The MaCh3 Oscillation Analysis Fitter

Thomas Holvey

thomas.holvey@physics.ox.ac.uk

Institute of Physics, OWAN 2021
10th November 2021

Outline

- Bayesian inference
- Markov Chain Monte Carlo methods
- The MaCh3 analysis framework
- Current and future plans for MaCh3

Bayesian Inference

- MaCh3 uses a Bayesian interpretation of statistics
- From a Bayesian perspective, there is no fundamental difference between the data (D) and model parameters (θ)
- To make inferences about model parameters, we must find a model that describes both the data and model parameters, this is known as the **joint probability distribution:**

$$P(D,\theta) = P(D|\theta)P(\theta)$$

Joint Probability Distribution = Likelihood of measuring data assuming set of parameters × Prior (previous knowledge about a parameter)

Aim of a Bayesian analysis: what is the probability for each parameter to have a certain value given our data, i.e. $P(\theta|D)$, the **posterior distribution**

Calculate posterior from joint probability distribution using Bayes' Theorem →

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{\boxed{\int P(D|\theta)P(\theta)d\theta}}$$

Hard to evaluate!

How do calculate the posterior..?

Markov Chain Monte Carlo

- MaCh3 uses a Markov Chain Monte Carlo (MCMC) method to calculate the posterior
- Chain carries out semi-random walk through parameter space, builds up discrete points whose density is proportional to the posterior
- For a chain to always have a stationary distribution, a chain must be:
  - **Irreducible** – chain has non-zero probability to reach all other potential states
  - **Recurrent** – once stationary distribution reached, all subsequent steps must be samples from the stationary distribution
  - **Aperiodic** – chain must not be periodic (e.g. can't alternate back and forth between states)
- Therefore, we need a chain that:
  1. Satisfies the **'regularity'** conditions listed above, and
  2. Has the posterior $P(\theta|D)$ as its stationary distribution
- Can ensure these conditions are met using the **Metropolis-Hastings algorithm**

