Efficient sampling for the precision era of neutrino experiments

IOP Joint APP and HEPP Annual Conference, Edinburgh

Yuhan (Hank) Hua

April 9, 2026

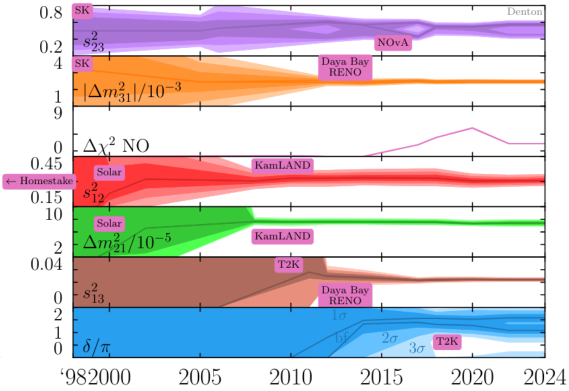
Outline

1. Motivation and background
2. Current challenges in Bayesian oscillation analysis
3. Adaptive MCMC and downsampling
4. Performance on DUNE
5. Summary and outlook

Introduction

Precision era

- Joint fits of existing experiments
  - T2K+NOvA
  - T2K+Super-K
- Next-generation neutrino experiments
  - Deep Underground Neutrino Experiment (DUNE)
  - Hyper-Kamiokande (Hyper-K)
  - Jiangmen Underground Neutrino Observatory (JUNO; data taking started 2025)
- Will be making precise measurements of oscillation parameters

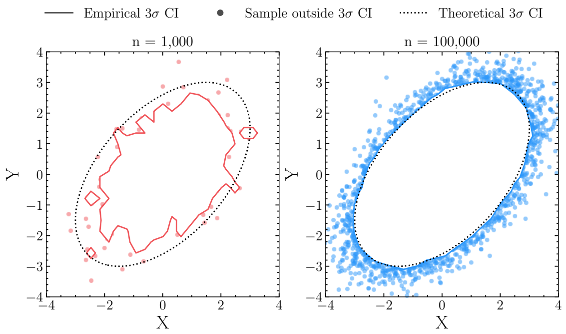


E.g. 3σ exclusion of $\delta_{\mathrm{CP}} = 0$ by T2K in 2018. Credit: Peter B. Denton[1]

[1] arXiv:2212.00809; NuFact 2024

The Bayesian approach

- Goal: construct credible interval (CI) for oscillation parameters by sampling from posterior distribution
- Need sufficient independent samples outside of CI to construct it
  - 3σ: 1 in ~ 300
  - 5σ: 1 in ~ 1.7M ($6000\times$ compared to 3σ)



CI estimation on a correlated 2D Gaussian

Introduction

Sampling the distribution

- Markov Chain Monte Carlo (MCMC) is a common way of sampling
  - Propose a new random step based on a proposal function
  - Evaluate the likelihood at the new step
  - Accept with a probability proportional to the likelihood ratio
- MaCh3 is an analysis framework that includes a Bayesian fitter[2]
  - Used by DUNE, Hyper-K, T2K, T2K-NOvA, T2K-SK, etc.
  - Used in this talk

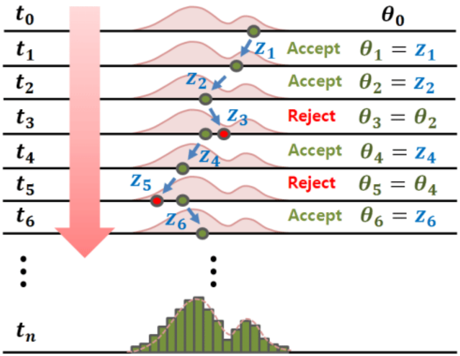


Diagram of MCMC[3]

[2]Github: https://github.com/mach3-software/MaCh3

[3]Jin et al. *Structure and Infrastructure Engineering, 15*(11), 1548–1565, 2019

Challenges with MCMC

- Running MCMC is nontrivial:

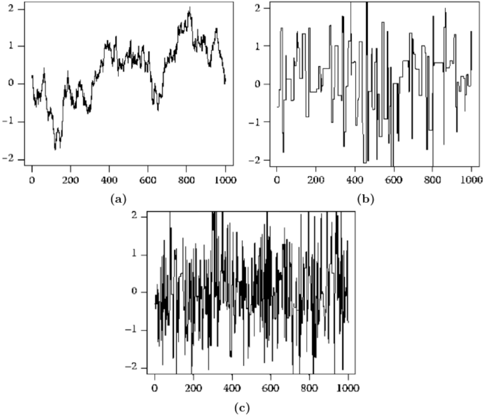
1. Computation resources
   - Sampling in high dimensional, highly correlated and degenerate space
   - Independent samples require more steps
2. Human time
   - Step size tuning needed to get a good chain
   - New tuning + fit every time something changes (e.g. new systematics models, new samples)

[4] Brooks et al. *Handbook of Markov Chain Monte Carlo.* CRC press, 2011

Challenges with MCMC

✦ Running MCMC is nontrivial:

1. Computation resources
   - Sampling in high dimensional, highly correlated and degenerate space
   - Independent samples require more steps
2. Human time
   - Step size tuning needed to get a good chain
   - New tuning + fit every time something changes (e.g. new systematics models, new samples)

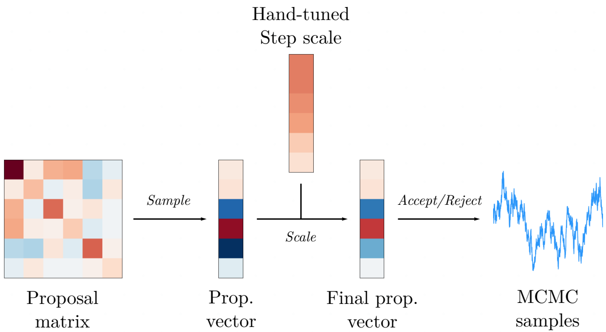


Step size (a) too small, (b) too big, (c) just right[4]

[4]Brooks et al. *Handbook of Markov Chain Monte Carlo.* CRC press, 2011

Step size tuning in MaCh3

1. Propose parameters from proposal matrix (usually prior)
2. Scale each proposed parameters by hand-picked step scale
3. Run MCMC and check performance
4. Tweak step scales[5]
5. Repeat



[5] "It is a bit of dark magic" – MaCh3 Wiki

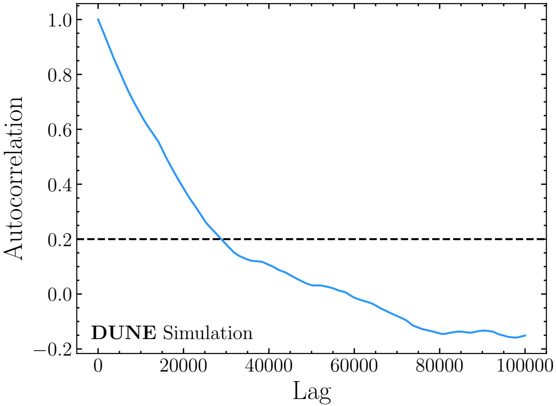
Challenges

Quantifying performance

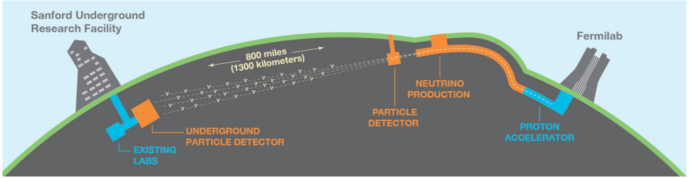
- MCMC does not always produce independent samples – steps are correlated
- Autocorrelation: a metric to determine when we get an independent sample

$$\rho(k) = \frac{1}{N-k} \sum_{i}^{N-k} \mathrm{corr}(X_i, X_{i+k})$$

  - Low autocorrelation → less affected by previous steps → more independent samples
- Typically use autocorrelation < 0.2 as the threshold for an independent sample

