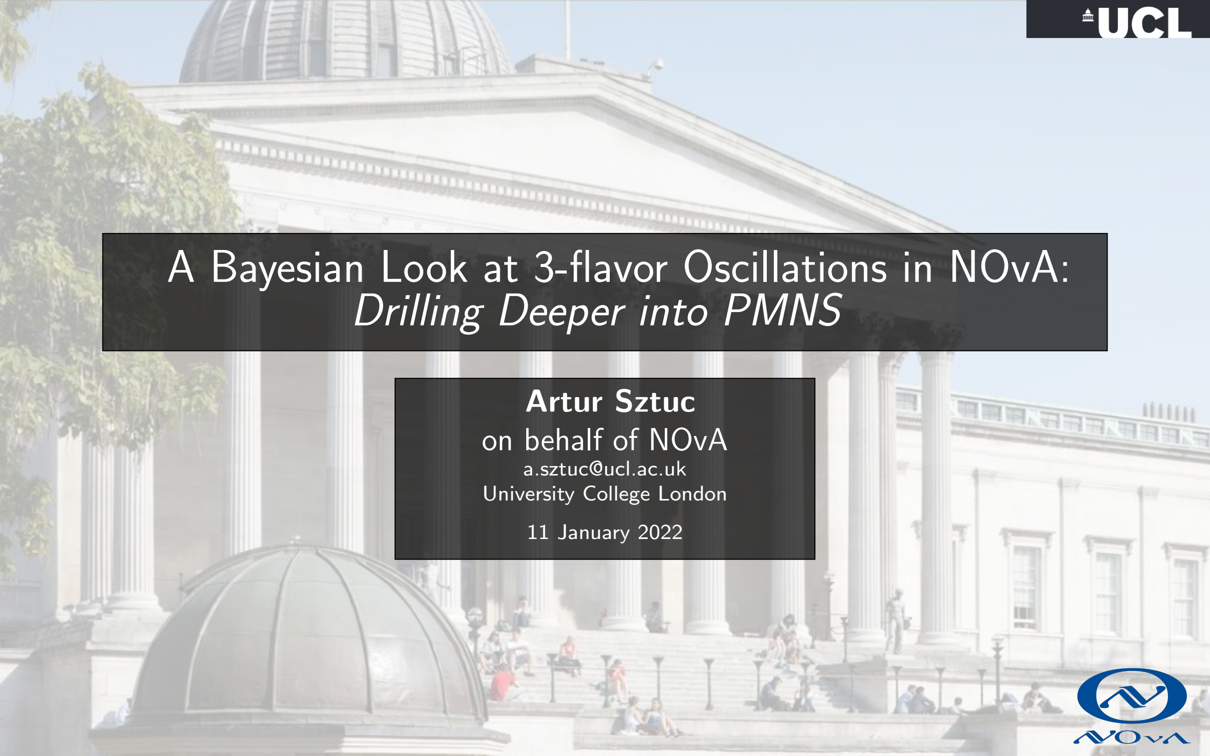


A Bayesian Look at 3-flavor Oscillations in NOvA: *Drilling Deeper into PMNS*

Artur Sztuc
on behalf of NOvA
a.sztuc@ucl.ac.uk
University College London

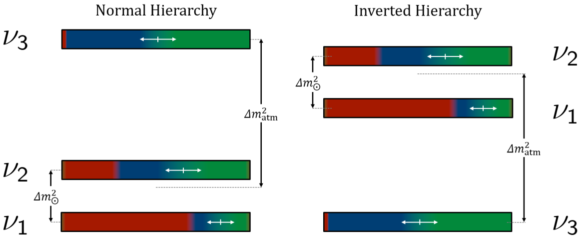
11 January 2022

Contents

1. Neutrino Oscillations: what and why?
2. Measuring Neutrino Oscillations with NOvA
3. Neutrino Events, Data, Predictions
4. Bayesian Inference into the PMNS model
5. (No) Reactor constraint

Neutrino Oscillations: what and why?

Neutrino oscillation physics



- Flavour eigenstates; ν_e, ν_μ and ν_τ (interact)
- Mass eigenstates; ν_1, ν_2 and ν_3 (propagate)

$$\begin{pmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{pmatrix} = \underbrace{\begin{pmatrix}1 & 0 & 0\\ 0 & c_{23} & s_{23}\\ 0 & -s_{23} & c_{23}\end{pmatrix}}_{\text{atmospheric, beam}} \underbrace{\begin{pmatrix}c_{13} & 0 & s_{13}e^{-i\delta_{\rm CP}}\\ 0 & 1 & 0\\ -s_{13}e^{i\delta_{\rm CP}} & 0 & c_{13}\end{pmatrix}}_{\text{reactor, beam}} \underbrace{\begin{pmatrix}c_{12} & s_{12} & 0\\ -s_{12} & c_{12} & 0\\ 0 & 0 & 1\end{pmatrix}}_{\text{solar, reactor}} \begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}$$

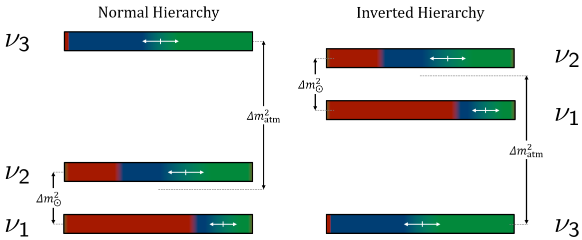
$$s_{ij} = \sin\theta_{ij}$$
$$c_{ij} = \cos\theta_{ij}$$

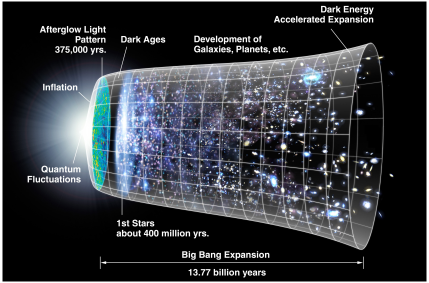
Super-K, IceCube, Opera, **NOvA**, T2K

Double Chooz, Daya Bay, RENO, **NOvA**, T2K

Super-K, SNO, KamLAND

Neutrino oscillation physics





- δ_{CP}: Charge-Parity violation in neutrino sector. Potential contribution to matter-antimatter asymmetry in the universe.
- Mass Ordering: Symmetries in neutrino physics, is ν_1 the lightest and ν_3 the heaviest? Has consequences for double-beta decay search.
- θ_{23}: Larger or smaller than 45? Important for $\nu_\tau - \nu_\mu$ symmetries.

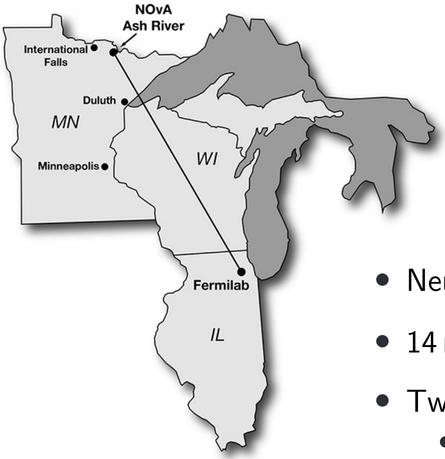
Neutrino oscillation physics

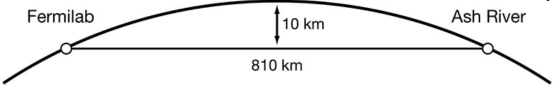
- How can we tell if PMNS is correct?
- First, see if our data fits well with the PMNS model.
- Second, cross-check with experiments at different baselines.
- Third, cross-check with different neutrino oscillations. Reactors measure $P(\bar{\nu}_e \rightarrow \bar{\nu}_e)$ and accelerators $P(\nu_\mu \rightarrow \nu_e)$ & $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$.
    - Both sensitive to θ_{13}.