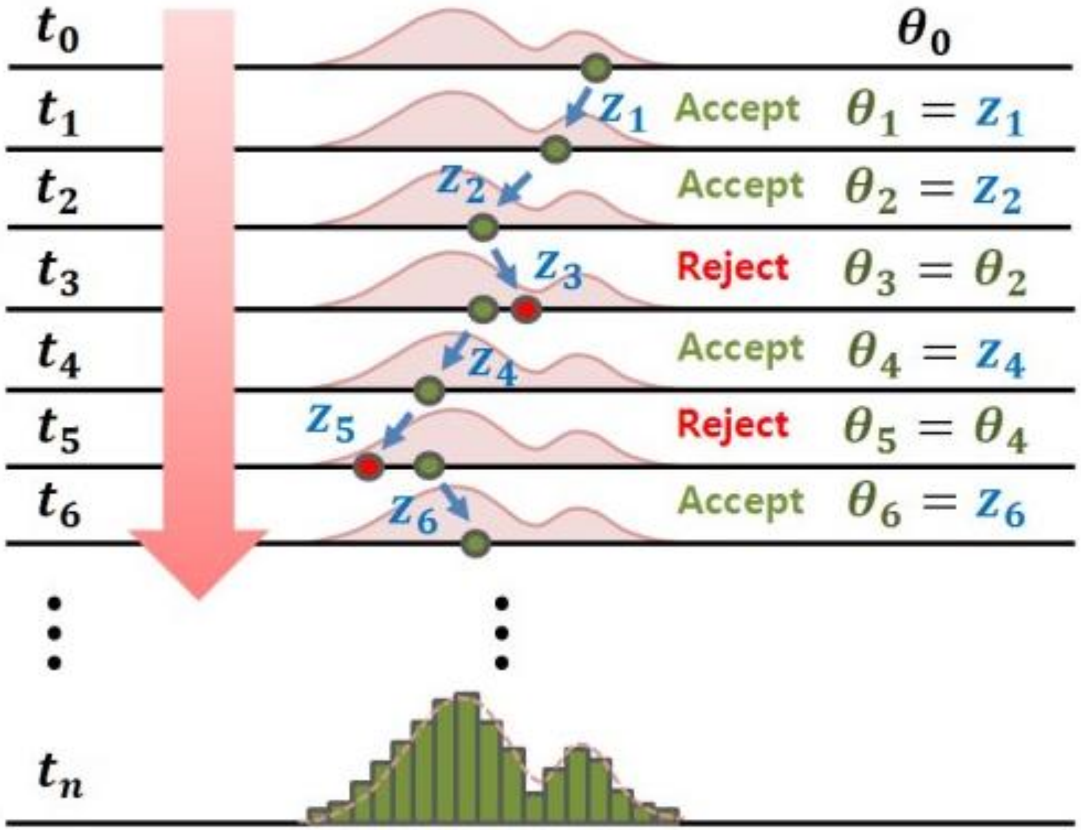
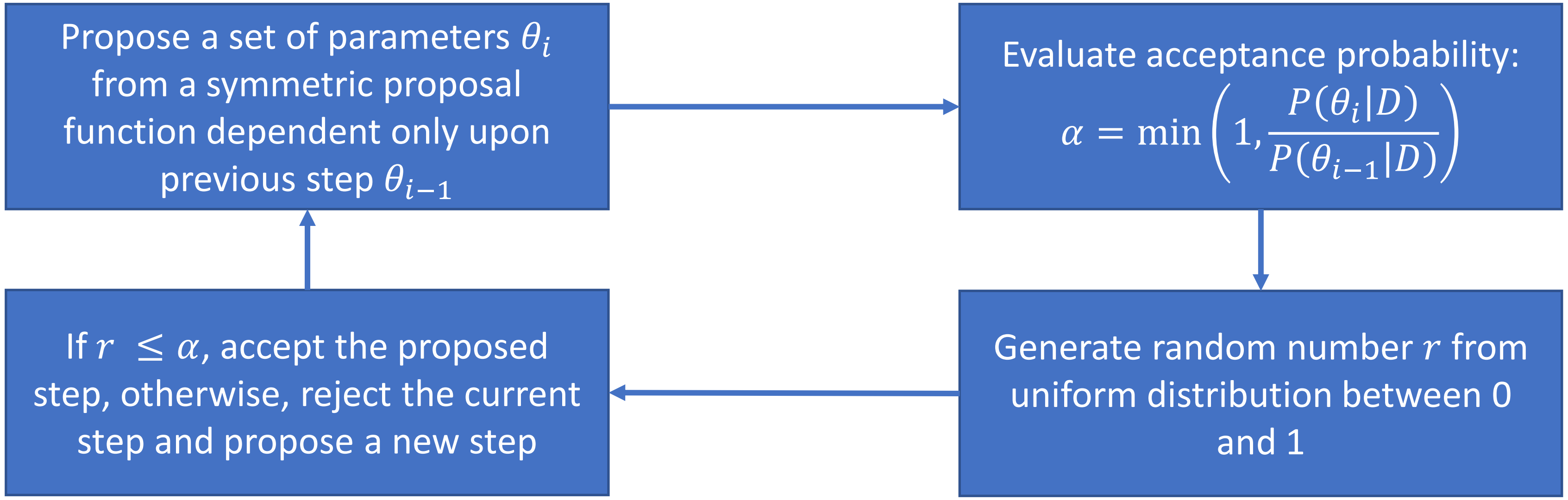


Illustration of MCMC method. Source: S.-S. Jin, H. Ju, and H.-J. Jung, Structure and Infrastructure Engineering, vol. 15, no. 11, pp. 1548–1565, 2019.