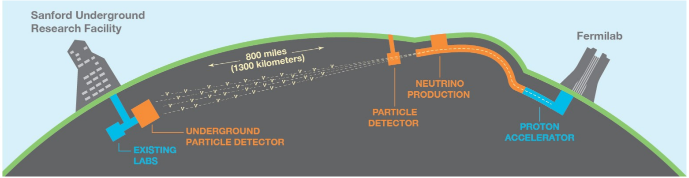


Challenges on DUNE



- DUNE – upcoming experiment under construction:
  - Long baseline
  - High intensity, wide-band neutrino beam
  - Liquid Argon Time Projection Chamber
  - Moveable Near Detector (ND): $\sim$ 100M neutrino events a year
- Enables precision measurements of e.g. δ_{CP}, θ_{23}, neutrino mass ordering

Challenges on DUNE



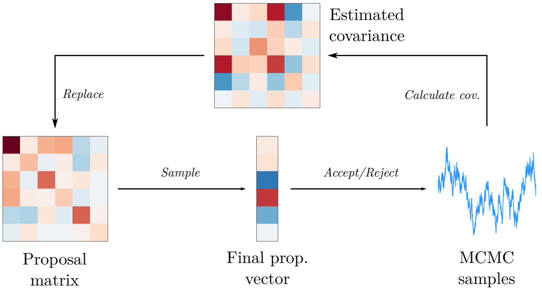
- MaCh3 was used for the Bayesian sensitivity study of DUNE using the existing samples and systematics[6]
  - ~ 300 parameters, some highly-correlated, non-Gaussian
  - Near detector (ND) has $\mathcal{O}(100\text{M})$ simulated events $\implies$ each MCMC step takes a long time
  - Step size initially hand-tuned, took $\mathcal{O}(\text{months})$
- Unfeasible when DUNE starts taking data

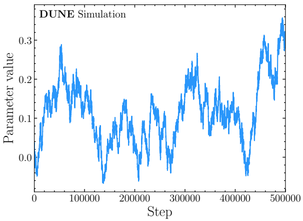
[6] arXiv:2002.03005

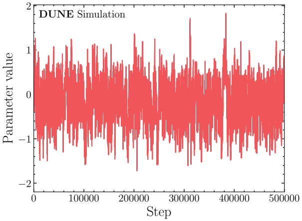
Solutions

Adaptive MCMC

- Adaptive MCMC can automatically adjust the step size as the chain converges
- Calculates covariance of the chain and adapts the proposal matrix
  - Optimal proposal $\propto$ posterior covariance matrix
- No human intervention needed
- Better proposal $\rightarrow$ more efficient



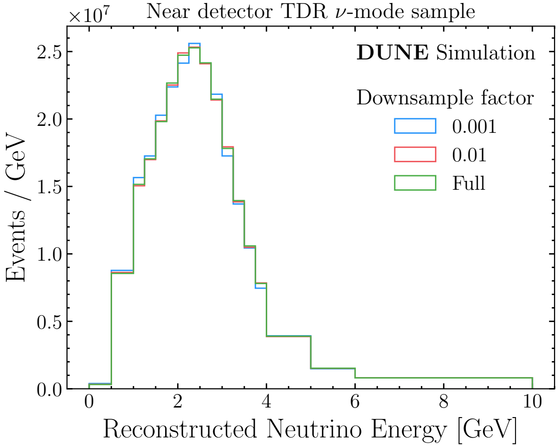




Y. Hua (h.hua23@imperial.ac.uk)

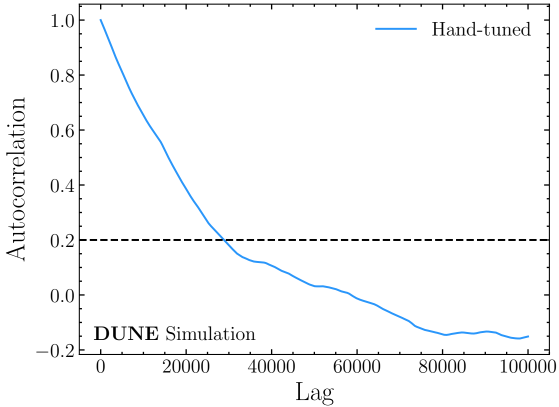
Downsampling events

- Problem: A lot of events $\implies$ adaptive MCMC is slow
- Downsampling: take a fraction of events, and scale accordingly
- Only needs approximate likelihood for adaptive MCMC
- Downsampling will only be used to obtain an adapted proposal matrix
- Fewer events $\to$ faster MCMC