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{\substack{\text{atmospheric, beam} \\ \text{Super-K, IceCube,} \\ \text{Opera, \textbf{NOvA}, T2K}}} \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{pmatrix}}_{\substack{\text{reactor, beam} \\ \text{Double Chooz, Daya Bay,} \\ \text{RENO, \textbf{NOvA}, T2K}}} \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\substack{\text{solar, reactor} \\ \text{Super-K, SNO,} \\ \text{KamLAND}}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

$$s_{ij} = \sin\theta_{ij}$$
$$c_{ij} = \cos\theta_{ij}$$

NOvA Experiment

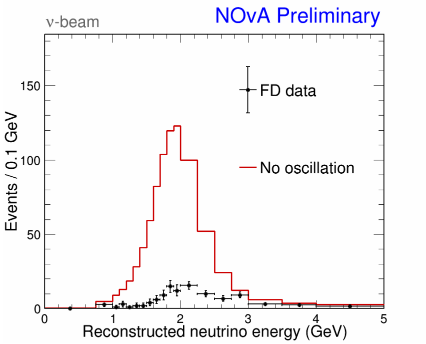




- Neutrino beam from Fermilab's NuMI Beamline.
- 14 mrad off-axis beam narrowly peaked at ∼2 GeV.
- Two functionally identical detectors:
  - Near Detector (ND), 0.3 kton, 1 km baseline.
  - Far detector (FD), 14 kton, 810 km baseline.

Neutrino oscillations with accelerators

ν_μ signal



Location of the dip: $|\Delta m^2_{32}|$

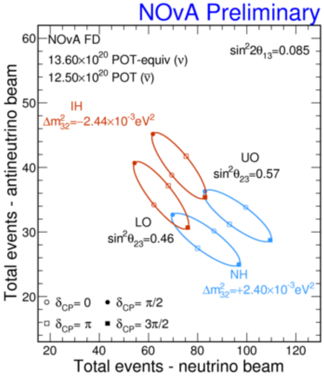
(does not depend on the sign)

Depth of the dip: $\sin^2(2\theta_{23})$

Difficult to separate $\theta_{23} > 45$ and $\theta_{23} < 45$

Is $\nu_\mu = \nu_\tau$ in ν_3 mass state?

ν_e signal



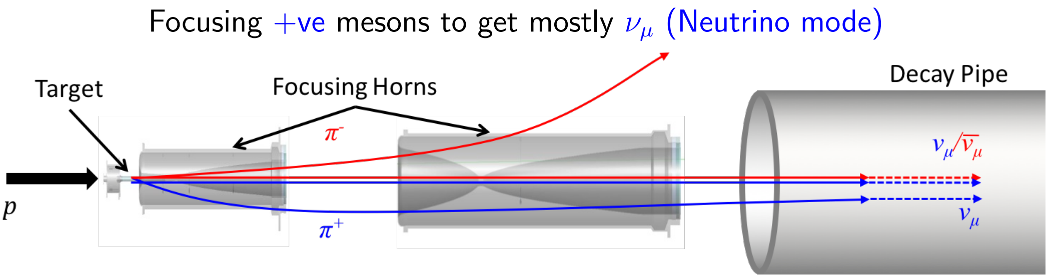
Combination of ν_e and $\bar{\nu}_e$ excess;

$\sin^2(\theta_{23})$, $\sin^2(\theta_{13})$, δ_{CP}

Good dependence on the sign of Δm^2_{32}

Channel for CP violation detection

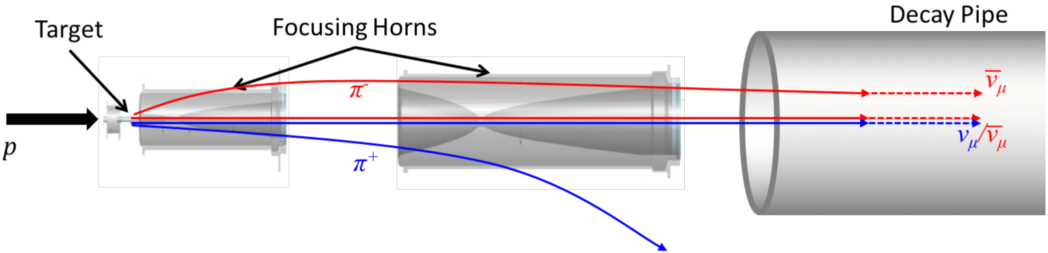
NuMI Neutrino beamline



- Beam of 120 GeV protons incident on carbon target.
- Focusing +ve or -ve mesons to obtain mostly ν_μ or $\bar{\nu}_\mu$.
  - Achieved by reversing the polarity of the magnetic horns.
- Neutrinos appear from the decaying mesons. 675 m decay pipe.

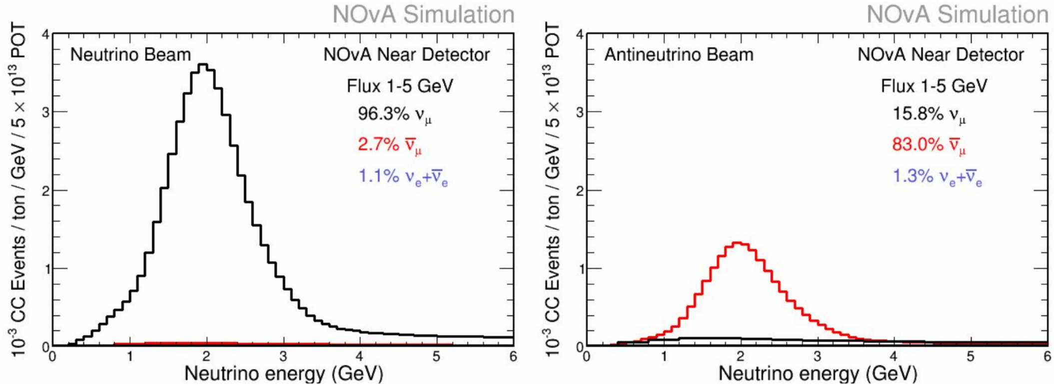
NuMI Neutrino beamline

Focusing -ve mesons to get mostly $\bar{\nu}_\mu$ (Antineutrino mode)



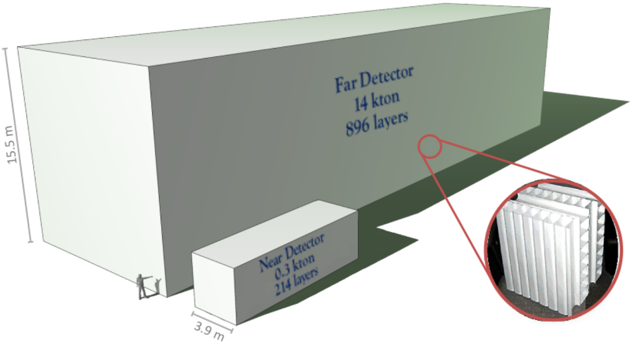
- Beam of 120 GeV protons incident on carbon target.
- Focusing +ve or -ve mesons to obtain mostly ν_μ or $\bar{\nu}_\mu$.
  - Achieved by reversing the polarity of the magnetic horns.
- Neutrinos appear from the decaying mesons. 675 m decay pipe.

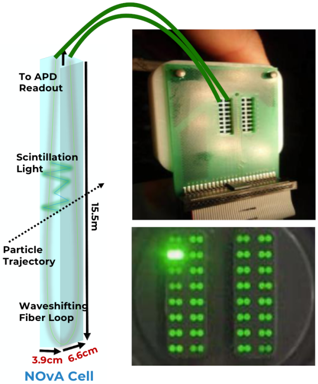
NuMI Neutrino beamline



- Very low-contamination with in both neutrino and anti-neutrino mode.
- Collected 37×10^{20} protons-on-target. Thank you Fermilab!
- Recent power record: 893 kW!

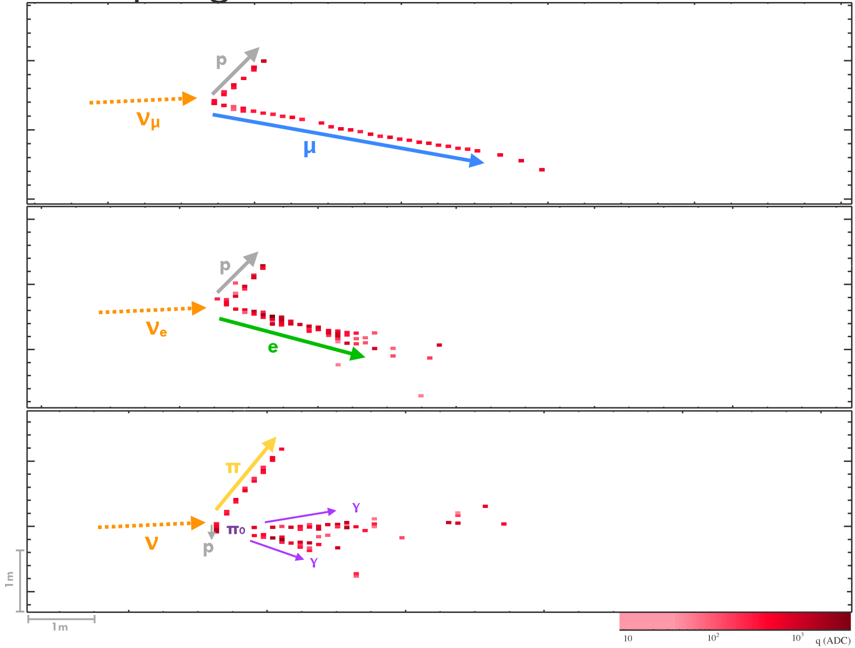
NOvA Detectors





- Extruded cells filled with liquid scintillator, with 62% active volume.
- Wavelength-shifting fiber collects and transports light to Avalanche photodiode.
  - Each APD sees 32 NOvA cells.
- Cells with alternating horizontal & vertical planes for 3D reconstruction.
- Optimized for electron showers.

