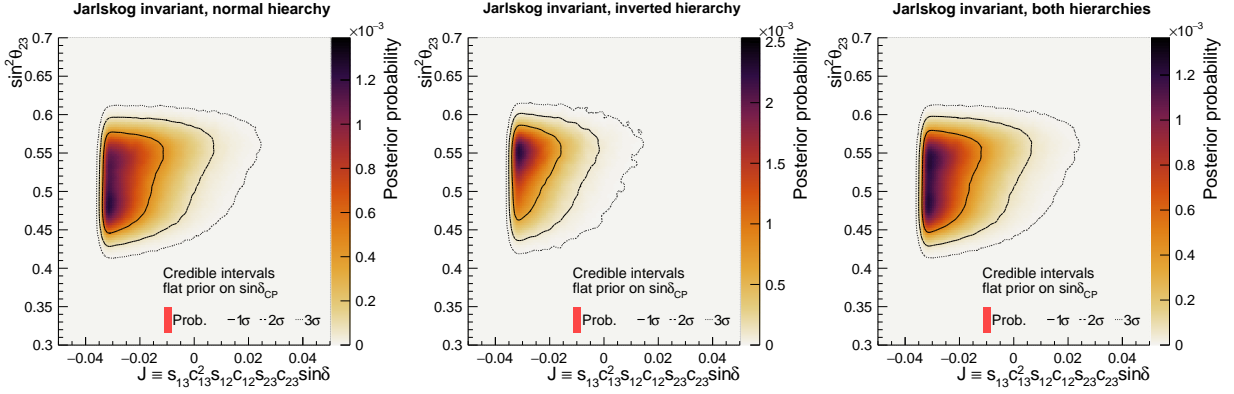


Figure 59: $J - \sin^2 \theta_{23}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

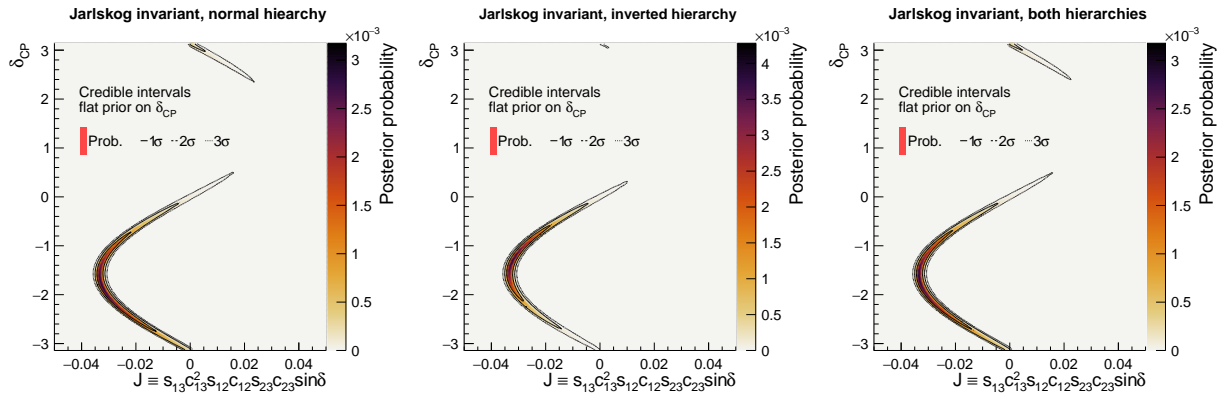


Figure 60: $J - \delta_{\text{CP}}$ for the real data analysis, with the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

372 **4.1.3 Triangle plots**

373 **Fig. 61**, **Fig. 62**, and **Fig. 63** show the one and two-dimensional oscillation parameters'
374 posteriors, displayed in triangular form. These all apply the reactor constraint and use a flat
375 prior on δ_{CP} .

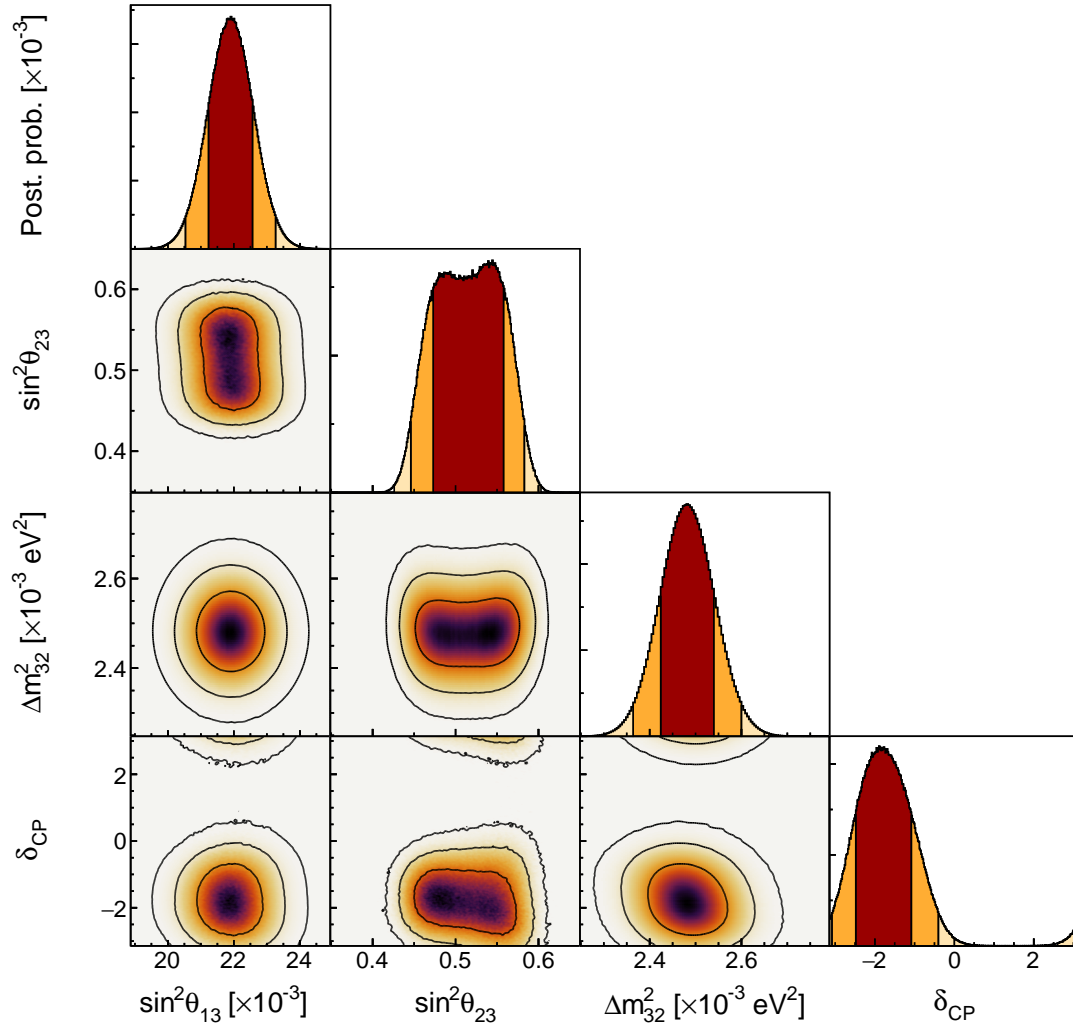


Figure 61: Triangle plots for the real data analysis, with the reactor constraint applied, for the normal ordering, with a flat prior on δ_{CP} . The 1, 2, and 3σ credible intervals are overlaid on the posterior.

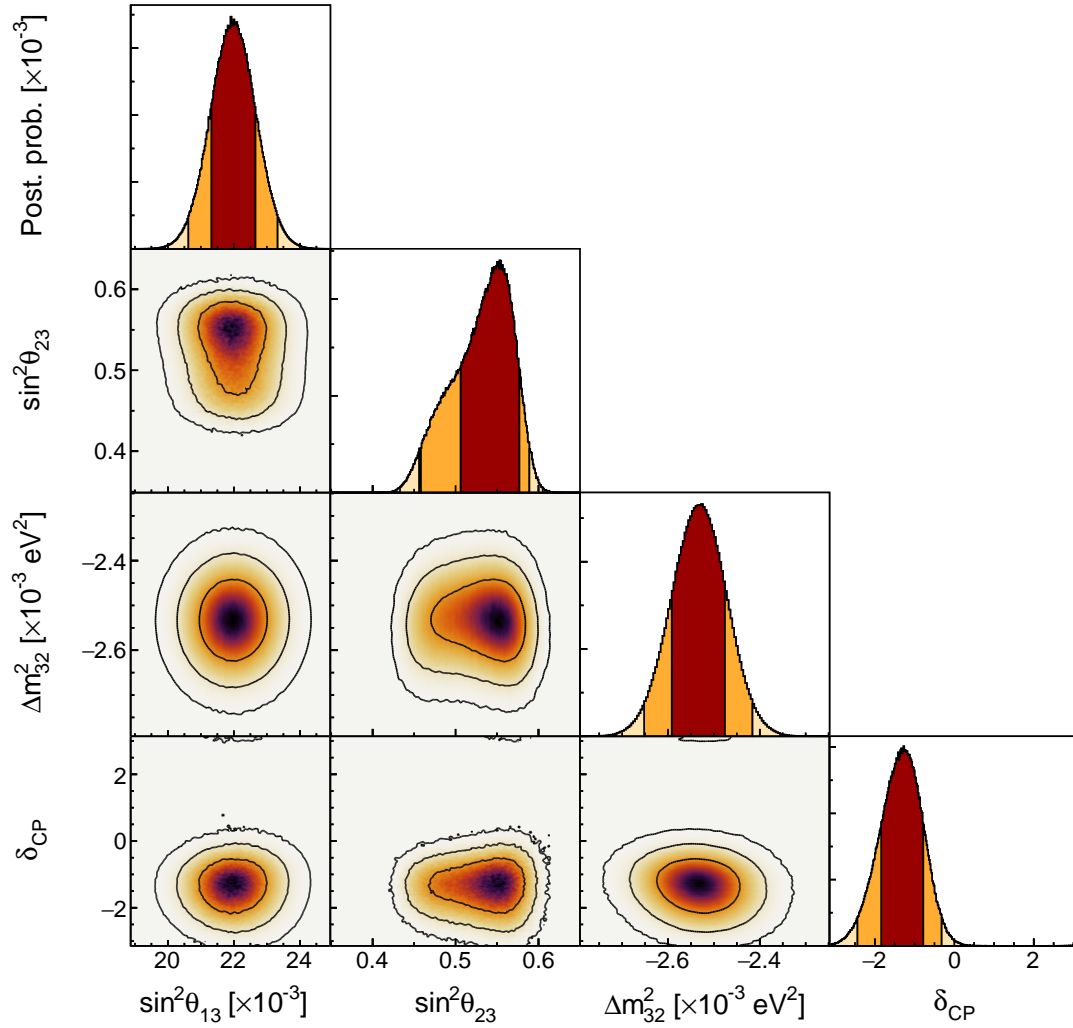


Figure 62: Triangle plots for the real data analysis, with the reactor constraint applied, for the inverted ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

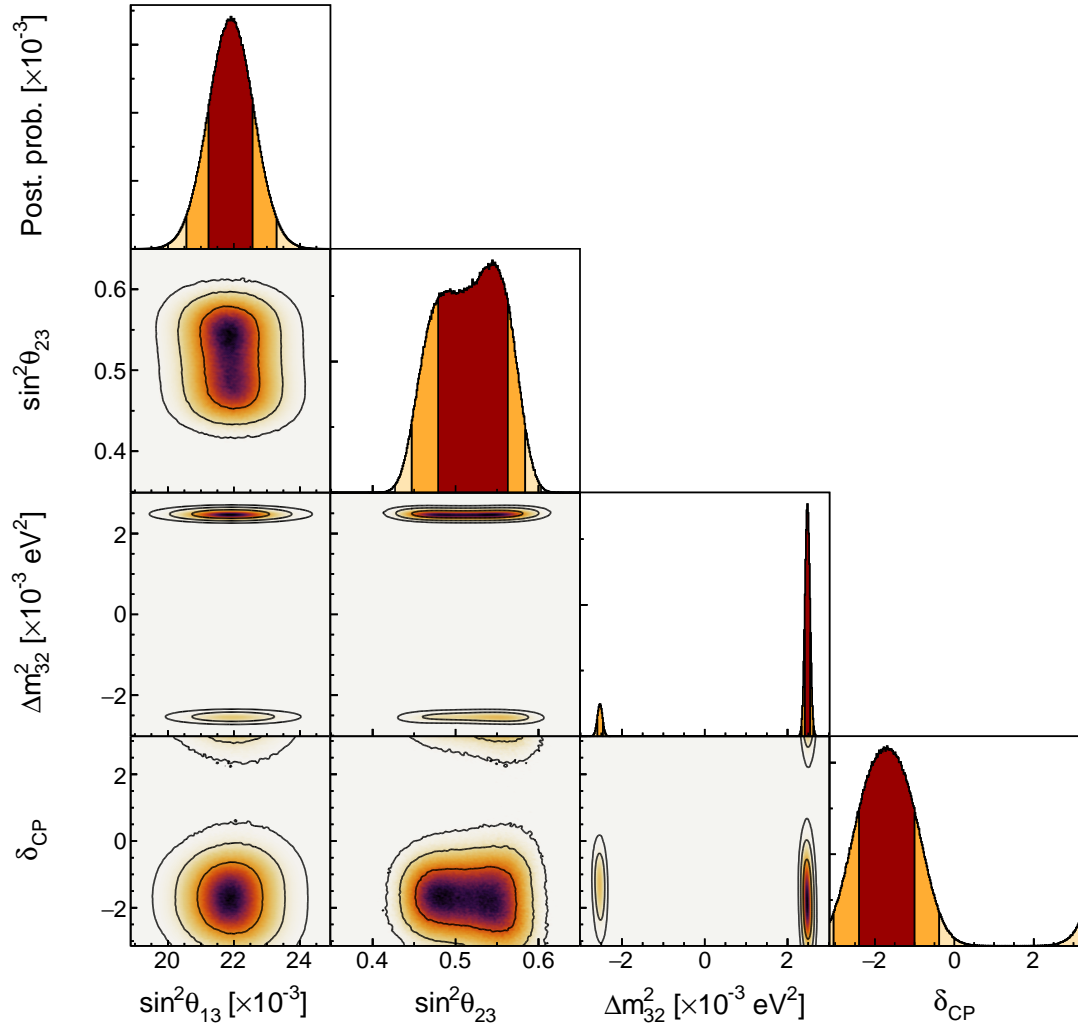


Figure 63: Triangle plots for the real data analysis, with the reactor constraint applied, for both orderings, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

376 **4.1.4 Bayes factors for atmospheric parameters**

377 Posterior probabilities from the data fit with reactor constraint for different hypothesis ap-
 378 plying the reactor constraint are calculated and summarized in **Tab. 2**. The Bayes factor
 379 of upper octant over lower octant is $B(\text{UO}/\text{LO}) = 1.78$ and the Bayes factor of normal
 380 ordering over inverted ordering is $B(\text{NO}/\text{IO}) = 7.33$. The combined Bayes factor for UO
 381 and IO is weak, with $B([\text{UO}+\text{NO}]/\text{Other}) = 1.17$. These posterior probabilities are compat-
 382 ible with P-Theta’s MCMC results in the data fit TN459 [12], and the mass ordering Bayes
 383 factor is notably stronger and the octant preference is weaker than the T2K-only analysis in
 TN393 [1]. Therein, T2K finds $B(\text{NO}/\text{IO}) = 4.2$ and $B(\text{UO}/\text{LO}) = 3.36$.

	$\sin^2 \theta_{23} < 0.5$	$\sin^2 \theta_{23} > 0.5$	Sum
NO	0.33	0.54	0.88
IO	0.03	0.10	0.12
Sum	0.36	0.64	1.00

Table 7: Posterior probabilities for the mass ordering and the octant in the real data analysis, with the reactor constraint applied.

384

385 4.2 Results without applying the reactor constraint

386 This section is dedicated to the data fit results without applying the reactor constraint on
 387 $\sin^2 \theta_{13}$. The same data fit chains are used with a burn-in of 80,000 steps per independent
 388 chain.

389 4.2.1 Parameter constraints

390 Each parameter constraint in this section is shown in normal ordering ($\Delta m_{32}^2 > 0$), inverted
 391 ordering ($\Delta m_{32}^2 < 0$), and over both orderings.

392 **One-dimensional posteriors** Fig. 64, Fig. 65, Fig. 66, and Fig. 67 show the marginalised
 393 one-dimensional posterior for δ_{CP} , Δm_{32}^2 , $\sin^2 \theta_{13}$ and $\sin^2 \theta_{23}$, respectively.

394 As for the results with reactor constraint applied, the constraint on δ_{CP} is weaker in the
 395 normal ordering in Fig. 64. $\delta_{\text{CP}} = 0$ is just excluded at 2σ credible interval, and $\delta_{\text{CP}} = -\pi$
 396 is included. The result in normal ordering is compatible with inverted ordering, although
 397 the inverted ordering highest posterior density is near $\delta_{\text{CP}} \sim -1$ whereas in normal ordering
 398 $\delta_{\text{CP}} \sim -2$.

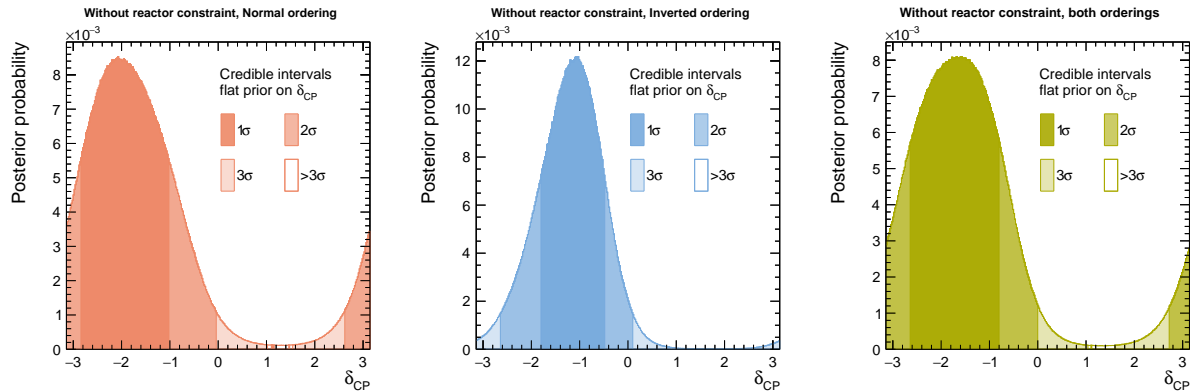


Figure 64: δ_{CP} for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

399 As for the mass ordering, Fig. 65 shows a similar but weaker preference for normal
 400 ordering.

401 The constraint on $\sin^2 \theta_{13}$ in Fig. 66 is compatible with the reactor constraint, and is the
 402 impact of the reactor constraint is shown in more detail in Sec. 4.4.

403 The constraint on $\sin^2 \theta_{23}$ without the reactor constraint consistently weakly prefers the
 404 lower octant for both orderings individually and combined. The constraint is weak, and the

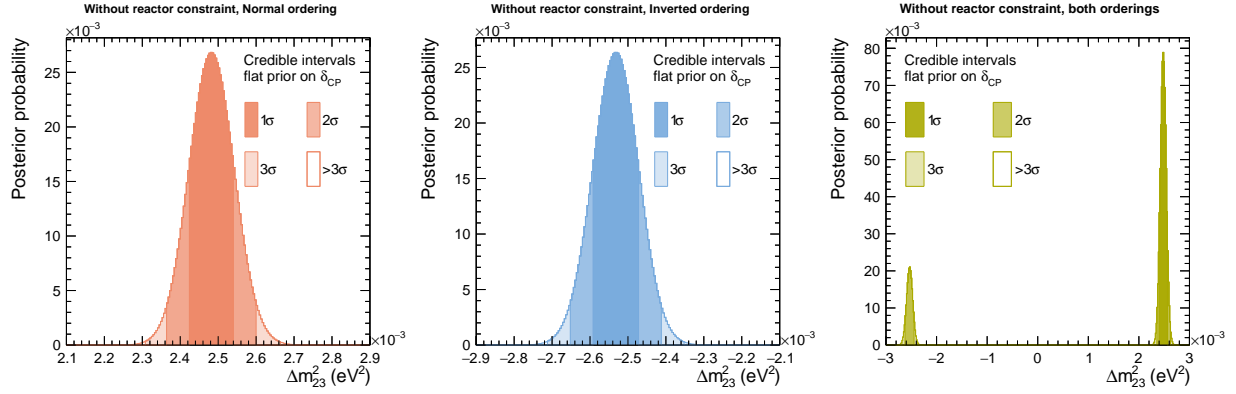


Figure 65: Δm_{32}^2 for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

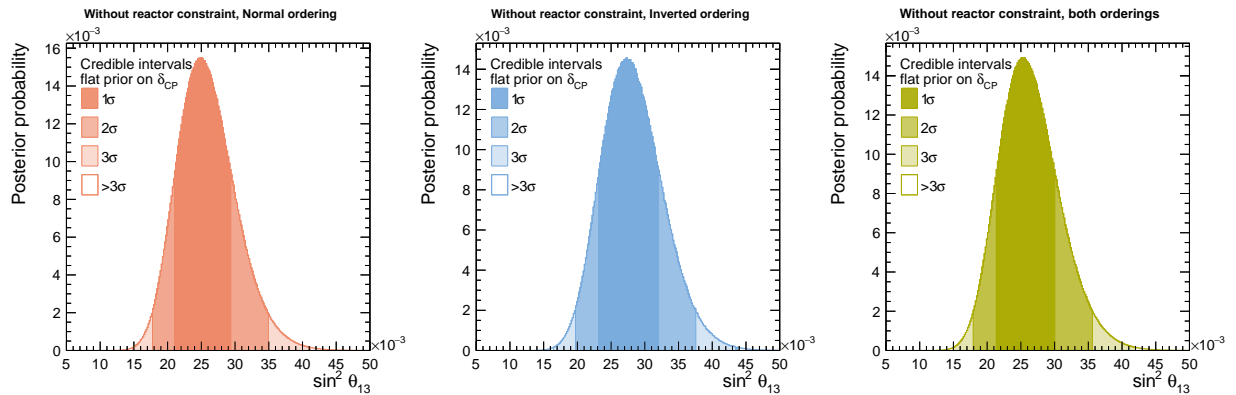


Figure 66: $\sin^2 \theta_{13}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

405 upper octant is also inside the 1σ credible interval, although the highest posterior density is
 406 in the lower octant. This is tabulated in the section dedicated to Bayes factors, [Sec. 4.2.3](#).

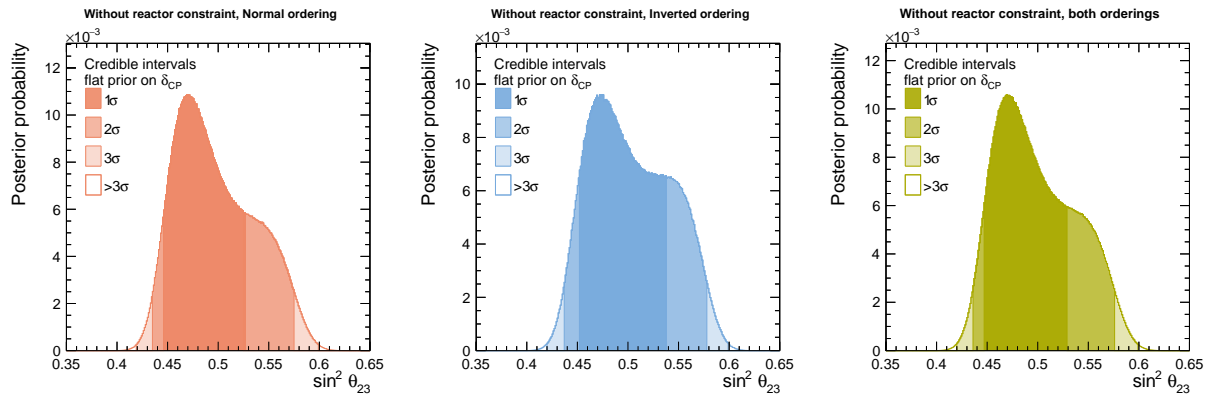


Figure 67: $\sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

407 **The Jarlskog invariant** Fig. 68 and Fig. 69 show the marginalised Jarlskog invariant (J)
 408 posterior for normal ordering, inverted ordering and both orderings with flat prior on δ_{CP}
 409 and $\sin \delta_{\text{CP}}$, respectively.

410 $J = 0$ is excluded between $1-2\sigma$ for both priors.

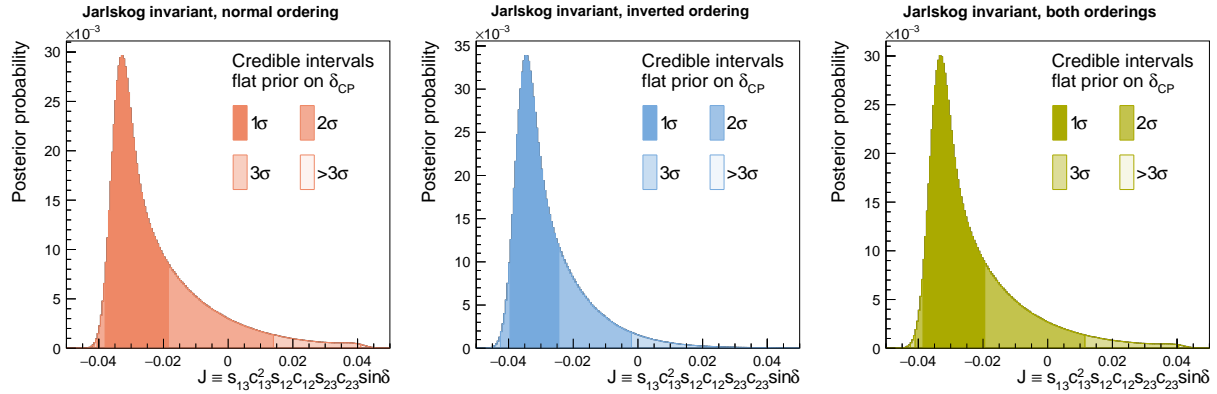


Figure 68: J for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

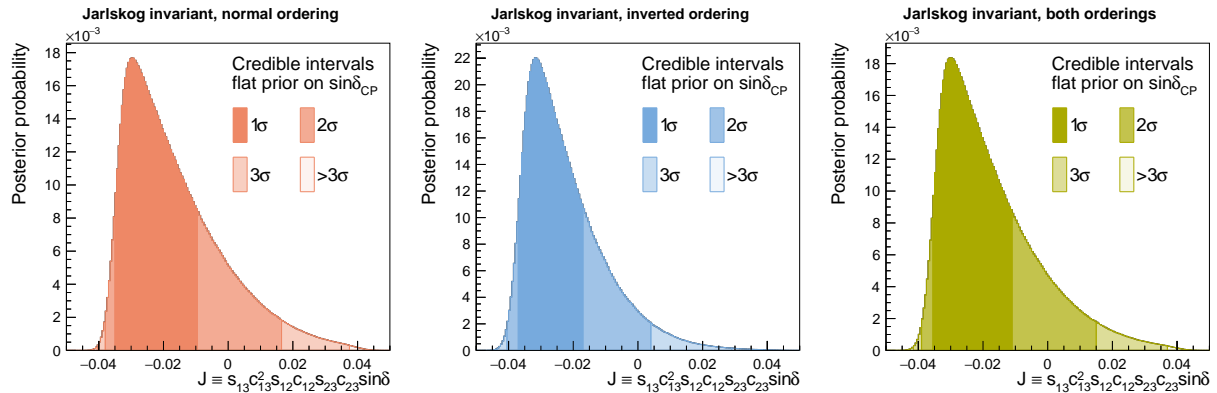


Figure 69: J for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{\text{CP}}$.