Metropolis-Hastings Algorithm



Propose a set of parameters θ_i from a symmetric proposal function dependent only upon previous step θ_{i-1}

Evaluate acceptance probability:

$$\alpha = \min\left(1, \frac{P(\theta_i|D)}{P(\theta_{i-1}|D)}\right)$$

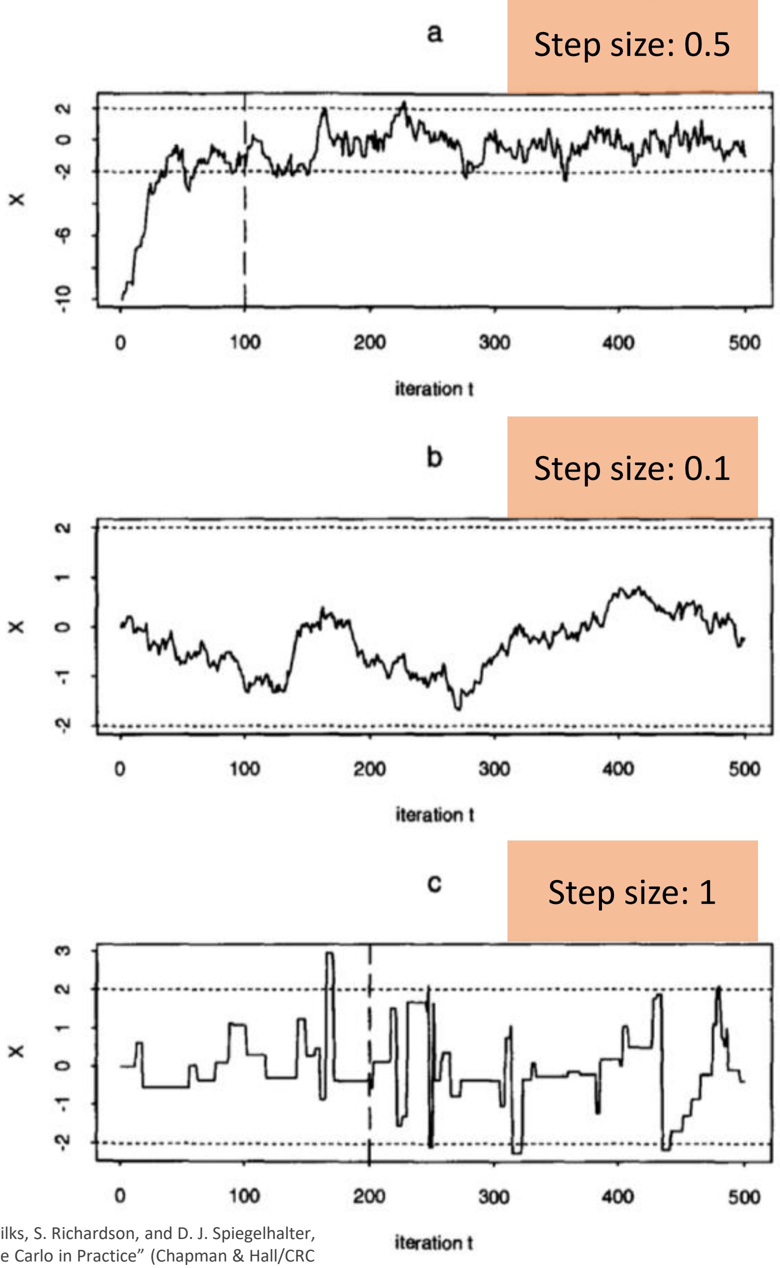
Generate random number r from uniform distribution between 0 and 1

If $r \ \leq \alpha$, accept the proposed step, otherwise, reject the current step and propose a new step

- As algorithm progresses, it builds up distribution of discrete points, with more points in high probability regions than low probability regions
- Highly probably regions are always accepted, other regions are sometimes accepted (semi-random aspect due to random number generation)
- The MH algorithm ensures our chain will always reach the stationary distribution (proportional to the posterior), but ideally it will converge as quickly as possible

The Art of MCMC

- MCMC is a powerful method, but with great power comes great responsibility…
- The efficiency and speed of your chain depends most on your choice of:
    - **Proposal function:** gives us the area of parameter space our chain will move to next
    - **Step size:** how large the jumps are between adjacent steps
- Step sizes should be tuned to achieve optimal performance/speed whilst also maintaining a reasonable acceptance rate