Event topologies



ν_μ CC

ν_e CC

NC

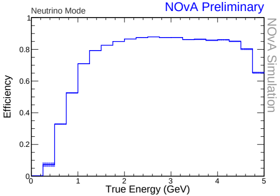
Event identification

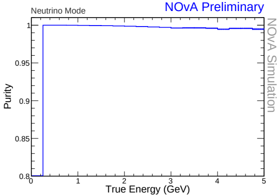
Pre-selections:

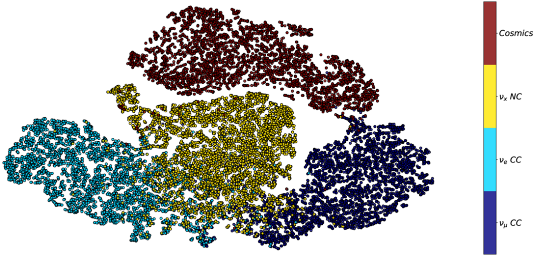
- Contained inside the detector.
- Inside of the beam spill-window.
- Cosmic particles rejection via BDT.

Event Identification:

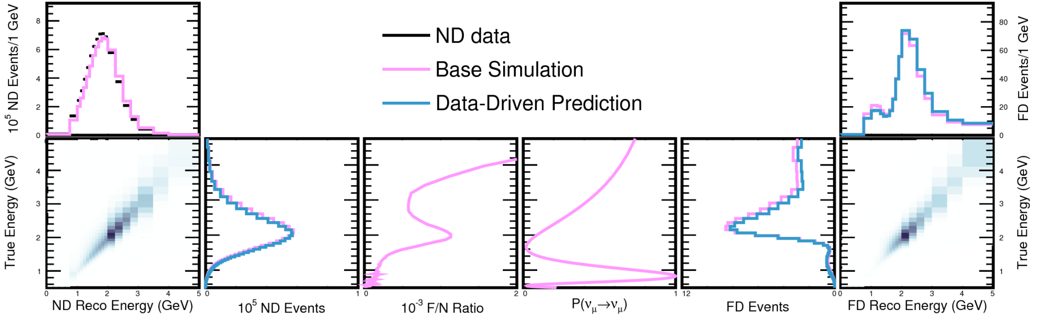
- Modern CNN techniques used to identify neutrino flavour.
- Learns features of different event topologies.
- Data-driven validations based on ND and FD control samples.
- Results in high purity samples.



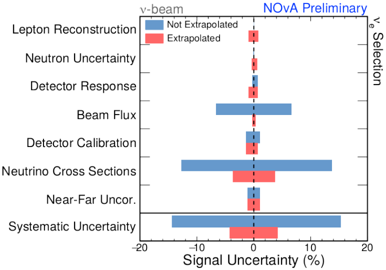




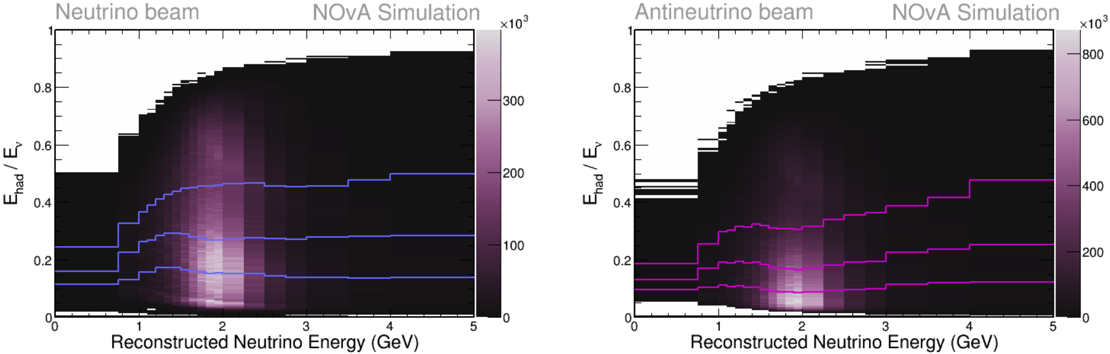
ND→FD Extrapolation



- Take advantage of detector similarity to extrapolate ND predictions to FD.
- Many systematic effects e.g. cross-sections, flux and efficiency are shared.
- Helps dealing with the “unknown unknowns”.
- Extrapolate different kinematic samples separately to deal with Near/Far acceptance differences.

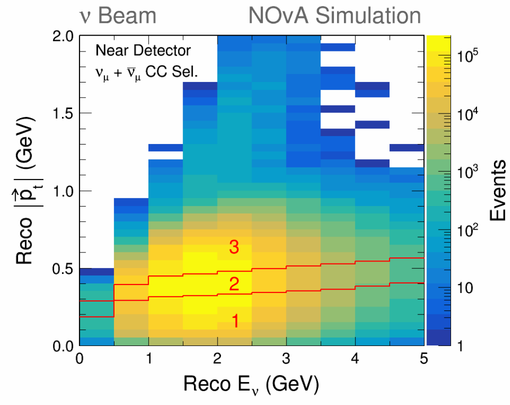


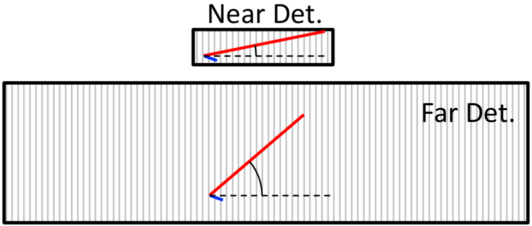
ND→FD Extrapolation: E_{had}



- The energy resolution varies between the detectors.
- Extrapolation split in four E_{had}/E_{ν} quartiles.
- Matches the Hadronic energy resolution between ND and FD.

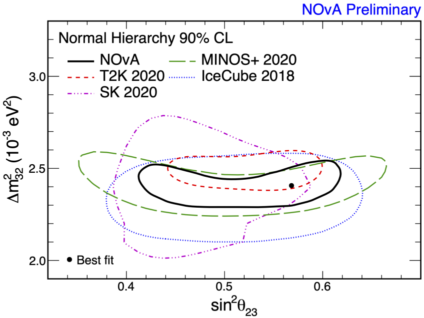
ND→FD Extrapolation: Lepton $|p_T|$





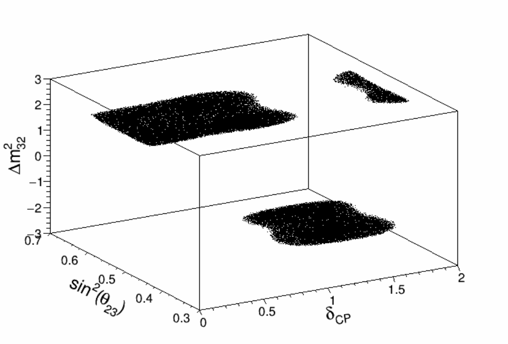
- Different lepton angle distributions due to the difference in detectors' size.
- Extrapolation split in three ranges of lepton transverse momenta.
- Done separately for each E_{had} quartile.
- Matches the detector acceptances between ND and FD.

Fitting/Sampling techniques



- NOvA fits 10 data samples: $4\nu_\mu$, $4\bar{\nu}_\mu$, $1\nu_e$ and $1\bar{\nu}_e$.
- All the previous NOvA results were Frequentist with use of profiling.
- New Bayesian frameworks implemented in NOvA.
- New studies now easier: Jarlskog-Invariant, NOvA-only θ_{13}, Bayes factors and possibly more!
- Other experiments often provide Marginalized and/or Bayesian results.