411 **Two-dimensional posteriors** Fig. 70, Fig. 71, Fig. 72, Fig. 73, Fig. 74, Fig. 75 and Fig. 76
 412 show the two-dimensional marginal posteriors for $\delta_{\text{CP}} - \sin^2 \theta_{13}$, $\delta_{\text{CP}} - \sin^2 \theta_{23}$, $\delta_{\text{CP}} - \Delta m_{32}^2$,
 413 $\Delta m_{32}^2 - \sin^2 \theta_{13}$, $\Delta m_{32}^2 - \sin^2 \theta_{23}$, $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ and $\sin^2 \theta_{23} - \Delta m_{32}^2$, respectively. These
 414 are shown separately for normal ordering, inverted ordering and both orderings, all without
 415 applying the reactor constraint.

416 Firstly, in Fig. 70 we notice a different δ_{CP} and $\sin^2 \theta_{13}$ correlation for normal and inverted
 417 orderings, and a different maximum of the posterior density. The normal ordering result
 418 favours a more negative δ_{CP} alongside a smaller $\sin^2 \theta_{13}$ with a negative correlation, and the
 419 inverted ordering favours a less negative value of δ_{CP} alongside a larger $\sin^2 \theta_{13}$ with a positive
 correlation.

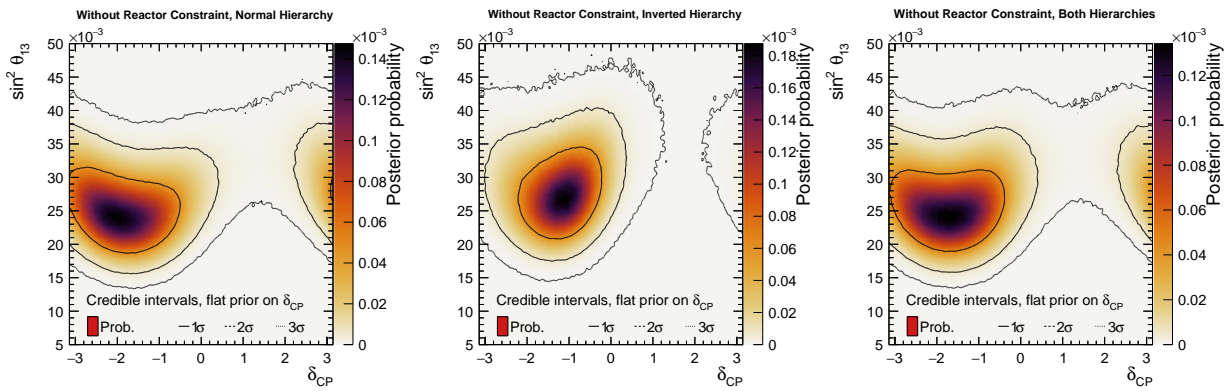


Figure 70: $\delta_{\text{CP}} - \sin^2 \theta_{13}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

420

421 For the $\delta_{\text{CP}} - \sin^2 \theta_{23}$ result in Fig. 71, the constraint on δ_{CP} is weakest in the lower octant
 422 (which has the highest posterior density), especially in the normal ordering. For inverted
 423 ordering, the two are barely correlated.

424 For the $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ posterior in Fig. 75, there is now an evident anti-correlation,
 425 whereas the analysis applying the reactor constraint showed barely any correlation. This
 426 stems from the external constraint on $\sin^2 \theta_{13}$ decoupling the two. The correlation is
 427 roughly the same in both hierarchies.

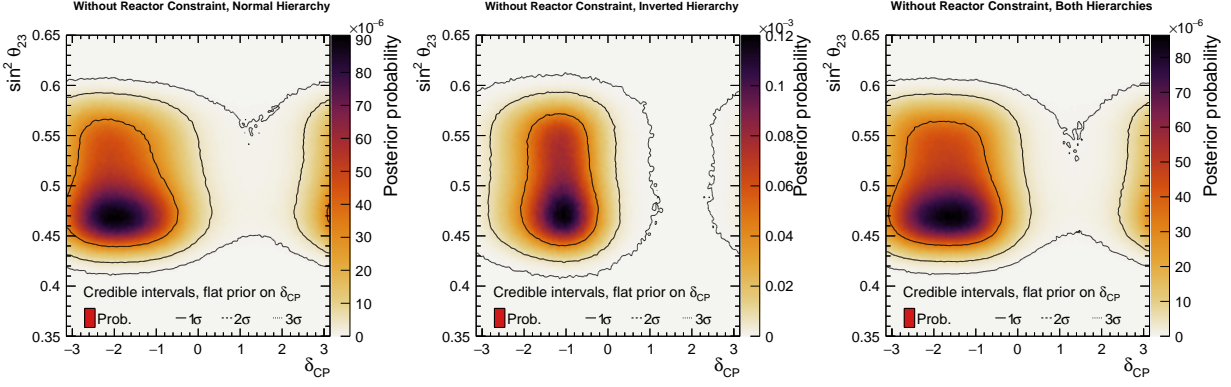


Figure 71: $\delta_{CP} - \sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

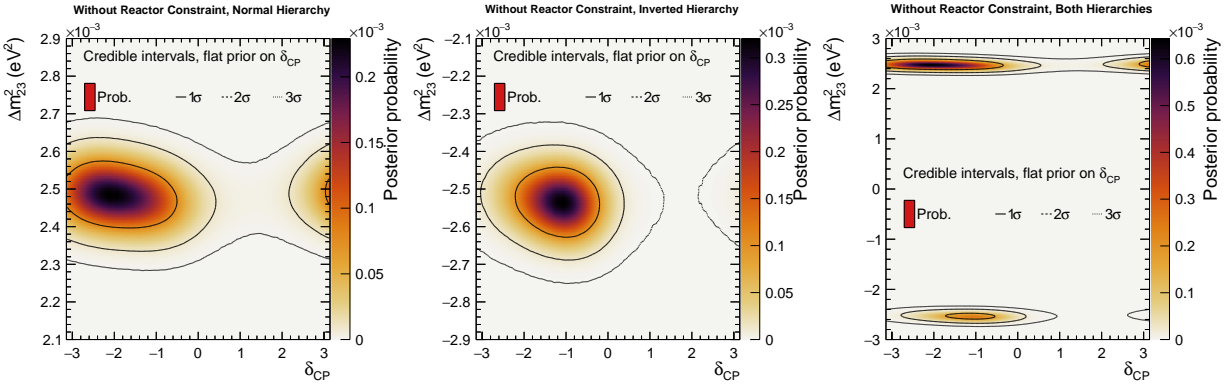


Figure 72: $\delta_{CP} - \Delta m_{32}^2$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

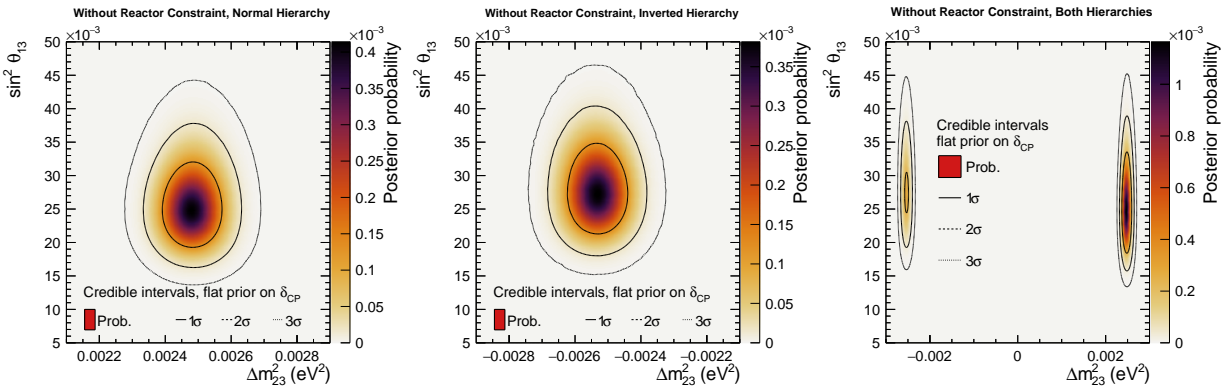


Figure 73: $\Delta m_{32}^2 - \sin^2 \theta_{13}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

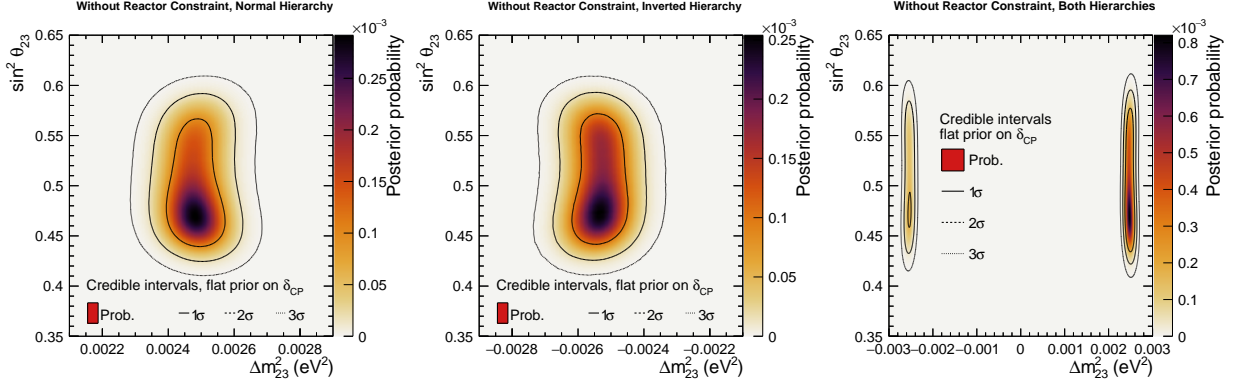


Figure 74: $\Delta m_{32}^2 - \sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

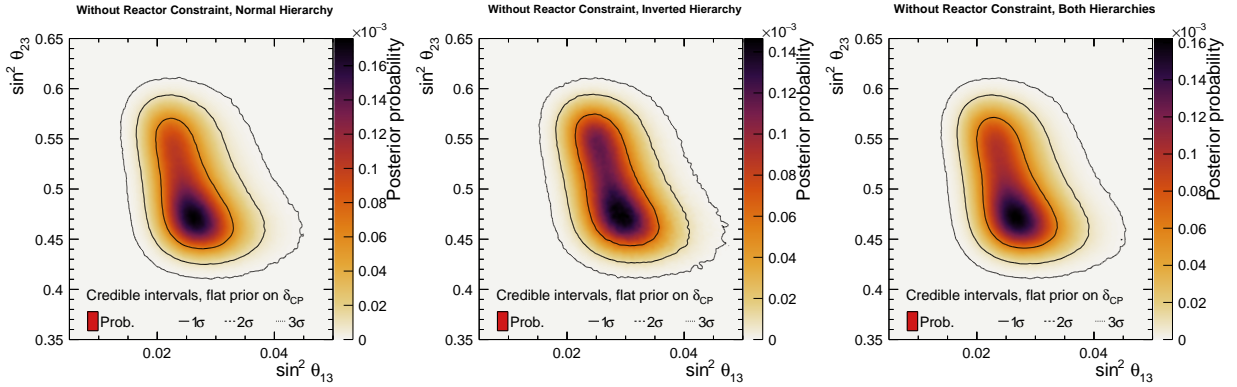


Figure 75: $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

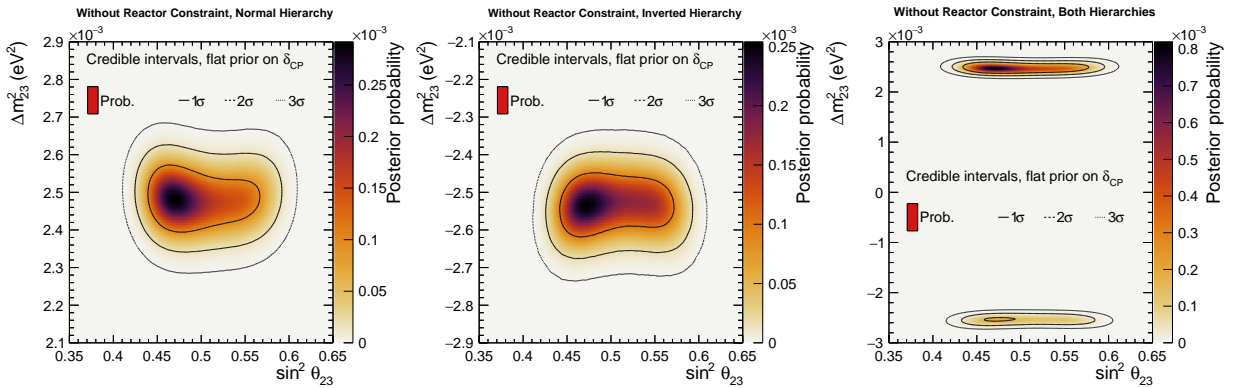


Figure 76: $\sin^2 \theta_{23} - \Delta m_{32}^2$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings.

428 **The Jarlskog invariant** Fig. 77 and Fig. 78 show the two-dimensional $J - \sin^2 \theta_{23}$ poste-
 429 riors for normal ordering, inverted ordering and both orderings, with a flat prior on δ_{CP} and
 430 $\sin \delta_{\text{CP}}$ respectively. Fig. 79 shows the two-dimensional $J - \delta_{\text{CP}}$ posteriors for the orderings,
 431 with a flat prior on δ_{CP} .

432 As for the two-dimensional $\delta_{\text{CP}} - \sin^2 \theta_{23}$ posteriors, J is more strongly constrained in
 433 the upper octant when the reactor constraint isn't applied. For instance, $J = 0$ is just
 434 barely excluded at 1σ in the lower octant for the analysis using a prior flat in $\sin \delta_{\text{CP}}$ when
 435 marginalising over both orderings and the normal ordering only.

436 As seen in both the Asimov A analysis and the data analysis, the prior flat in δ_{CP} or
 437 $\sin \delta_{\text{CP}}$ mainly impact the 1σ credible interval, and the 2 and 3σ intervals are very similar
 438 regardless of prior choice. The positive J solution weakly appears in the analysis with a flat
 439 prior on δ_{CP} , but there is less probability there than in the Asimov A study.

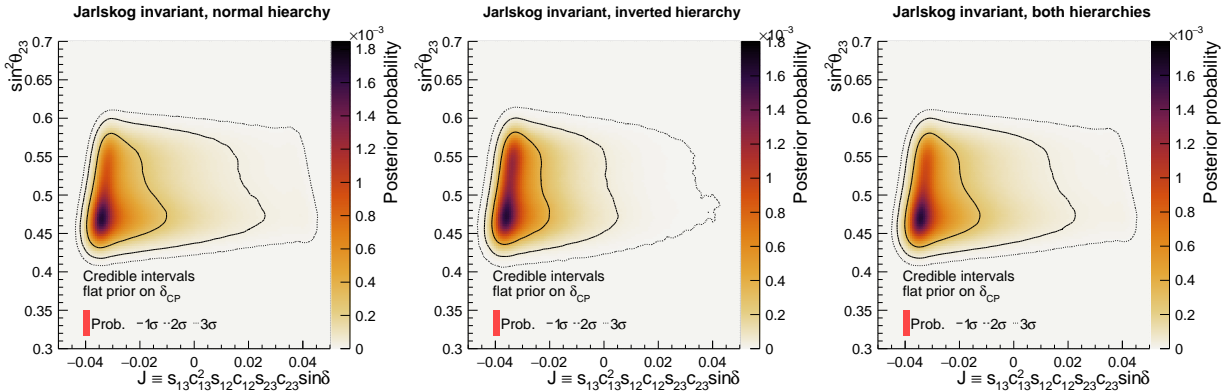


Figure 77: $J - \sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

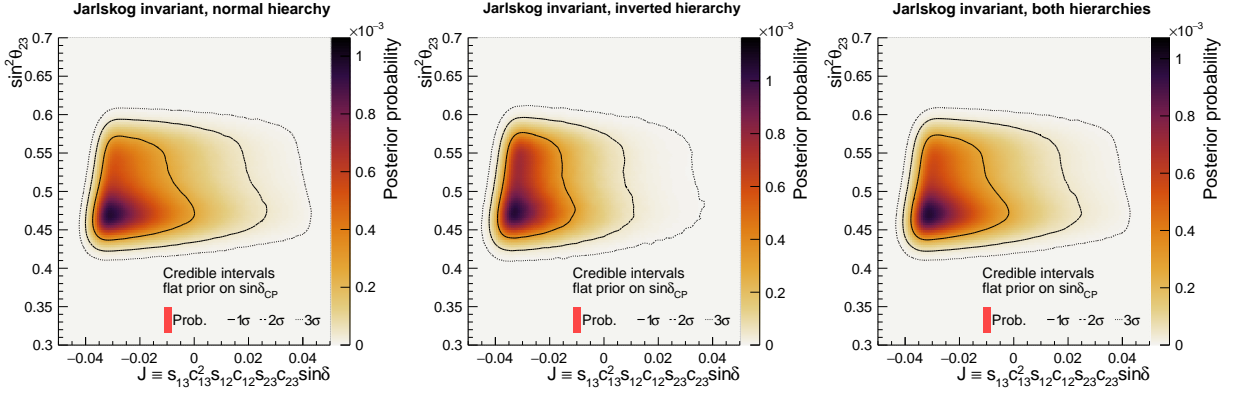


Figure 78: $J - \sin^2 \theta_{23}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on $\sin \delta_{CP}$.

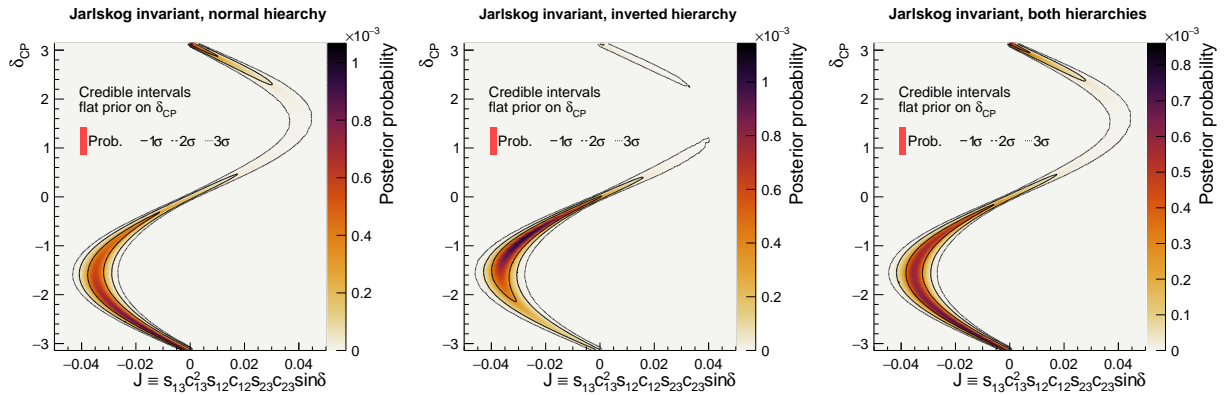


Figure 79: $J - \delta_{CP}$ for the real data analysis, without the reactor constraint applied, for normal (left), inverted (center), and both (right) orderings, with a flat prior on δ_{CP} .

440 4.2.2 Triangle plots

441 Fig. 80, Fig. 81, and Fig. 82 show the one and two-dimensional oscillation parameters'
442 posteriors, displayed in triangular form, summarising the main results. These do not apply
443 the reactor constraint, and use a flat prior on δ_{CP} .

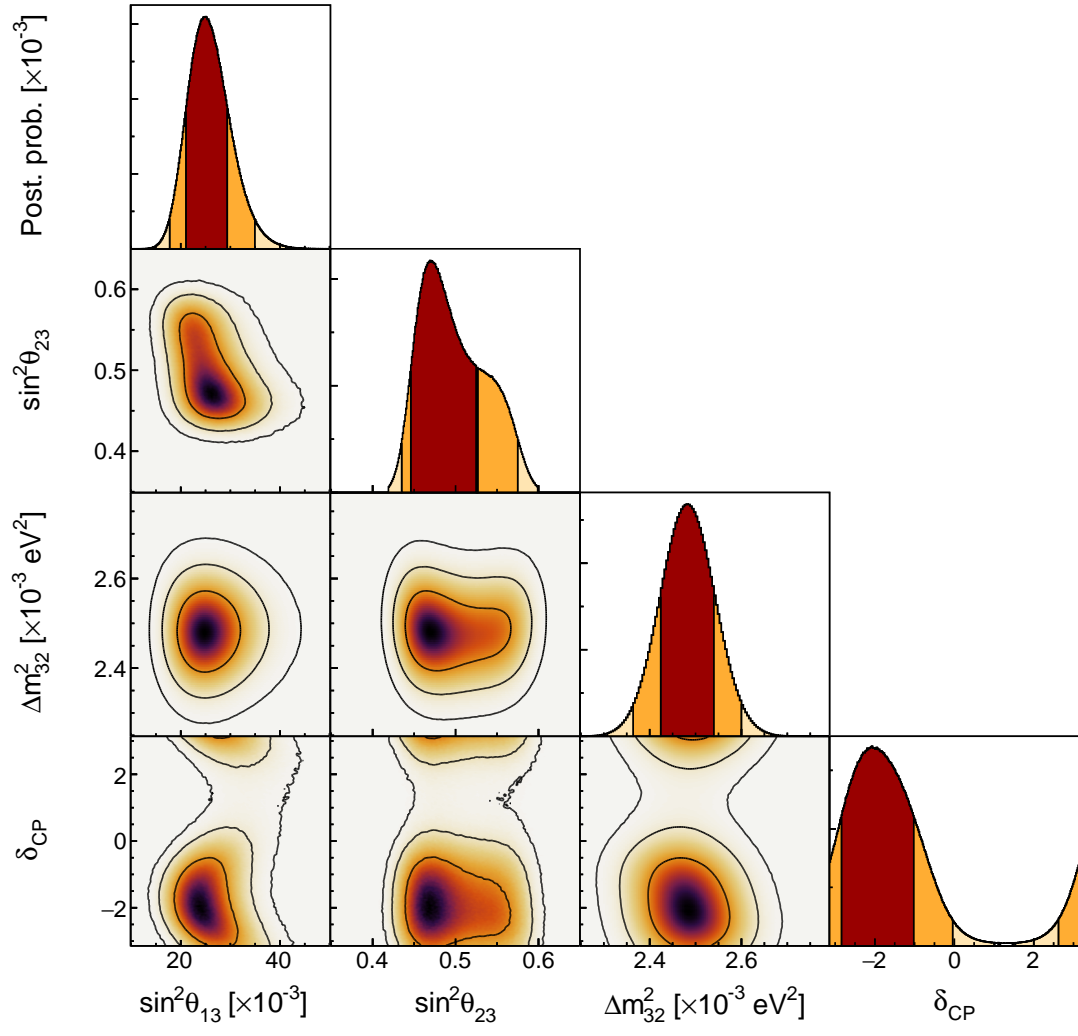


Figure 80: Triangle plots for the real data analysis, without the reactor constraint applied, for the normal ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

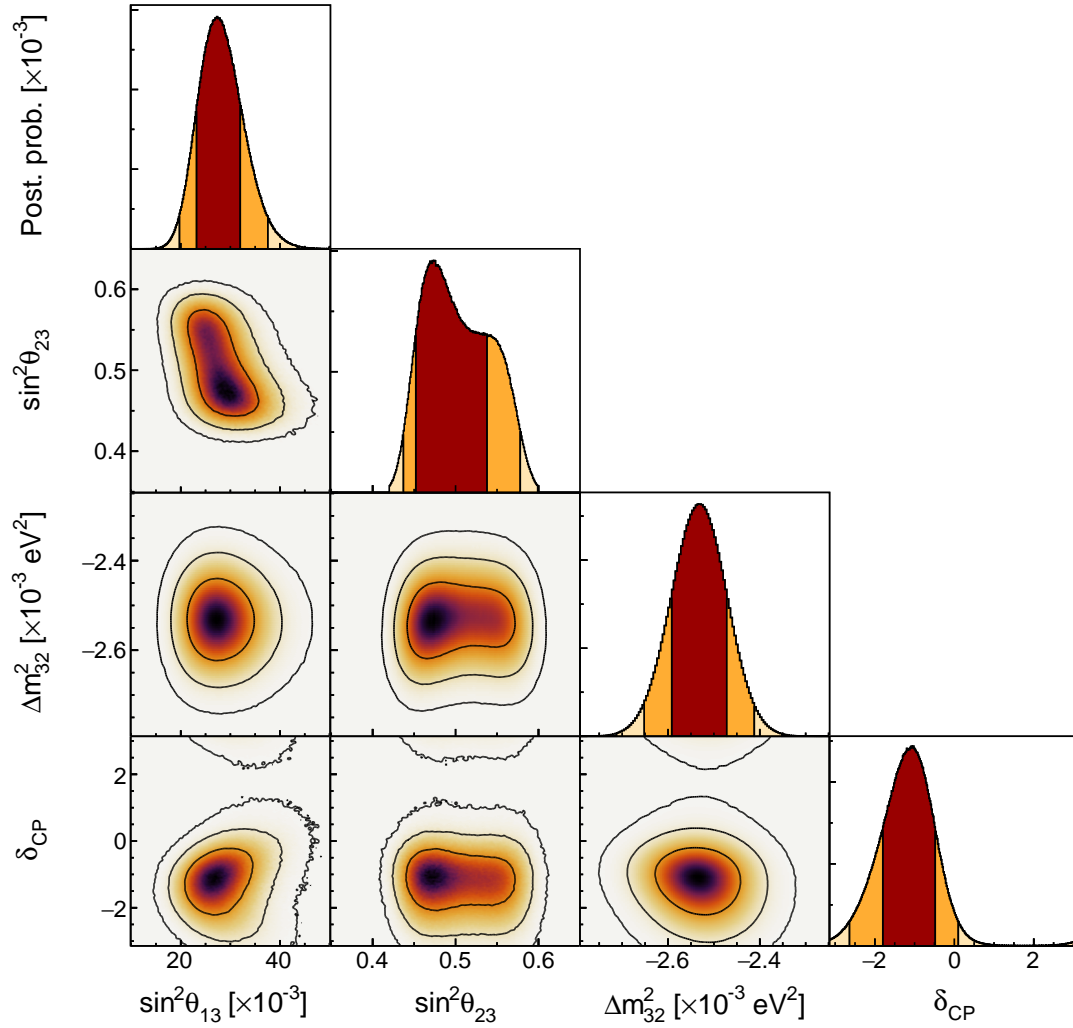


Figure 81: Triangle plots for the real data analysis, without the reactor constraint applied, for the inverted ordering, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