Step size: 0.5

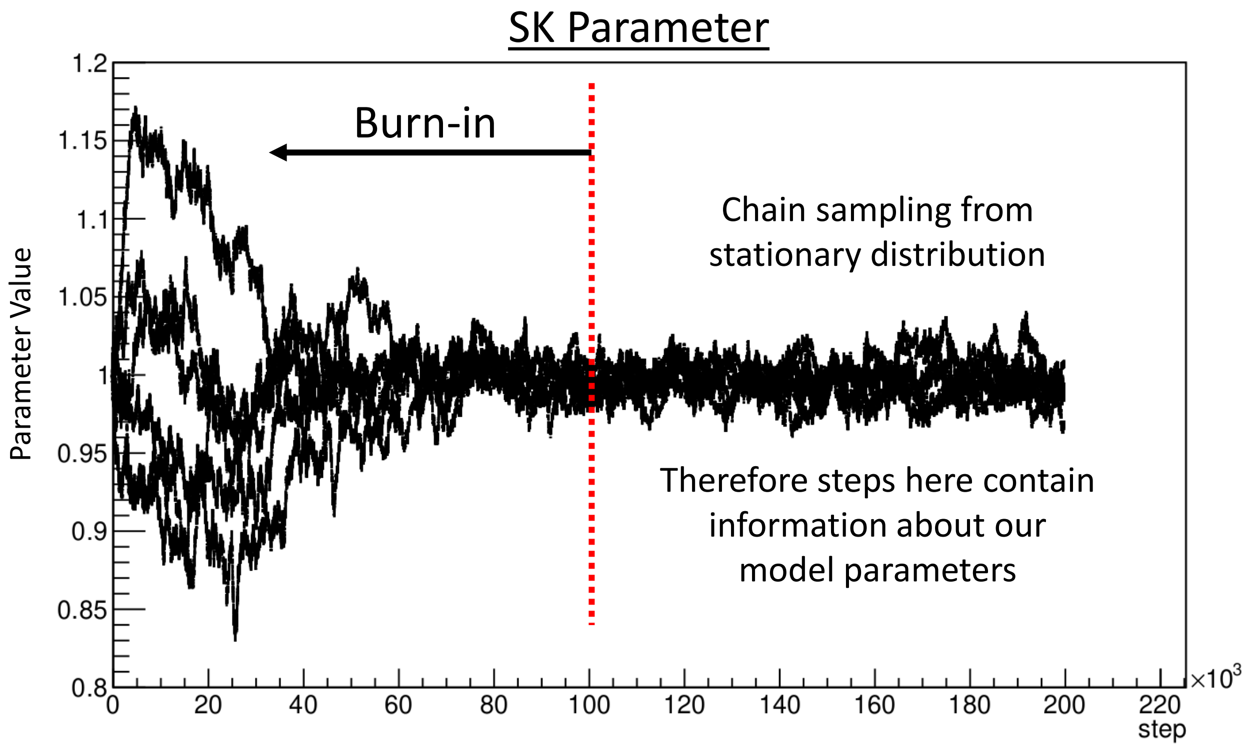
Step size: 0.1

Step size: 1

Figure from: W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, “Markov Chain Monte Carlo in Practice” (Chapman & Hall/CRC Interdisciplinary Statistics), Chapman and Hall/CRC (1995)

Burn-in

- Markov Chains don’t usually start in a region of high probability, you need to let them adequately explore the parameter space
- This usually takes a significant number of steps to achieve, known as the ‘burn-in’ phase
- These steps won’t contain any information on your underlying stationary distribution, so these should be discarded before any analysis is done
- From the ‘regularity’ conditions, the chain must ‘forget’ its starting point, therefore can safely discount the burn-in phase of the chain