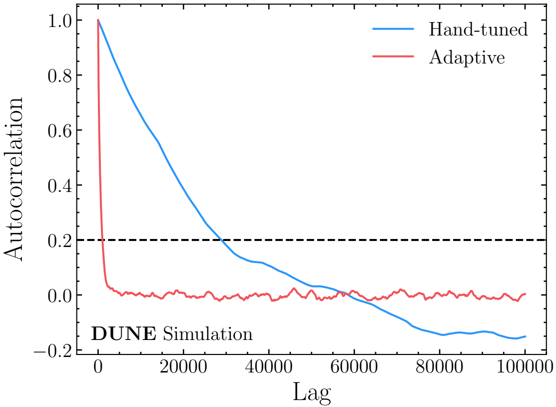
Autocorrelation

- Autocorrelation of a parameter from a hand-tuned chain



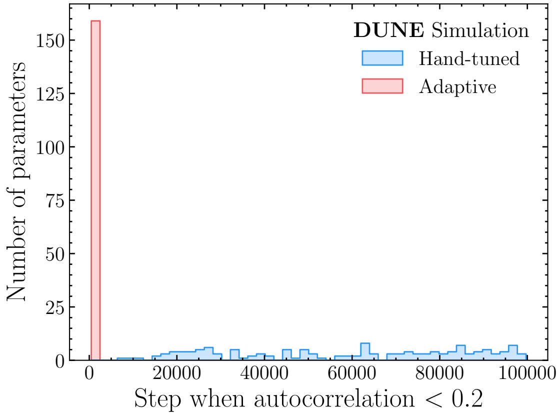
Autocorrelation

- Autocorrelation of a parameter from a hand-tuned chain
- Using adapted proposal matrix in a normal fit, autocorrelation drops to < 0.2 very quickly for this parameter



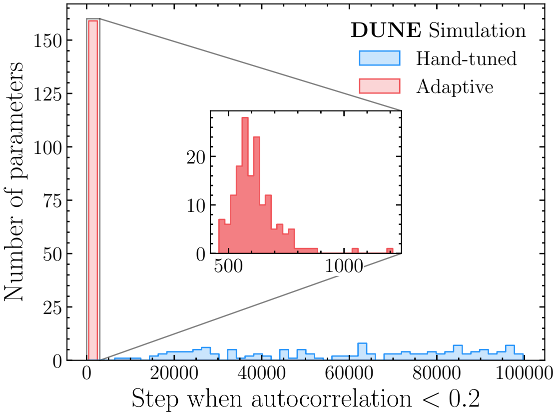
Autocorrelation

- Autocorrelation of a parameter from a hand-tuned chain
- Using adapted proposal matrix in a normal fit, autocorrelation drops to < 0.2 very quickly for this parameter



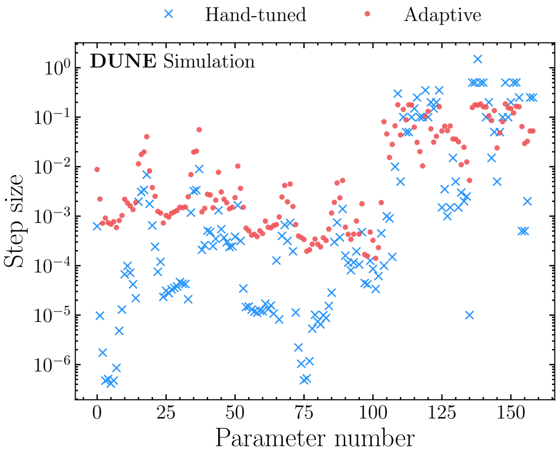
Autocorrelation

- Autocorrelation of a parameter from a hand-tuned chain
- Using adapted proposal matrix in a normal fit, autocorrelation drops to < 0.2 very quickly for this parameter
- Largest step before adaptation: ~ 100000; largest step after adaptation ~ 2000 $\implies$ $50\times$ more independent samples