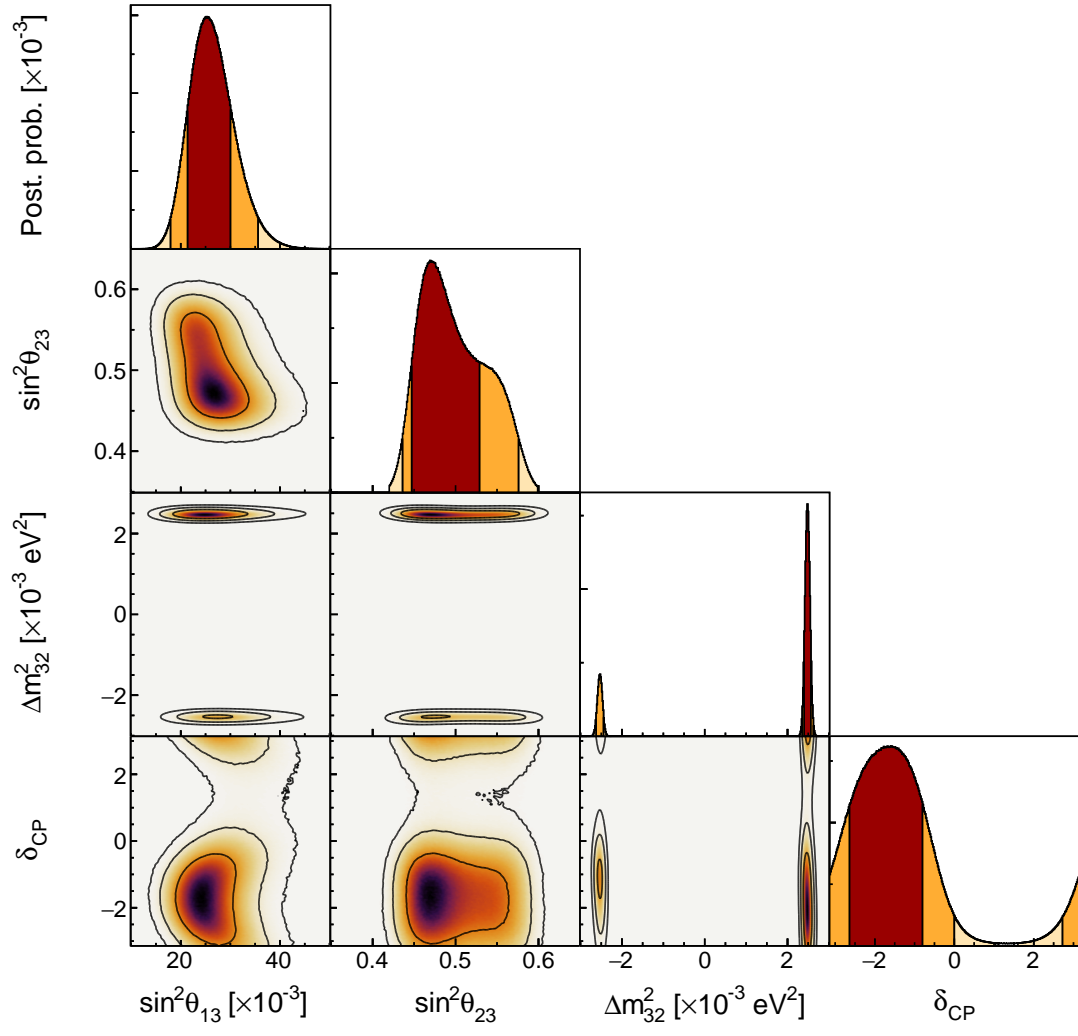


Figure 82: Triangle plots for the real data analysis, without the reactor constraint applied, for both mass orderings, with a flat prior on δ_{CP} . The 1, 2, and 3 σ credible intervals are overlaid on the posterior.

444 **4.2.3 Bayes factors for atmospheric parameters**

445 Posterior probabilities for the data fit without reactor constraint are calculated and sum-
 446 marized in **Tab. 8**. The Bayes factor for upper octant over lower octant $B(\text{UO}/\text{LO})$ is 0.82,
 447 notably weaker than the with reactor constraint case. This is expected due to the corre-
 448 lations between $\sin^2 \theta_{23}$ and $\sin^2 \theta_{13}$. The Bayes factor for normal ordering over inverted
 449 ordering $B(\text{NO}/\text{IO})$ is 3.76, which is comparable to T2K’s main result using reactor con-
 450 straint ($B(\text{NO}/\text{IO}) = 4.2$). The Bayes factor for upper octant and normal ordering over its
 451 competing hypotheses $B([\text{UO}+\text{NO}]/\text{Other})$ is 0.52.

	$\sin^2 \theta_{23} < 0.5$	$\sin^2 \theta_{23} > 0.5$	Sum
NO	0.44	0.34	0.79
IO	0.11	0.11	0.21
Sum	0.55	0.45	1.00

Table 8: Posterior probabilities for the mass ordering and the octant in the real data analysis, without the reactor constraint applied.

4.3 Impact of a flat prior on δ_{CP} or $\sin \delta_{\text{CP}}$

This section overlays the one-dimensional posterior probabilities using a prior flat in δ_{CP} (identical to results in Sec. 4) with when applying a prior flat in $\sin \delta_{\text{CP}}$. This section applies the reactor constraint for all results. Fig. 83 shows the posterior probabilities and 1, 2, 3 σ credible intervals for the two choices of prior. The 2 σ interval just contains $+\pi$ with a prior flat in $\sin \delta_{\text{CP}}$, whereas it excludes it with a prior flat in δ_{CP} . Thus, a statement on 2 σ exclusion of $\delta_{\text{CP}} = 0, \pi$ is prior dependent, and should be avoided.

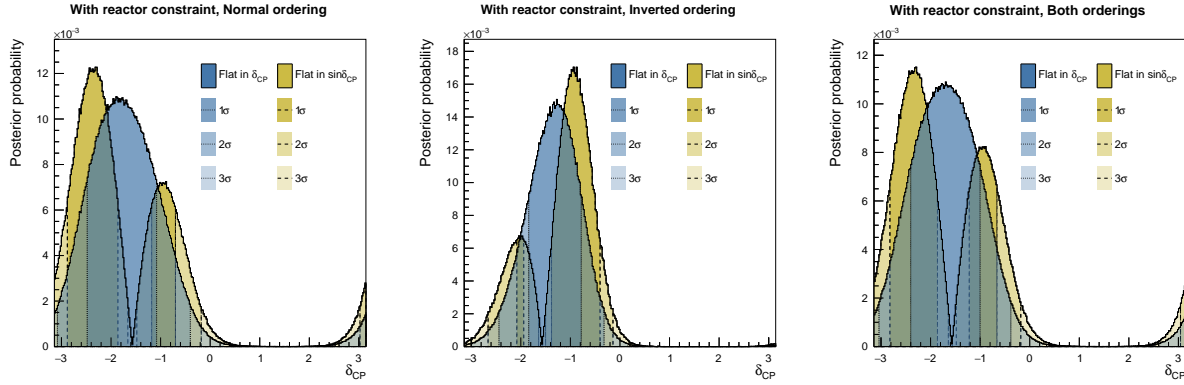


Figure 83: δ_{CP} from real data fit with a prior flat in δ_{CP} (blue shaded) and flat in $\sin \delta_{\text{CP}}$ (yellow shaded) applied.

Fig. 84 shows the same result but for Δm_{32}^2 , which has close to no impact on both the credible intervals and the mass ordering preference. This is expected from the weak correlation between Δm_{32}^2 and δ_{CP} in Sec. 4.

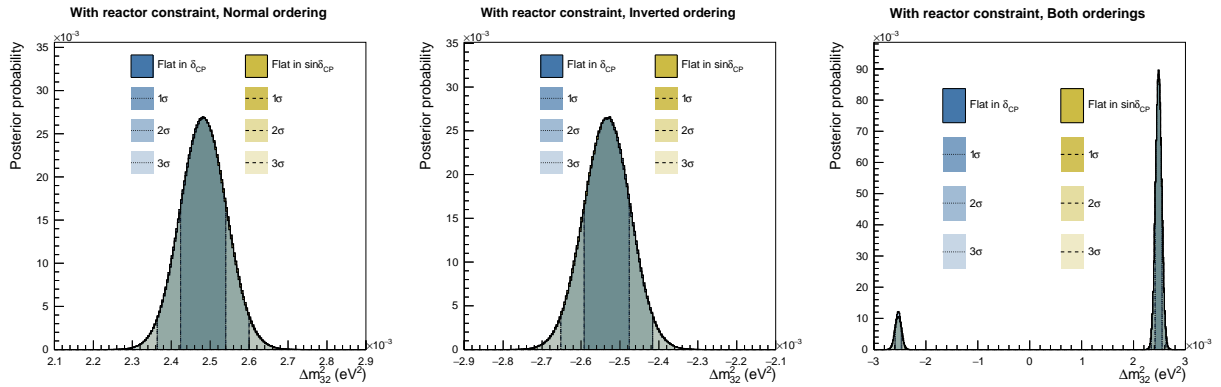


Figure 84: Δm_{32}^2 from real data fit with a prior flat in δ_{CP} (blue shaded) and flat in $\sin \delta_{\text{CP}}$ (yellow shaded) applied.

Similarly, the posteriors for $\sin^2 \theta_{13}$ in Fig. 85 are identical.

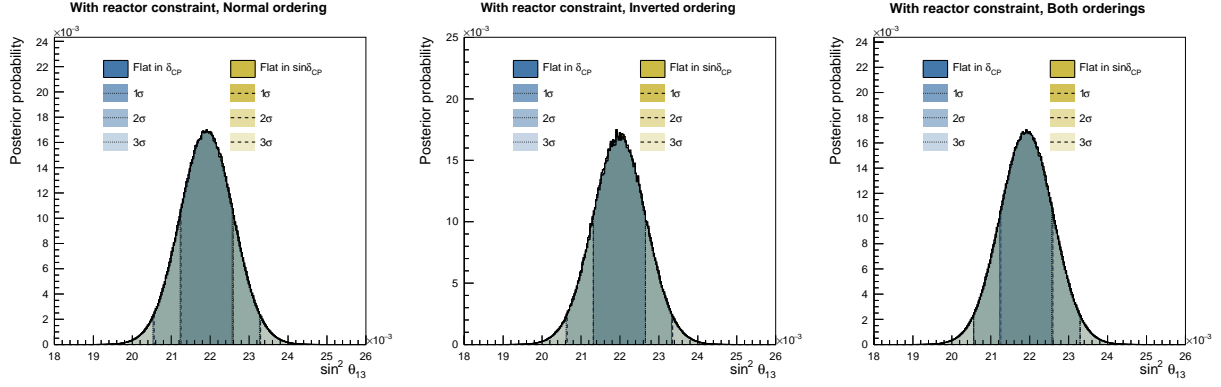


Figure 85: $\sin^2 \theta_{13}$ from real data fit with a prior flat in δ_{CP} (blue shaded) and flat in $\sin \delta_{CP}$ (yellow shaded) applied.

463 Finally, the $\sin^2 \theta_{23}$ is shown in Fig. 86, where there is a marginally stronger preference
 464 for the upper octant when a flat prior on $\sin \delta_{CP}$ is used in the normal ordering. For inverted
 ordering the posterior is largely the same.

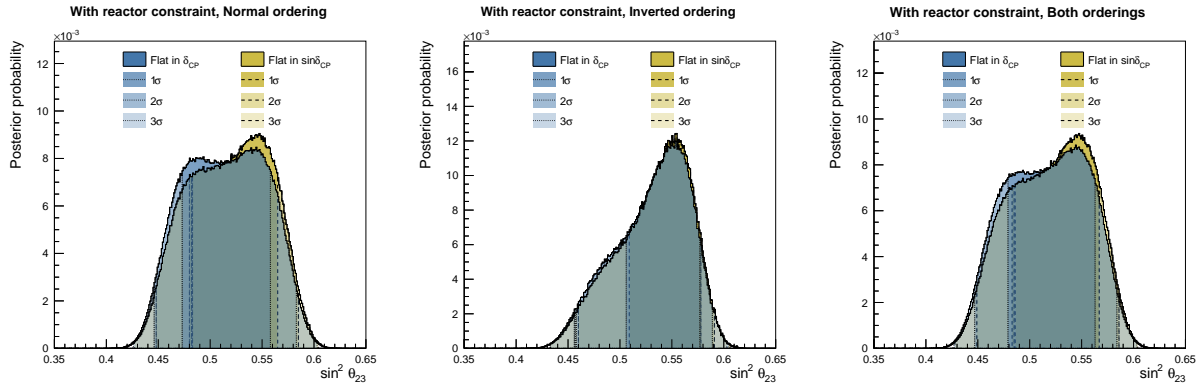


Figure 86: $\sin^2 \theta_{23}$ from real data fit with a prior flat in δ_{CP} (blue shaded) and flat in $\sin \delta_{CP}$ (yellow shaded) applied.

465

4.4 Impact of reactor constraint

Fig. 87, Fig. 88, Fig. 89 and Fig. 90 are the comparison of one dimensional contours for results with reactor constraint (blue shaded) and without reactor constraint (yellow shaded) of δ_{CP} , $\sin^2 \theta_{13}$, $\sin^2 \theta_{23}$ and Δm_{32}^2 respectively.

In Fig. 87, similar contour shape differences are observed in OA2020 with/without reactor constraint comparison (see Figure 88 in TN393 [1]). For normal ordering and inverted ordering, contours with reactor constraint have better constraint on δ_{CP} and pull the best-fit points slightly away from contours without reactor constraint, but in different directions. Results without reactor constraint can barely exclude $\delta_{\text{CP}} = 0$ at 2σ for normal ordering, but cannot exclude it for inverted ordering. Results with reactor constraint can easily exclude it at 2σ for both orderings.

Fig. 88 shows the direct effect of reactor constraint on $\sin^2 \theta_{13}$. The posterior probability is narrowed and is approximately a Gaussian with the central value close to 0.0218.

In Fig. 89, contrarily to the comparison in OA2020 (Fig. 88 in TN393 [1]), the results without reactor constraint show a preference of the lower octant while the results with reactor constraint show a much weaker preference of one octant over another, highlighting the impact of the SK atmospheric selections.

Fig. 90 shows that the constraint on Δm_{32}^2 is largely independent of the reactor constraint on $\sin^2 \theta_{13}$.

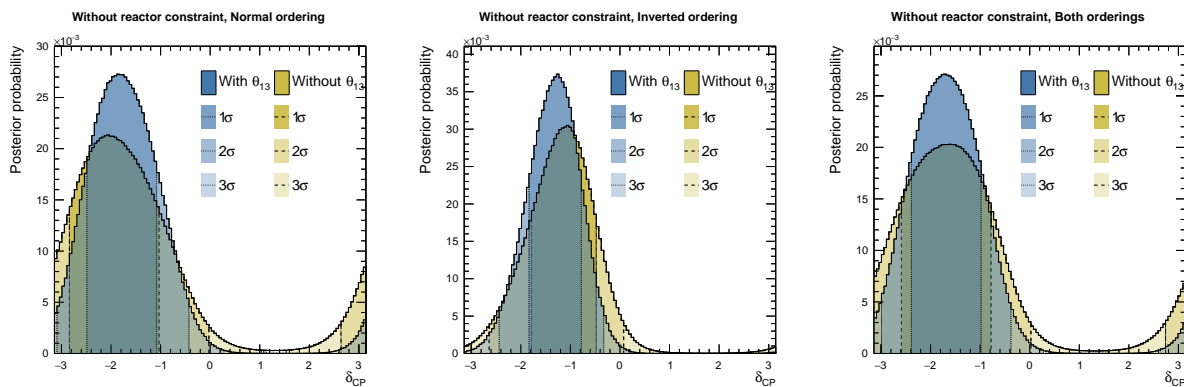


Figure 87: δ_{CP} from real data fit with (blue shaded) and without (yellow shaded) reactor constraint applied.

Bayes factors Tab. 9 summarises the posterior probabilities of $\sin \delta_{\text{CP}} \geq 0$ and $\sin \delta_{\text{CP}} < 0$ without and with reactor constraint respectively. The Bayes factor of $\sin \delta_{\text{CP}} < 0$ over

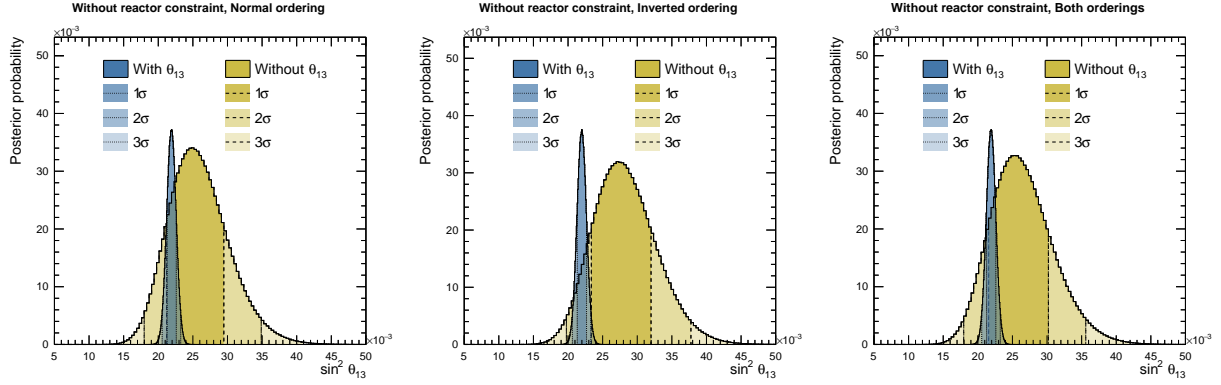


Figure 88: $\sin^2 \theta_{13}$ from real data fit with (blue shaded) and without (yellow shaded) reactor constraint applied.

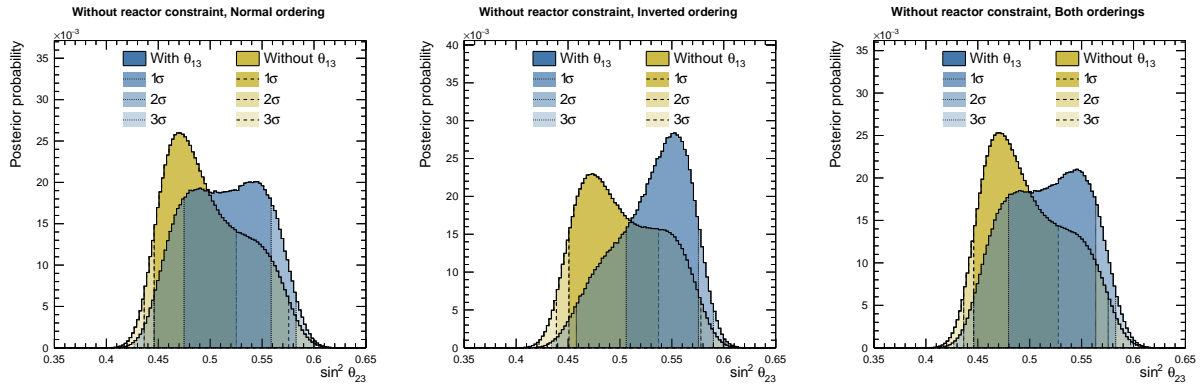


Figure 89: $\sin^2 \theta_{23}$ from real data fit with (blue shaded) and without (yellow shaded) reactor constraint applied.

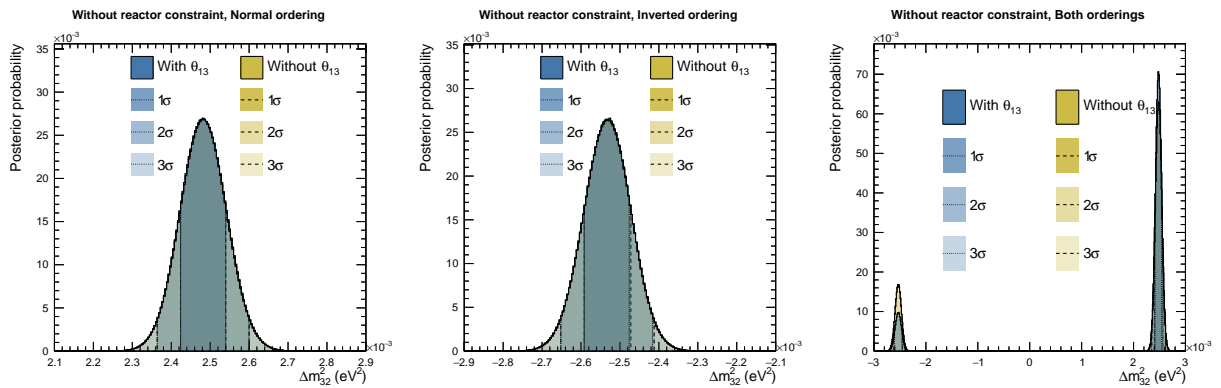


Figure 90: Δm_{32}^2 from real data fit with (blue shaded) and without (yellow shaded) reactor constraint applied.

487 $\sin \delta_{\text{CP}} \geq 0$ is calculated to be 10.5 without reactor constraint and 50.0 with reactor con-
 488 straint.

	$\sin \delta_{\text{CP}} \geq 0$	$\sin \delta_{\text{CP}} < 0$	Sum
No RC	0.086	0.914	1.000
RC	0.019	0.981	1.000

Table 9: Posterior probabilities for data analysis with and without reactor constraint

488

489 **Tab. 10** summarises the Bayes factors of upper octant over lower octant and normal
 490 ordering over inverted ordering for results with reactor constraint and without reactor con-
 491 straint. Compared to results without reactor constraint, with reactor constraint has a slightly
 492 stronger preference of the upper octant. And for mass ordering, results with reactor con-
 493 straint has a stronger preference of normal ordering.

	B(UO/LO)	B(NO/IO)
No RC	0.82	3.76
RC	1.78	7.33

Table 10: Summary table of the Bayes factors of upper octant over lower octant and normal ordering over inverted ordering for data fit results with reactor constraint and without reactor constraint.

494 4.5 Posterior Predictive Spectra and p -values

495 4.5.1 Posterior predictive spectra

496 Posterior predictive spectra are constructed by randomly selecting 3000 throws from the post
497 burn-in steps in the data fit. The mean of each reconstructed kinematic bin is calculated
498 with the RMS of the 3000 throws distribution representing the uncertainty. The data (black
499 markers) are also shown in the same plot. The posterior predictive spectra are shown with
500 (blue markers) and without (red markers) reactor constraint. The posterior predictive spec-
501 tra with reactor constraint are obtained by applying the reactor constraint weights to each
502 throw.

503 To focus on the region where the oscillations occur, the spectra of T2K mu-like beam
504 samples (Fig. 91) binned in reconstructed energy are only shown below 2 GeV, even though
505 higher E_{rec} bins are included in the overall analysis. The e-like T2K samples are binned in
506 both reconstructed neutrino energy and lepton angle, and the projections on the lepton angle
507 are shown in Fig. 92.

508 Fig. 93, Fig. 95 and Fig. 97 are spectra binned in reconstructed lepton momentum
509 for atmospheric SubGeV, MultiGeV and PC&UpMu samples respectively. Though UpMu-
510 nonshower and UpMu-shower samples only have one bin in reconstructed lepton momentum,
511 the overlay plots are shown. Fig. 94, Fig. 96 and Fig. 98 are the spectra binned in recon-
512 structed cosine zenith angle for atmospheric SubGeV, MultiGeV and PC&UpMu samples
513 respectively. Though SubGeV1Re1dcy, SubGeV1Rmu2dcy and SubGeV2Rpi0 samples only
514 have one bin in reconstructed lepton angle, the overlay plots are shown.

515 4.5.2 Impact of uncertainty sources

516 Tab. 11 summarises different groups of systematics uncertainty contributions to the total
517 uncertainty for each SK sample without the reactor constraint applied. The uncertainty is
518 calculated as the change in the total number of events in the sample using to the post-fit
519 constraints on the systematics. The uncertainty on each group is calculated by randomly
520 drawing 3000 throws of systematics and using them to produce sample spectra from post
521 burn-in MCMC chains while keeping the other groups fixed at nominal (AsimovA). When
522 extracting a specific group of systematics contribution, other groups are fixed at their nom-
523 inal values. Hence, the correlation between groups are not considered and thus, the total
524 uncertainty is not naively a sum of different groups in quadrature. The total systematic
525 uncertainties (labelled as “Total Syst.”) are extracted by varying all systematics including

526 near-detector systematics and fixing all oscillation parameters to their Asimov A values. The
527 total uncertainties (labelled as “Total”) are extracted by varying all systematics including
528 near-detector systematics and oscillation parameters.