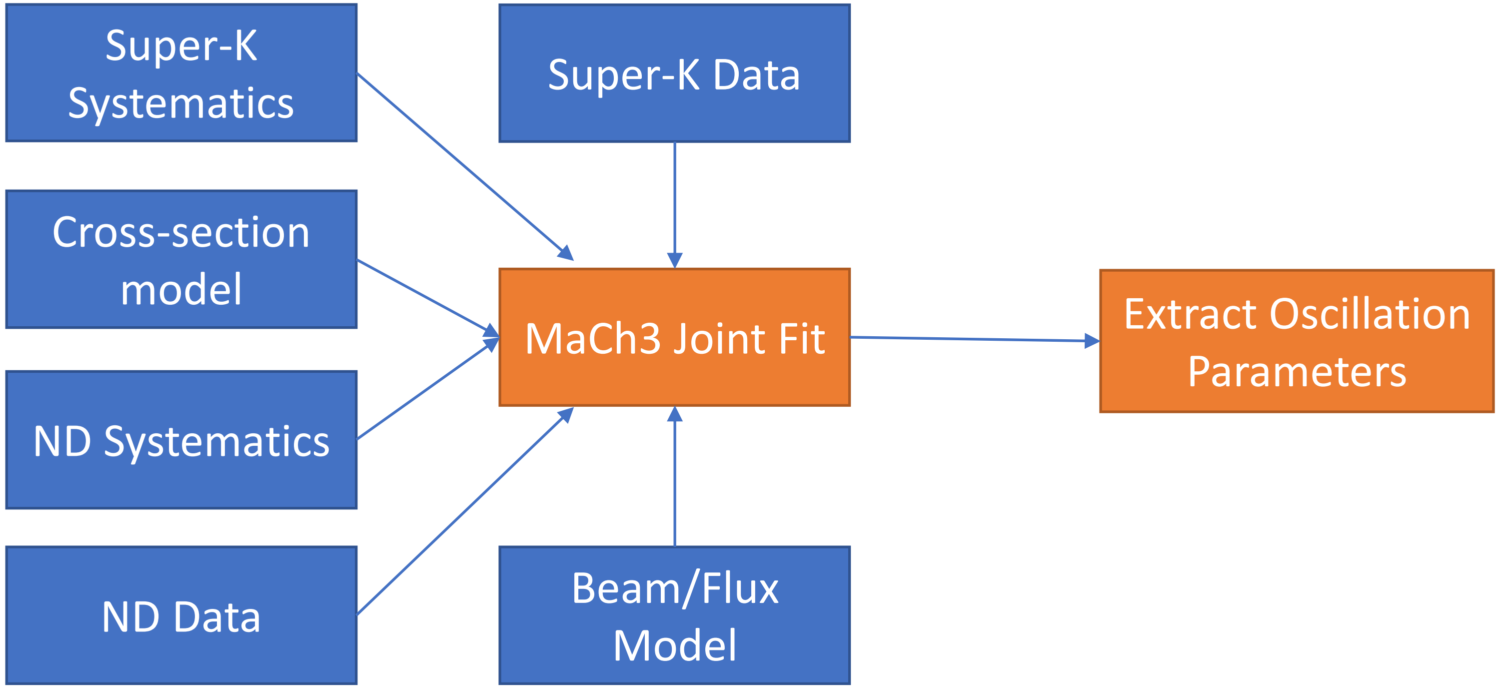


Chain sampling from stationary distribution

Therefore steps here contain information about our model parameters

MaCh3

- MaCh3 uses a MCMC method to sample the posterior distribution, from this conclusions can be drawn about the oscillation parameters
- MaCh3 performs a joint-fit of both near- and far-detector parameters:
- Output of MaCh3 fit is a ~750 dimensional posterior – far too large to visualise or interpret
- Posterior is marginalised (integrated) onto fewer dimensions and filled into a histogram