Markov Chain Monte Carlo for NOvA



Bayes Theorem:

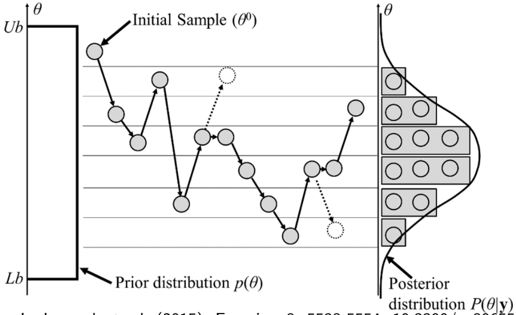
$$\mathcal{P}(\vec{\theta}|D) \approx \mathcal{P}(D|\vec{\theta})\mathcal{P}(\vec{a})$$

$$Posterior \approx Likelihood \times Prior$$

$$Posterior \approx e^{-\frac{\chi^2}{2}} \times Prior$$

- Bayesian results given in terms of posterior probability distributions.
- Need to produce N-dimensional probability distribution for marginalized results.
- MCMC generates samples on N-dimensional.
    - Sample density corresponds to posterior probability density.

Markov Chain Monte Carlo for NOvA



L. Jaewook et. al. (2015). Energies. 8. 5538-5554. 10.3390/en8065538.

Bayes Theorem:

$$\mathcal{P}(\vec{\theta}|D) \approx \mathcal{P}(D|\vec{\theta})\mathcal{P}(\vec{a})$$

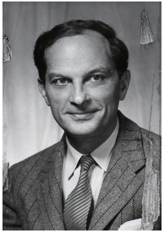
$$Posterior \approx Likelihood \times Prior$$

$$Posterior \approx e^{-\frac{\chi^2}{2}} \times Prior$$

- MCMC generates samples by iteratively deviating parameters from their previous values.
- At each iteration we can either accept, or reject the step.
  - Accept: new step added to the end of the chain.
  - Reject: previous values repeated at the end of the chain.
- Over time, this “chain” ensemble starts resembling posterior probability.

Markov Chain Monte Carlo for NOvA

Arianna Rosenbluth



Stanislaw Ulam

- Two algorithms in NOvA: Metropolis-Hastings and Hamiltonian MCMC.
- Hamiltonian MCMC is based on Stan library (https://mc-stan.org).
  - Stanislaw Ulam invented the methods of Monte-Carlo.
- Metropolis-Hastings was written from scratch in-house.
  - Named Aria after Arianna Rosenbluth, who first implemented the method.
- Importantly, both algorithms produce identical results.

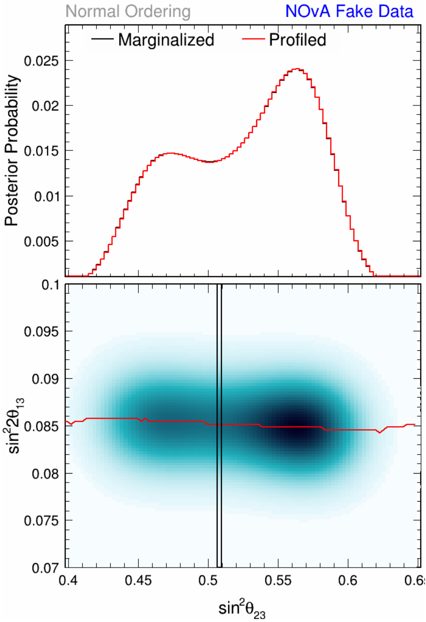
References: Metropolis-Hastings doi:10.1063/1.1699114, Hamiltonian doi:10.1016/0370-2693(87)91197-X

Bayesian vs Frequentist, Marginalization vs Profiling

- MCMC uses marginalization rather than profiling.
    - Not necessarily reserved to Bayesian methods!
    - Profiling: Maximize parameters not shown.
    - Marginalization: Integrate over parameters not shown.

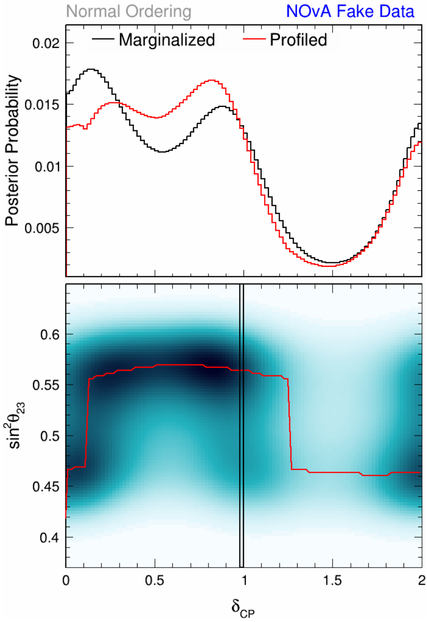
- Example: marginalizing/profiling over $\sin^2 2\theta_{13}$.
    - Line of best fit to profile over $\sin^2 2\theta_{13}$.
    - Box with probabilities to sum over for marginalization.

- Use posterior probability densities, not χ^2.
    - Bayes. Credible Intervals vs Freq. Confidence Levels.



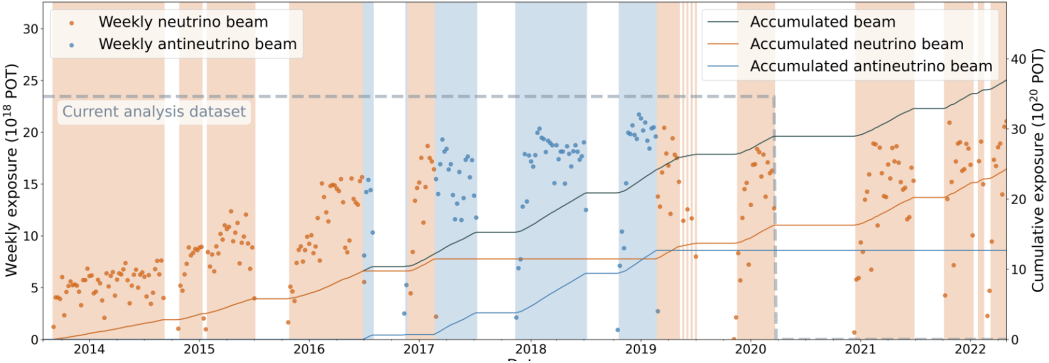
Bayesian vs Frequentist, Marginalization vs Profiling

- MCMC uses marginalization rather than profiling.
    - Not necessarily reserved to Bayesian methods!
    - Profiling: Maximize parameters not shown.
    - Marginalization: Integrate over parameters not shown.
- Example: marginalizing/profiling over $\sin^2\theta_{23}$.
    - Line of best fit to profile over $\sin^2\theta_{23}$.
    - Box with probabilities to sum over for marginalization.
- Use posterior probability densities, not χ^2.
    - Bayes. Credible Intervals vs Freq. Confidence Levels.



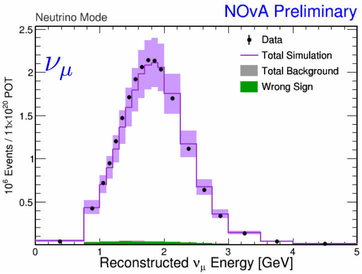
Collected data

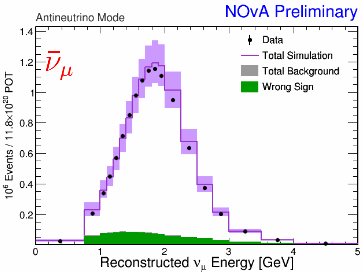
Collected beam data

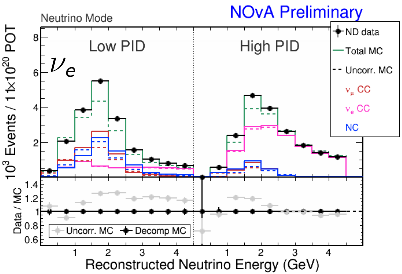


- Collected 37×10^{20} protons-on-target up to date.
- Data up to early 2020 included in the analysis shown here.
  - 13.6×10^{20} in ν-beam mode.
  - 12.5×10^{20} in $\bar{\nu}$-beam mode.