529 We see that for the T2K samples the cross-section uncertainties and detector uncertainties
530 play the largest role, and are the largest for the FHC 1Re1de and RHC 1Re samples (7%,
531 4%, respectively). The uncertainties are largely the same as in T2K’s analysis, with the
532 exception of the total uncertainty on the FHC 1Re1de, which is 9.2% compared to $\sim 14\%$
533 on T2K, likely due to correlations in the systematics with the atmospheric selections. For
534 the SK atmospheric samples, the atmospheric flux uncertainties are $\sim 3\%$ for the SubGeV
535 samples, which is similar to the impact of the other classes of systematics. As for T2K, the
536 e-like SK samples often have large cross-section systematics than the mu-like equivalents.
537 For the higher energy samples, the uncertainty increases to $\sim 4 - 6\%$ and the detector
538 systematics are the smallest source of uncertainty. Interestingly, the overall uncertainty on
539 the SK atmospheric samples is smaller than for the T2K samples.

540 The uncertainty from all sources (“Total”) may be less than “Total syst” due to correla-
541 tions between oscillation parameters and systematic parameters, which are not accounted for
542 when only calculating the “Total syst”. These correlations can lessen the overall uncertainty
543 on the number of events at SK.

544 4.5.3 Posterior predictive p -values

545 Two p -values are calculated to quantify the agreement between the model and the data. One
546 is calculated using a χ^2 based on total event rate for each sample (single bin analysis) and
547 the other is calculated based on the shape (i.e. bin-by-bin) for each sample. We repeat these
548 studies with and without applying the reactor constraint on $\sin^2 \theta_{13}$.

549 The p -value calculation is based on making random draws of all the parameters (system-
550 atics and oscillation parameters) from the chain after burn in, and building the prediction of
551 the samples for this specific draw. The χ^2 between the draw and the data is calculated, and
552 compared to the χ^2 between the draw and a statistical fluctuation of the draw. The p -value is
553 extracted as the percentage of draws for which the $\chi^2(\text{data}, \text{draw}) < \chi^2(\text{fluct. draw}, \text{draw})$.
554 The p -value calculation compares the expected χ^2 for a given draw ($\chi^2(\text{fluct. draw}, \text{draw})$)
555 against the actual realised χ^2 for the data and the given draw ($\chi^2(\text{data}, \text{draw})$). If the
556 model is insufficiently parametrised, the $\chi^2(\text{data}, \text{draw})$ is expected to be much larger
557 than $\chi^2(\text{fluct. draw}, \text{draw})$. If the model is sufficient, you’d expect $\chi^2(\text{data}, \text{draw})$ to
558 be similarly sized to $\chi^2(\text{fluct. draw}, \text{draw})$. Finally, in the case where we have an over-

559 parametrised model with too much freedom, we expect $\chi^2(\text{data}, \text{draw})$ to be much smaller
 560 than $\chi^2(\text{fluct. draw}, \text{draw})$. Hence, a p -value close to 1 is not necessarily a good thing with
 561 posterior predictive p -values, and should also be investigated.

562 The p -values are investigated in details in [Tab. 12](#). First, we look at the p -value from
 563 both the total event rate and the shape of the spectrum. For the shape of the spectrum, we
 564 project the T2K samples onto the reconstructed neutrino energy E_{rec} and the SK samples
 565 on the reconstructed lepton momentum, p_{lep} , except for the UpThru samples, which are
 566 only binned in $\cos\theta_z$. Secondly, we also restrict the shape p -value to the region relevant to
 567 neutrino oscillations, that is $E_{\text{rec}} < 2$ GeV for the T2K samples and $p_{\text{lep}} < 10$ GeV/c for the
 568 SK samples. Thirdly, we also investigate the impact of the reactor constraint on $\sin^2\theta_{13}$.

569 The overall rate-only p -value is 0.387 without reactor constraint, and 0.420 with the
 570 reactor constraint. For the rate-only p -value, we note the T2K FHC 1Rmu (0.195) and SK
 571 MultiGeV-elike-nue (0.220) fall below $p = 0.25$. As previously noted, these are relatively
 572 strong p -values, and we conclude the rate for all samples to be well described. The reactor
 573 constraint has the largest impact on the T2K RHC 1Re sample (0.475 \rightarrow 0.673), T2K FHC
 574 1Re1de (0.444 \rightarrow 0.351). As expected, the reactor constraint has a much smaller effect on
 575 SK samples.

576 The overall shape p -value falls to 0.106 (0.111) when (not) applying the reactor constraint.
 577 The shape p -value does a better job at exposing weaknesses in modelling the samples. The
 578 T2K FHC 1Rmu (0.240), T2K FHC 1Re (0.209), SK SubGeV-mulike-1dcy (0.162), SK
 579 SubGeV-pi0like (0.162), SK MultiGeV-elike-nuebar (0.207), SK MultiRingOther-1 (0.063),
 580 and SK UpThruNonShower-mu (0.03) samples fall below $p = 0.25$, with the latter two
 581 samples being particularly bad.

582 Breaking up the total shape p -value into the T2K and SK part, the former has $p = 0.471$
 583 and latter has $p = 0.039$. Splitting the SK p -values into SubGeV and MultiGeV+MultiRing,
 584 the former has $p = 0.251$ and the latter has $p = 0.037$ when imposing the 10 GeV/c cut, and
 585 $p = 0.001$ when looking at the entire p_{lep} phase space. Hence, the poor shape p -value comes
 586 primarily from the MultiGeV and MultiRing samples. For the two set of samples that share
 587 systematics, the T2K and SK SubGeV, the p -values are good for both the event-rate (0.536)
 588 and shape (0.365), indicating that the correlated uncertainty model is robust in describing
 589 both set of samples.

590 [Fig. 99](#) complements [Tab. 12](#) by showing the distributions which determine the rate and
 591 shape-based p -values. We see that the underlying distribution is more circular for the rate-
 592 based p -value (implying there is roughly the equal spread in uncertainty from systematics as

593 there are from statistics if the data is similar to the model), whereas the shape-based p -value
594 is more oblong.

595 To illustrate the by-sample p -values, the same underlying distributions are shown by
596 sample in Fig. 100 for SK atmospheric and Fig. 101 for T2K beam samples. We note that
597 for T2K samples the distributions are more horizontal, which comes from the low statistics
598 making the χ^2 from a statistical variation large. The relative contributions to the χ^2 from
599 the samples can also be seen, where the χ^2 from the SK Sub-GeV samples is often below 10,
600 whereas for T2K they are often above 20.

601 To further investigate the poor shape-based posterior predictive p -values of the multi-
602 GeV and multi-ring samples at SK, we instead project the selections on the reconstructed
603 zenith angle, $\cos \theta_z$, in Tab. 13. For these calculations, we keep the T2K samples projected
604 onto the reconstructed neutrino energy E_{rec} and the SK SubGeV samples projected onto
605 the reconstructed lepton momentum p_{lep} . The overall total p -value markedly improves from
606 $p = 0.111$ to $p = 0.322$, although by-sample the change is more nuanced. The SK MultiGeV-
607 like-nue selection goes from relatively well described to acceptable ($p = 0.313 \rightarrow 0.098$) and
608 the SK PCStop selection goes from 0.682 to 0.360. However, the MultiRingOther-1, which
609 was poorly described in p_{Lep} , is now well described in $\cos \theta_z$ with $p = 0.755$. Finally, the
610 p -value from MultiGeV and MultiRing selections, which was $p = 0.037$ in p_{lep} , improves to
611 $p = 0.398$ in $\cos \theta_z$.

612 CW: (JJ)Syst constraint plots need slightly adjustment, will do

613 **This section has been considerably altered from that circulated in draft: v2.0**

614 4.6 Systematic constraints

615 This section shows the prior and posterior constraints on nuisance parameters in the
616 joint-fit. Reactor constraint is applied to the posterior results. Fig. 102, Fig. 103, Fig. 104
617 and Fig. 105 show the overlay (upper panel) and the error band ratio (lower panel) of the
618 prior (red bands) and the posterior (black markers) for flux group, BANFF and other cross-
619 section group and correlated far detector group respectively. Note that the classification of
620 systematics is based on the convenience of their implementation in MaCh3 and has some
621 differences compared to the groups in MaCh3 vs PTheta posterior constraints comparison.
622 However, it is important to note that despite the order and the grouping when presenting
623 results, two fitters use the same set of systematics. The posterior constraints are extracted
624 from the posterior probability distributions which are produced from 1 million randomly

625 drawn steps from the total burn-in steps. Due to limited time, we are not able to utilise all
626 post burn-in steps but will do so in the near future.

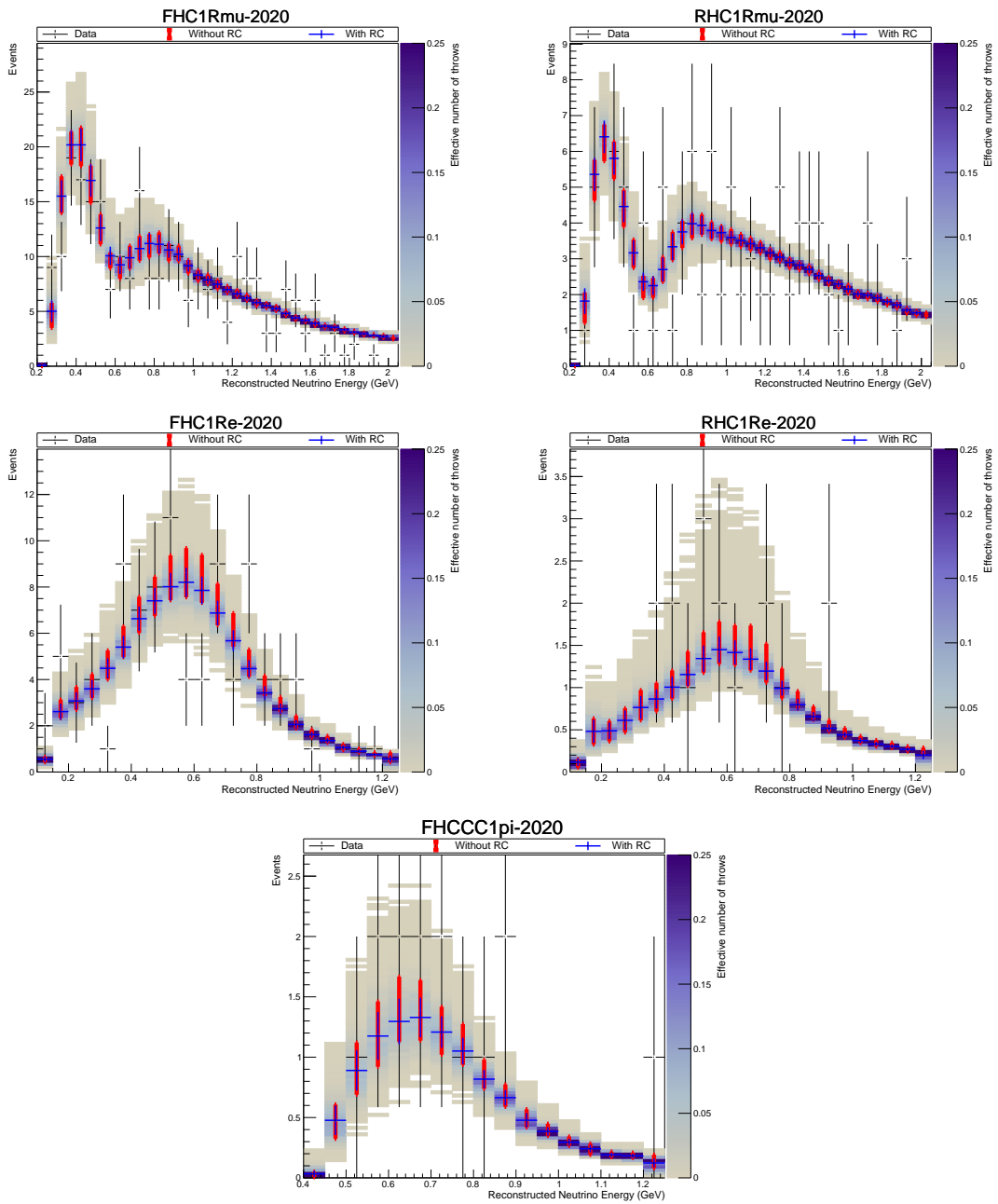


Figure 91: Posterior predictive spectra from the data fit for the T2K beam samples, binned in reconstructed energy. Spectra with (blue) and without (red) reactor constraint are shown.

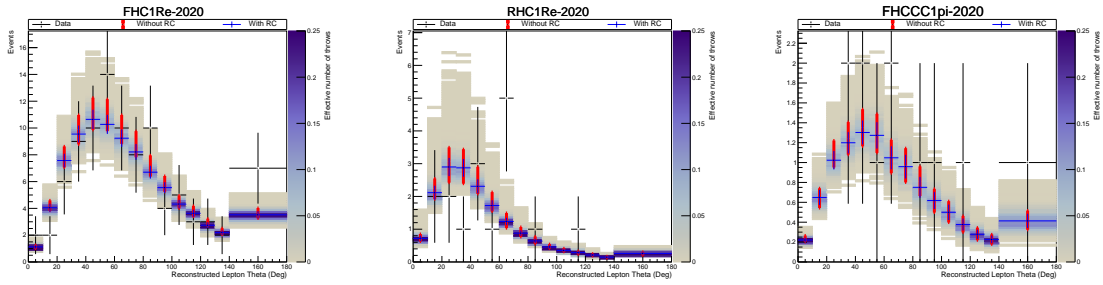


Figure 92: Posterior predictive spectra from the data fit for the T2K e-like beam samples, binned in reconstructed theta angle. Spectra with (blue) and without (red) reactor constraint are shown.

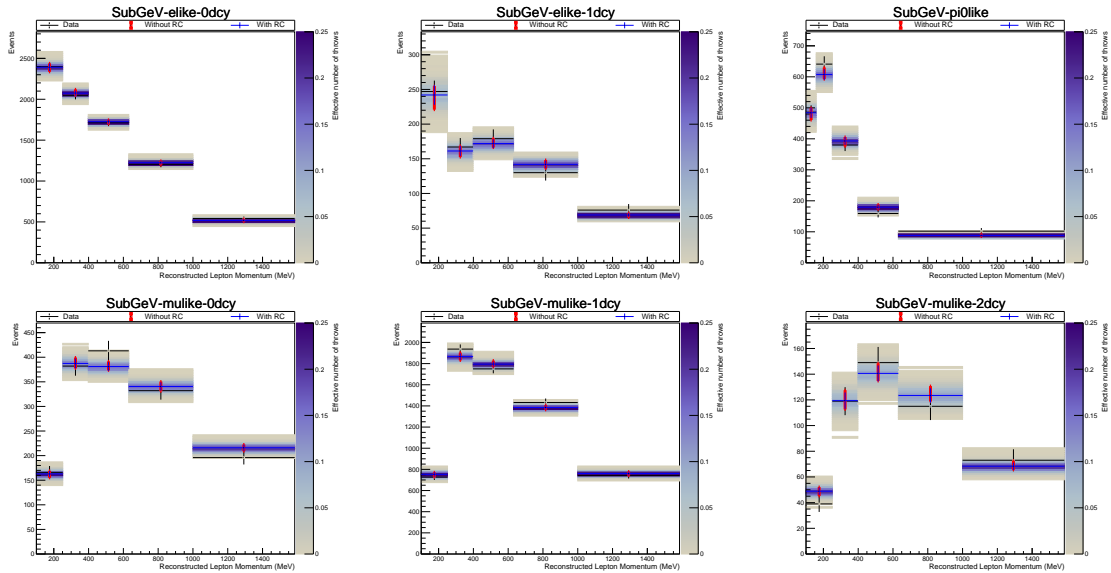


Figure 93: Posterior predictive spectra from the data fit for the SK SubGeV atmospheric samples binned in reconstructed lepton momentum. Spectra with (blue) and without (red) reactor constraint are shown.

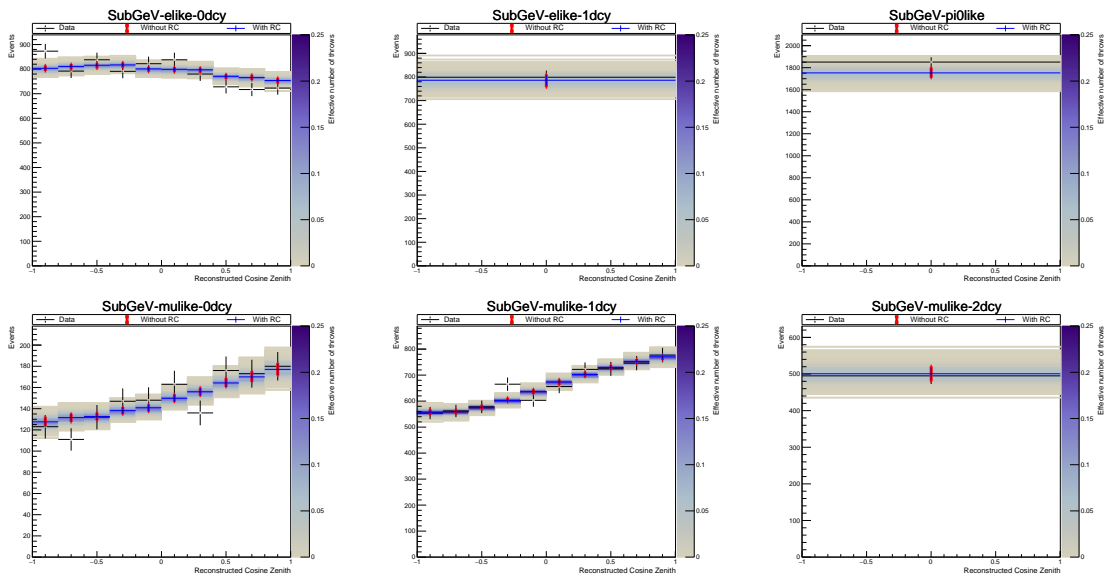


Figure 94: Posterior predictive spectra from the data fit for the SK SubGeV atmospheric samples binned in reconstructed cosine zenith angle. Spectra with (blue) and without (red) reactor constraint are shown.

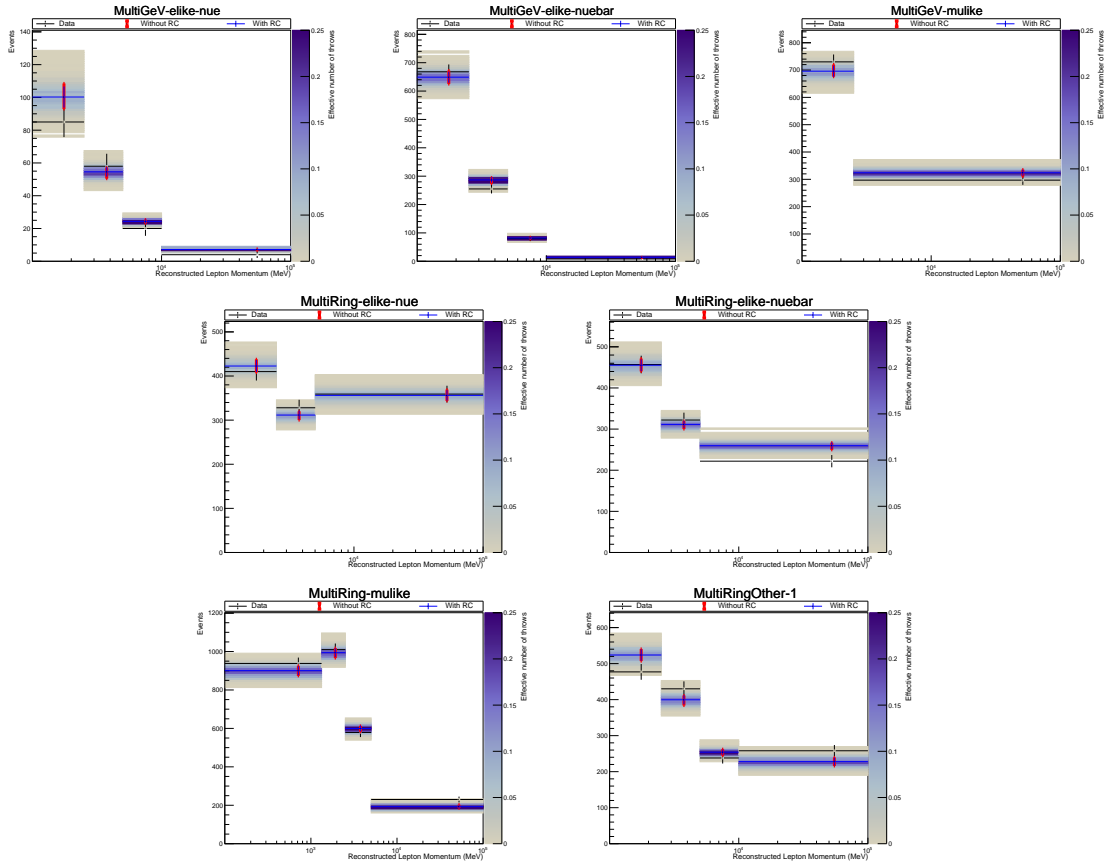


Figure 95: Posterior predictive spectra from the data fit for the SK MultiGeV atmospheric samples binned in reconstructed lepton momentum. Spectra with (blue) and without (red) reactor constraint are shown.

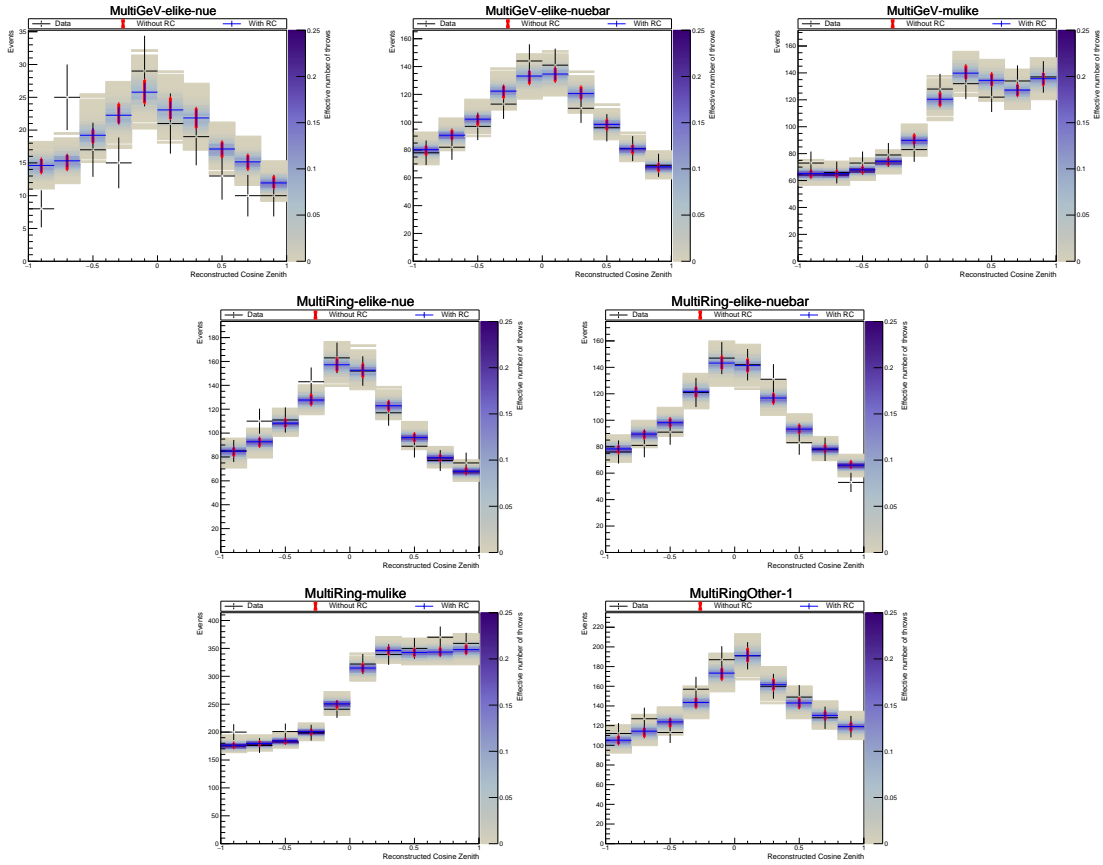


Figure 96: Posterior predictive spectra from the data fit for the SK MultiGeV atmospheric samples binned in reconstructed cosine zenith angle. Spectra with (blue) and without (red) reactor constraint are shown.

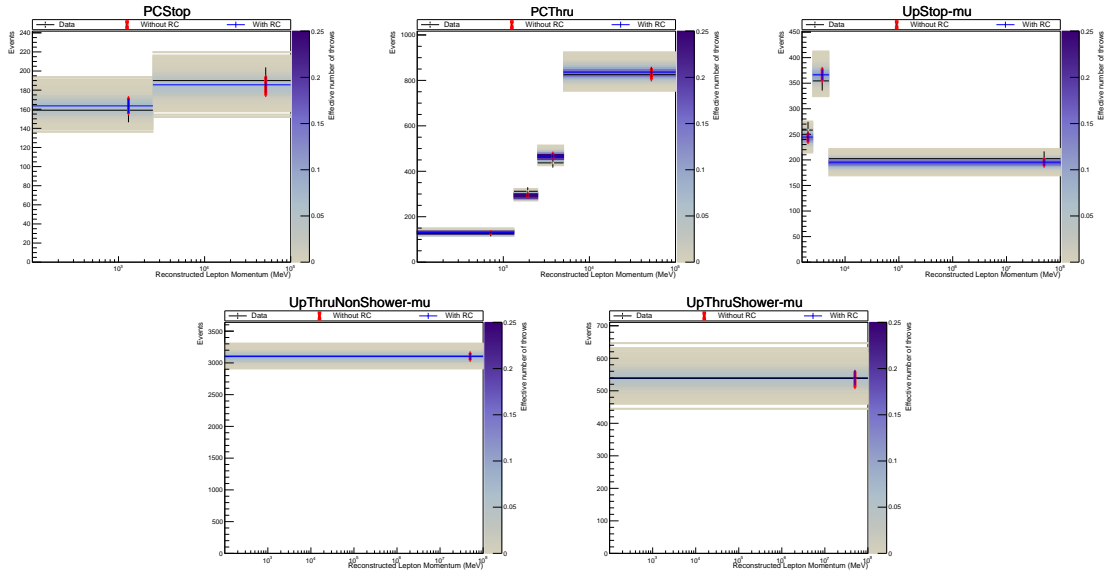


Figure 97: Posterior predictive spectra from the data fit for the SK PC (partially-contained) and UpMu atmospheric samples binned in reconstructed lepton momentum. Spectra with (blue) and without (red) reactor constraint are shown.