Step sizes

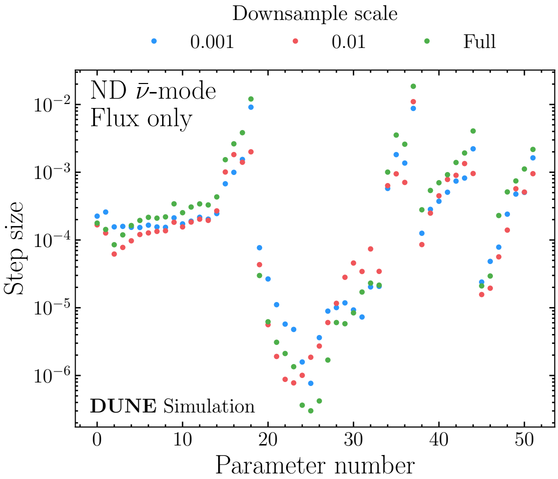
- "Step size" – diagonal of the Cholesky decomposed proposal matrix
- Adapted MCMC reached larger step sizes, but has similar trend – where improvements in autocorrelation come from
- Hand-tuning took $\mathcal{O}$(months) of human time, adaptive MCMC with downsampling took $\mathcal{O}$(days)
- Downsampling reached a similar level of step size as full MC sample
  - ND RHC, flux parameters only – not directly comparable with the other plot



Results

Step sizes

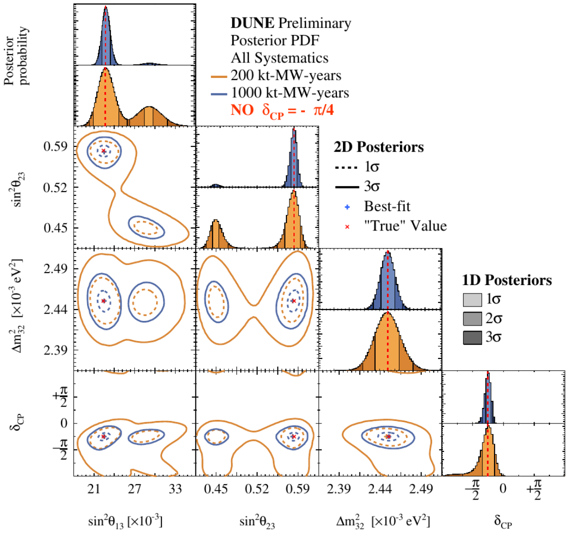
- "Step size" – diagonal of the Cholesky decomposed proposal matrix
- Adapted MCMC reached larger step sizes, but has similar trend – where improvements in autocorrelation come from
- Hand-tuning took $\mathcal{O}$(months) of human time, adaptive MCMC with downsampling took $\mathcal{O}$(days)
- Downsampling reached a similar level of step size as full MC sample
  - ND RHC, flux parameters only – not directly comparable with the other plot



Y. Hua (h.hua23@imperial.ac.uk)

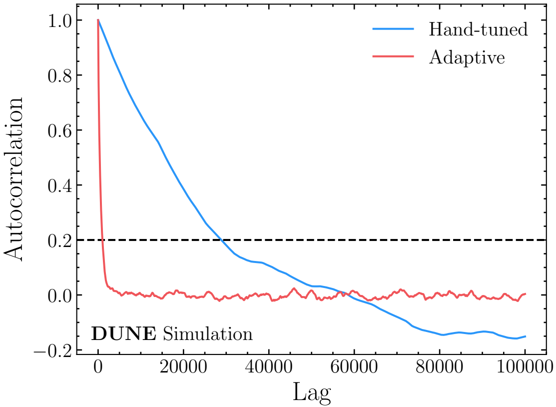
Adaptive MCMC applications

- Bayesian sensitivity study for DUNE
  - 50× more independent samples per chain
- Technique currently used in T2K-NOvA joint analysis



Summary

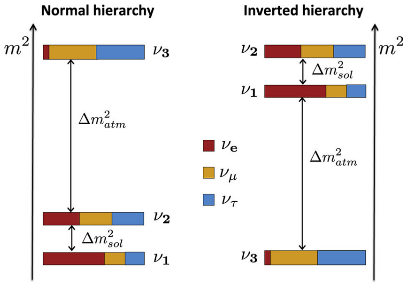
- Finding the right MCMC step size has been a challenge for oscillation analysis, requires both computational and human time
- We used adaptive MCMC and downsampling on DUNE systematics and samples
- Computational improvement: $\sim 50\times$ more independent samples per chain
- Human time saved: $\mathcal{O}$(months) $\rightarrow$ $\mathcal{O}$(days)



Neutrino oscillation

$$U_{\mathrm{PMNS}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Neutrino flavour states are mixtures of mass states → change in flavour as neutrino propagates
- Is this model correct? Is there anything new?
- Precise measurements of values can reveal what's missing / what's unexpected

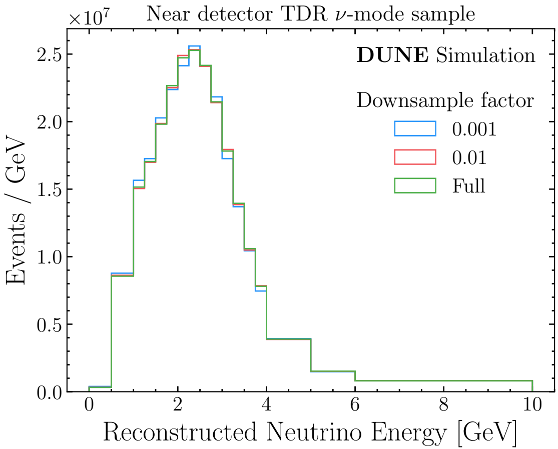


Adaptive MCMC for DUNE

- Henry Wallace initially implemented this in T2K MaCh3, now in MaCh3 core; Michael Reh used it for SK atmospherics
- Using adaptive MCMC in DUNE can:
  - Reduce human labour required to run a chain (hand-tuning takes months for given systematics and samples, + need to retune)
  - Produce more independent samples for the same chain length, necessary for high-significance results & reduce compuation requirements
- Goal: given any new systematics + samples, use adaptive MCMC to get adapted proposal matrix, use this to run full (non-adaptive) chains

Downsampling events

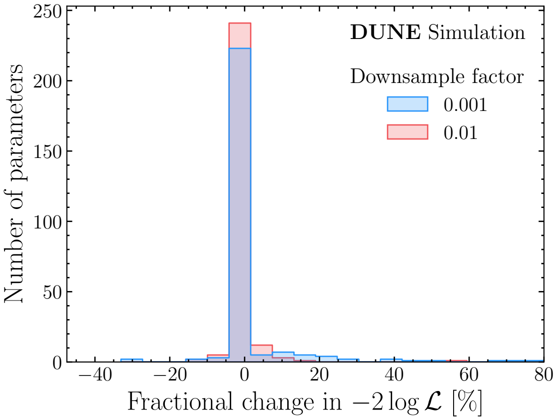
- Problem: A lot of events $\implies$ adaptive MCMC is slow
- A solution: take a fraction of events, and scale POT correspondingly
- Downsampling will only be used to obtain an adapted proposal matrix
- Event rates look similar, run takes $\sim$ hours



Backup

Downsampling

- Problem: A lot of events at ND $\implies$ adaptive MCMC is slow
- Downsampling: take a fraction of events, and scale accordingly
- Calculated the change in $-2\log\mathcal{L}$ as each parameter deviates $\pm1\sigma$ from nominal
- Fewer events $\rightarrow$ faster MCMC step
- Downsampling will only be used to obtain an adapted proposal matrix, not during fits



Fractional change:
$(\mathrm{LLH}_{\mathrm{downsampled}} - \mathrm{LLH}_{\mathrm{full}})/\mathrm{LLH}_{\mathrm{full}}$