Near detector data

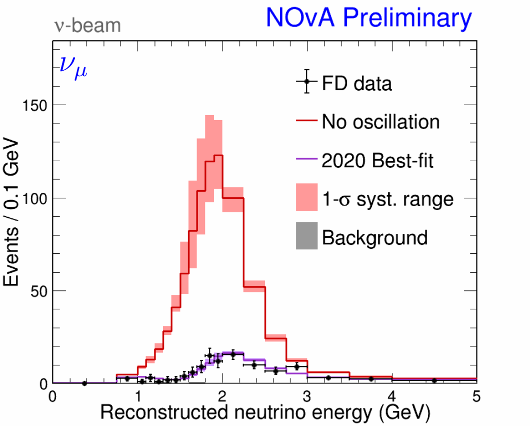


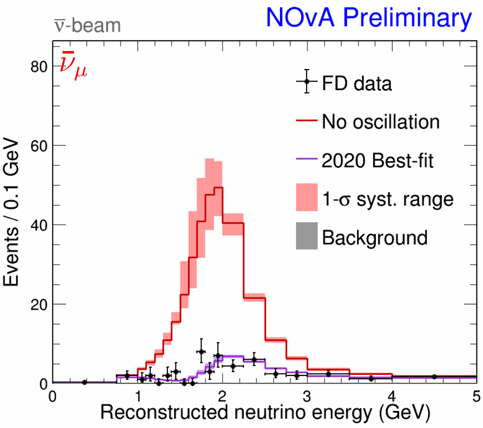




- ν_μ and $\bar{\nu}_\mu$ ND samples are used to correct the FD unoscillated predictions via extrapolation.
- We can then apply the $P(\nu_\mu \to \nu_e)$ curve to the corrected predictions.
- The ν_e samples are used to correct the irreducible ν_e background in the beam at the FD.

Far detector data Muon neutrinos

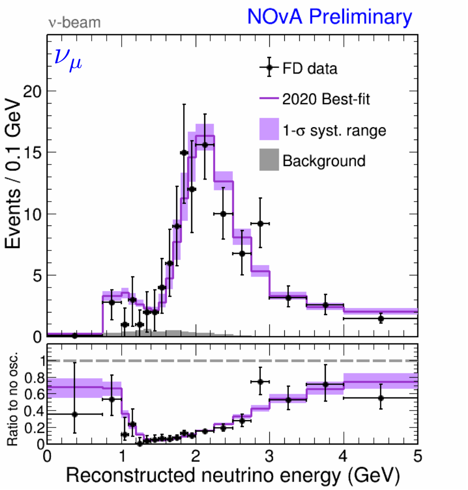


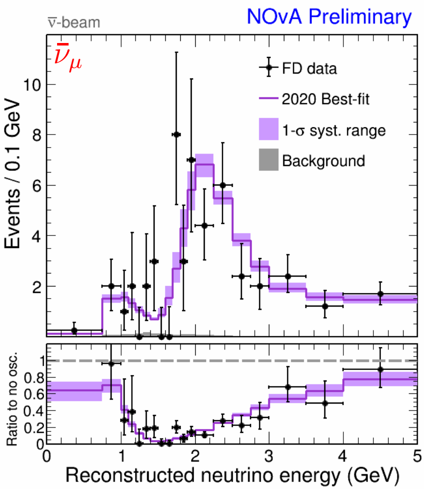


- Observed: 211
- Best Fit Prediction: 222.3
- Background: 8.2

- Observed: 105
- Best Fit Prediction: 105.4
- Background: 2.1

Far detector data Muon neutrinos

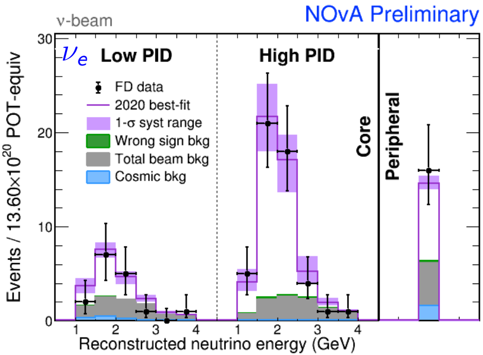


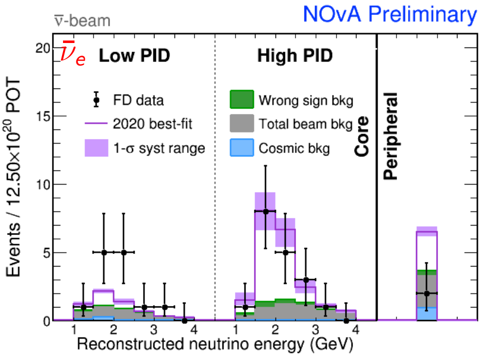


- Observed: 211
- Best Fit Prediction: 222.3
- Background: 8.2

- Observed: 105
- Best Fit Prediction: 105.4
- Background: 2.1

Far detector data Electron neutrinos





- Observed: 82
- Best Fit Prediction: 85.8
- Background: 26.8

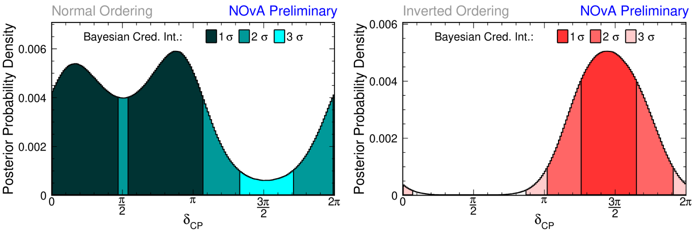
- Observed: 33
- Best Fit Prediction: 33.2
- Background: 14

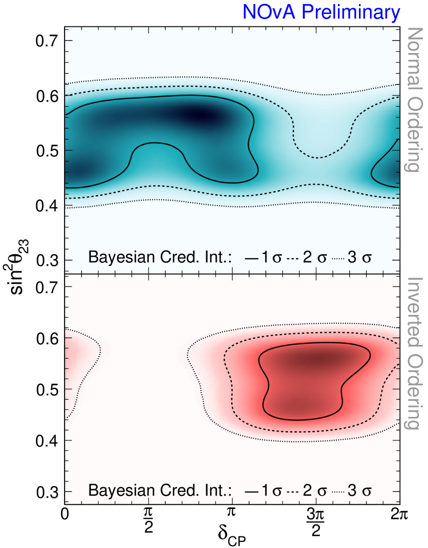
$> 4\,\sigma$ evidence of electron antineutrino appearance

Bayesian Inference into the PMNS model

Appearance parameters' results

🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering

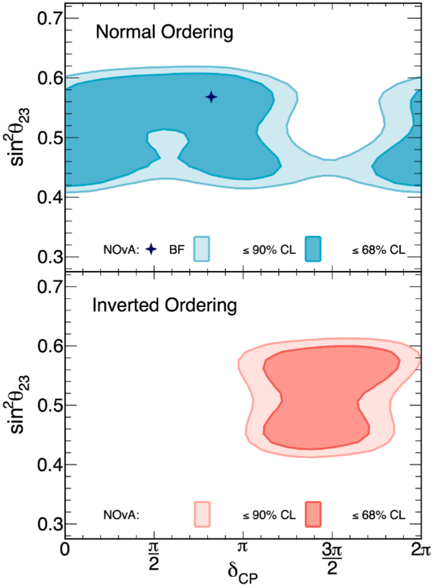




- General conclusions the same as in the 2020 Frequentist analysis.
  - $\delta_{CP} = 1.5\pi$ outside of 2σ credible intervals (NO).
  - $\delta_{CP} = 0.5\pi$ outside of 3σ credible intervals (IO).
  - Prefer Upper Octant of θ_{23}.