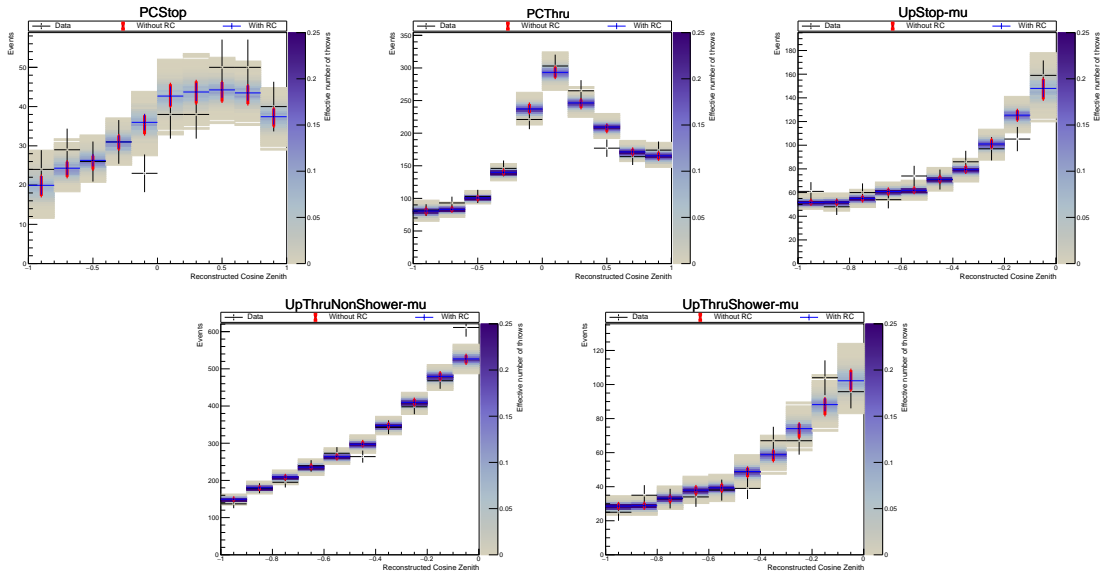


Figure 98: Posterior predictive spectra from the data fit for the SK PC (partially-contained) and UpMu atmospheric samples binned in reconstructed cosine zenith angle. Spectra with (blue) and without (red) reactor constraint are shown.

Sample	SK atm. flux	T2K beam flux	SK det.	Cross sections	Total Syst.	Total
T2K FHC 1Rmu	0.00%	2.54%	2.39%	2.65%	2.81%	3.13%
T2K RHC 1Rmu	0.00%	2.64%	2.28%	3.58%	3.69%	3.80%
T2K FHC 1Re	0.00%	2.55%	2.69%	3.95%	3.99%	8.87%
T2K RHC 1Re	0.00%	2.72%	4.35%	4.37%	6.35%	14.75%
T2K FHC 1Re1de	0.00%	2.53%	7.87%	6.37%	9.21%	12.14%
SK SubGeV- <i>elike-0dcy</i>	3.53%	0.00%	2.55%	3.47%	1.42%	1.07%
SK SubGeV- <i>elike-1dcy</i>	3.00%	0.00%	3.39%	4.38%	3.03%	2.77%
SK SubGeV- <i>mulike-0dcy</i>	3.02%	0.00%	3.19%	2.72%	2.29%	2.16%
SK SubGeV- <i>mulike-1dcy</i>	3.08%	0.00%	2.58%	2.70%	1.28%	1.13%
SK SubGeV- <i>mulike-2dcy</i>	3.01%	0.00%	2.72%	4.34%	3.32%	3.28%
SK SubGeV- <i>pi0like</i>	2.72%	0.00%	2.39%	3.61%	2.32%	2.34%
SK MultiGeV- <i>elike-nue</i>	4.43%	0.00%	3.26%	7.09%	5.49%	5.61%
SK MultiGeV- <i>elike-nuebar</i>	4.23%	0.00%	3.21%	4.43%	2.97%	2.87%
SK MultiGeV- <i>mulike</i>	4.20%	0.00%	2.73%	4.23%	2.87%	2.90%
SK MultiRing- <i>elike-nue</i>	4.30%	0.00%	3.18%	4.26%	2.76%	2.73%
SK MultiRing- <i>elike-nuebar</i>	4.22%	0.00%	3.37%	4.24%	2.73%	2.63%
SK MultiRing- <i>mulike</i>	4.22%	0.00%	2.28%	4.12%	1.76%	1.75%
SK MultiRing- <i>Other-1</i>	4.15%	0.00%	3.84%	5.03%	2.61%	2.56%
SK PCStop	4.37%	0.00%	4.80%	3.61%	4.45%	4.50%
SK PCThru	3.17%	0.00%	2.24%	3.82%	2.09%	2.09%
SK UpStop- <i>mu</i>	4.51%	0.00%	2.00%	3.77%	2.92%	2.99%
SK UpThruNonShower- <i>mu</i>	4.33%	0.00%	1.66%	3.90%	1.80%	1.76%
SK UpThruShower- <i>mu</i>	5.78%	0.00%	5.16%	3.72%	4.19%	4.22%

Table 11: Table of different groups of systematics uncertainty contributions to the joint analysis. The uncertainty is calculated as the change in the total number of events in the sample using to the post-fit constraints on the systematics. “SK atm. flux” refers to the group of atmospheric flux systematics. “T2K beam flux” refers to the group of T2K beam flux systematics. “SK det.” refers to the group of correlated far detector systematics. Note that the two newly developed systematics associated with e/μ PID at low pion momentum region are included in “SK det.”. “Cross sections” refers to the group of cross-section systematics, including the new ad-hoc Adler angle parameter. “Total Syst.” refers to the total error values calculated by varying all the systematics including near-detector systematics and fixing the oscillation parameters at Asimov A. “Total” refers to the total error values calculated by varying all the systematics, including near-detector systematics, and oscillation parameters. All the uncertainties are calculated from using the standard deviations of the posterior predictive distributions. No reactor constraint is applied.

Sample	woRC rate	woRC shape	wRC rate	wRC shape
T2K FHC 1Rmu	0.195	0.240	0.191	0.241
T2K RHC 1Rmu	0.725	0.643	0.742	0.640
T2K FHC 1Re	0.515	0.209	0.498	0.222
T2K RHC 1Re	0.475	0.625	0.673	0.658
T2K FHC 1Re1de	0.444	0.912	0.351	0.902
SK SubGeV-eliike-0dcy	0.491	0.574	0.492	0.580
SK SubGeV-eliike-1dcy	0.500	0.553	0.533	0.571
SK SubGeV-mulike-0dcy	0.559	0.356	0.559	0.338
SK SubGeV-mulike-1dcy	0.499	0.162	0.490	0.159
SK SubGeV-mulike-2dcy	0.559	0.491	0.591	0.486
SK SubGeV-pi0like	0.481	0.162	0.482	0.162
SK MultiGeV-eliike-nue	0.220	0.313	0.236	0.314
SK MultiGeV-eliike-nuebar	0.454	0.207	0.431	0.201
SK MultiGeV-mulike	0.512	0.271	0.550	0.308
SK MultiRing-eliike-nue	0.518	0.400	0.526	0.407
SK MultiRing-eliike-nuebar	0.422	0.657	0.431	0.641
SK MultiRing-mulike	0.276	0.389	0.255	0.415
SK MultiRingOther-1	0.518	0.063	0.495	0.047
SK PCStop	0.545	0.682	0.554	0.689
SK PCThru	0.462	0.378	0.452	0.369
SK UpStop-mu	0.544	0.564	0.537	0.569
SK UpThruNonShower-mu	0.513	0.030	0.530	0.026
SK UpThruShower-mu	0.508	0.429	0.515	0.423
Total	0.387	0.111	0.420	0.106
T2K only	0.465	0.471	0.491	0.505
SK only	0.384	0.039	0.414	0.036
SK SubGeV only	0.569	0.251	0.591	0.251
T2K+SK SubGeV only	0.536	0.365	0.569	0.384
SK Multi-GeV/Ring, rest.	0.334	0.037	0.355	0.059
SK Multi-GeV/Ring, full	0.321	0.001	0.328	0.000

Table 12: Table of different p -values from the data fit broken down in to each SK sample. The χ^2 calculation is based on both total event rate and shape, and the impact of the reactor constraint is shown (“wRC” being with reactor constraint, and “woRC” being without reactor constraint). A kinematic cut of $E_{\text{rec}} < 2$ GeV is applied for T2K samples and $p_{\text{lep}} < 10$ GeV/c for SK samples.

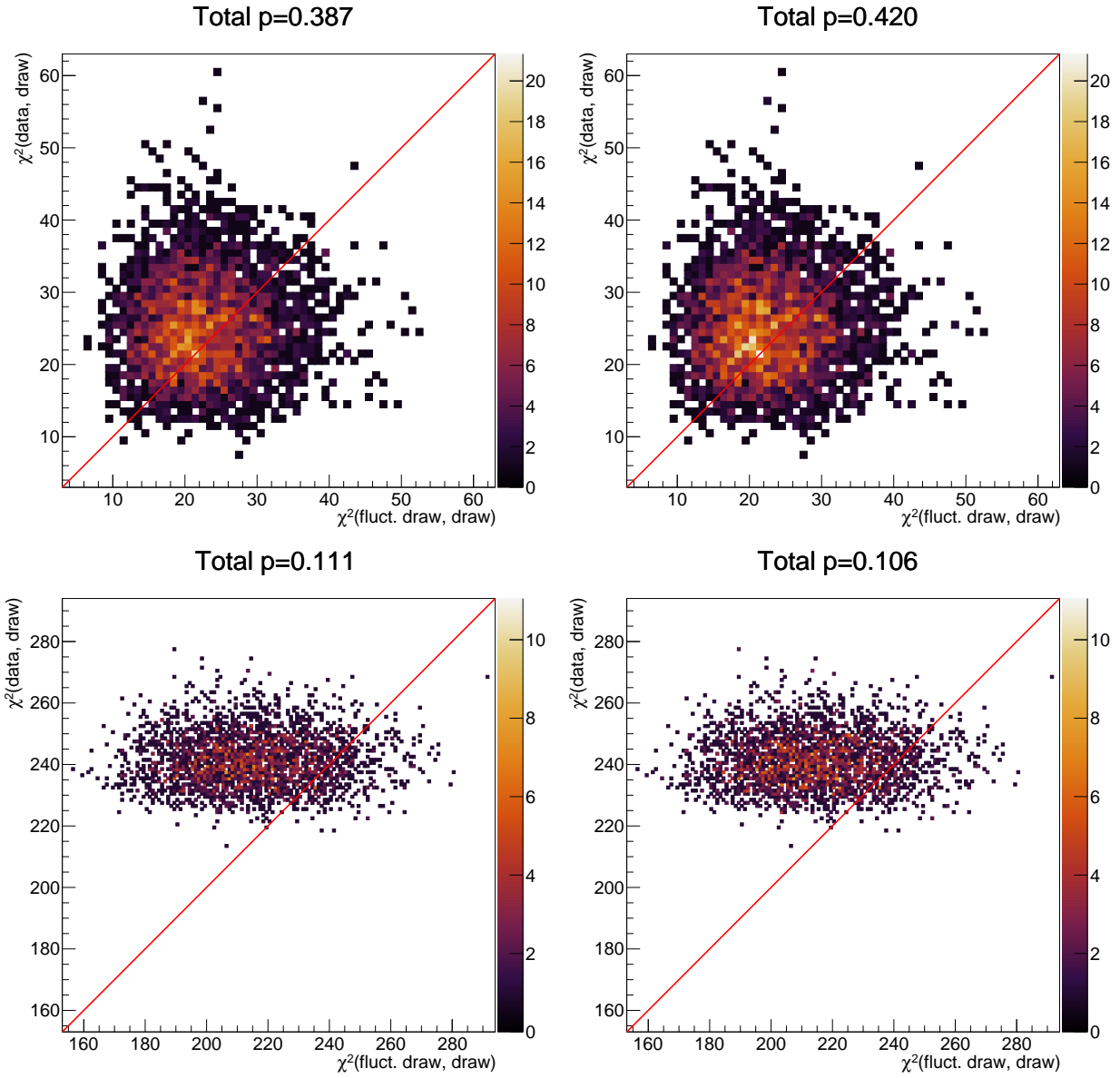


Figure 99: Rate (top) and shape (bottom) based posterior predictive p -values without reactor constraint (left) and with reactor constraint (right), using $E_{\text{rec}} < 2$ GeV for the T2K samples, and $p_{\text{lep}} < 10$ GeV/c for the SK atmospheric samples.

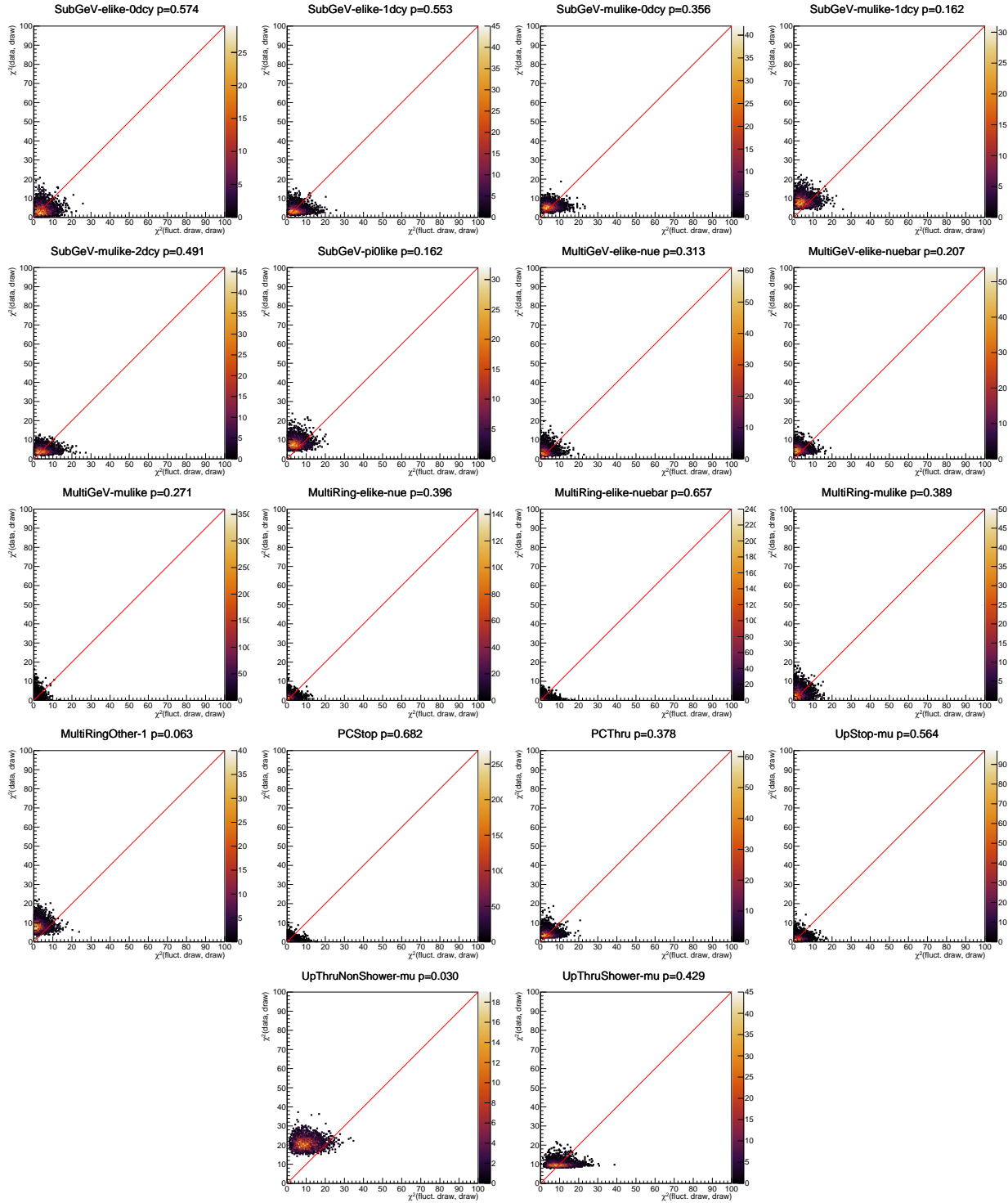


Figure 100: Shape-based posterior predictive p -value without reactor constraint for the SK atmospheric samples, using $p_{lep} < 10$ GeV/c.

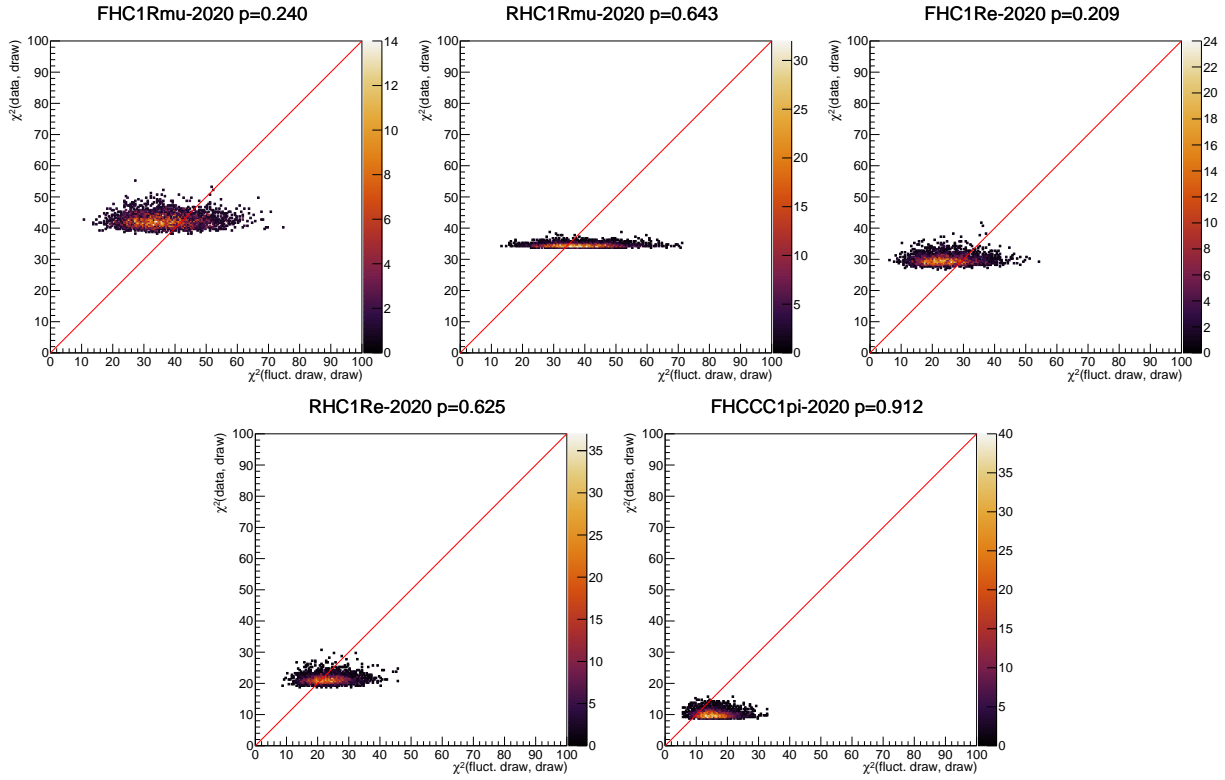


Figure 101: Shape-based posterior predictive p -value without reactor constraint for the T2K beam samples, using $E_{\text{rec}} < 2$ GeV.

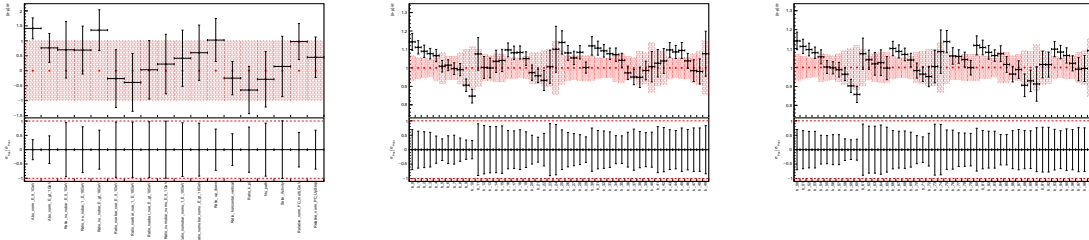


Figure 102: Comparison of prior and posterior constraints on flux parameters. From left to right: atmospheric flux, beam near detector flux and beam far detector flux. For each plot, the upper panel shows the overlay of central values and error bands of prior constraints (red bands) used in joint-fit and posterior constraints (black markers) from joint-fit with reactor constraint applied. The lower panel shows the error bands of the posterior constraints normalised by the prior constraints. The red dashed line indicate the self-normalised prior constraints. All the posterior errors are the standard deviations of the posterior probability distributions.

Sample	Shape, p_{lep}	Shape, $\cos \theta_z$
T2K FHC 1Rmu		0.240
T2K RHC 1Rmu		0.643
T2K FHC 1Re		0.209
T2K RHC 1Re		0.625
T2K FHC 1Re1de		0.912
SK SubGeV-elike-0dcy		0.574
SK SubGeV-elike-1dcy		0.553
SK SubGeV-mulike-0dcy		0.356
SK SubGeV-mulike-1dcy		0.162
SK SubGeV-mulike-2dcy		0.491
SK SubGeV-pi0like		0.162
SK MultiGeV-elike-nue	0.313	0.098
SK MultiGeV-elike-nuebar	0.207	0.870
SK MultiGeV-mulike	0.271	0.841
SK MultiRing-elike-nue	0.400	0.622
SK MultiRing-elike-nuebar	0.657	0.630
SK MultiRing-mulike	0.389	0.588
SK MultiRingOther-1	0.063	0.755
SK PCStop	0.682	0.360
SK PCThru	0.378	0.321
SK UpStop-mu	0.564	0.336
SK UpThruNonShower-mu		0.030
SK UpThruShower-mu		0.429
Total	0.111	0.322
SK MultiGeV/ring rest.	0.037	0.398

Table 13: Table of shape-based p -values from the data fit for each far-detector sample, where T2K and SK SubGeV samples are projected on E_{rec} and p_{Lep} , respectively, and SK multi-ring and MultiGeV samples are projected on $\cos \theta_z$. The reactor constraint is not applied. For the T2K samples, a $E_{\text{rec}} < 2$ GeV cut is applied, and for the SK samples a $p_{\text{Lep}} < 10$ GeV/c cut is applied.

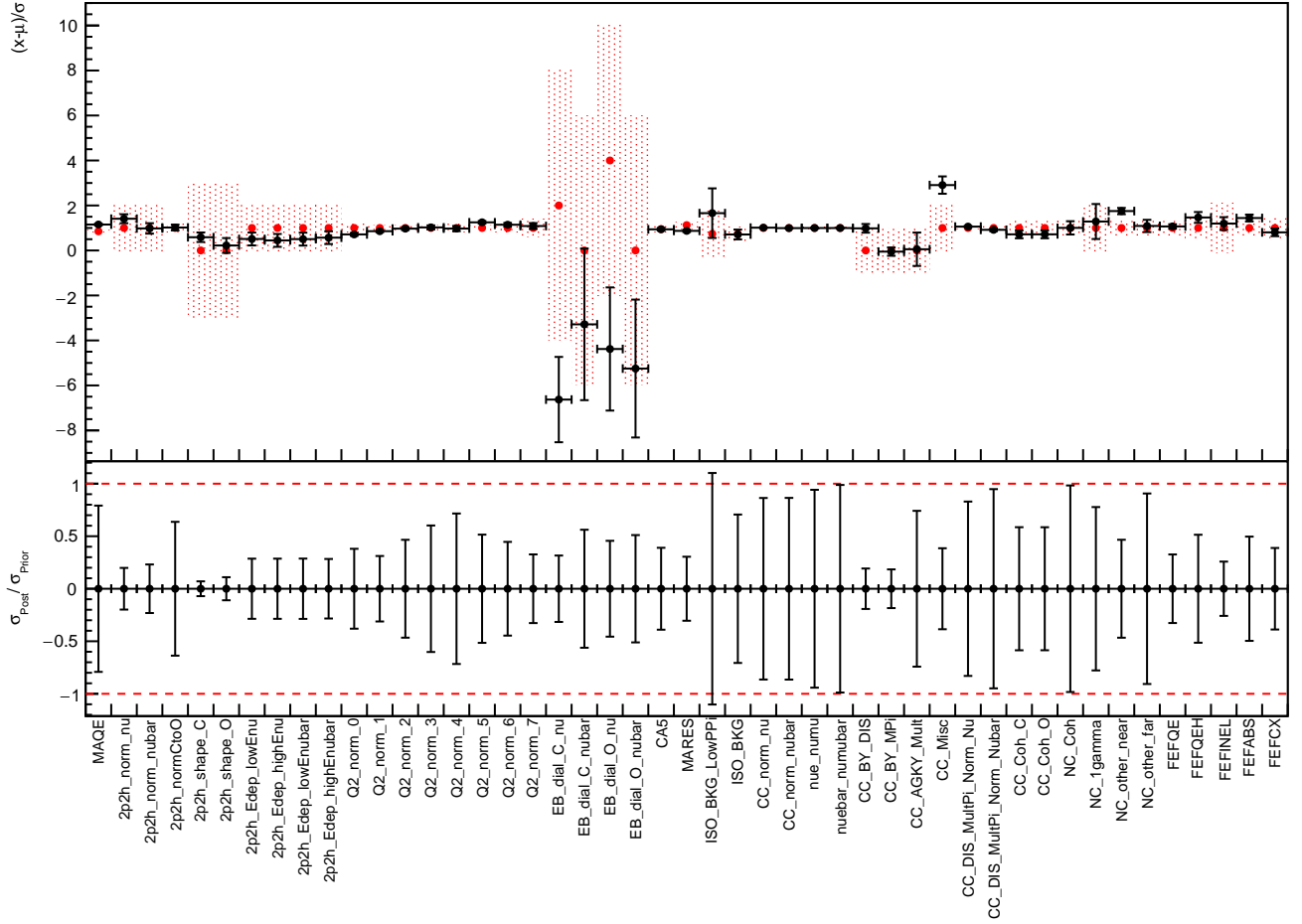


Figure 103: Comparison of prior and posterior constraints on BANFF cross-section parameters. In the plot, the upper panel shows the overlay of central values and error bands of prior constraints (red bands) used in joint-fit and posterior constraints (black markers) from joint-fit with reactor constraint applied. The lower panel shows the error bands of the posterior constraints normalised by the prior constraints. The red dashed line indicate the self-normalised prior constraints. All the posterior errors are the standard deviations of the posterior probability distributions.