Advantages of MaCh3

- MaCh3 has full access to event-by-event kinematic information, which gives the fitter several advantages:
    - Allows functional re-weights to be applied to any event variable
    - Enables shifts to be applied to kinematic variables (e.g. binding energy corrections applied to lepton momentum)
    - Can construct any variable combinations that are needed
- MaCh3 has full implementation of the near-detector, allowing a simultaneous of ND+FD data and systematics – avoids assumptions about ND constraints
- Can apply different priors by simply re-weighting the posterior distribution; equivalent to re-running a new MCMC with different priors
    - For T2K, it is useful to re-weight the posterior using the PDG world-average for the θ_{13} mixing angle from reactor experiments, which have superior sensitivity

Configurable uncertainty model

- MaCh3 has been developed to allow uncertainty models to be easily configurable by users
- Can specify any response function or normalisation systematic and its correlations using a simple xml config file
  - For normalization parameters we have extended this to allow user to choose which variables (e.g. interaction modes, kinematic regions, etc.) are affected without touching compiled code

```xml
<!-- Single pion parameters -->
<!-- Priors and correaltion updated in 2021 https://www.t2k.org/docs/technotes/414 -->
<!-- See chapter 5.2 New priors on Rein-Sehgal parameters -->
  <parameter name="CA5" nom="1.01" prior="1.06" lb="0" ub="9999" error="0.1" renorm="0" type="spline" splineind="28" detid="985" sk_splin
    <correlation par="MARES">-0.11</correlation>
    <correlation par="ISO_BKG">-0.03</correlation>
    <sk_mode>1 5 6 14</sk_mode>
  </parameter>
  <parameter name="MARES" nom="0.95" prior="0.91" lb="0" ub="9999" error="0.1" renorm="0" type="spline" splineind="29" detid="985" sk_spl
    <correlation par="CA5">-0.11</correlation>
    <sk_mode>1 5 6 14</sk_mode>
  </parameter>
  <parameter name="ISO_BKG_LowPPi" nom="1.3" prior="1.3" lb="0" ub="9999" error="1.30" renorm="0" type="spline" splineind="30" detid="984
    <sk_mode>1 5 6 14</sk_mode>
  </parameter>
  <parameter name="ISO_BKG" nom="1.3" prior="1.21" lb="0" ub="9999" error="0.27" renorm="0" type="spline" splineind="31" detid="985" sk_s
    <correlation par="CA5">-0.03</correlation>
    <sk_mode>1 5 6 14</sk_mode>
  </parameter>
```

Current/Future Plans for MaCh3

- MaCh3 is being used extensively in the T2K Oscillation Analysis and is also being used in several other areas within **T2K:**
  - MaCh3 has been modified to allow for joint-fit between T2K beam data and SK atmospheric data (see Dan Barrow's talk)
  - MaCh3 is also being used in T2K-NOvA joint-fit effort
- Work is also underway to adapt MaCh3 for use in next-generation neutrino experiments, DUNE and Hyper-K
  - **Hyper-K:** MCMC step-size tuning and validations with other fitters
  - **DUNE:** Currently reproducing DUNE Far Detector Technical Design Report analysis – also restructuring and redesigning MaCh3 to more easily support multiple experiments

Current MaCh3 Personnel


Asher Kaboth


Patrick Dunne


Henry Israel


Tom Holvey


Alex Carter


Liban Warsame


Balint Radics


Kevin Wood


Dan Barrow

Ed Atkin


Clarence Wret


Kamil Skwarczynski


Abraham Teklu

~20 people actively working on MaCh3 across many institutions

Summary

- MaCh3 uses a Bayesian Markov Chain Monte Carlo method to draw conclusions about oscillation parameters
- It has full access to event-by-event kinematic information, which allows re-weighting and shifts to be easily applied
- MaCh3 fits both ND+FD data/systematics, therefore avoiding assumptions about ND constraints
- MaCh3 is widely used in many LBL neutrino experiments, both current and next-generation
- Plenty of scope to get involved in MaCh3, contact Patrick Dunne (p.dunne12@imperial.ac.uk) to find out more!