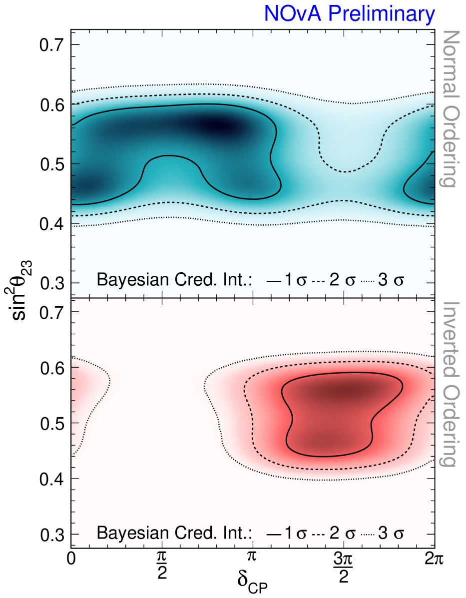
🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering

Frequentist result



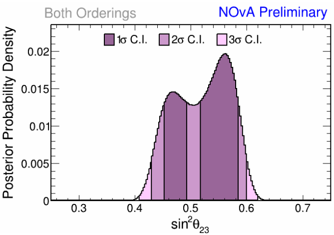
Frequentist results reference: arXiv:2108.08219

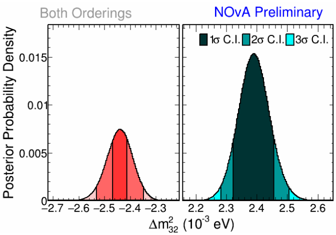
Bayesian result



Disappearance parameters' results

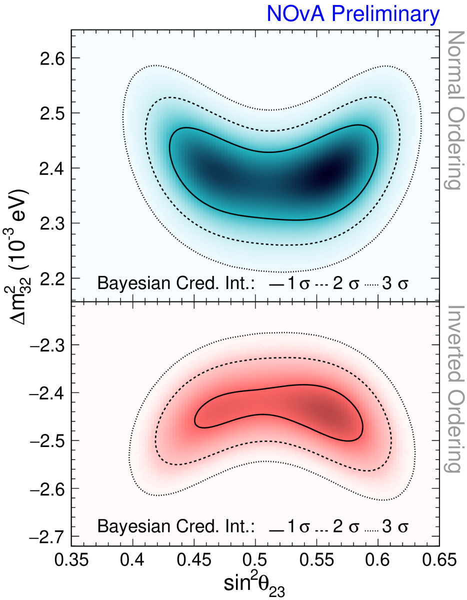
🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering





- Prefer upper octant and normal mass ordering.
- Neither preference is significant, below $\sim 1\,\sigma$.

	N. Ordering	I. Ordering	
U. Octant	41.7%	20.9%	62.6%
L. Octant	25.8%	11.5%	37.4%
	67.5%	32.5%	

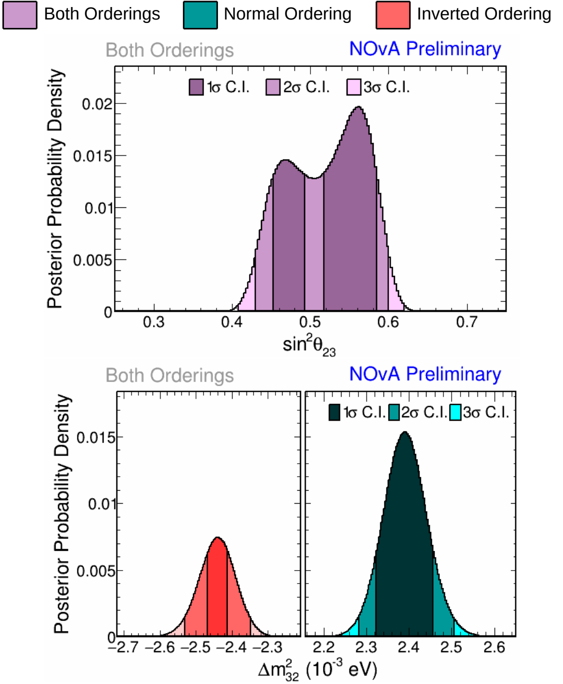


Bayes Factors

	N. Ordering	I. Ordering	
U. Octant	41.7%	20.9%	62.6%
L. Octant	25.8%	11.5%	37.4%
	67.5%	32.5%	

- Bayes Factors: odds ratio, how much more likely one model is than another.
- NO/IO: 2.1, UO/LO: 1.7
- Both can be interpreted as below $1\,\sigma$ or "not worth more than a bare mention" according to Jeffreys and Raftery & Kass scales.

References: Jeffreys ISBN:9780191589676,
Raftery & Kass doi:10.2307/2291091

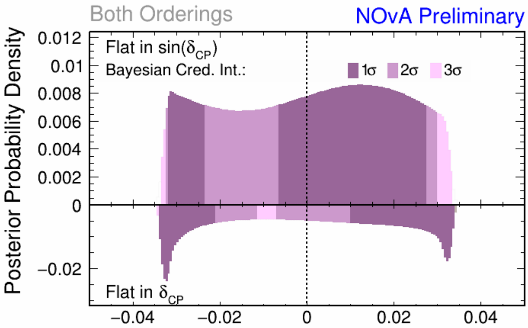


Jarlskog-Invariant

What? Why? Priors?

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering

- $J = c_{12} c_{13}^2 c_{23} s_{12} s_{13} s_{23} s_{CP}$
- Jarlskog-Invariant:
  measure of CP-violation
  independent of parametrization.
- Also used in the Quark sector



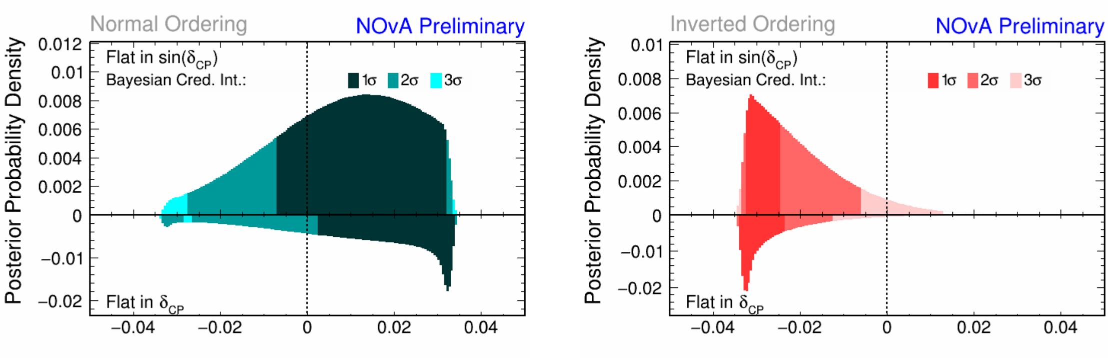
- J=0: CP-Conservation. J≠0: CP-Violation
- A prior flat in sin(δ_{CP}) provides data-only preference (upper half).
- A prior flat in δ_{CP} has a bias away from minimal CPV (lower half).
  - There's some theoretical motivation (Neutrino Mixing Anarchy) for this.

Reference: Neutrino Mixing Anarchy arXiv:1204.1249

Jarlskog-Invariant

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering

Results

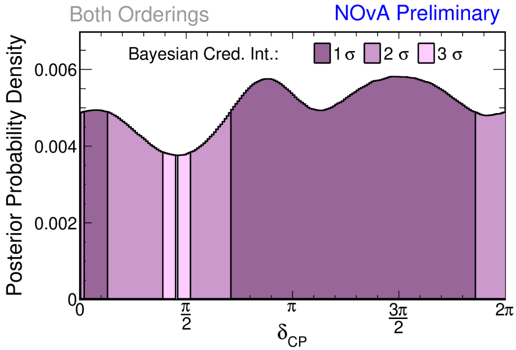


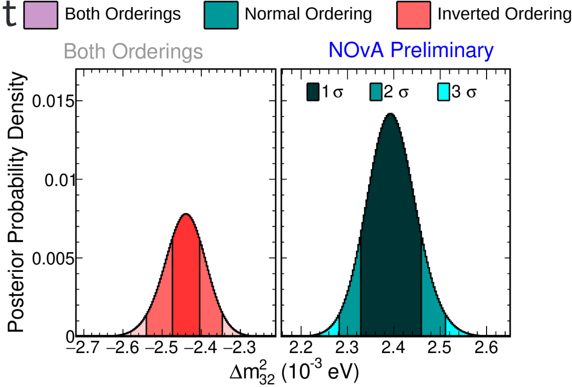
- CP-Conservation (J=0) within 1 σ interval in NO, within 3 σ in IO.
  - Disfavoured more with a prior uniform in δ_{CP}.
- Bayes factor for J$\neq$0 using Savage-Dickey method: 1.5 for **both priors**.
  - Less than 1 σ significance, or "not worth more than a bare mention".
- Slight, but not significant preference for CP-violation.

Reference: Savage-Dickey arXiv:2004.09899

(No) Reactor constraint on θ_{13}

Results without Reactor Constraint

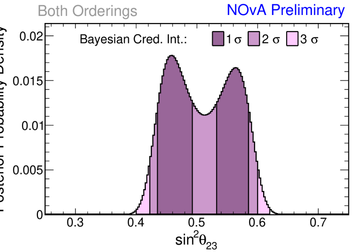


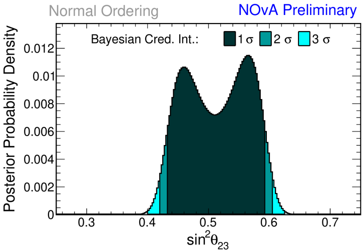


- All previous Frequentist results shown with exteranl θ_{13} constraint.
- But NOvA has sensitivity to θ_{13}! How does it affect our results?
  - Do we agree with the Reactors? Tensions in the PMNS model?
- Allowing unconstrained θ_{13} to give NOvA-only preferences:
  - δ_{CP} preferences don't change much.
  - Prefer normal mass ordering.
  - General conclusions similar to Reactor-constrained θ_{13}.