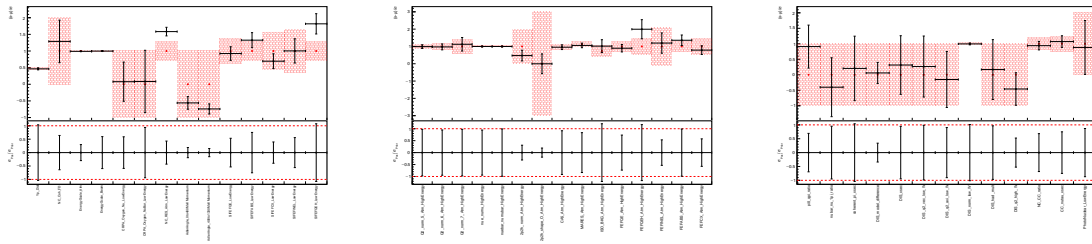


Figure 104: Comparison of prior and posterior constraints on other cross-section parameters. The middle plot is the high energy group. For each plot, the upper panel shows the overlay of central values and error bands of prior constraints (red bands) used in joint-fit and posterior constraints (black markers) from joint-fit with reactor constraint applied. The lower panel shows the error bands of the posterior constraints normalised by the prior constraints. The red dashed line indicate the self-normalised prior constraints. All the posterior errors are the standard deviations of the posterior probability distributions.

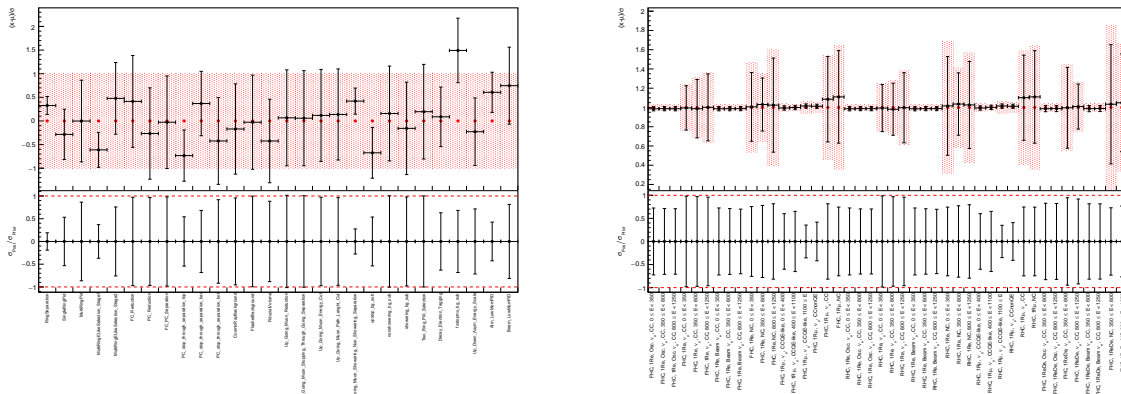


Figure 105: Comparison of prior and posterior constraints on the correlated far detector parameters. The left plot is the SK detector group plus the Up_Down_Asym_EnergyScale and two newly developed systematics associated with e/μ PID in the low lepton momentum region. For each plot, the upper panel shows the overlay of central values and error bands of prior constraints (red bands) used in joint-fit and posterior constraints (black markers) from joint-fit with reactor constraint applied. The lower panel shows the error bands of the posterior constraints normalised by the prior constraints. The red dashed line indicate the self-normalised prior constraints. All the posterior errors are the standard deviations of the posterior probability distributions.

627 4.7 Comparison to OA2020 MaCh3 results

628 This section compares parameter constraints and Bayes factors between the MaCh3 joint-fit
629 with SK and T2K's OA2020 result, on which it is based. The same reactor constraint on
630 $\sin^2 \theta_{13}$ is applied to the joint-fit as the T2K OA2020 analysis TN393 [1] from the PDG [10].
631 However, a different smearing of Δm_{32}^2 is used: for joint-fit, the applied smearing in the joint
632 fit is $3.6 \times 10^{-5} \text{ eV}^2$ and for OA2020 was $1.4 \times 10^{-5} \text{ eV}^2$ (Sec. 1.1 in TN393 [1]).

633 The post burn-in steps in OA2020 is about 170 million and in joint-fit is about 230
634 million. All results presented in this section are obtained from post burn-in steps.

635 4.7.1 Parameter constraints

636 This section presents both the one dimensional and two dimensional posterior comparisons.
637 For each oscillation parameter, results in the normal, inverted and both orderings are over-
638 laid. The red lines are contours and intervals from joint-fit and the black lines are contours
639 and intervals from OA2020. Different credible intervals are presented in different line styles.

640 **One-dimensional posteriors** Fig. 106 is the one dimensional contour comparison of δ_{CP} .
641 The joint-fit has a slightly stronger constraint on δ_{CP} than OA2020, especially around $\delta_{\text{CP}} \sim \pi$.
642 For the inverted ordering, both the joint-fit and OA2020 can exclude δ_{CP} at 0 and $-\pi$ between
643 2σ and 3σ . For normal ordering and both ordering, joint-fit still can exclude both 0 and $-\pi$
644 between 2σ and 3σ and OA2020 can only exclude 0 between 2σ and 3σ .

645 Fig. 107 is the one dimensional contour comparison of $\sin^2 \theta_{23}$. For the normal ordering
646 and both orderings, the joint-fit shows a weaker preference for one octant over another,
647 consistent with the Bayes factor calculations presented earlier. For the inverted ordering,
648 compared with OA2020, joint-fit shows a slightly weaker preference of upper octant. This is
649 consistent with SK's publication of atmospheric samples (see SK PTEP 2019).

650 Fig. 108 is the one dimensional contour comparison of Δm_{32}^2 . Compared with OA2020,
651 the joint-fit constraint on Δm_{32}^2 is slightly weaker. The main reason is that the two results
652 use different Δm_{32}^2 smearing values and the one used by the joint-fit is larger than OA2020
653 by ~ 2.57 times.

654 Fig. 109 is the one dimensional contour comparison of $\sin^2 \theta_{13}$. Two results agree very
655 well which is expected since both analyses used the same reactor constraint on $\sin^2 \theta_{13}$.

656 **Two-dimensional posteriors** Fig. 110 is the two dimensional contour comparison of
657 $\delta_{\text{CP}} - \sin^2 \theta_{13}$. Since the same reactor constraint is applied in the two analyses, the constraint

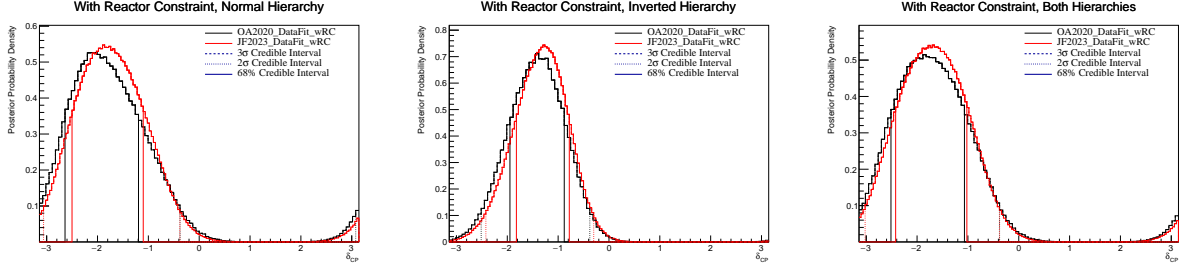


Figure 106: δ_{CP} for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

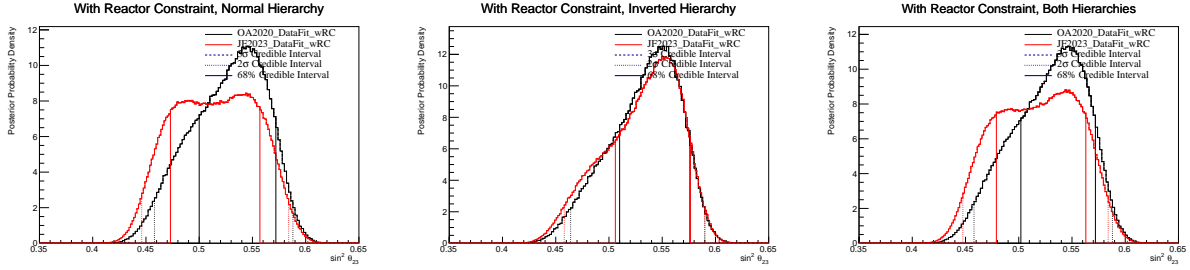


Figure 107: $\sin^2 \theta_{23}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

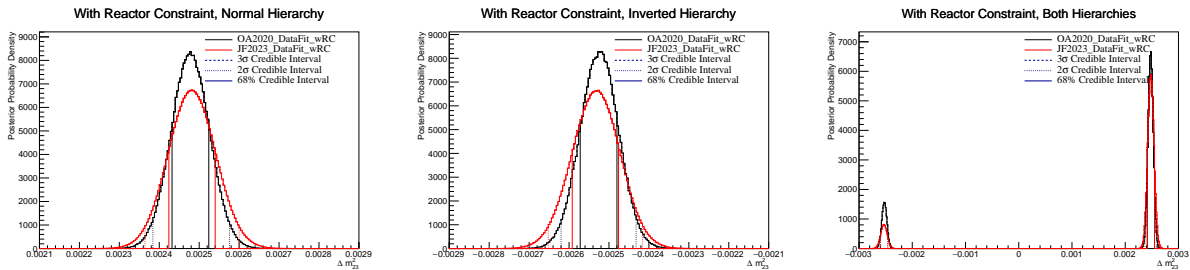


Figure 108: Δm_{32}^2 for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

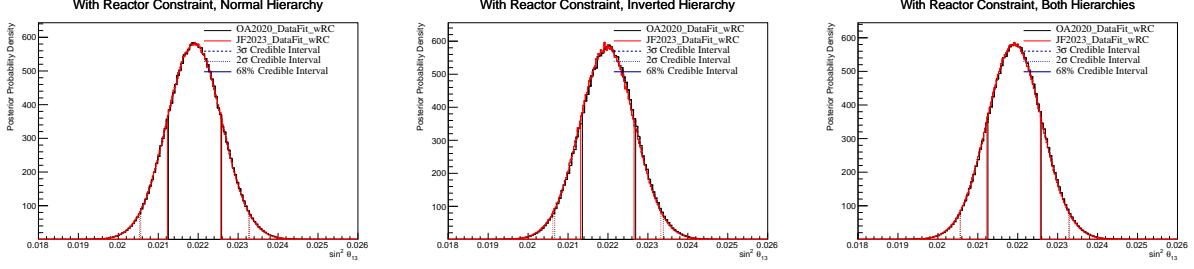


Figure 109: $\sin^2 \theta_{13}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

658 on $\sin^2 \theta_{13}$ are very similar. The better constraint on δ_{CP} from joint-fit is more obvious from
 659 [Fig. 110](#) than [Fig. 106](#).

660 [Fig. 111](#) is the two dimensional contour comparison of $\delta_{CP} - \sin^2 \theta_{23}$. A better constraint
 661 on δ_{CP} and a preference for maximal mixing from joint-fit compared with OA2020 can be
 662 seen clearly.

663 [Fig. 112](#) is the two dimensional contour comparison of $\delta_{CP} - \Delta m_{32}^2$. A slight relaxation
 664 on the constraint of Δm_{32}^2 can be seen from joint-fit compared with OA2020. As explained
 665 in the one dimensional contour comparison section, the main reason is the different Δm_{32}^2
 666 smearing value used in the two analyses.

667 [Fig. 113](#) is the two dimensional contour comparison of $\Delta m_{32}^2 - \sin^2 \theta_{13}$. Similar to [Fig. 112](#),
 668 a lightly weaker constraint on Δm_{32}^2 is seen from joint-fit compared with OA2020.

669 [Fig. 114](#) and [Fig. 116](#) are the two dimensional contour comparison of Δm_{32}^2 against
 670 $\sin^2 \theta_{23}$. The preference of a maximal mixing from joint-fit is very obvious in normal ordering
 671 comparison. The relaxation of Δm_{32}^2 from joint-fit can be seen in both normal ordering and
 672 inverted ordering.

673 [Fig. 115](#) is the two dimensional contour comparison of $\sin^2 \theta_{13} - \sin^2 \theta_{23}$. A stretch of
 674 the contours towards the lower octant of $\sin^2 \theta_{23}$ from joint-fit is quite obvious in normal
 675 ordering. For inverted ordering, the stretch is smaller.

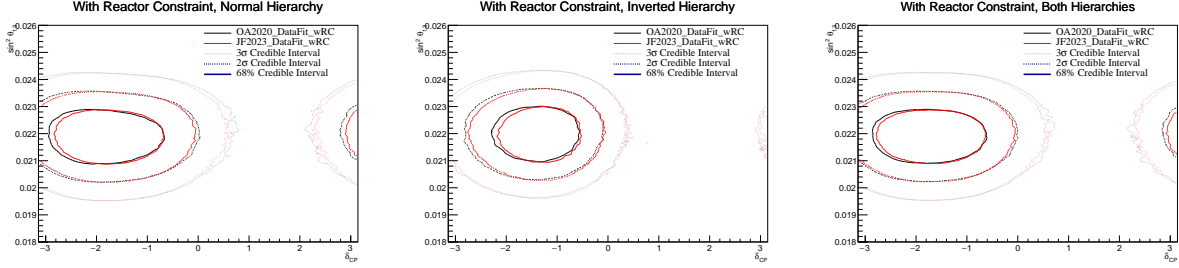


Figure 110: $\delta_{\text{CP}} - \sin^2 \theta_{13}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

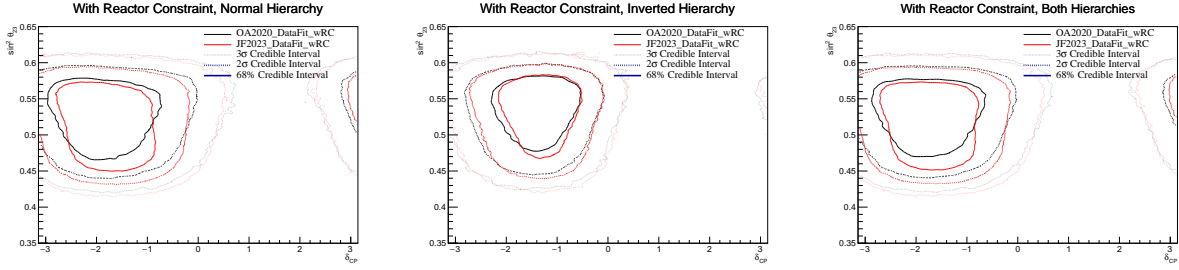


Figure 111: $\delta_{\text{CP}} - \sin^2 \theta_{23}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

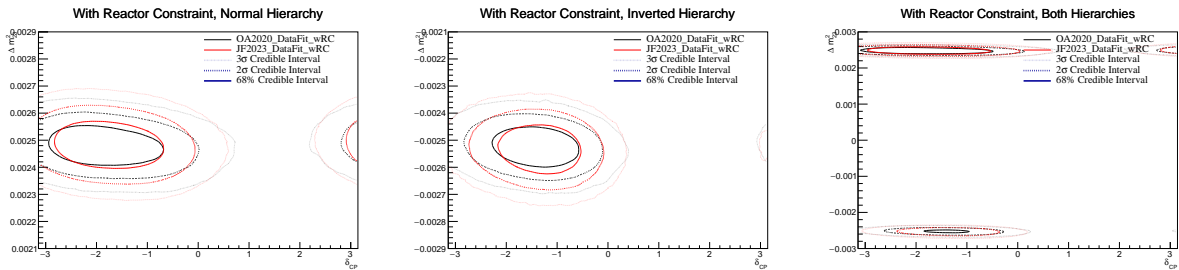


Figure 112: $\delta_{\text{CP}} - \Delta m_{32}^2$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

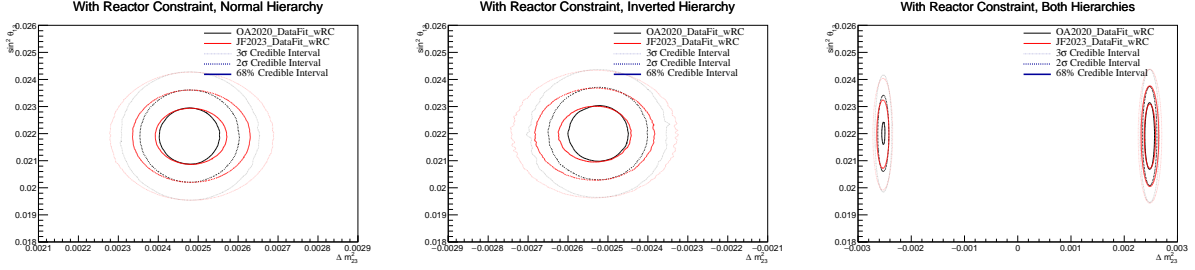


Figure 113: $\Delta m_{32}^2 - \sin^2 \theta_{13}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

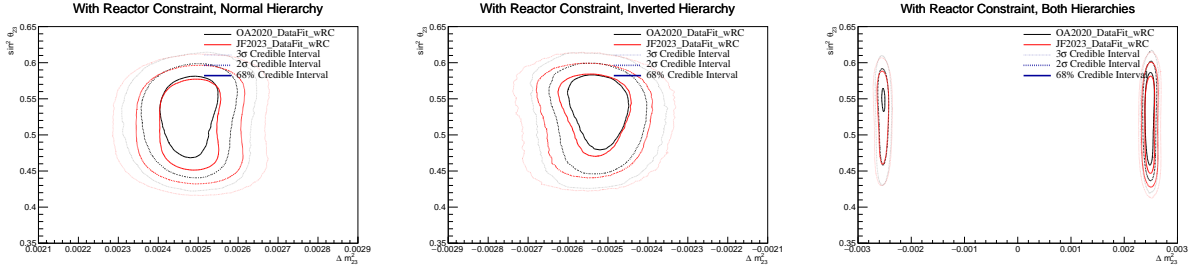


Figure 114: $\Delta m_{32}^2 - \sin^2 \theta_{23}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

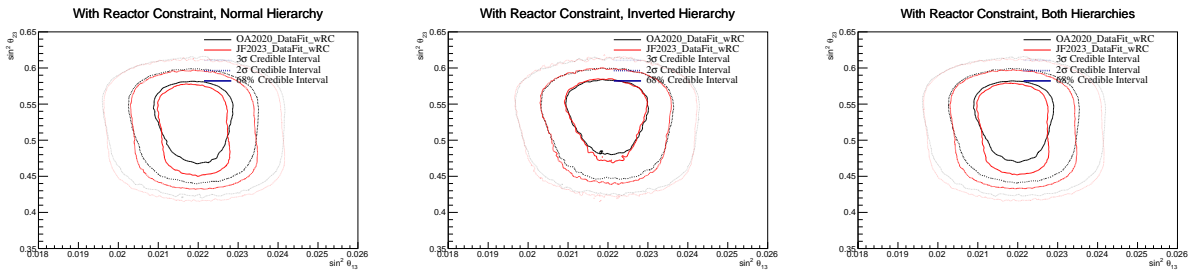


Figure 115: $\sin^2 \theta_{13} - \sin^2 \theta_{23}$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

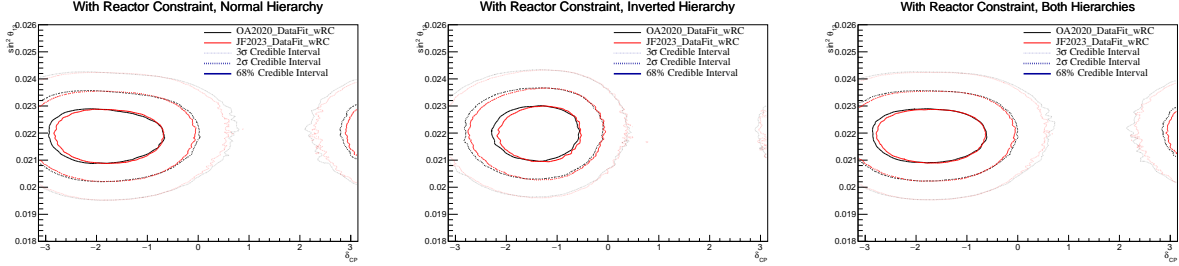


Figure 116: $\sin^2 \theta_{23} - \Delta m_{32}^2$ for the real data analysis comparison between joint-fit (red lines) and OA2020 (black lines), with same reactor constraint and different Δm_{32}^2 smearing values applied, for normal (left), inverted (center), and both (right) orderings.

676 4.7.2 Bayes factors for atmospheric parameters

677 The joint-fit data fit results with reactor constraint is presented in Tab. 7 and OA2020 results
 678 with reactor constraint in presented in Table 13 of TN393 [1]. Tab. 14 summarises the figures
 679 in one table. As seen from the previous comparison of contours, the joint-fit has a smaller
 680 Bayes factor for upper octant over lower octant compared with OA2020 and a larger Bayes
 681 factor for normal over inverted ordering despite different Δm_{32}^2 smearing values used. The
 682 Bayes factors from both analyses do not reach a level of significance where conclusions could
 683 be drawn to exclude either possibilities of the mass hierarchy or the octants. However, both
 analyses find consistent preferences of upper octant and normal ordering.

	T2K+SK atm.	T2K 2020
B(UO/LO)	1.78	3.36
B(NO/IO)	7.33	4.20

Table 14: Comparison of Bayes factors for the joint fit between T2K beam and SK atmospheric neutrinos, and the main results from T2K. Same reactor constraint is applied to both results. Different smearing of $\sin^2 \theta_{23}$ is applied.

684

5 Results comparisons

This section is dedicated to show the comparison of the joint-fit results from MaCh3 and PTheta, the comparison of joint-fit results and T2K official analysis (OA2020) from both MaCh3 and PTheta, and also the investigation of joint-fit results differences between MaCh3 and PTheta. All MaCh3 joint-fit results are produced from post burn-in steps (burn-in being 80,000 steps). All distributions shown in this section have the same reactor constraint applied, which was the same as OA2020.

5.1 Joint-fit results comparison between MaCh3 and PTheta

Parameter constraints Fig. 117, Fig. 118, Fig. 119 and Fig. 120 show the overlay of joint-fit results for δ_{CP} , $\sin^2 \theta_{23}$, Δm_{32}^2 ($|\Delta m_{31}^2|$) and $\sin^2 \theta_{13}$ from MaCh3 and PTheta respectively. Fig. 121 and Fig. 122 are the overlay comparison of Jarlskog invariants using flat prior on δ_{CP} and $\sin \delta_{\text{CP}}$ respectively. Results from PTheta are produced from PTheta MCMC method. Hence, for $\Delta\chi^2$ representation, MCMC results are converted for the purpose of comparisons.

In Fig. 117, PTheta has a stronger constraint on δ_{CP} than MaCh3 for normal ordering with very similar δ_{CP} value for the highest-posterior density. Although the posterior at this value is different, they are almost indistinguishable in the regions of no CP violation ($\delta_{\text{CP}} = 0, \pi$). This was also observed in OA2020 TN393 [1] (Fig. 61). For inverted ordering, the PTheta result shows a shift to higher Δm_{32}^2 compared to MaCh3. For either orderings, both fitters exclude $\sin \delta_{\text{CP}} = 0$ at 2σ . A similar observation can be drawn from Fig. 121 and Fig. 122. Fig. 121 shows both fitters exclude $\delta_{\text{CP}} = 0$ at 2σ (see Fig. 49) with a flat prior on δ_{CP} , and PTheta has a stronger constraint. Fig. 122 shows a similar but weaker tendency when there's a flat prior on $\sin \delta_{\text{CP}}$, and PTheta has a slightly stronger constraint.

In Fig. 118, for normal ordering hypothesis, no strong preference to upper/lower octant is seen in either fitter results. For normal ordering, compared with the slightly bimodal contour shape from PTheta, MaCh3 shows a stronger preference for maximal mixing. For inverted ordering, both fitters show a weak preference to the upper octant. However, both results are consistent with each other at 1σ .

For normal ordering in Fig. 119, a consistent shift of the best-fit and credible intervals of PTheta is observed relative to MaCh3. Since PTheta has only $|\Delta m_{31}^2|$ for inverted ordering, MaCh3 results of Δm_{32}^2 are converted to $|\Delta m_{31}^2|$ by subtracting $|\Delta m_{21}^2|$ from $|\Delta m_{23}^2|$. A similar shift of Δm^2 also exists for inverted ordering. Potential explanations are documented in Sec. 5.3.

From Fig. 120, it is clear that the results from two fitters use the same reactor constraint.