Backup

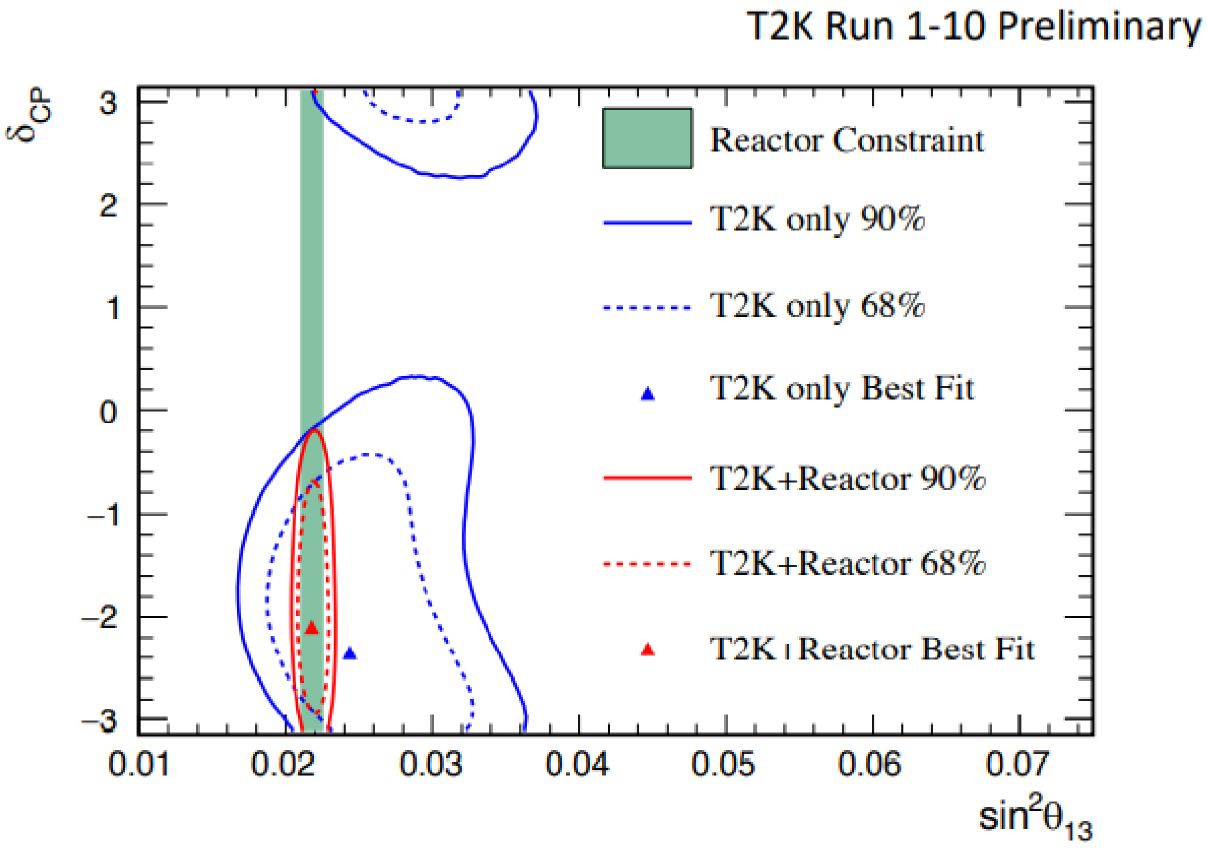
Marginalisation

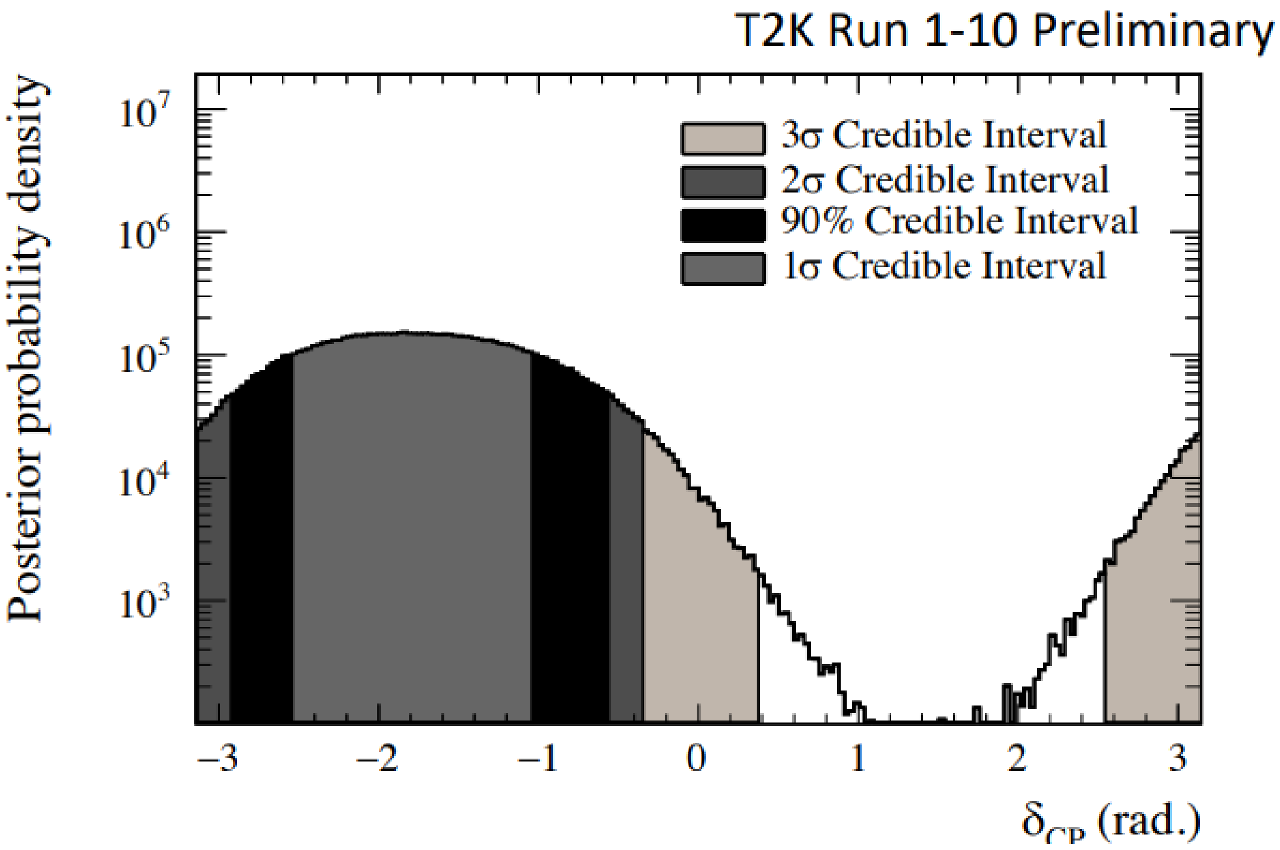
- MaCh3 involves fitting a ~750 parameter distribution, but only interested in 4 oscillation parameters
- Need to ensure the uncertainty introduced due to other model parameters are appropriately taken into account
- To do this, we marginalise (integrate) over all other ‘nuisance’ parameters (i.e. parameters that aren’t of specific interest to an oscillation analysis)
- Doing this yields the marginal posterior:

$$P\left(\theta_{poi}\middle|D\right) = \frac{\int P\left(D\middle|\theta_{poi},\theta_n\right)P\left(\theta_{poi},\theta_n\right)d\theta_n}{\int P(D|\theta)P(\theta)d\theta}$$

- In T2K MaCh3, the posterior is often marginalised onto 1 or 2 dimensions, and filled into a histogram

MaCh3 Appearance Analysis





35% of δ_{CP} values excluded at 3σ
CP conserving values $(0, \pi)$ excluded at 90% CL