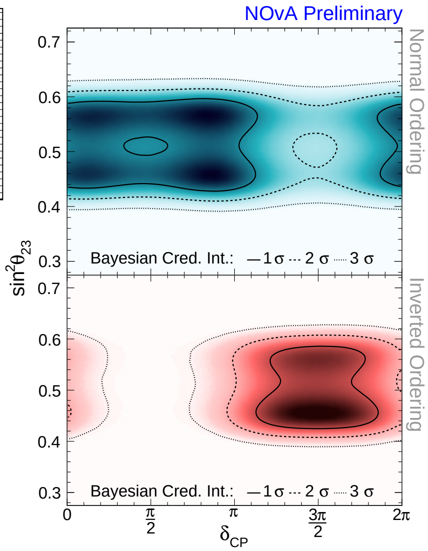
Results without Reactor Constraint

🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering





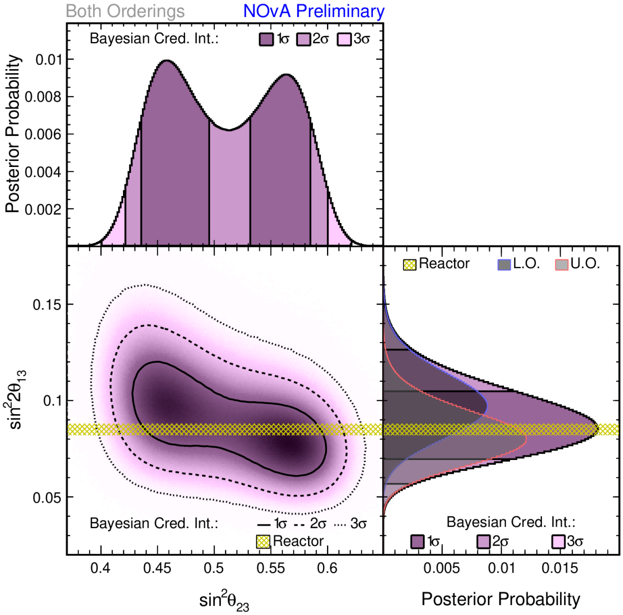
Lower octant: $\sin^2\theta_{23} < 0.5$, Upper Octant: $\sin^2\theta_{23} > 0.5$



- Prefer Lower Octant overall with NOvA-only θ_{13}
- Slight preference for Upper Octant in Normal Ordering.
- Higher preference for Lower Octant in Inverted Ordering.
- We need to look at θ_{13} to understand this.

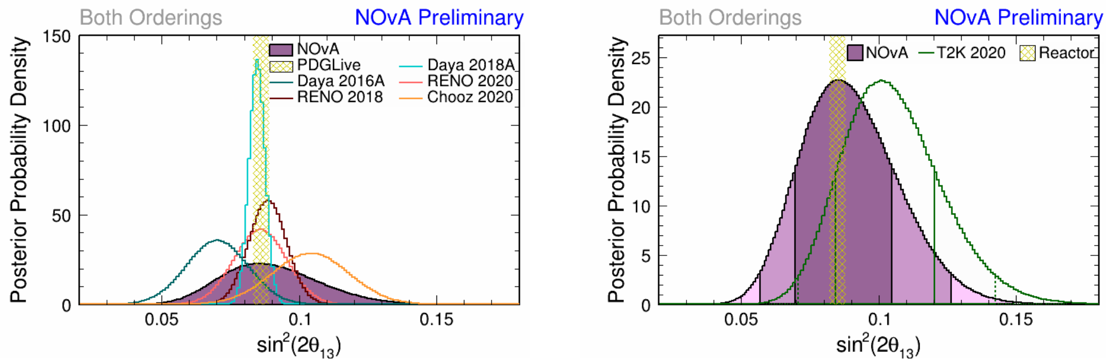
First NOvA-only θ_{13} measurement

🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering



- $\mathbf{\sin^2(2\theta_{13}) = 0.085^{+0.020}_{-0.016}}$
- NOvA in a good agreement with the reactor experiments.
- θ_{13} strongly linked with θ_{23}.
- Each θ_{23} octant and mass ordering prefers slightly different central value.
- Reactor's θ_{13} value, when used, causes higher preference for upper octant.

NOvA-only θ_{13}

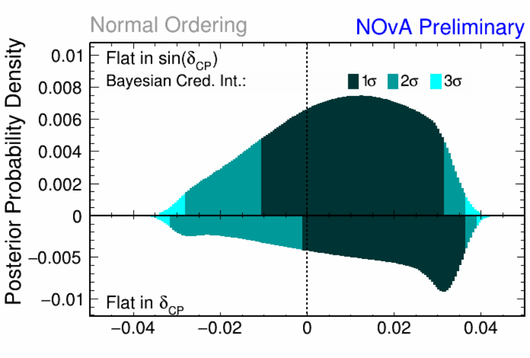


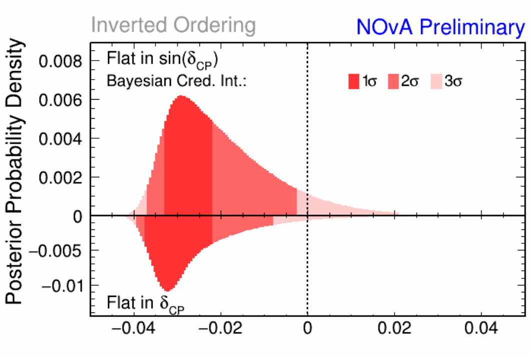
- NOvA (purple) in agreement with the reactor experiments.
- Also in agreement with T2K.
- No tensions between short-distance $P(\bar{\nu}_e \to \bar{\nu}_e)$ and long-distance $P(\nu_\mu \to \nu_e)$ & $P(\bar{\nu}_\mu \to \bar{\nu}_e)$.
- Gives the PMNS model extra credibility.

NOvA-only CP-violation

Jarlskog-Invariant

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering

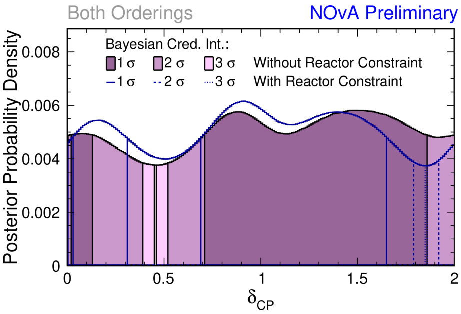


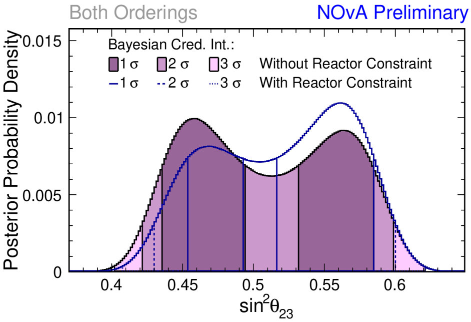


- The shape changes because θ_{13} is allowed to take more values.
- Nevertheless, the general conclusions about CP-conservation are similar.

Reactor constraint comparisons

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering





- Setting θ_{13} free does change our results slightly.
- Prefer lower octant with free θ_{13}, upper octant when constrained.
- These differences are low, however.
  - $1\,\sigma$ intervals in both octants.
  - Low Bayes Factors.

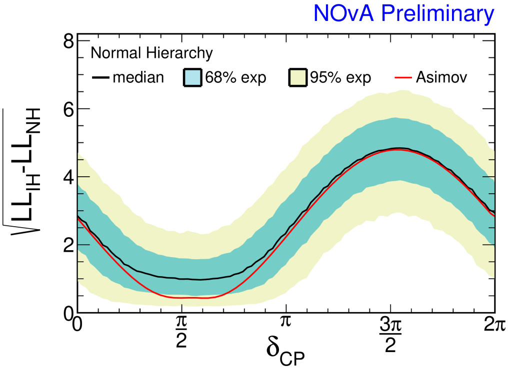
Bright Future

- NOvA Test-Beam to measure detector response.

- MW-capable horn and target already installed.
  - New power record reached last month!