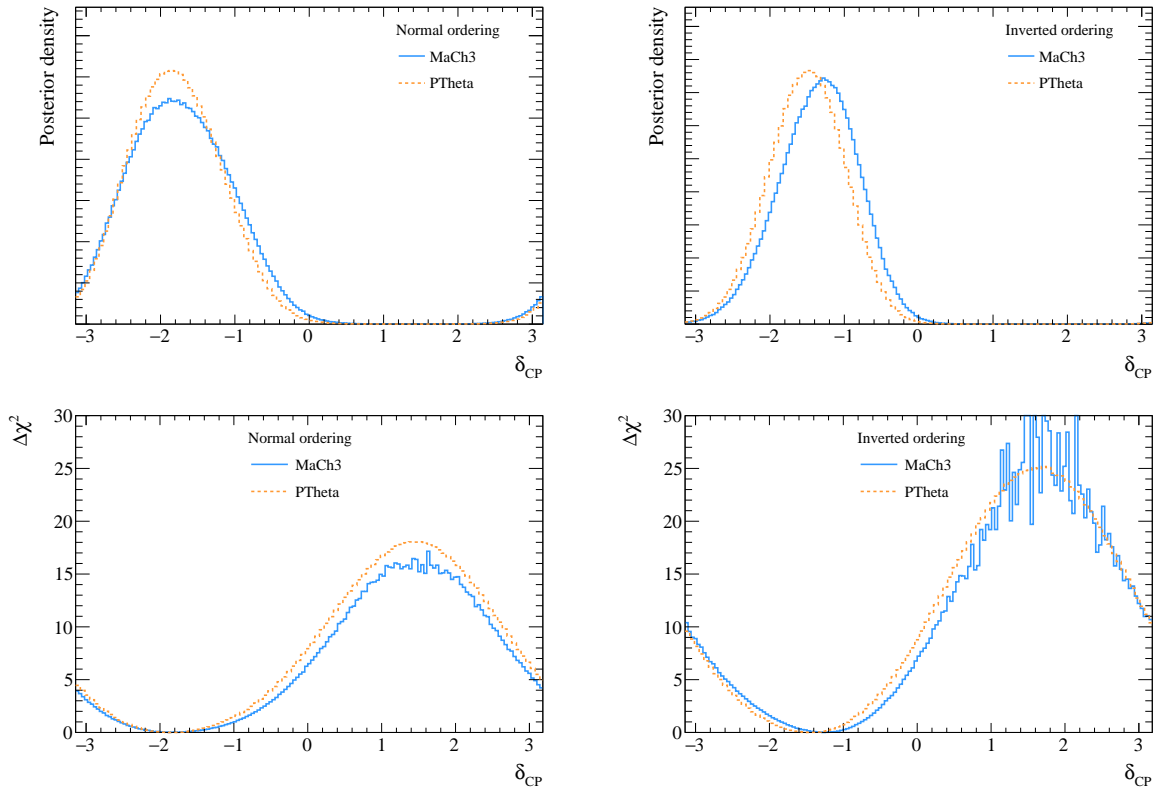


Figure 117: Overlay of δ_{CP} from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Upper row shows the posterior probability distributions and lower row shows $\Delta\chi^2$ distributions. Left column shows the normal ordering results and right column shows the inverted ordering results. Same reactor constraint is applied.

718 **Bayes factors for atmospheric parameters** Tab. 15 summarises the Bayes factors for
 719 octant hypotheses and mass ordering from the joint fit results of MaCh3 and P-Theta. The
 720 Bayes factors from the two fitters are quite consistent. In terms of Bayes factor value, both
 721 fitters show a weaker preference of one octant over another and both fitters show a stronger
 722 preference to normal ordering compared to OA2020 results [1]. The Bayes factor values from
 723 both fitters are not sufficient to make a conclusion on hypotheses.

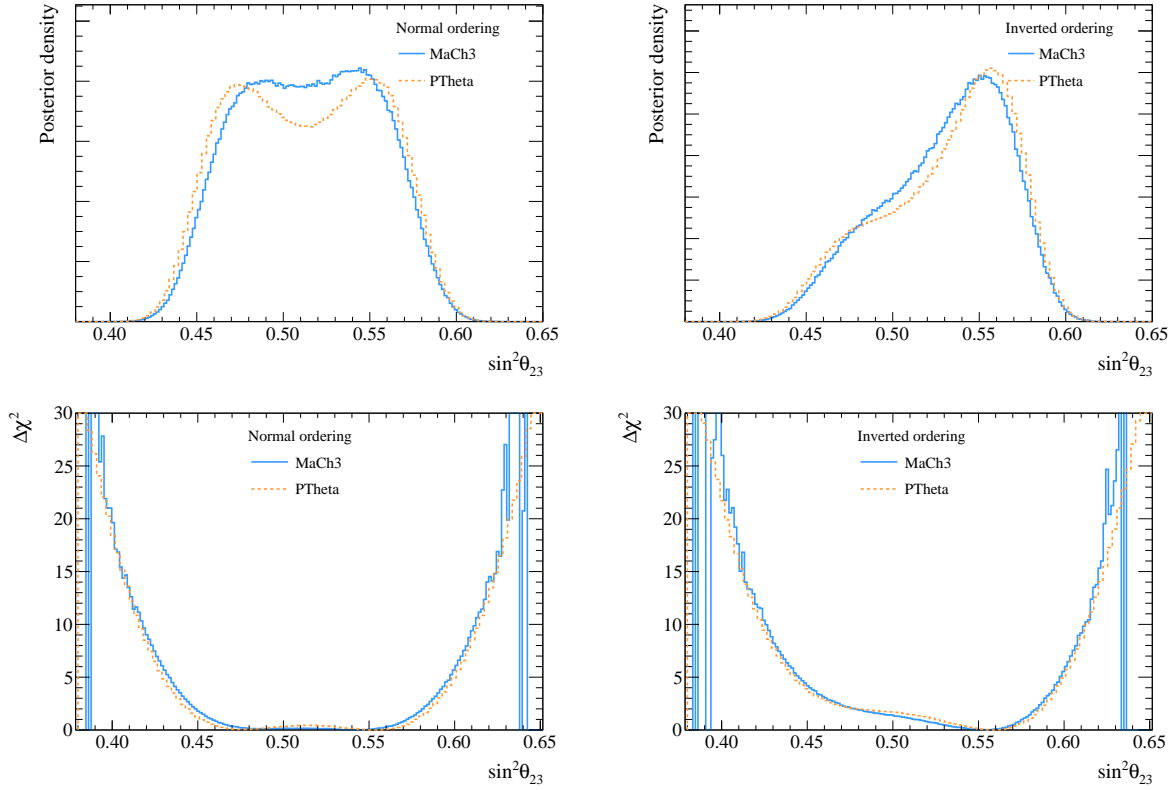


Figure 118: Overlay of $\sin^2 \theta_{23}$ from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Upper row shows the posterior probability distributions and lower row shows $\Delta\chi^2$ distributions. Left column shows the normal ordering results and right column shows the inverted ordering results. Same reactor constraint is applied.

	MaCh3	P-Theta
B(UO/LO)	1.78	1.57
B(NO/IO)	7.33	8.98

Table 15: Comparison of Bayes factors for the joint fit between MaCh3 and P-Theta. Same reactor constraint and same smearing of $\sin^2 \theta_{23}$ are applied to both results. P-Theta results are documented in TN459 [12] Table 5 in v1.1.

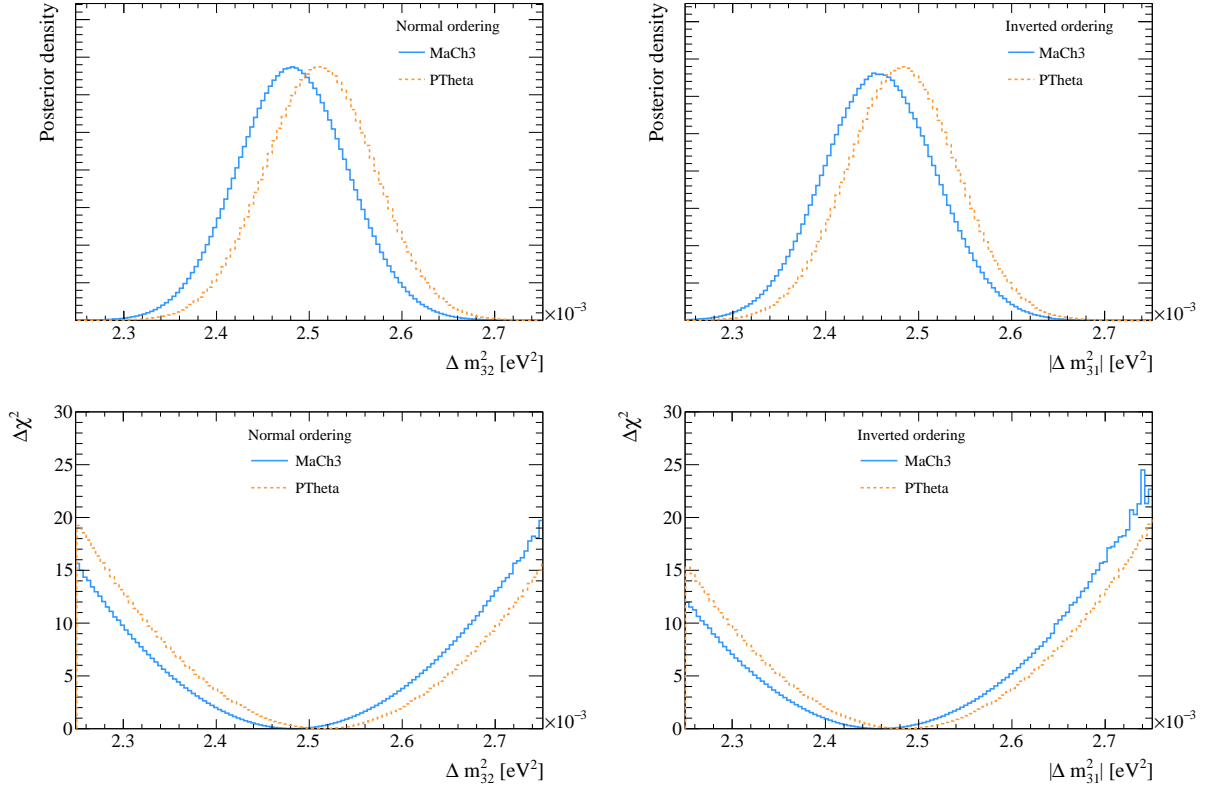


Figure 119: Overlay of Δm_{32}^2 and $|\Delta m_{31}^2|$ from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Upper row shows the posterior probability distributions and lower row shows $\Delta\chi^2$ distributions. Left column shows the normal ordering results of Δm_{32}^2 and right column shows the inverted ordering results of $|\Delta m_{31}^2|$. Same reactor constraint is applied.

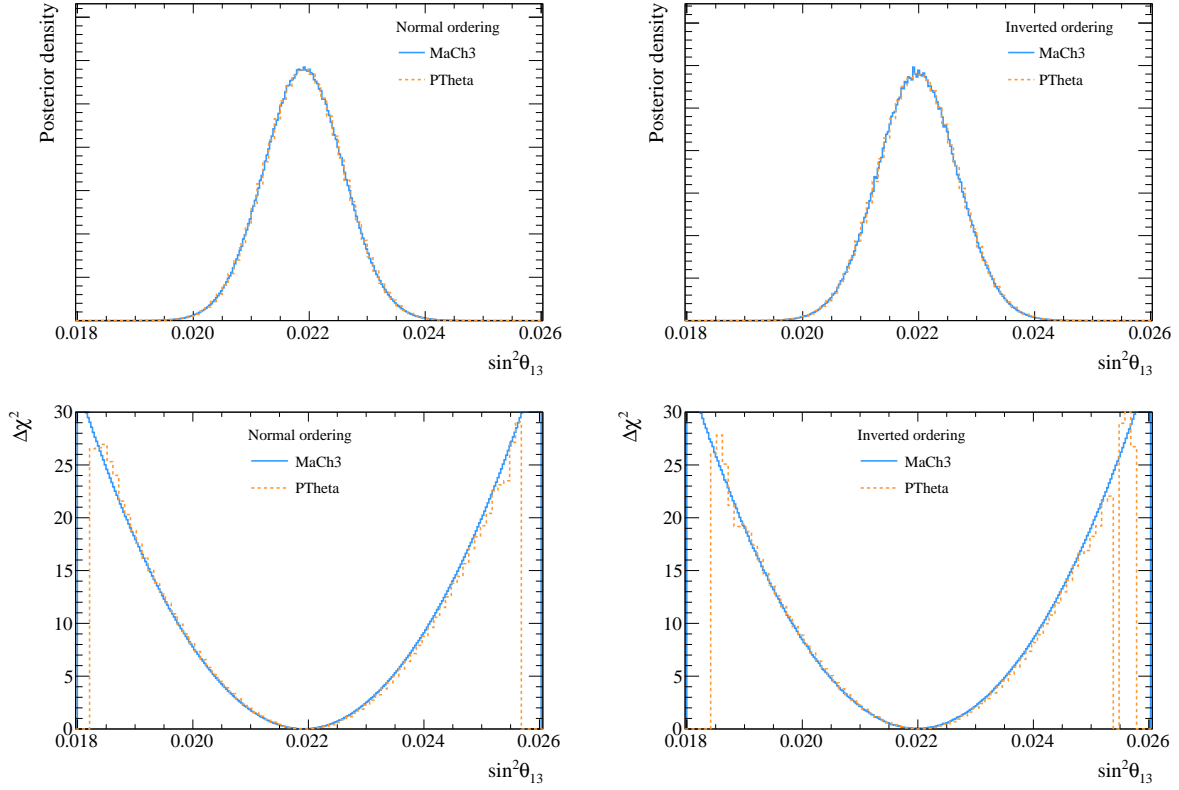


Figure 120: Overlay of $\sin^2 \theta_{13}$ from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Upper row shows the posterior probability distributions and lower row shows $\Delta\chi^2$ distributions. Left column shows the normal ordering results and right column shows the inverted ordering results. Same reactor constraint is applied.

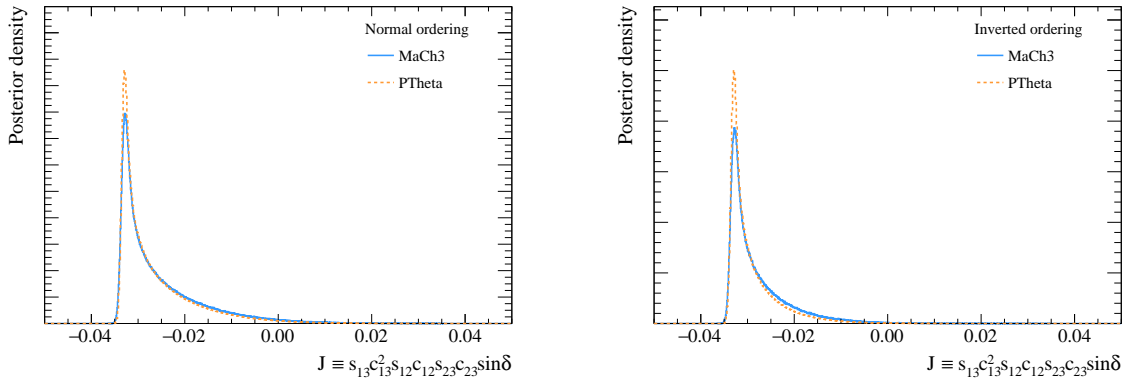


Figure 121: Overlay of J from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Left plot shows the normal ordering results and right plot shows the inverted ordering results. Same reactor constraint is applied. Flat prior on δ_{CP} is assumed.

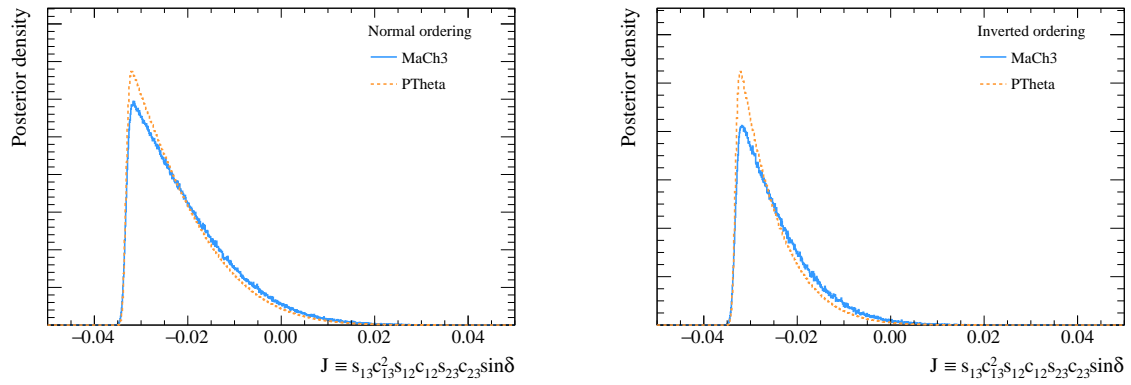


Figure 122: Overlay of J from data fits of MaCh3 (blue solid line) and PTheta (orange dashed line). Left plot shows the normal ordering results and right plot shows the inverted ordering results. Same reactor constraint is applied. Flat prior on $\sin \delta_{\text{CP}}$ is assumed.

5.2 Joint-fit results and OA2020 from MaCh3 and PTheta

Fig. 123, Fig. 124, Fig. 125 and Fig. 126 show the joint-fit results overlaid with T2K official results (OA2020) from MaCh3 and PTheta. MCMC results from MaCh3 are converted to $\Delta\chi^2$ distributions for comparison purpose.

For normal ordering comparison in Fig. 123, both fitters show stronger constraints on δ_{CP} in the joint-fit compared to T2K-only. MaCh3 joint-fit results are closer to MaCh3 OA2020 results because MaCh3 intended to always strictly follow OA2020 prescriptions, documented in TN393 [1]. The differences of joint-fit P-Theta and OA2020 can be found in TN459 [12] (see Appendix A.1). For normal ordering, both fitters' results from joint-fit can exclude $\sin\delta_{\text{CP}} = 0$ at 2σ which shows better constraint compared to OA2020. For inverted ordering, similarly, both fitter results show better constraint than OA2020.

In Fig. 124, for normal ordering and inverted ordering, there is a weaker preference for the upper octant in the joint-fit results (both fitters) compared to OA2020 (both fitters). This behaviour mainly comes from the SK atmospheric samples, as discussed in the P-Theta joint-fit results TN459 [12] (see Fig 21 in v1.1), and can also be seen in 2019 SK oscillation analysis [13]. Especially in normal ordering, both fitters in the joint-fit show a preference of maximal mixing.

For Δm^2 in Fig. 125, the two fitters from OA2020 are very close and MaCh3 results from the joint-fit have the highest posterior density point very close to that from OA2020, but show a weaker constraint than OA2020. This might due to that the joint-fit results adopt a larger smearing on Δm_{32}^2 than OA2020. Interestingly, PTheta's results from the joint-fit show a non-negligible shift compared to both OA2020 and joint-fit MaCh3 results. Detailed discussion of this shift of Δm_{32}^2 can be found in TN459 [12] (see Appendix A in v1.1).

Fig. 120 shows that the same reactor constraint is applied to the results shown in this section.

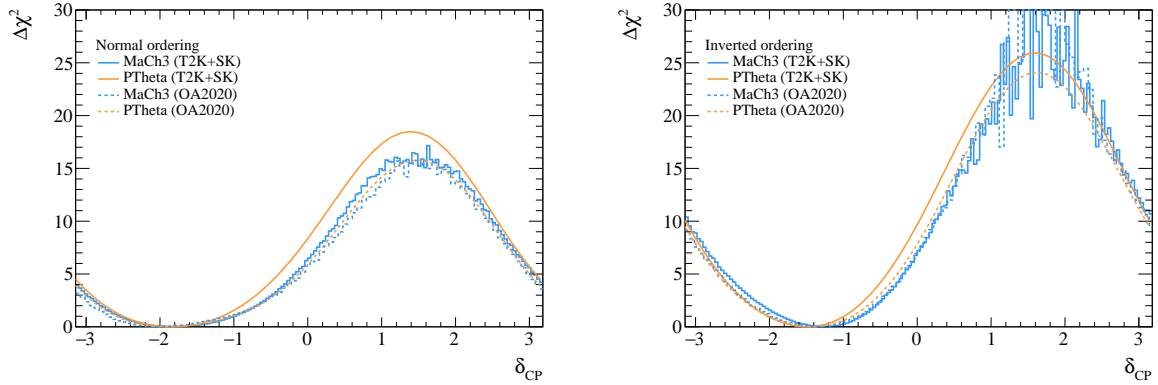


Figure 123: Overlay of δ_{CP} from data fits of joint-fit and T2K official analysis. The solid lines are results from joint-fit and the dashed lines are results from OA2020. Blue lines are MaCh3 results and orange lines are PTheta results. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

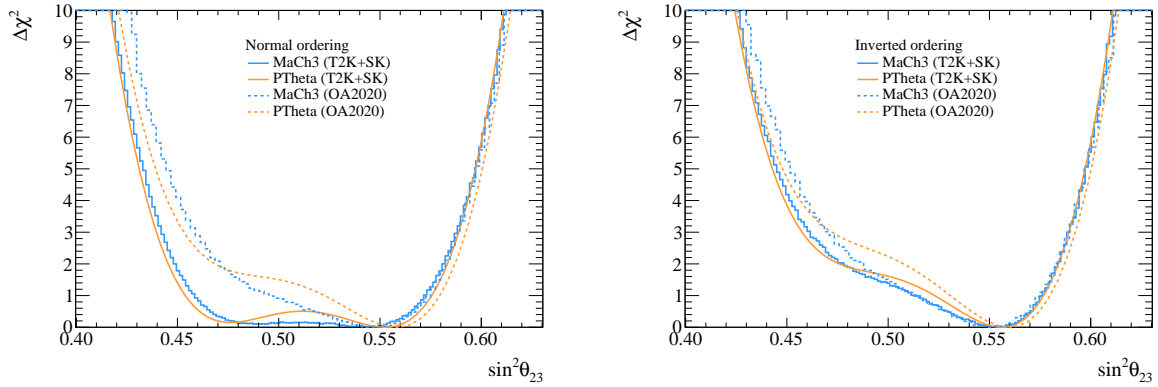


Figure 124: Overlay of $\sin^2\theta_{23}$ from data fits of joint-fit and T2K official analysis. The solid lines are results from joint-fit and the dashed lines are results from OA2020. Blue lines are MaCh3 results and orange lines are PTheta results. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

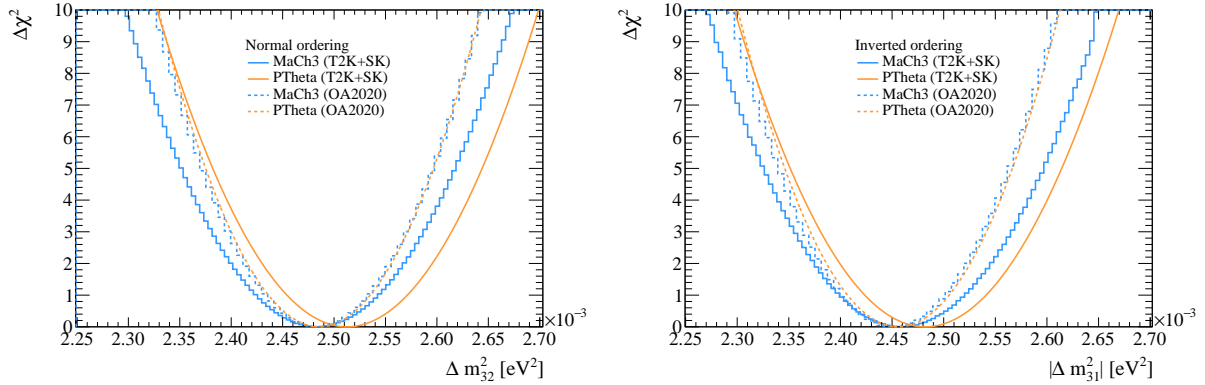


Figure 125: Overlay of Δm_{32}^2 and $|\Delta m_{31}^2|$ from data fits of joint-fit and T2K official analysis. The solid lines are results from joint-fit and the dashed lines are results from OA2020. Blue lines are MaCh3 results and orange lines are PTheta results. Left plot shows the normal ordering $\Delta\chi^2$ distributions of Δm_{32}^2 and right plot shows the inverted ordering $\Delta\chi^2$ distributions of $|\Delta m_{31}^2|$. Same reactor constraint is applied.

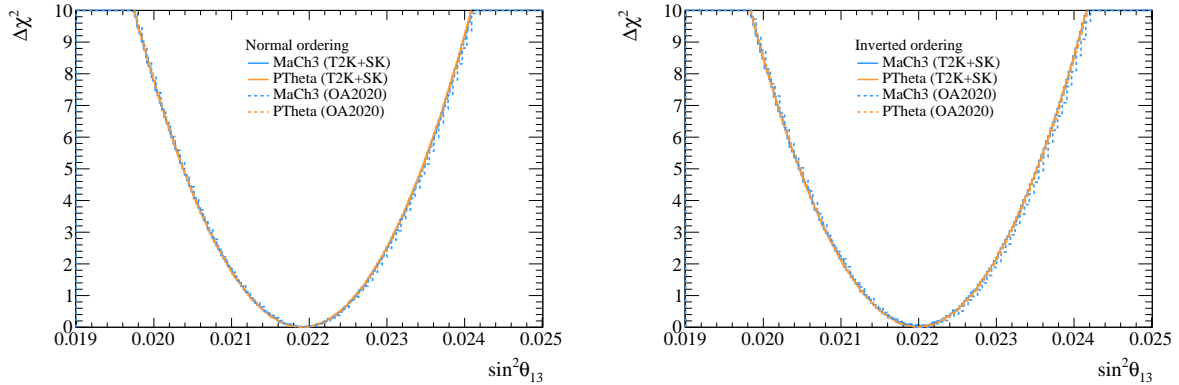


Figure 126: Overlay of $\sin^2\theta_{13}$ from data fits of joint-fit and T2K official analysis. The solid lines are results from joint-fit and the dashed lines are results from OA2020. Blue lines are MaCh3 results and orange lines are PTheta results. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

5.3 Potential explanation of results differences

In this section, implementation differences are investigated to understand the differences between the joint-fit results from MaCh3 and PTheta.

Kinematic cut: the order of applying the cut and applying the shifts (e.g. Coulomb correction) to the reconstructed kinematics in the event selections is different in the two fitters. MaCh3 follows the OA2020 style and applies the cut before any corrections, whilst PTheta applies the cut afterwards following the OA2021 approach. When testing this option, PTheta uses the same order as MaCh3 while keeps other implementation unchanged.

E_{rec} binning: MaCh3 uses E_{rec} binning for beam μ -like samples and $E_{rec} - \theta$ binning for beam e-like samples while PTheta uses $E_{rec} - \theta$ binning for beam μ -like samples and $p_{rec} - \theta$ for beam e-like samples. When testing this option, PTheta uses the same binning as MaCh3 whilst keeping other implementation choices unchanged.

Kinematic cut + E_{rec} binning: to investigate the combined effect of the above two implementation differences, PTheta uses the same implementation as MaCh3 whilst keeping other implementation choices unchanged.

Near detector fit: PTheta and MaCh3 fit near detector samples in different ways. MaCh3 fits near detector samples simultaneously with far detector samples while PTheta uses BANFF inputs to propagate near detector constraints to far detector samples. This has been previously studied in OA2020, as documented in TN393 [1]. Its impact will be studied in the next round of TNs.

Fig. 127, Fig. 128, Fig. 129 and Fig. 130 show the overlays of joint-fit results for the oscillation parameters from MaCh3 and PTheta with different options of implementation described above.

In Fig. 127, PTheta with the same beam sample binning pulls the results towards MaCh3 results while the results with the same kinematic cut implementation cannot explain the fitter results differences very well. The **kinematic cut + E_{rec} binning** option shows very similar results compared to **E_{rec} binning** option.

In Fig. 128, similar to Fig. 127, **E_{rec} binning** option pulls PTheta results towards MaCh3 and can thus better explain the differences. The **kinematic cut + E_{rec} binning** option shows very similar results compared to **E_{rec} binning** option for inverted ordering.

In Fig. 129, unlike in Fig. 127 and Fig. 128, **kinematic cut** option pulls PTheta results towards MaCh3 the most for both normal and inverted orderings. The **kinematic cut + E_{rec} binning** option shows very similar effect as **kinematic cut**. Although the three listed options cannot perfectly explain the shift in Δm^2 for both orderings, they still show

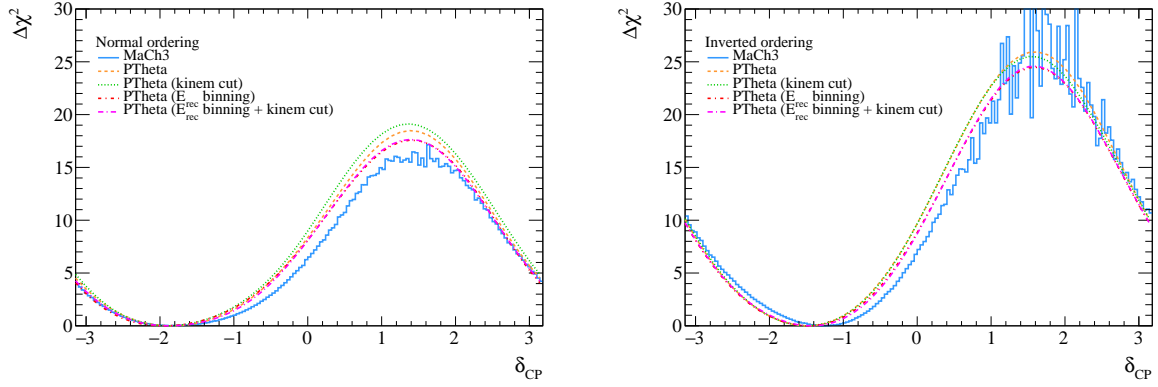


Figure 127: Overlay of δ_{CP} from joint-fit of MaCh3 and PTheta testing different framework implementation options. The solid blue lines are results from MaCh3 joint-fit. The orange dashed lines are results from PTheta joint-fit. The green dashed lines are results from PTheta joint-fit using the same kinematic cut implementation as MaCh3 (i.e. applying the kinematic cut before applying any corrections to the kinematic). The red dashed lines are results from PTheta joint-fit using the same T2K beam sample binnings as MaCh3 (i.e. using E_{rec} rather than p_{rec} for e-like samples). The magenta dashed lines are results from PTheta joint-fit using both the same kinematic cut and same beam sample binnings as MaCh3. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

783 non-negligible influence on the contours.

784 **Fig. 130** confirms the same reactor constraint used in all results shown in this section.

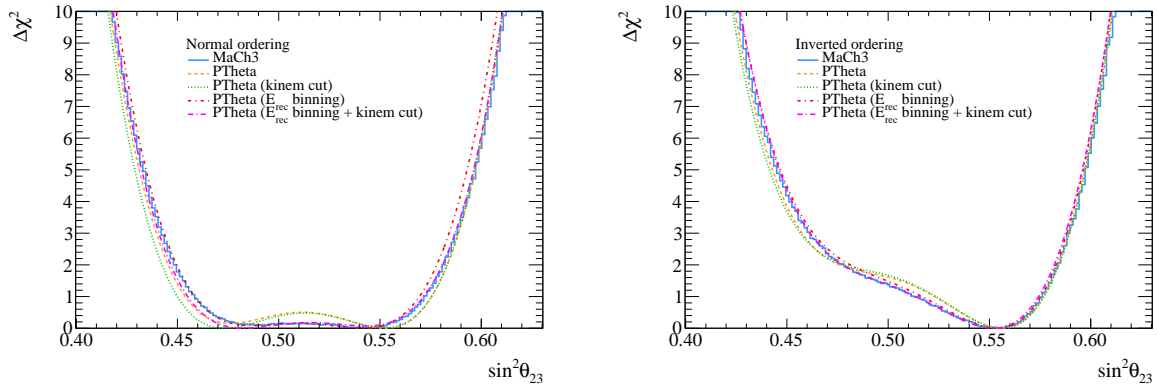


Figure 128: Overlay of $\sin^2 \theta_{23}$ from joint-fit of MaCh3 and PTheta testing different framework implementation options. The solid blue lines are results from MaCh3 joint-fit. The orange dashed lines are results from PTheta joint-fit. The green dashed lines are results from PTheta joint-fit using the same kinematic cut implementation as MaCh3 (i.e. applying the kinematic cut before applying any corrections to the kinematic). The red dashed lines are results from PTheta joint-fit using the same T2K beam sample binnings as MaCh3 (i.e. using E_{rec} rather than p_{rec} for e-like samples). The magenta dashed lines are results from PTheta joint-fit using both the same kinematic cut and same beam sample binnings as MaCh3. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

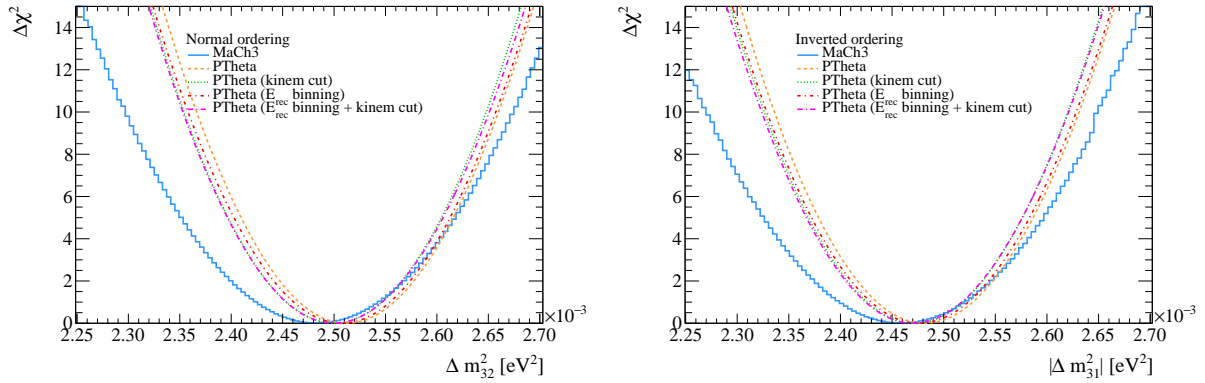


Figure 129: Overlay of Δm_{32}^2 and $|\Delta m_{31}^2|$ from joint-fit of MaCh3 and PTheta testing different framework implementation options. The solid blue lines are results from MaCh3 joint-fit. The orange dashed lines are results from PTheta joint-fit. The green dashed lines are results from PTheta joint-fit using the same kinematic cut implementation as MaCh3 (i.e. applying the kinematic cut before applying any corrections to the kinematic). The red dashed lines are results from PTheta joint-fit using the same T2K beam sample binnings as MaCh3 (i.e. using E_{rec} rather than p_{rec} for e-like samples). The magenta dashed lines are results from PTheta joint-fit using both the same kinematic cut and same beam sample binnings as MaCh3. Left plot shows the normal ordering $\Delta\chi^2$ distributions of Δm_{32}^2 and right plot shows the inverted ordering $\Delta\chi^2$ distributions of $|\Delta m_{31}^2|$. Same reactor constraint is applied.

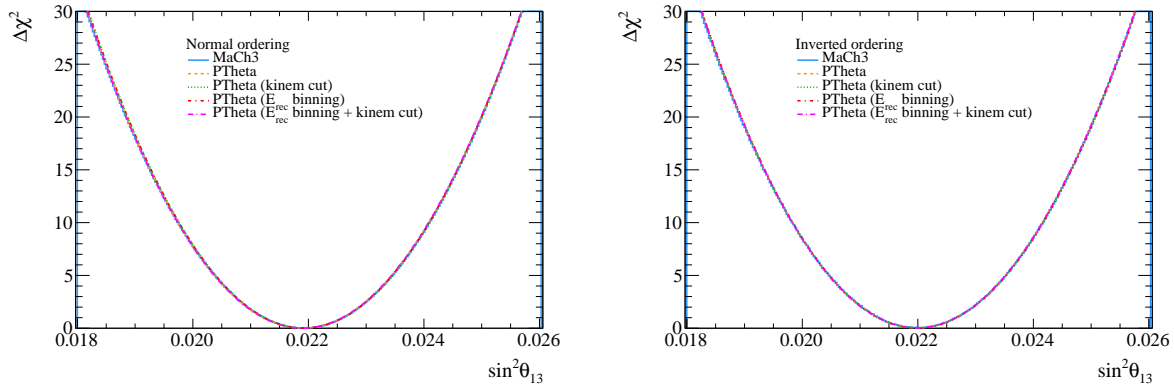


Figure 130: Overlay of $\sin^2 \theta_{13}$ from joint-fit of MaCh3 and PTheta testing different framework implementation options. The solid blue lines are results from MaCh3 joint-fit. The orange dashed lines are results from PTheta joint-fit. The green dashed lines are results from PTheta joint-fit using the same kinematic cut implementation as MaCh3 (i.e. applying the kinematic cut before applying any corrections to the kinematic). The red dashed lines are results from PTheta joint-fit using the same T2K beam sample binnings as MaCh3 (i.e. using E_{rec} rather than p_{rec} for e-like samples). The magenta dashed lines are results from PTheta joint-fit using both the same kinematic cut and same beam sample binnings as MaCh3. Left plot shows the normal ordering $\Delta\chi^2$ distributions and right plot shows the inverted ordering results. Same reactor constraint is applied.

5.4 Comparison of the systematic posterior distributions

In this section, we compare the post-fit systematic constraints between MaCh3 and P-theta. We should note that the comparisons shown in this section do not show the effect of correlations between systematic parameters. In addition, some parameters have intrinsic differences at the implementation level. Therefore, even if there are differences in the individual parameter's distribution, it does not necessarily mean that we have a problem in either of the fitters.

To perform this comparison, we construct the MCMC posterior distributions for each systematic using the 1 million steps randomly sampled from the post-burn-in chains for MaCh3 and 20 million steps for P-theta. The summary of the mean and variance of each systematic parameter's posterior distribution is shown in Fig. 131.

For the T2K and SK flux systematic parameters, MaCh3 and P-theta distributions agree within 1 post-fit σ ranges. We can see a trend that MaCh3 has larger values of flux systematic parameters in the low-energy region (< 1 GeV). A similar trend was seen in the comparison of near detector constraints of these parameters between MaCh3 and BANFF (Appendix 3 of TN397 [7]). Since these parameters are well constrained by the near detector fit, the differences in the post-fit distributions could also be affected by the difference of near detector fit. This can be confirmed once the effect of the near detector fit is studied.

For the T2K detector systematic parameters, we only compare the ones relevant to the μ -like samples because the two fitters use the different binning for the e -like samples (MaCh3 uses E_{rec} binning while P-theta uses momentum binning) and we cannot simply compare the corresponding systematics. These T2K detector systematic parameters show very consistent results between the two fitters, except for the $\nu_e/\bar{\nu}_e$ background components in the μ -like samples (`SK_nu_dis_nue_nueb_CC`, `SK_nubar_nue_nueb_CC`) where MaCh3 has stronger constraints compared to P-theta. We should note that, as described in Section 6.6 of TN422 [5], P-theta includes the secondary interaction (SI) systematics not as the independent dials but as the combined effect in the detector systematics. On the other hand, MaCh3 implements these systematics as independent dials, and therefore the detector systematics do not include the effect of SI in MaCh3. These different treatments of SI systematics could cause the differences seen in the T2K detector systematic parameters. Among the other T2K and SK detector systematics, the SI systematics (`SI_*`) and the energy-scale systematics (`SK_p_scale`, `SK_atm_p_scale`, `UpDown_Asymmetry_Energy_Scale`) also show some differences. These systematic parameters are implemented in different ways and are known to behave differently due to the different implementation philosophies of both fitters.

819 The cross-section systematics in the high-energy model and the new cross-section system-
820 atics introduced in the joint analysis have very consistent distributions between MaCh3 and
821 P-theta. The low-energy cross-section systematics show some differences, but they mostly
822 agree within 1 post-fit σ ranges. The posterior distributions of selected cross-section system-
823 atics are shown in Fig. 132. Similarly to the T2K flux systematics, these distributions show
824 a similar trend with the near detector constraints of these parameters between MaCh3 and
825 BANFF (Appendix 3 of TN397 [7]). Therefore, these differences could be further understood
826 once the effect of the near detector fit is studied.

827 In conclusion, we have observed some differences in the systematic posterior distribu-
828 tions but these differences are thought to be understandable level considering the different
829 implementations of each systematic between MaCh3 and P-theta.

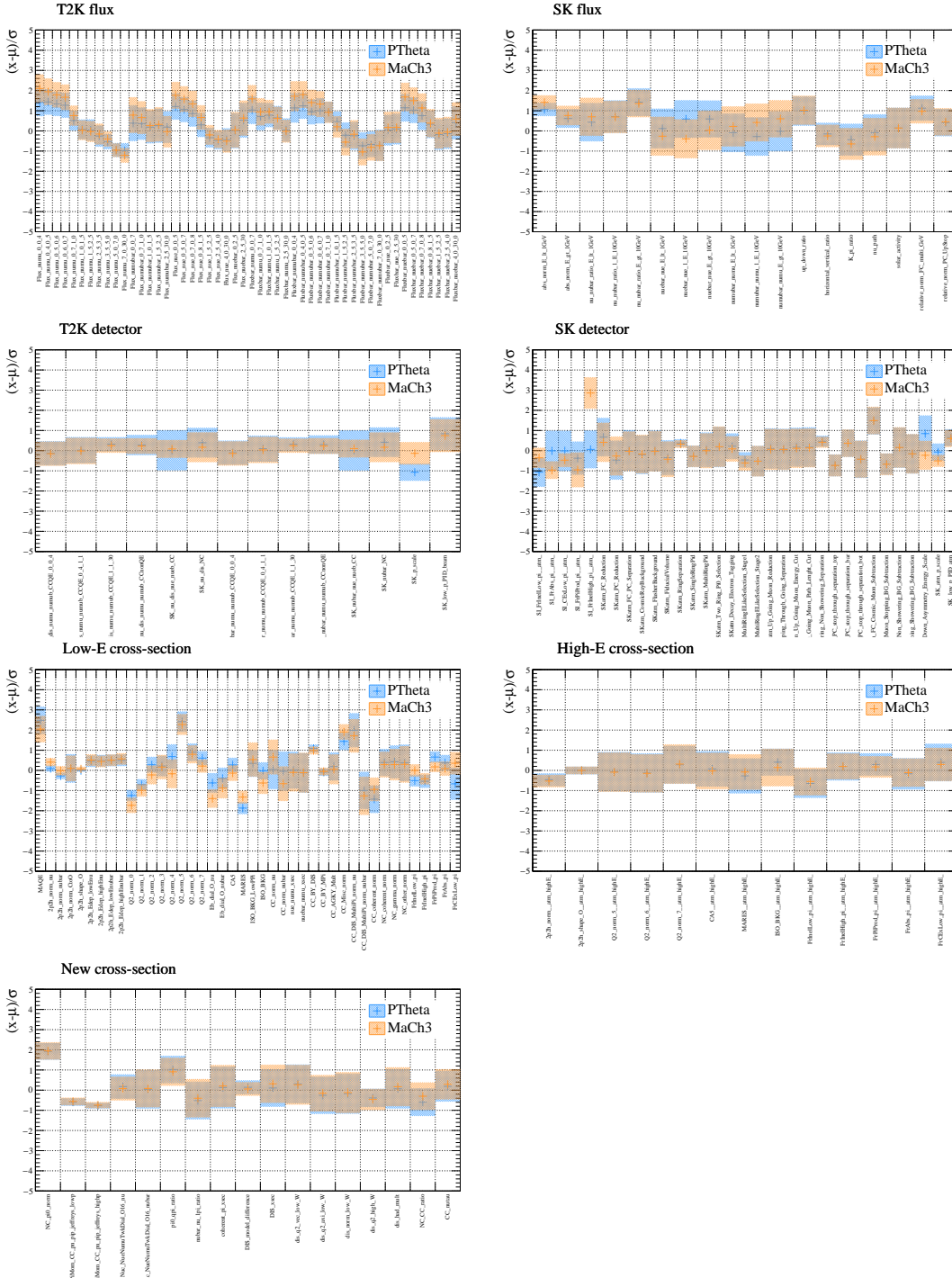


Figure 131: Comparison of the post-fit systematic constraints between MaCh3 and P-theta computed from the MCMC posterior distributions. The center and error of each parameter error band show the arithmetic mean and variance of posterior distributions. The parameter values are normalized using their prior mean and variance.

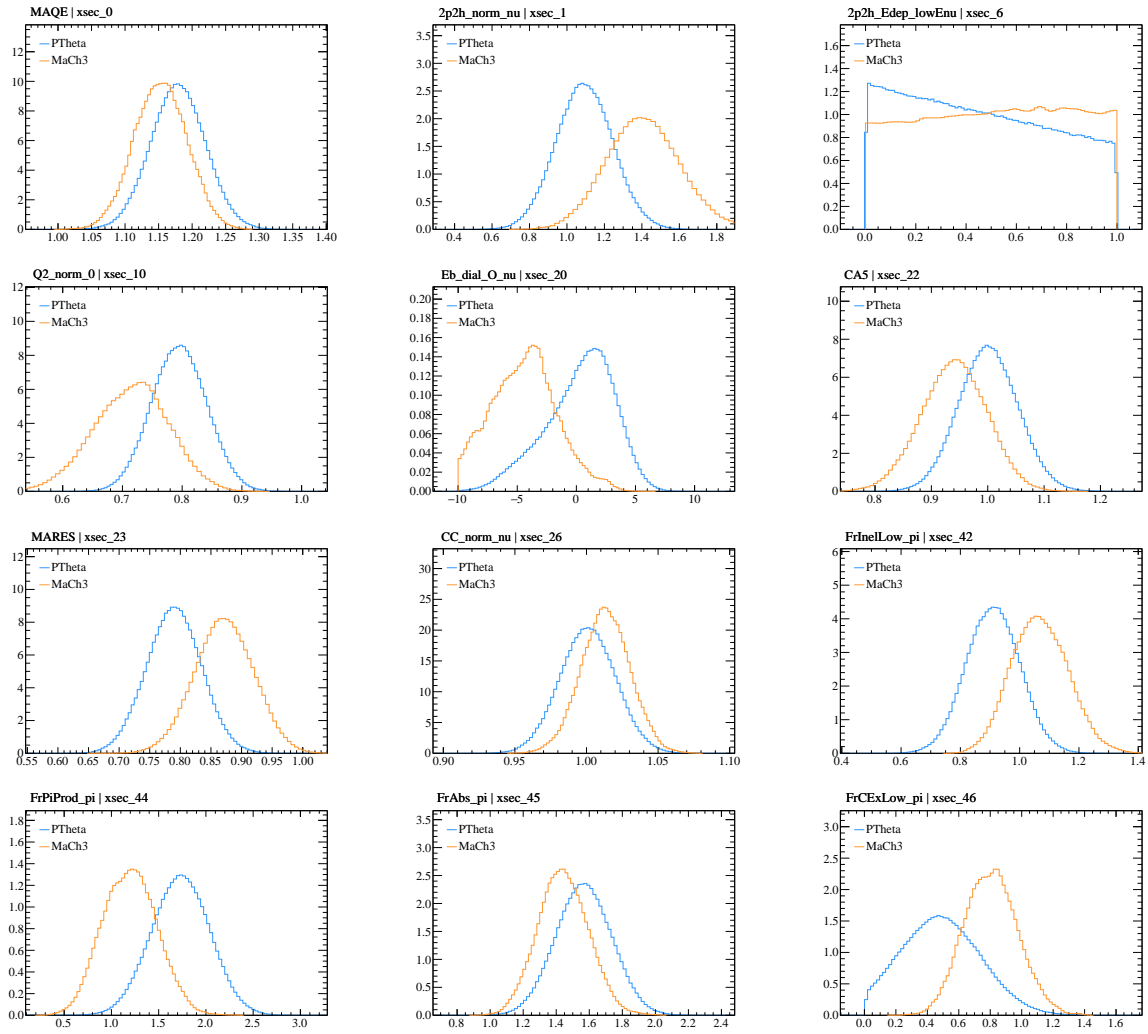


Figure 132: Comparison of the posterior distributions for selected cross-section systematic parameters between MaCh3 and P-theta. The parameter values are **not** normalized with their prior mean and variance.

6 Summary

This technical note presented the joint analysis of SK atmospheric neutrinos and T2K beam neutrinos with the MaCh3 framework, with model updates in TN456 [6]. Analyses of Asimov data (set A) and the real data were presented, simultaneously fitting the near detector samples and far detector samples (5 T2K beam samples and 18 SK atmospheric samples) using Markov chain Monte Carlo to sample the posterior probability in systematic phase space.

For both the Asimov A and real data analyses, results were presented with and without the reactor constraint on $\sin^2\theta_{13}$ applied. All results were smeared with Δm_{32}^2 by $3.6 \times 10^{-5} \text{ eV}^2$ from the fake-data studies. For each fit, one dimensional and two dimensional oscillation parameter contours were produced, and Bayes factors of octant hypotheses, mass ordering and CP violation were calculated. For the data fit, the posterior predictive spectra (both with and without reactor constraint) were compared to the data spectra and the corresponding posterior predictive p-values (total p-value and by-sample p-values) were extracted. The data fit results were also compared to the MaCh3 analysis in OA2020 and to P-Theta results in the joint fit and OA2020.

Mass ordering For the Asimov A fit, the Bayes factor of normal ordering over inverted ordering $B(\text{NO}/\text{IO})$ is 7.26(3.98) with(out) the reactor constraint applied. For the data fit, the Bayes factor of normal ordering over inverted ordering $B(\text{NO}/\text{IO})$ is 7.33(3.76) with(out) the reactor constraint applied. Compared to T2K OA2020 MaCh3 results, the Bayes factor $B(\text{NO}/\text{IO})$ is 4.2(1.8) with(out) the reactor constraint applied. Although none of the values reach a level of significance where conclusions could be drawn to exclude either possibilities of the mass ordering, we can still see an enhancement of substantial evidence of normal ordering over inverted ordering according to Jeffrey's scale [14].

CP violation For the Asimov A fit, the Bayes factor $B[(\sin\delta_{\text{CP}} < 0)/(\sin\delta_{\text{CP}} \geq 0)]$ is 24.5(6.3) with(out) the reactor constraint applied. For the data fit, the Bayes factor $B[(\sin\delta_{\text{CP}} < 0)/(\sin\delta_{\text{CP}} \geq 0)]$ is 50.0(10.5) with(out) the reactor constraint applied. From the δ_{CP} contours of data fit with reactor constraint, $\delta_{\text{CP}} = 0$ can be easily excluded between 2σ and 3σ credible intervals. Hence, from both the contours and Bayes factor values, joint-fit results show support for CP violation neutrinos.

860 $\sin^2 \theta_{23}$ **octant** For the Asimov A fit, the Bayes factor of upper octant over lower octant
861 $B(\text{UO}/\text{LO})$ is 2.72(1.67) with(out) the reactor constraint applied. For the data fit, the
862 Bayes factor of upper octant over lower octant $B(\text{UO}/\text{LO})$ is 1.78(0.82) with(out) the reactor
863 constraint applied. Compared to OA2020 MaCh3 results [1], joint-fit data fit results show a
864 weaker preference of upper octant.

References

- 865
- 866 [1] E. Atkin et al. “Measuring PMNS parameters in a joint ND280-SK analysis using
867 MCMC”. In: *T2K Technical Note* 393 (2020).
- 868 [2] D Barrow et al. “SK atmospheric T2K beam joint fit technical note, MaCh3 details”.
869 In: *T2K Technical Report* 426 (2022).
- 870 [3] Roger Wilking Michael Wilking Cris Vilela. “Event Selection and Detector Systematic
871 Uncertainties for the SK+T2K Joint Fit”. In: *T2K Technical Report* 428 (2021).
- 872 [4] D. Barrow et al. “Oscillation probability calculation for the T2K+SK atmospheric
873 joint fit”. In: *T2K Technical Note* 425 (2021).
- 874 [5] C. Bronner et al. “Flux and interaction models for the T2K-SK atmospheric joint fit”.
875 In: *T2K Technical Note* 422 (2021).
- 876 [6] L. Berns et al. “T2K+SK atmospheric joint analysis: model updates for the data fit”.
877 In: *T2K Technical Note* 456 (2023).
- 878 [7] L. Berns et al. “P-theta 2020 Run 1-10 Joint-fit Analysis”. In: *T2K Technical Note*
879 397 (2020).
- 880 [8] A. Eguchi et al. “Fitter validations for joint T2K+SK atmospheric Asimov fits”. In:
881 *T2K Technical Note* 471 (2023).
- 882 [9] L. Munteanu et al. “Constraining the Flux and Cross Section Models with Data from
883 ND280 using FGD1 and FGD2 for the 2020 Oscillation Analysis”. In: *T2K Technical*
884 *Note* 395 (2020).
- 885 [10] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (2018), p. 030001.
- 886 [11] A. Eguchi et al. “Fake data studies for the T2K+SK atmospheric joint analysis”. In:
887 *T2K Technical Note* 457 (2023).
- 888 [12] L. Berns et al. “P-theta SK+T2K joint analysis”. In: *T2K Technical Note* 459 (2023).
- 889 [13] M Jiang et al. “Atmospheric neutrino oscillation analysis with improved event re-
890 construction in Super-Kamiokande IV”. In: *Progress of Theoretical and Experimental*
891 *Physics* 2019.5 (May 2019). ISSN: 2050-3911. DOI: [10.1093/ptep/ptz015](https://doi.org/10.1093/ptep/ptz015). URL: <http://dx.doi.org/10.1093/ptep/ptz015>.
- 892
- 893 [14] H. Jeffreys. “The theory of probability”. In: *UOP Oxford* (1998).