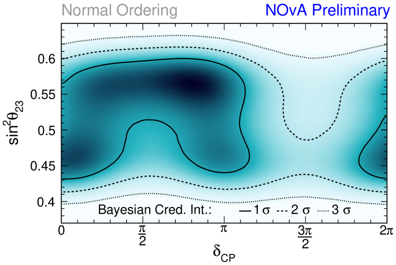
- Expect $> 2\times$ more in both ν and $\bar{\nu}$ data.
  - Analysed 26e20 POT.
  - 11e20 POT more collected since.
  - Goal by 2027: 67–72e9 POT.

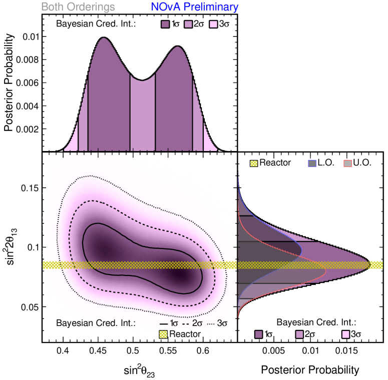
- NOvA-T2K effort to produce joint result.



Conclusions

- New techniques provide a different statistical interpretation of NOvA data.
- General agreement with the 2020 Frequentist results.
- First NOvA-only measurement of θ_{13}
  $\mathbf{\sin^2(2\theta_{13}) = 0.085^{+0.020}_{-0.016}}$
- **PMNS formalism explains NOvA data very well:**
  - No tension between Accelerator and Reactor neutrinos.
  - Jarlskog-Invariant: no high preference for CP-Violation or CP-Conservation.
- Similar results with and without external θ_{13} constraint.
  - Different preference for θ_{23} octant, but not significant.





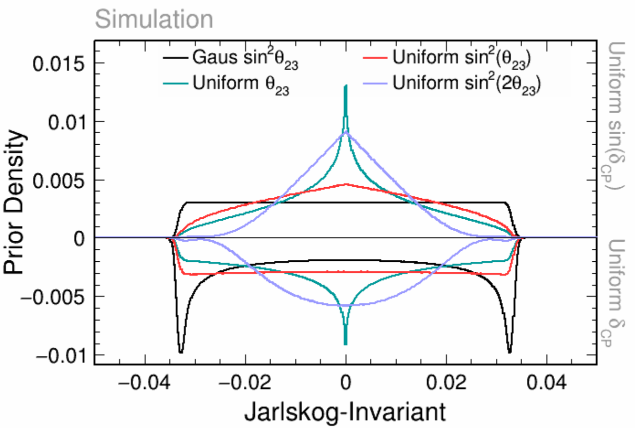

Patricia Vahle
Alex Himmel (he/him)
Queen Mary
574BOSTON406l
Jeffrey Hartnell (he/him)
Jozef Trokan-Tenorio (He/Him)
Louise Suter
Kevin John Vockerodt
Thomas Coan
Maria (she/her) Manrique Plata
W. Anthony Mann
Daniel Kaplan
Ryan B. Patterson
Mark Messier (he/him)
Prabhjot Singh (he/him/his)
David Dueñas
Reed Bowles (he/him)
Cathal Sweeney
Jeffrey Kleykamp (he/him)
Alexander Booth
Adam Aurisano
Robert Hatcher (he/him)
Yagmur Torun
Derek Doyle
Jon Paley (he/him)
Abhilash Yallappa Dombara
Kevin Mulder
Mualem Leon (he/him/his)
Miranda Elkins (she/her)
Mario A. Acero Ortega
Bindu Bambah
Alexandre B Sousa ((he/him))
Anna Hall (she)
Zoya (she/her)
Linda Cremonesi | QMUL (she/her)
Ryan Nichol
Karl Warburton (he/him)
Mike Wallbank
Akshay Chatla (He/Him)
Erica Smith (she/her)
Chris Backhouse
Erika (she/her) Catano-Mur
Adam Lister (he/him)
Anne Norrick (she/her)
Miriama Rajaoalisoa
Andrew Sutton
Enrique Arrieta Diaz
Steven Calvez
Connor Johnson
Harry Refael Hausner
Kirk Bays
Peter Shanahan
Lily Asquith
amit
Pierre Jean Joseph Lasorak (he/his)
Margot
Robert Kralik
Milos Lokajicek
Gao, Fan
Ayşe Bat
Dinesh Singha
Matt Judah
Artur Sztuc
Ashley Back (he/him)
Wilf Shorrock
Viktor Pec (he/him/his)
Rukmani Mohanta
Haejun (Stella) Oh
Tarak Thakore
v hewes ((they/them))
Gavin Davies
Travis Olson

April 2022

BACKUPS

Jarlskog-Invariant Priors

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering

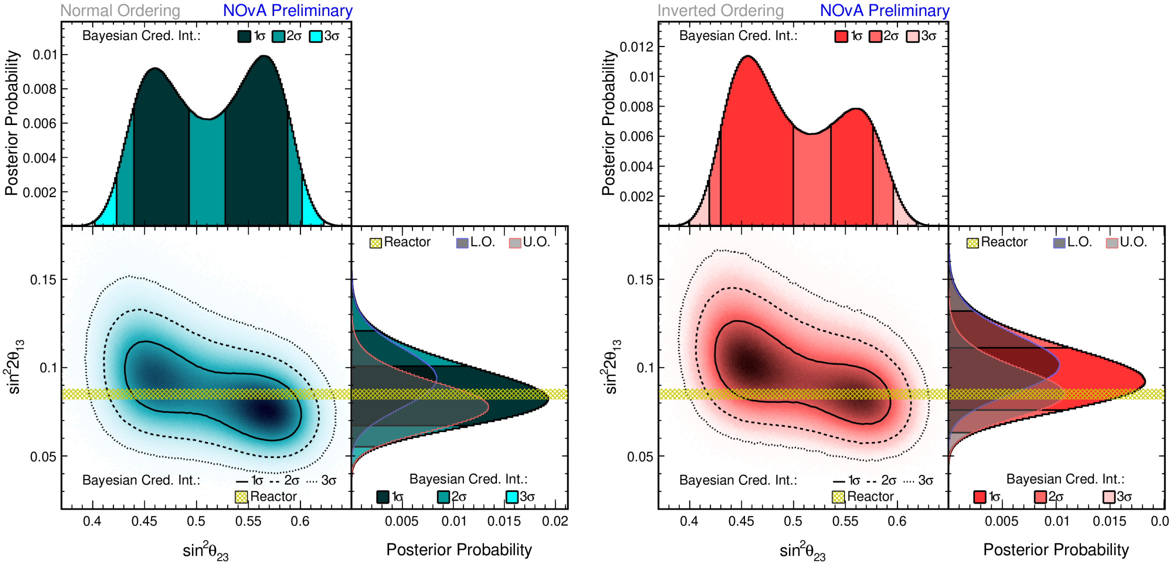


How about priors on other oscillation parameters for J?

- Changing θ priors to be uniform in θ, $\sin^2\theta$ changes the prior contribution.
- It does not, however, change the posterior – our results.
- Likelihood is stronger than the prior for high-stats data, overwhelming it.
- This does not happen for δ_{CP} because we don't constrain it well.

NOvA-only θ_{13} measurements

🟪 Both Orderings 🟦 Normal Ordering 🟥 Inverted Ordering

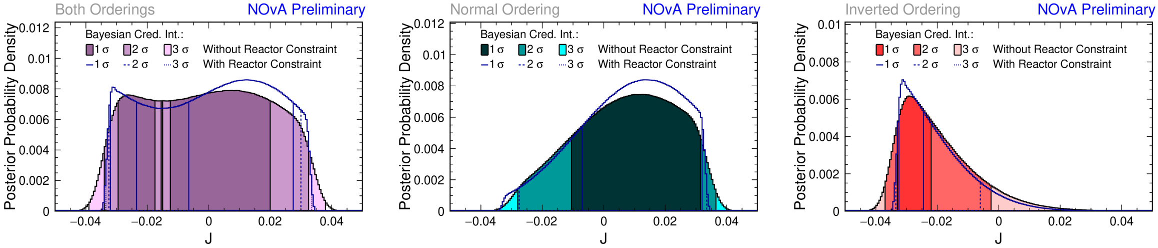


Jarlskog-Invariant

🟪 Both Orderings 🟩 Normal Ordering 🟥 Inverted Ordering

θ_{13} constraint comparisons

Prior uniform in $\sin(\delta_{CP})$:



Prior uniform in δ_{CP}: