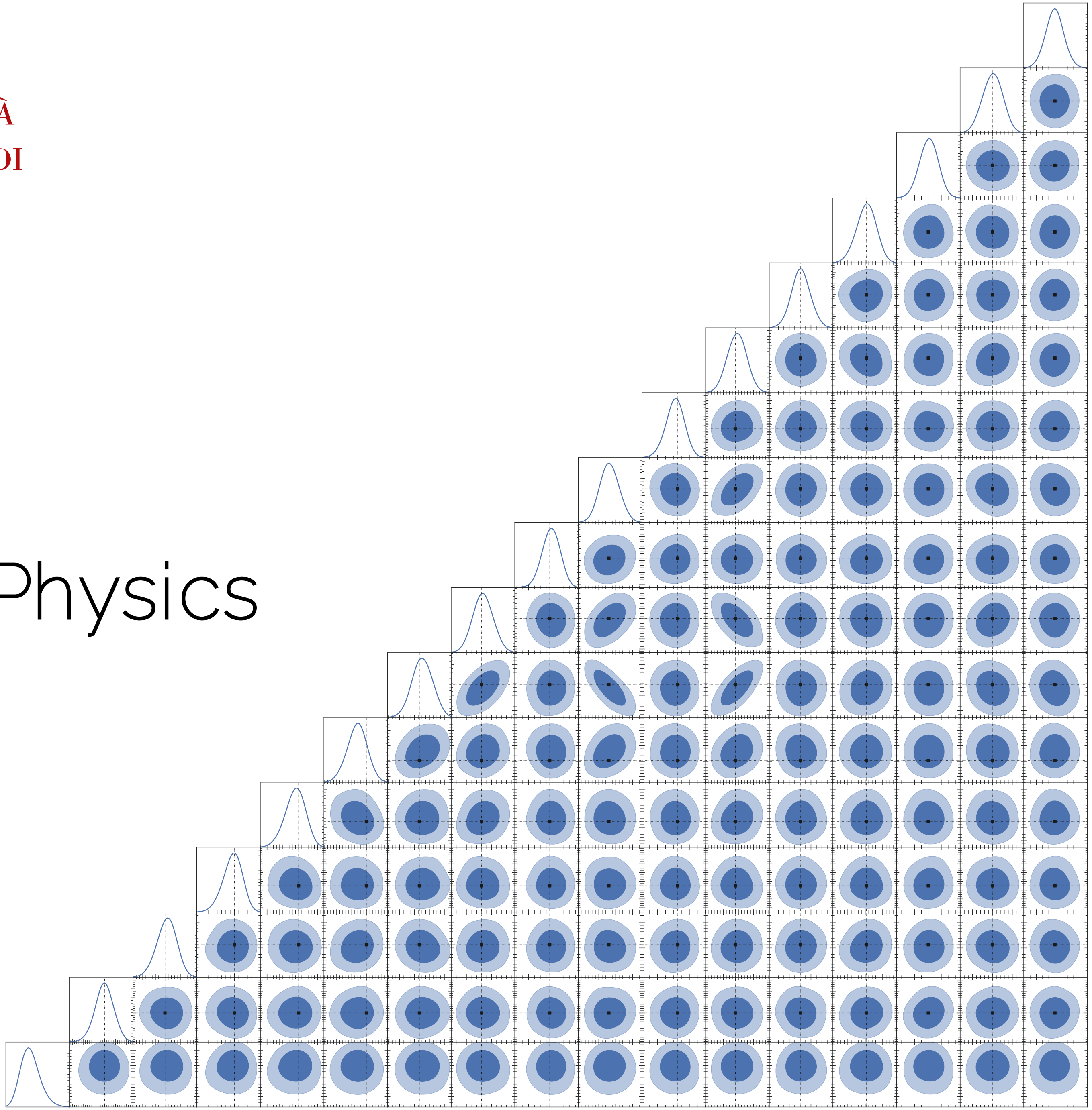
Istituto Nazionale di Fisica Nucleare

Università degli Studi di Padova

Accelerating Neutrino Physics Analysis with GPU Parallelization

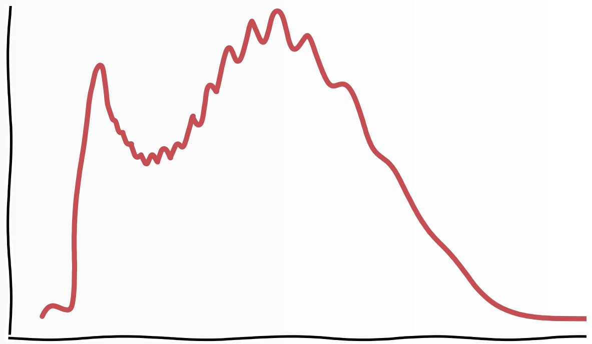
Andrea Serafini
on behalf of the Padova group
andrea.serafini@pd.infn.it

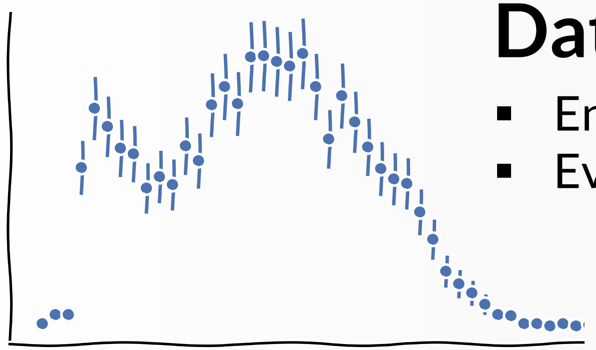


ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response



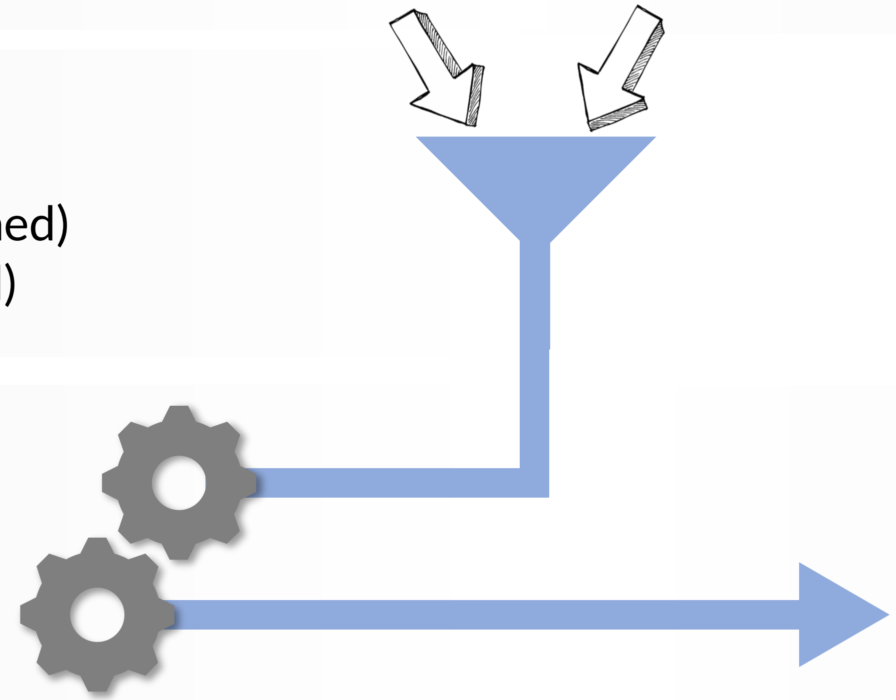


Data processing

- Energy reconstruction
- Event selection

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)



Fitter

- Minuit
- Markov Chains MC
- Nested sampling

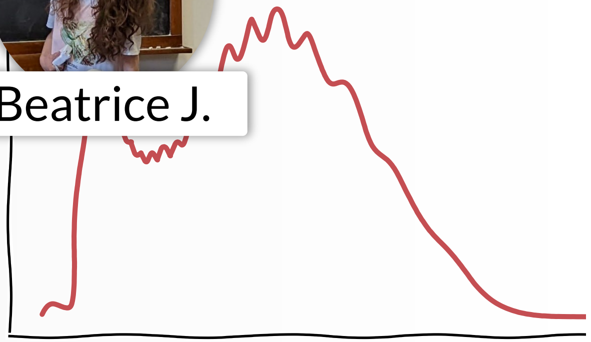
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response



Data processing

- Energy reconstruction
- Event selection

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

Fitter

- Minuit
- Markov Chains MC
- Nested sampling

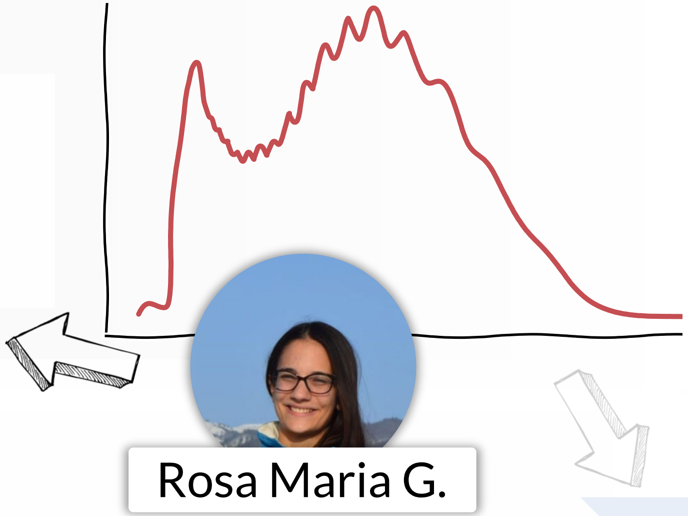
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- **Detector response**

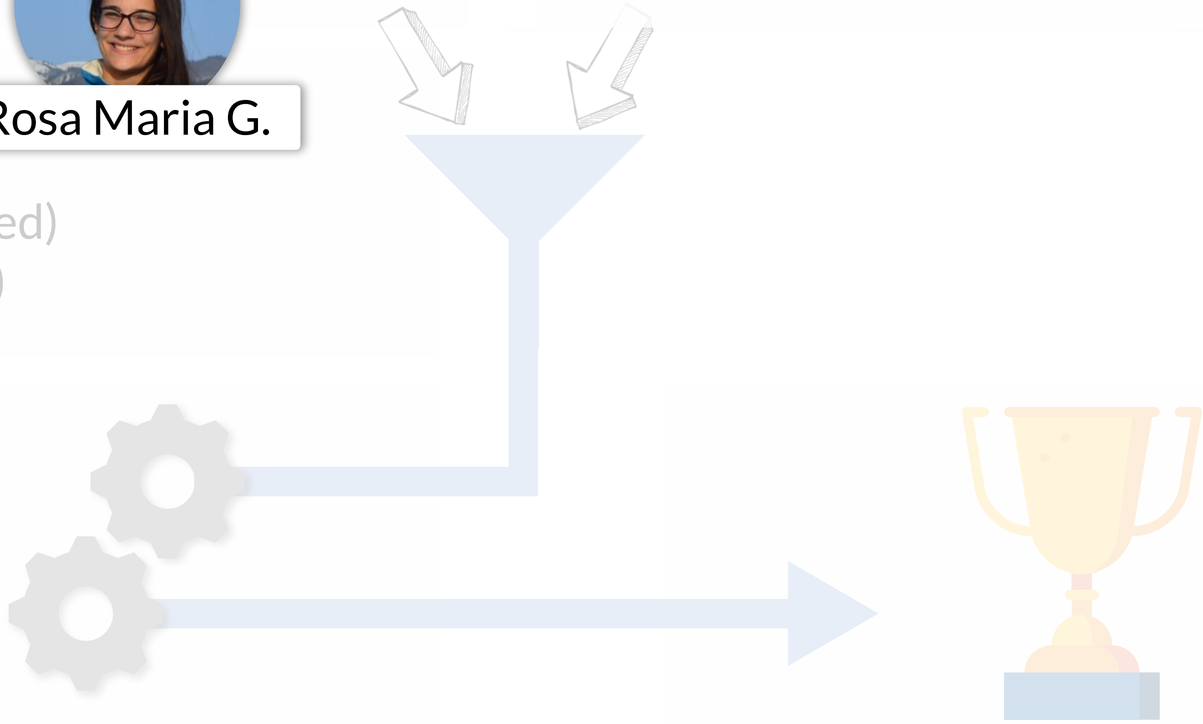


Data processing

- Energy reconstruction
- Event selection

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)



Fitter

- Minuit
- Markov Chains MC
- Nested sampling

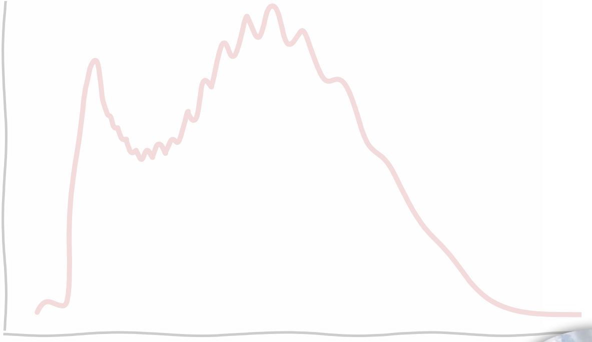
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response



Data processing

- Energy reconstruction
- Event selection

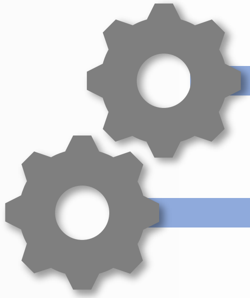
Andrea S.

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

Fitter

- Minuit
- Markov Chains MC
- Nested sampling



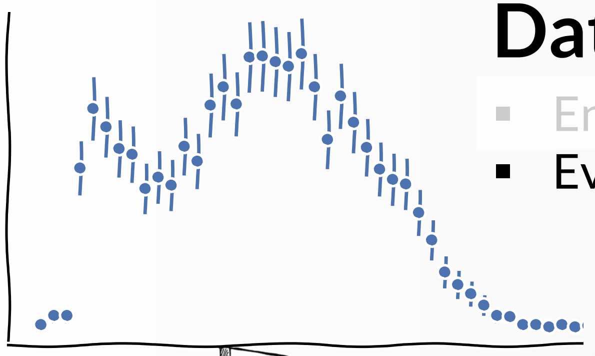
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

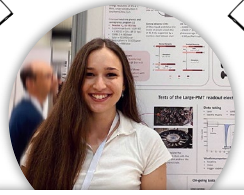
Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response



Data processing

- Energy reconstruction
- Event selection



Vanessa C.

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

Fitter

- Minuit
- Markov Chains MC
- Nested sampling

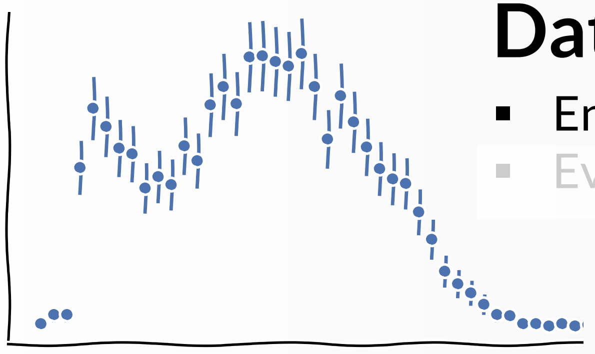
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response



Data processing

- Energy reconstruction
- Event selection

Arsenii G.

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

Fitter

- Minuit
- Markov Chains MC
- Nested sampling

Results

- Best fit values
- Posteriors
- Correlations
- Model selection

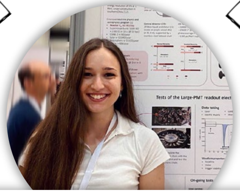
ORSA (Oscillated Reactor Spectrum Analysis)

Model
- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response

Beatrice J.

Rosa Maria G.

Andrea S.



Vanessa C.

Arsenii G.

Data processing
- Energy reconstruction
- Event selection

Cost function
- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

Fitter
- Minuit
- Markov Chains MC
- Nested sampling

Results
- Best fit values
- Posteriors
- Correlations
- Model selection

ORSA (Oscillated Reactor Spectrum Analysis)

Model

- Reactor model
- Backgrounds model
- Covariances
- SNIPER
- Detector response

Data processing

- Energy reconstruction
- Event selection

Cost function

- χ^2 (Pearson, Neyman, combined)
- Likelihood (binned, extended)

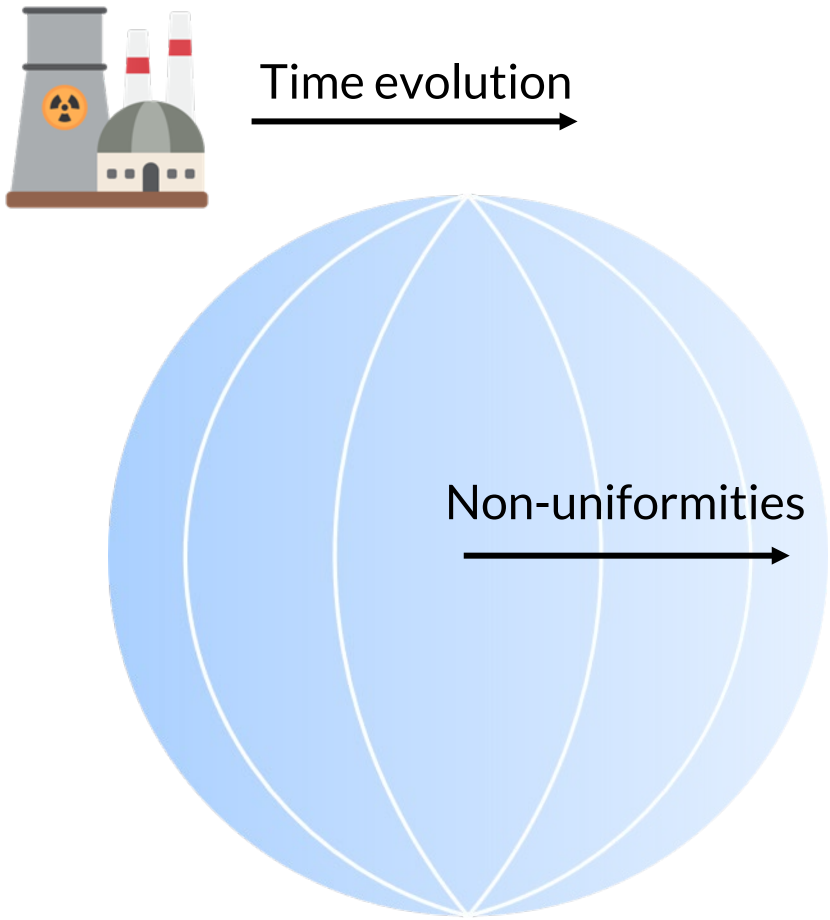
Fitter

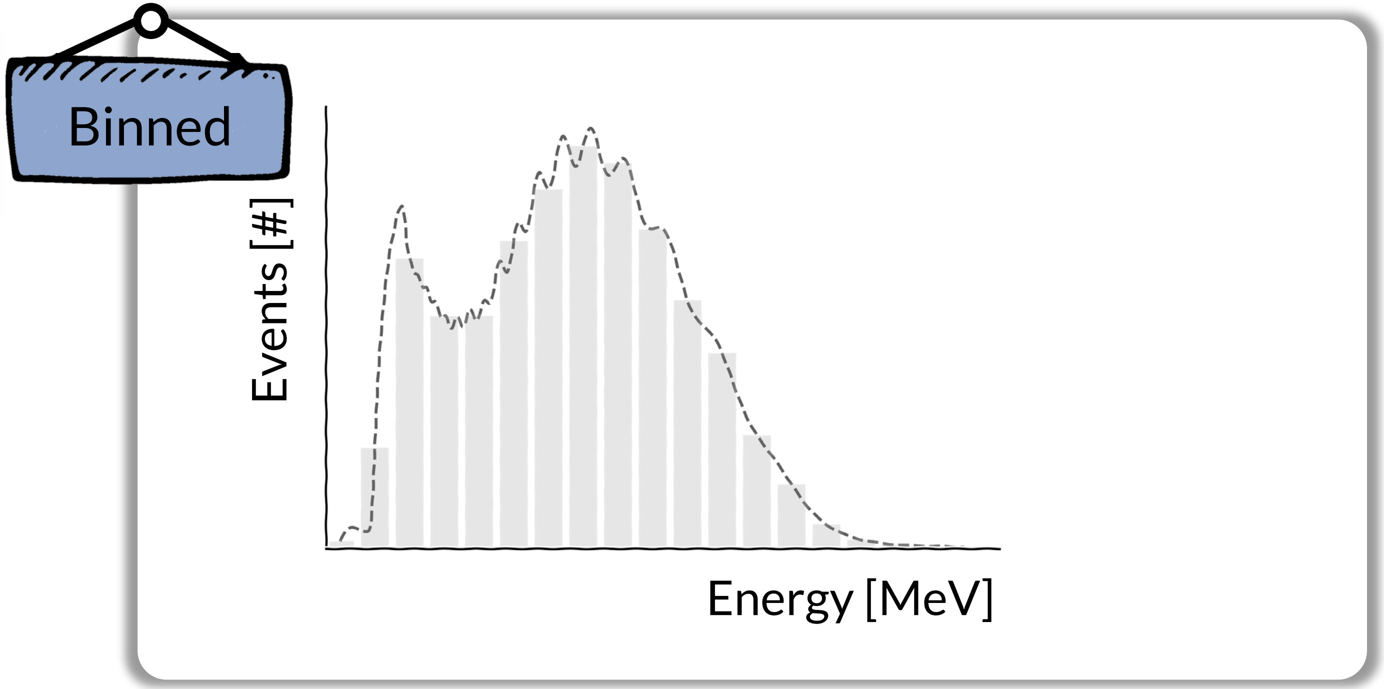
- Minuit
- Markov Chains MC
- Nested sampling

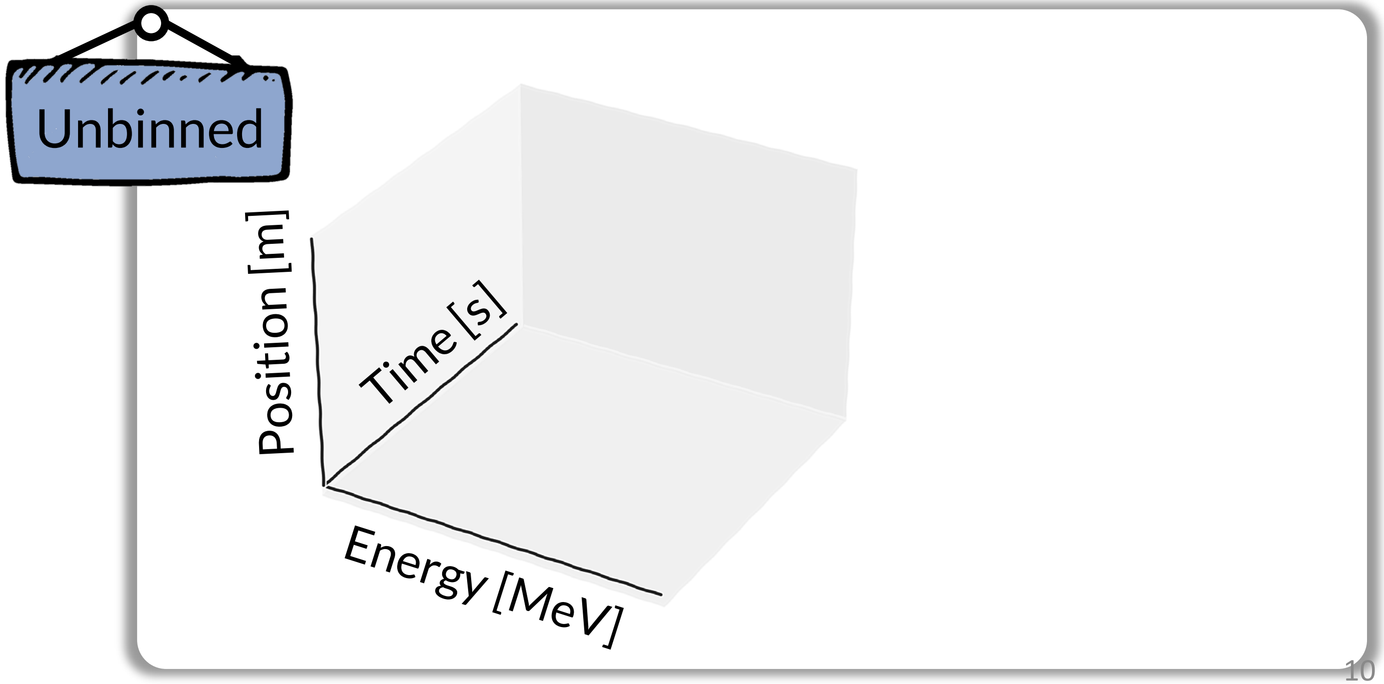
Results

- Best fit values
- Posteriors
- Correlations
- Model selection

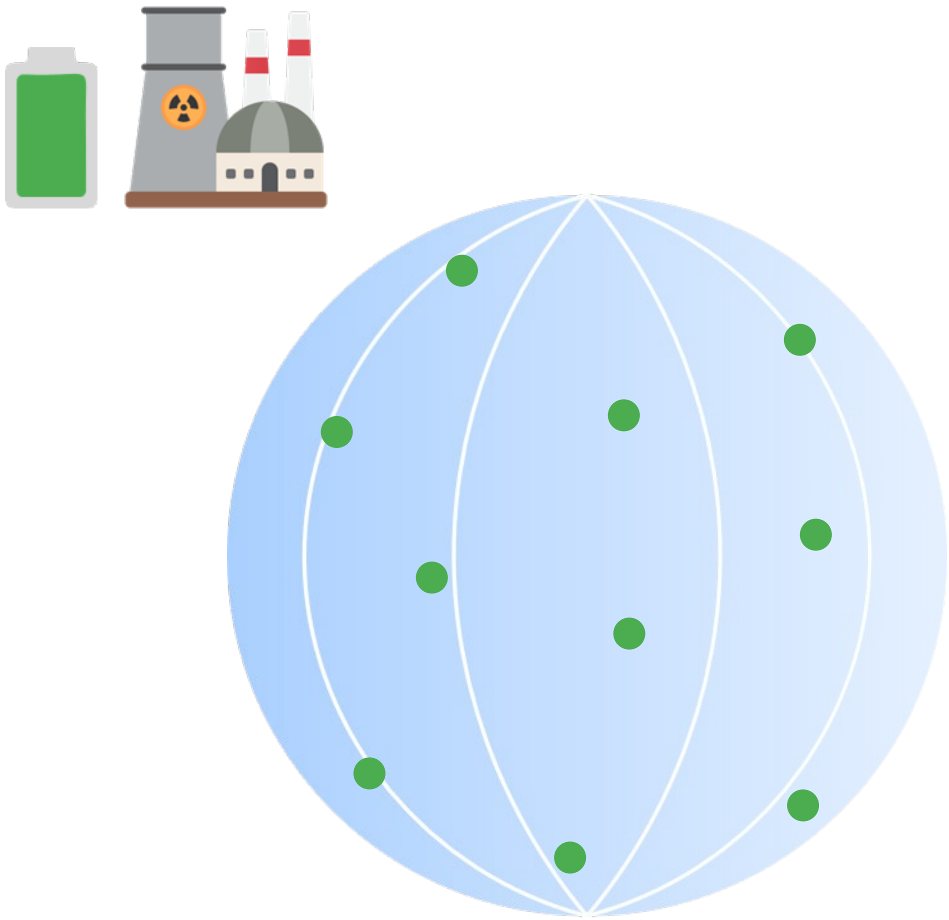
Why unbinned?

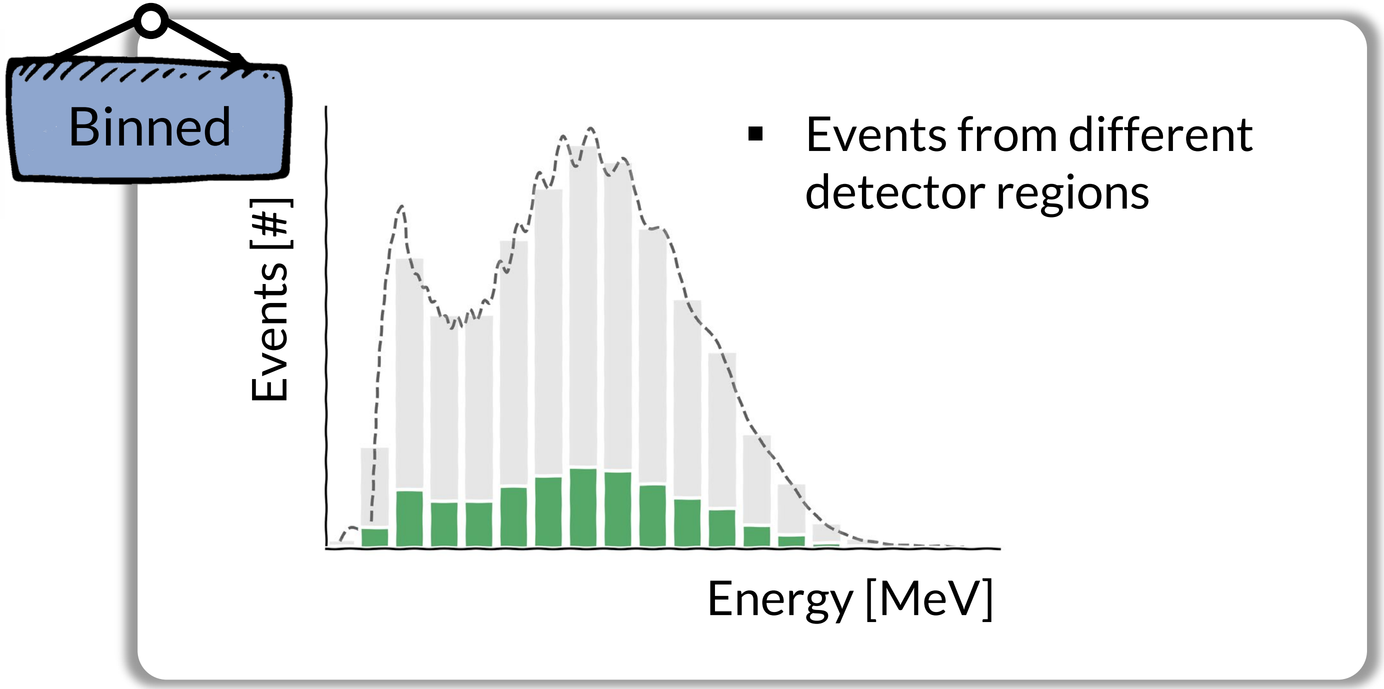




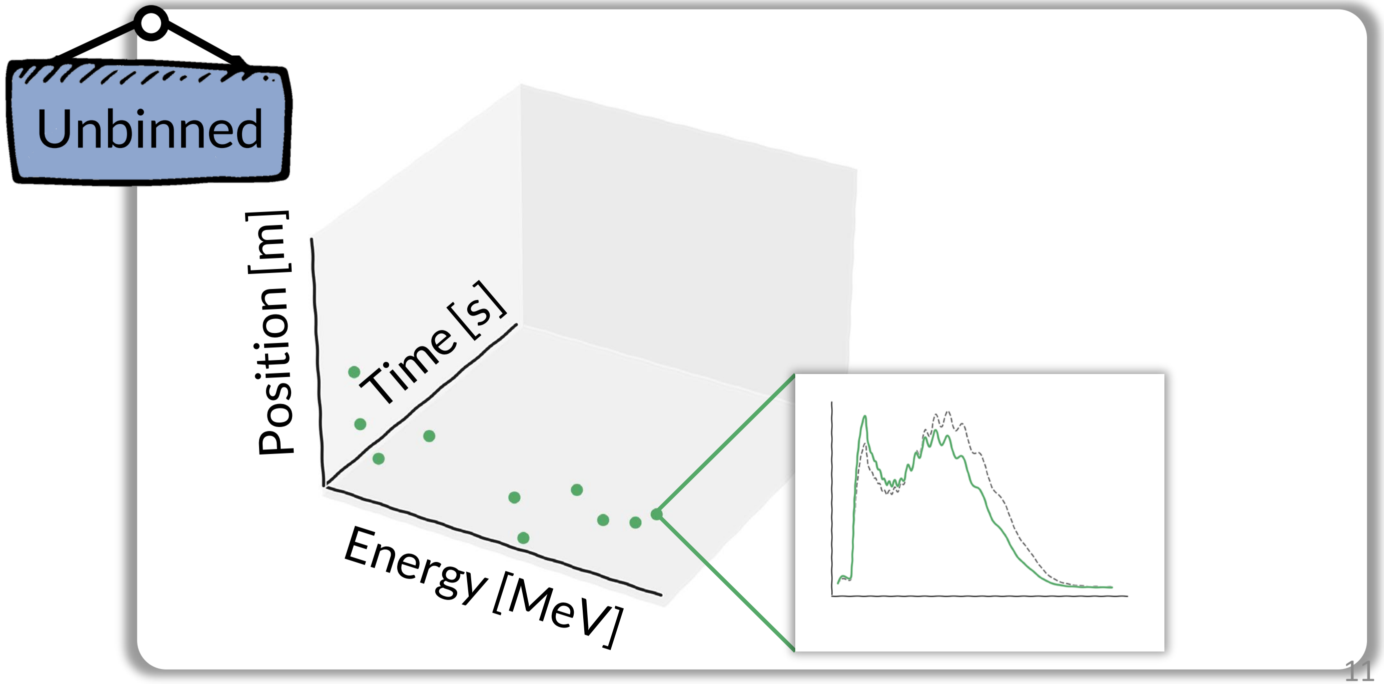


Why unbinned?

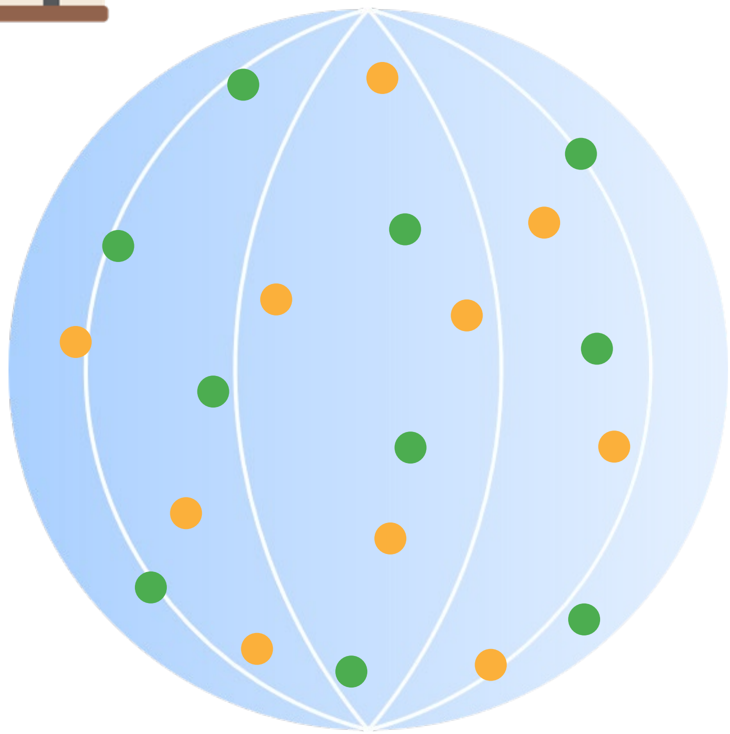


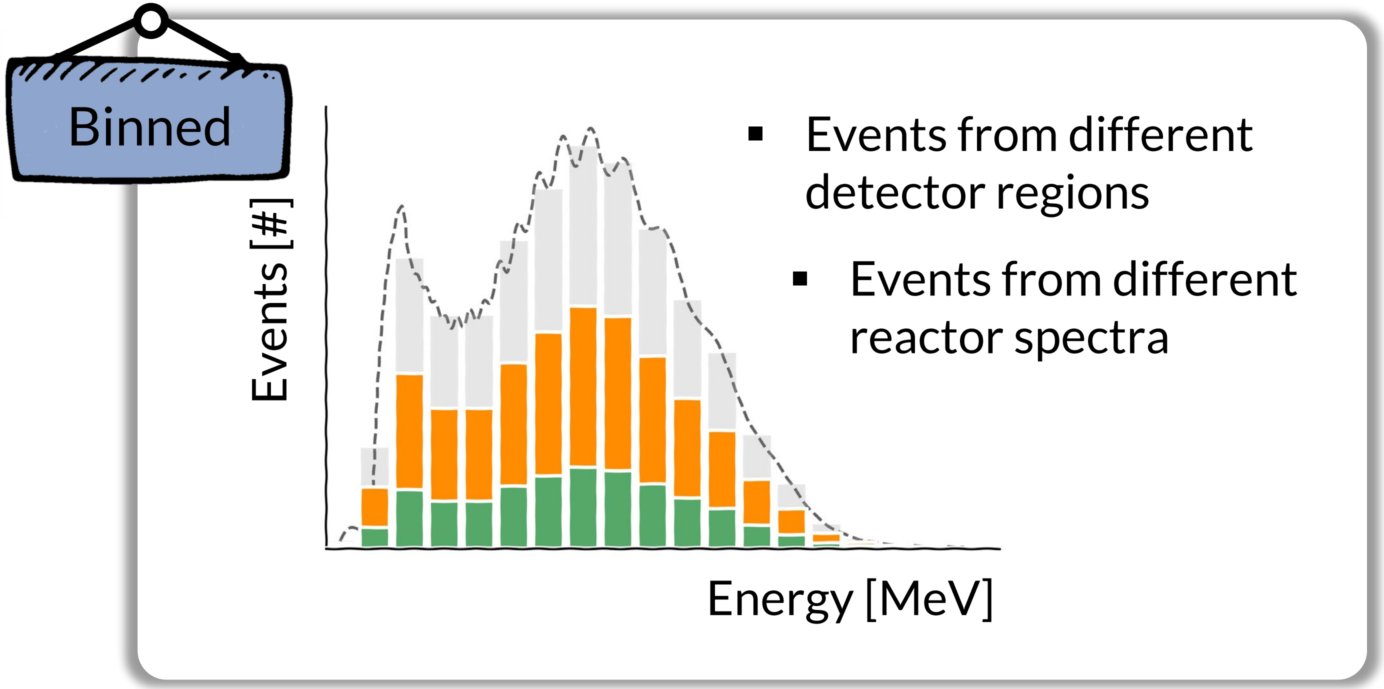


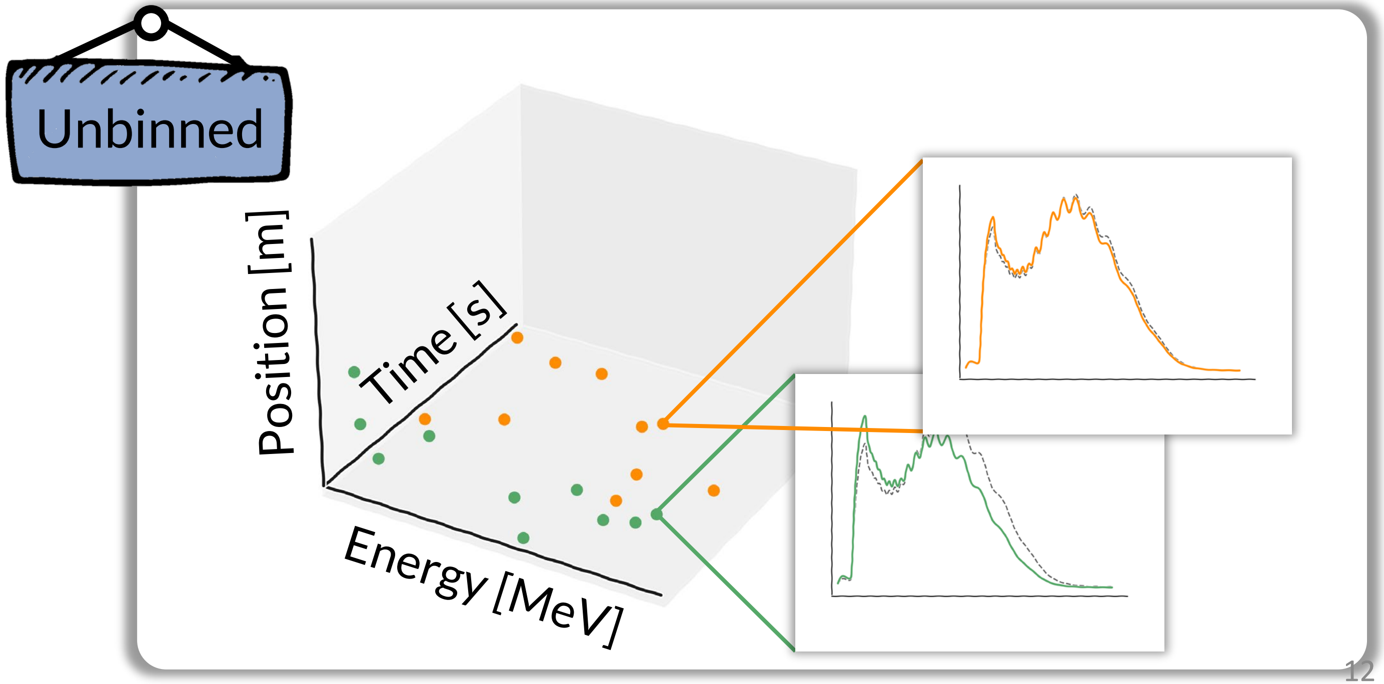
- Events from different detector regions



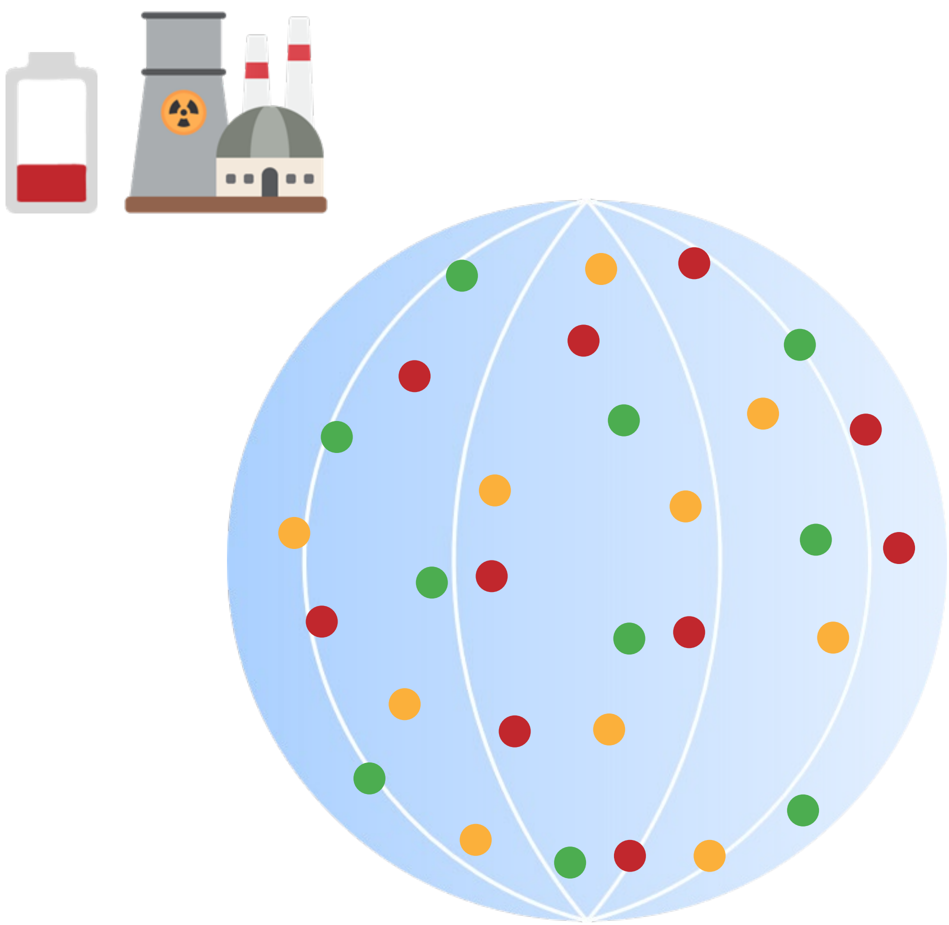
Why unbinned?

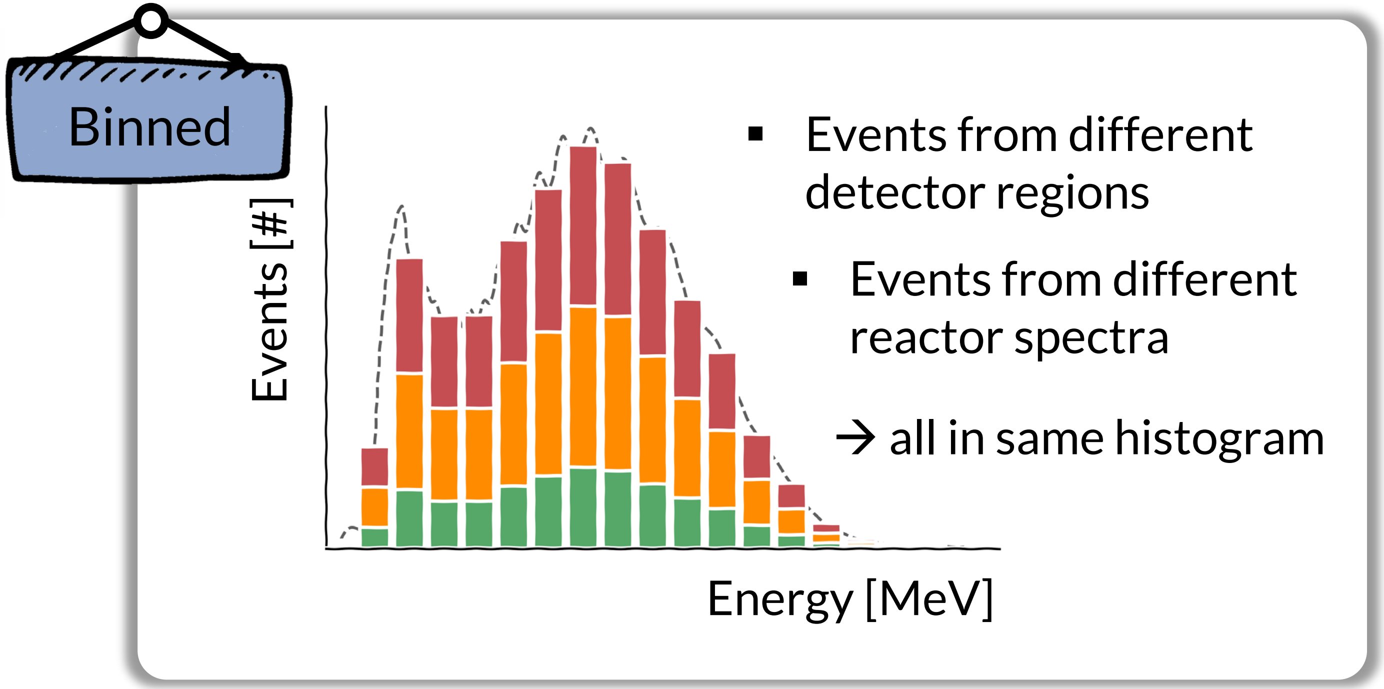






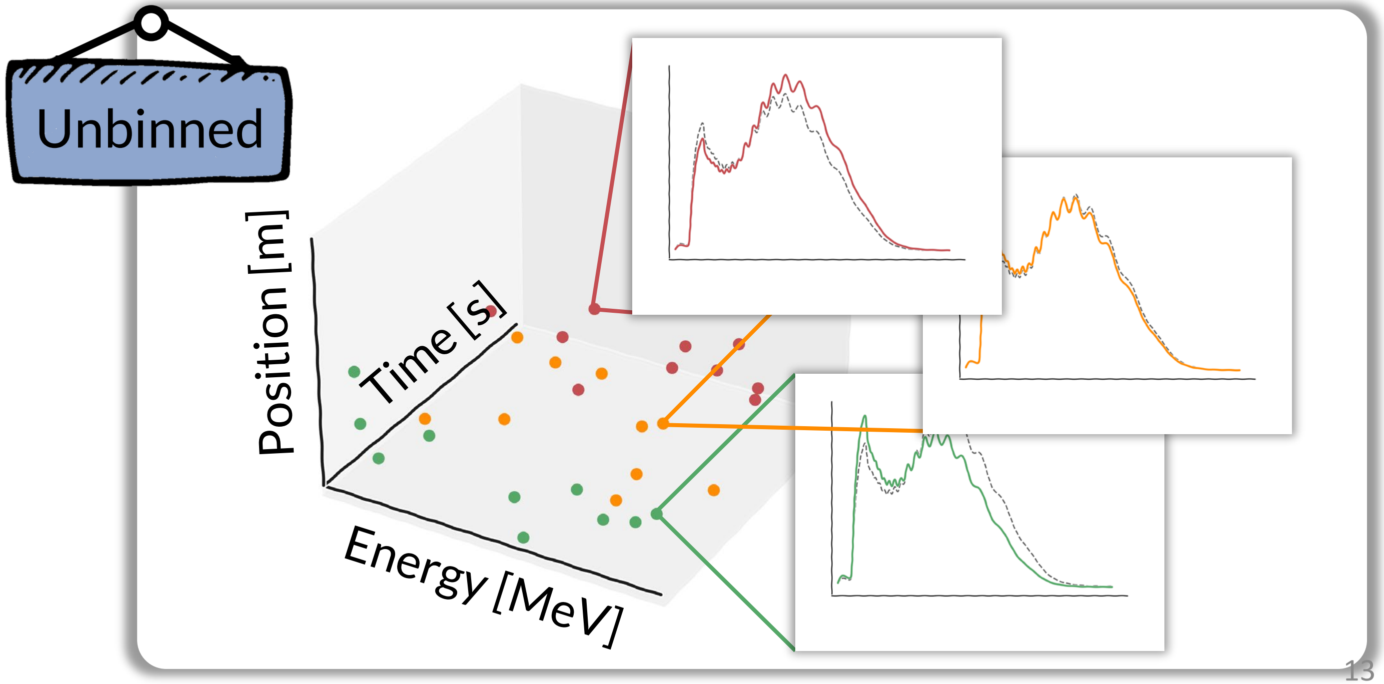
Why unbinned?





- Events from different detector regions
- Events from different reactor spectra

→ all in same histogram



Unbinned likelihood calculation

Poisson term — **Spectrum PDF**

$$\mathcal{L}(\theta) = P(N_{evts}|\theta) \cdot \prod_{i}^{N_{evts}} P(E_i, \bar{r}_i, t_i|\theta)$$

- Reactor antineutrinos
- Geoneutrinos U/Th
- Cosmogenic $^{9}Li/^{8}He$
- Fast neutrons
- α/n reactions
- Accidental coincidences
- Correlated radio events (e.g. Bi-Po)
- Global reactor antineutrinos
- Atmospheric (anti)neutrinos

Unbinned likelihood calculation

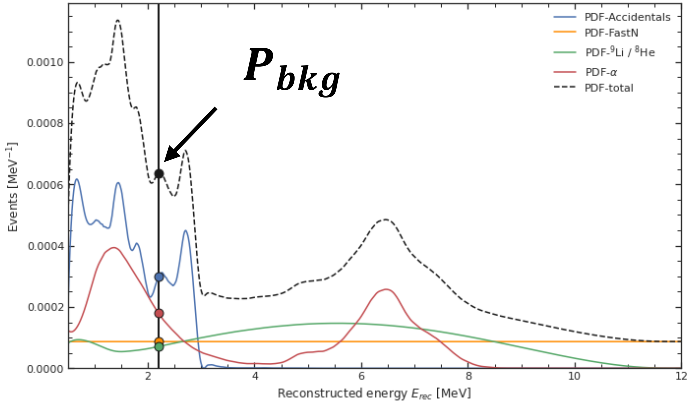
Poisson term

Spectrum PDF

$$\mathcal{L}(\theta) = P(N_{evts}|\theta) \cdot \prod_{i}^{N_{evts}} P(E_i, \bar{r}_i, t_i|\theta)$$

- Reactor antineutrinos
- Geoneutrinos U/Th
- Cosmogenic $^{9}Li/^{8}He$
- Fast neutrons
- α/n reactions
- Accidental coincidences
- Correlated radio events (e.g. Bi-Po)
- Global reactor antineutrinos
- Atmospheric (anti)neutrinos

Reconstructed energy E_{rec} $P_{rec}(E_i, \bar{r}_i, t_i|\theta) = f_{rec}(E_i, \bar{r}_i, t_i|\theta)$



Unbinned likelihood calculation

Poisson term ↑ **Spectrum PDF** ↑

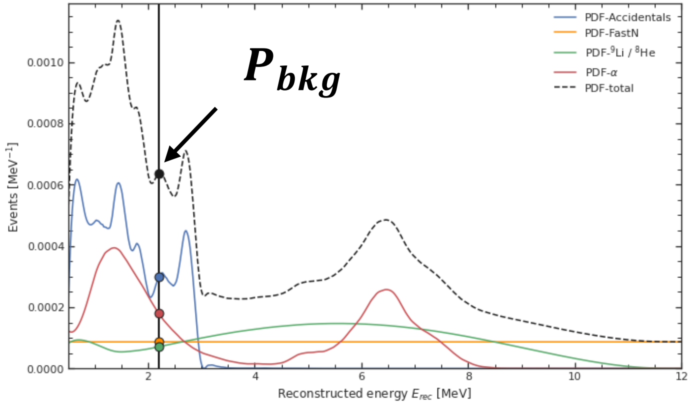
$$\mathcal{L}(\theta) = P(N_{evts}|\theta) \cdot \prod_{i}^{N_{evts}} P(E_i, \bar{r}_i, t_i|\theta)$$

- Reactor antineutrinos
- Geoneutrinos U/Th
- Cosmogenic $^9Li/^8He$
- Fast neutrons
- α/n reactions
- Accidental coincidences
- Correlated radio events (e.g. Bi-Po)
- Global reactor antineutrinos
- Atmospheric (anti)neutrinos

Reconstructed energy E_{rec} $\quad P_{rec}(E_i, \bar{r}_i, t_i|\theta) = f_{rec}(E_i, \bar{r}_i, t_i|\theta)$

Energy resolution ***G***

Visible energy E_{vis} $\quad P_{vis}(E_i, \bar{r}_i, t_i|\theta) = \int f_{vis}(E_{vis}, \bar{r}_i, t_i|\theta) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$



Unbinned likelihood calculation

Poisson term ↑ **Spectrum PDF** ↑

$$\mathcal{L}(\theta) = P(N_{evts}|\theta) \cdot \prod_{i}^{N_{evts}} P(E_i, \bar{r}_i, t_i|\theta)$$

- Reactor antineutrinos
- Geoneutrinos U/Th
- Cosmogenic $^{9}Li/^{8}He$
- Fast neutrons
- α/n reactions
- Accidental coincidences
- Correlated radio events (e.g. Bi-Po)
- Global reactor antineutrinos
- Atmospheric (anti)neutrinos

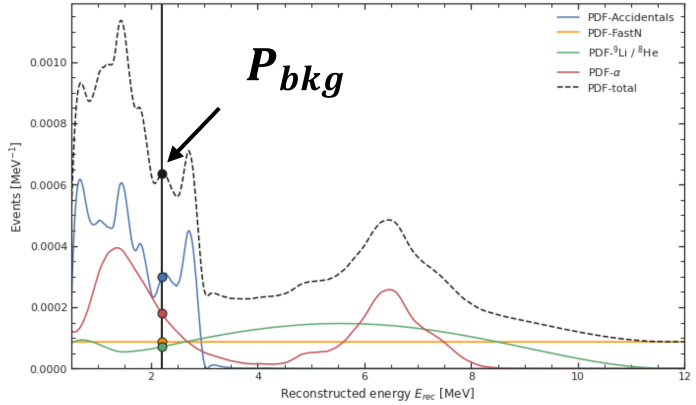
Reconstructed energy E_{rec} $$P_{rec}(E_i, \bar{r}_i, t_i|\theta) = f_{rec}(E_i, \bar{r}_i, t_i|\theta)$$

Energy resolution ***G***

Visible energy E_{vis} $$P_{vis}(E_i, \bar{r}_i, t_i|\theta) = \int f_{vis}(E_{vis}, \bar{r}_i, t_i|\theta) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$$

Scintillator nonlinearity ***nl***

Deposited energy E_{dep} $$P_{dep}(E_i, \bar{r}_i, t_i|\theta) = \int f_{dep}\big(nl(E_{dep}), \bar{r}_i, t_i|\theta\big) \cdot G\big(nl(E_{dep}), E_i\big) \cdot dE_{vis}$$



Unbinned likelihood calculation

Poisson term — **Spectrum PDF**

$$\mathcal{L}(\theta) = P(N_{evts}|\theta) \cdot \prod_{i}^{N_{evts}} P(E_i, \bar{r}_i, t_i|\theta)$$

- Reactor antineutrinos
- Geoneutrinos U/Th
- Cosmogenic $^{9}Li/^{8}He$
- Fast neutrons
- α/n reactions
- Accidental coincidences
- Correlated radio events (e.g. Bi-Po)
- Global reactor antineutrinos
- Atmospheric (anti)neutrinos

Reconstructed energy E_{rec} — $P_{rec}(E_i, \bar{r}_i, t_i|\theta) = f_{rec}(E_i, \bar{r}_i, t_i|\theta)$

Energy resolution ***G***

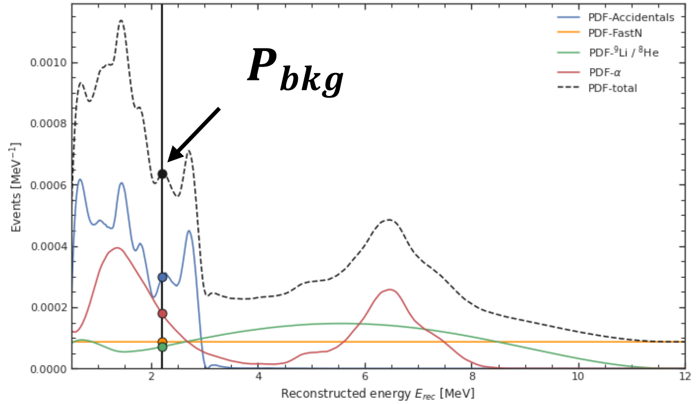
Visible energy E_{vis} — $P_{vis}(E_i, \bar{r}_i, t_i|\theta) = \int f_{vis}(E_{vis}, \bar{r}_i, t_i|\theta) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$

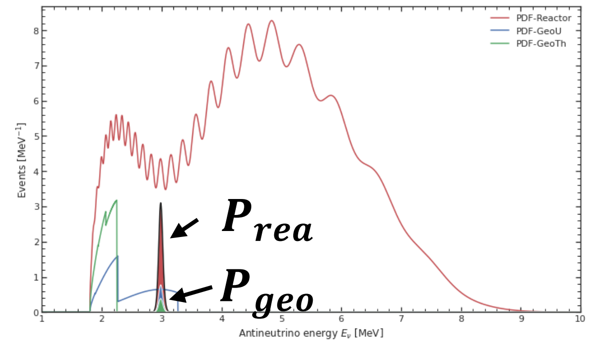
Scintillator nonlinearity ***nl***

Deposited energy E_{dep} — $P_{dep}(E_i, \bar{r}_i, t_i|\theta) = \int f_{dep}(nl(E_{dep}), \bar{r}_i, t_i|\theta) \cdot G(nl(E_{dep}), E_i) \cdot dE_{vis}$

IBD cross section ***xs***

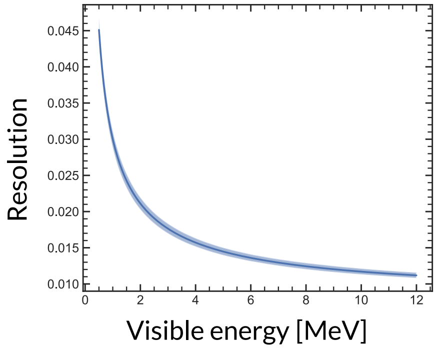
Antineutrino energy $E_{\bar{\nu}}$ — $P_{\bar{\nu}}(E_i, \bar{r}_i, t_i|\theta) = \int f_{\bar{\nu}}(nl(xs(E_{\bar{\nu}})), \bar{r}_i, t_i|\theta) \cdot G(nl(xs(E_{\bar{\nu}})), E_i) \cdot dE_{vis}$

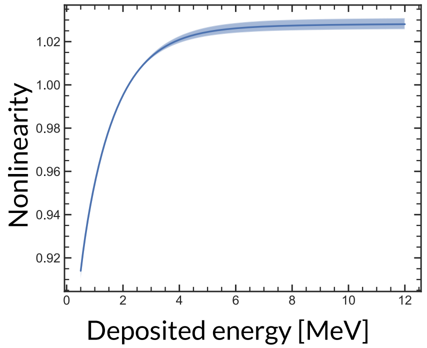




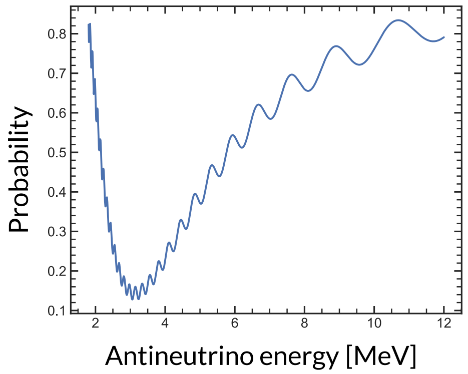
Our model

Detector response:
- IBD cross section (w/ recoil) (fixed)
- Energy resolution
- Liquid scintillator nonlinearity

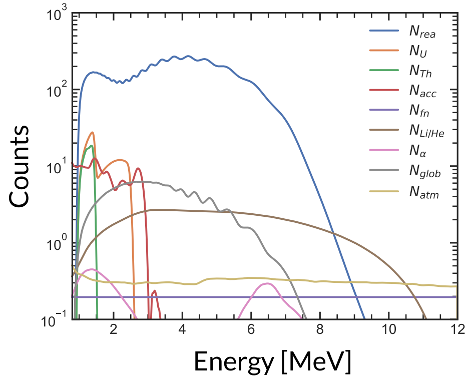




Neutrino oscillations:
- Matter density (fixed)
- Oscillation parameters



Signal components:
- PDFs (fixed)
- Normalizations



Resolution (0–2) | Non-linearity (3–6) | Osc. parameters (7–10) | Backgrounds (12–19)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
a	b	c	p_0	p_1	p_2	p_3	dm^2_{21}	dm^2_{31}	s^2_{12}	s^2_{13}	N_{rea}	N_U^{geo}	N_{Th}^{geo}	N_{acc}	N_{fn}	N_{lihe}	N_{α}	N_{glob}	N_{atm}

Constrained with prior at 0.5% (0–6) | Constrained with DayaBay prior (10) | Constrained with prior at 5% (14–19)

Unbinned likelihood numerical computation

In theory:

$$\mathcal{L}(\theta) \cong \prod_{i}^{N_{evts}} \int P(E_{vis}) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$$

Unbinned likelihood numerical computation

In theory:

$$\mathcal{L}(\theta) \cong \prod_{i}^{N_{evts}} \int P(E_{vis}) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$$

In practice:

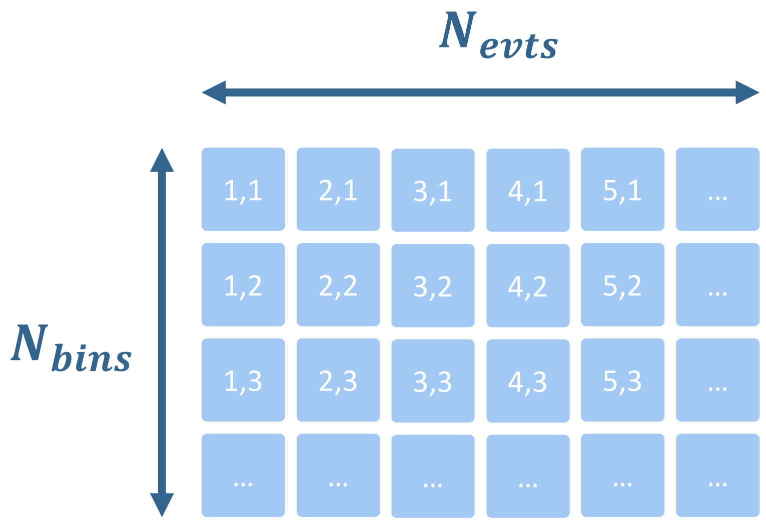
- $\mathcal{L}(\theta)$ → $\log \mathcal{L}(\theta)$ (Π → Σ)
- $\int$ → Σ

$$\log \mathcal{L}(\theta) \cong \sum_{i}^{N_{evts}} \log\left(\sum_{j}^{N_{bins}} P\left(E_j\right) \cdot G\left(E_j, E_i\right) \cdot \Delta E_j\right) = \sum_{i}^{N_{evts}} \log \sum_{j}^{N_{bins}} C_{i,j}$$

Unbinned likelihood numerical computation

In theory:

$$\mathcal{L}(\theta) \cong \prod_{i}^{N_{evts}} \int P(E_{vis}) \cdot G(E_{vis}, E_i) \cdot dE_{vis}$$



In practice:

- $\mathcal{L}(\theta) \rightarrow \log \mathcal{L}(\theta) \quad (\Pi \rightarrow \Sigma)$
- $\int \rightarrow \Sigma$

$$\log \mathcal{L}(\theta) \cong \sum_{i}^{N_{evts}} \log\left(\sum_{j}^{N_{bins}} P(E_j) \cdot G(E_j, E_i) \cdot \Delta E_j\right) = \sum_{i}^{N_{evts}} \log \sum_{j}^{N_{bins}} C_{i,j}$$

Complexity scales with $N_{evts} \times N_{bins}$ **!**

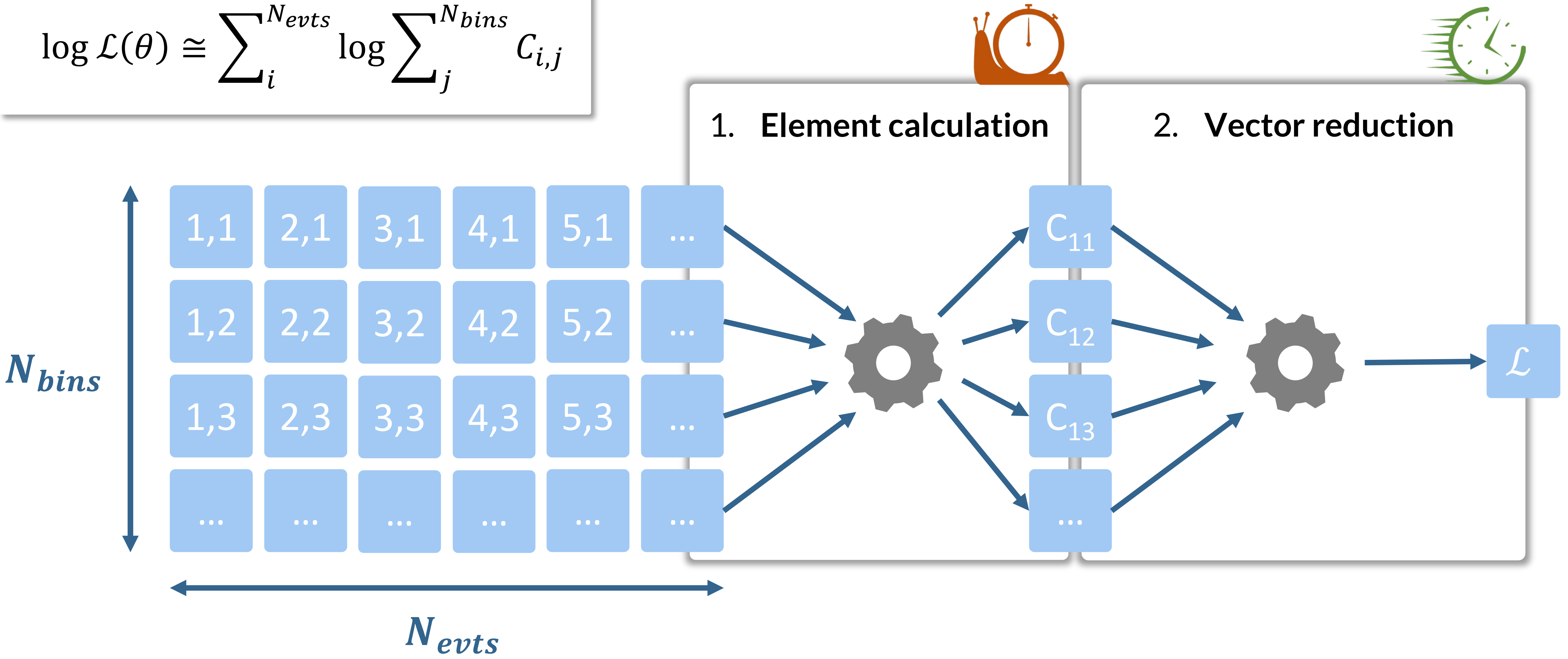
But the **computation** of the $C_{i,j}$ is **independent** → "**embarrassingly parallel**" computation

Implementing parallelization in likelihood calculation

Single-thread

$$\log \mathcal{L}(\theta) \cong \sum_{i}^{N_{evts}} \log \sum_{j}^{N_{bins}} C_{i,j}$$

$$\Delta t = \Delta t_1 + \Delta t_2$$

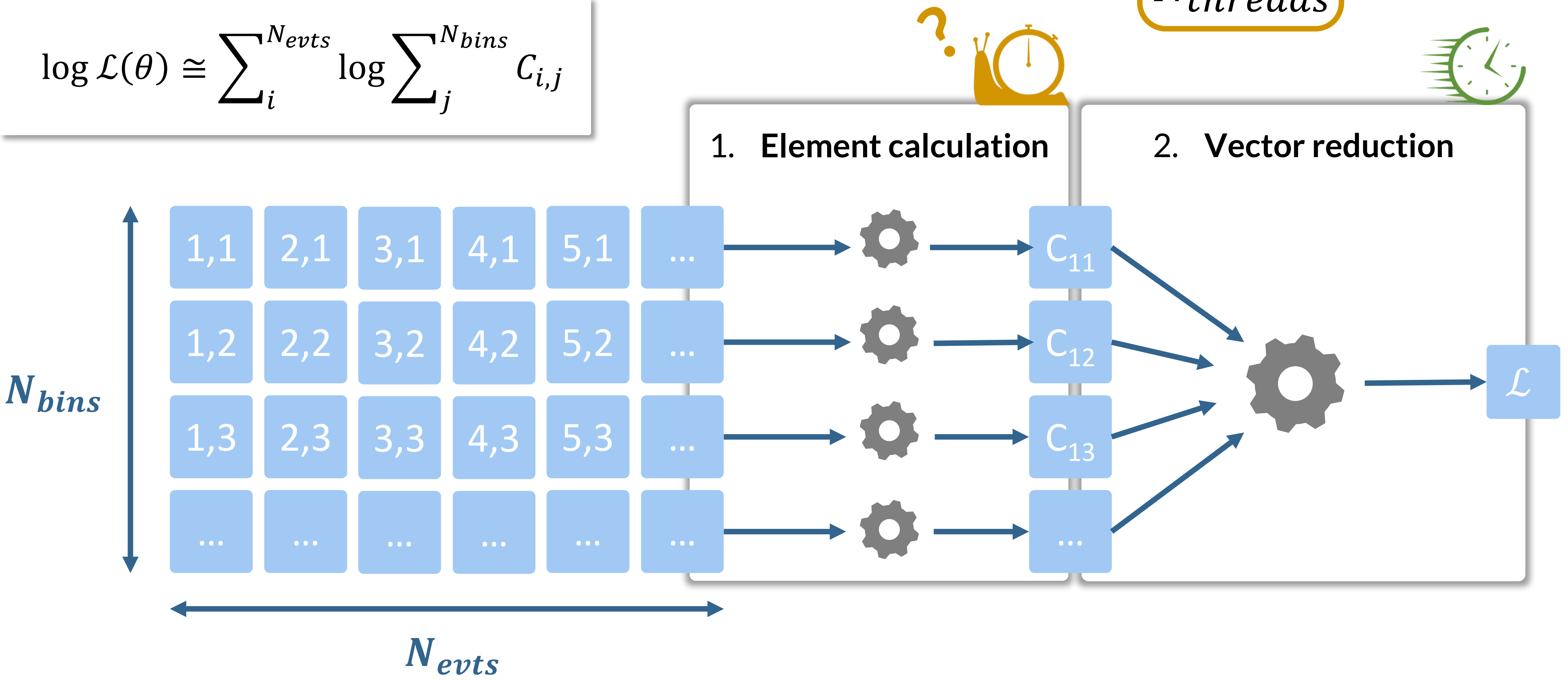


Implementing parallelization in likelihood calculation

Multi-thread

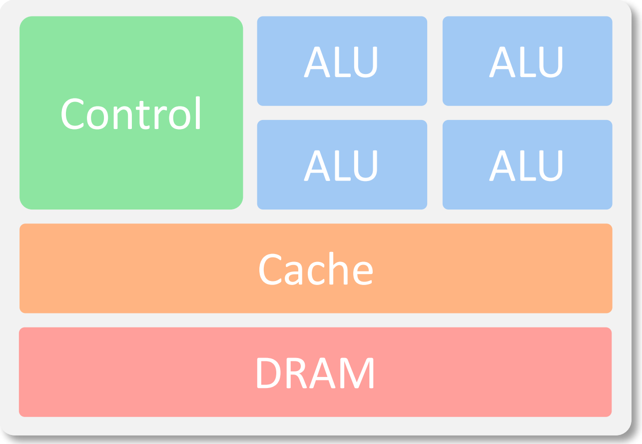
$$\Delta t = \frac{\Delta t_1}{N_{threads}} + \Delta t_2$$

$$\log \mathcal{L}(\theta) \cong \sum_i^{N_{evts}} \log \sum_j^{N_{bins}} C_{i,j}$$



GPU accelerators: taking parallelization to the next level

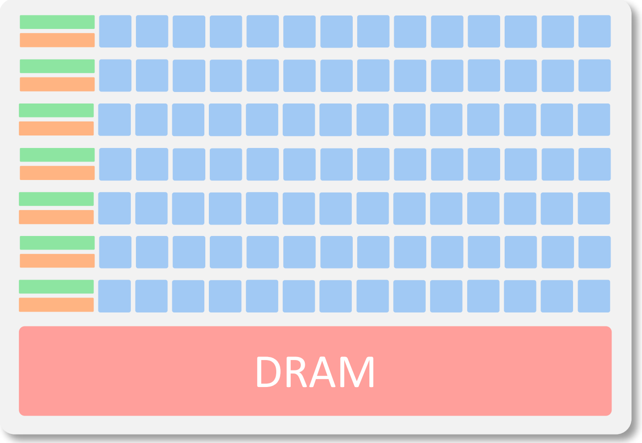
CPU



Pros:	**Cons:**
More memory	Less cores
Faster ALUs	

Model: Intel Xeon Gold 6238R CPU
N° of ALU*: 28 cores (56 threads)
Price: 3 k€

GPU



Pros:	**Cons:**
Thousands core	Less memory

Model: NVIDIA Tesla V100 16 GB SXM2
N° of ALU*: 5120 CUDA cores
Price: 14 k€

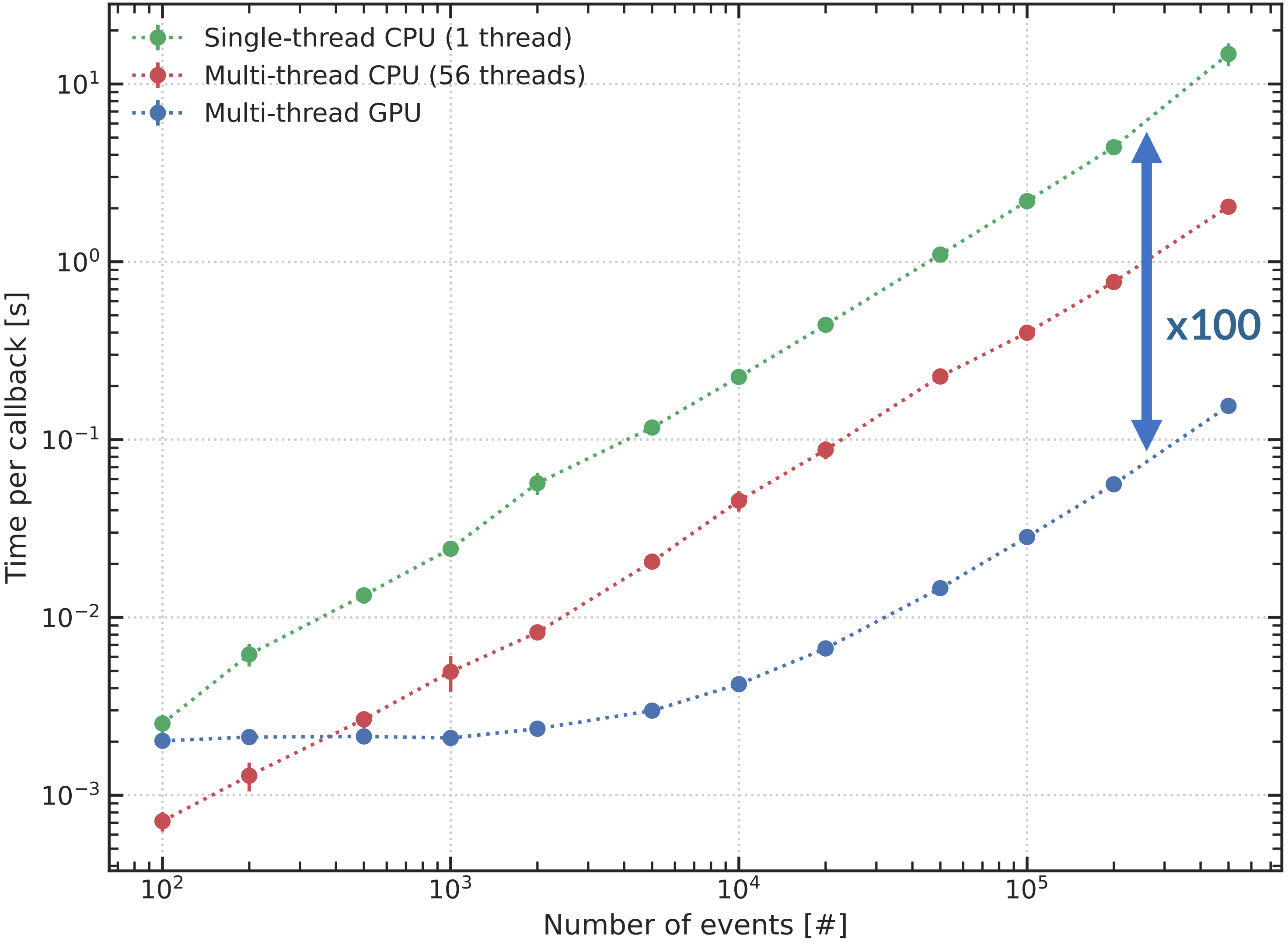
*ALU: arithmetic-logic unit

Performances: call-back time of cost function

Multi-threading implemented in both CPU and GPU!

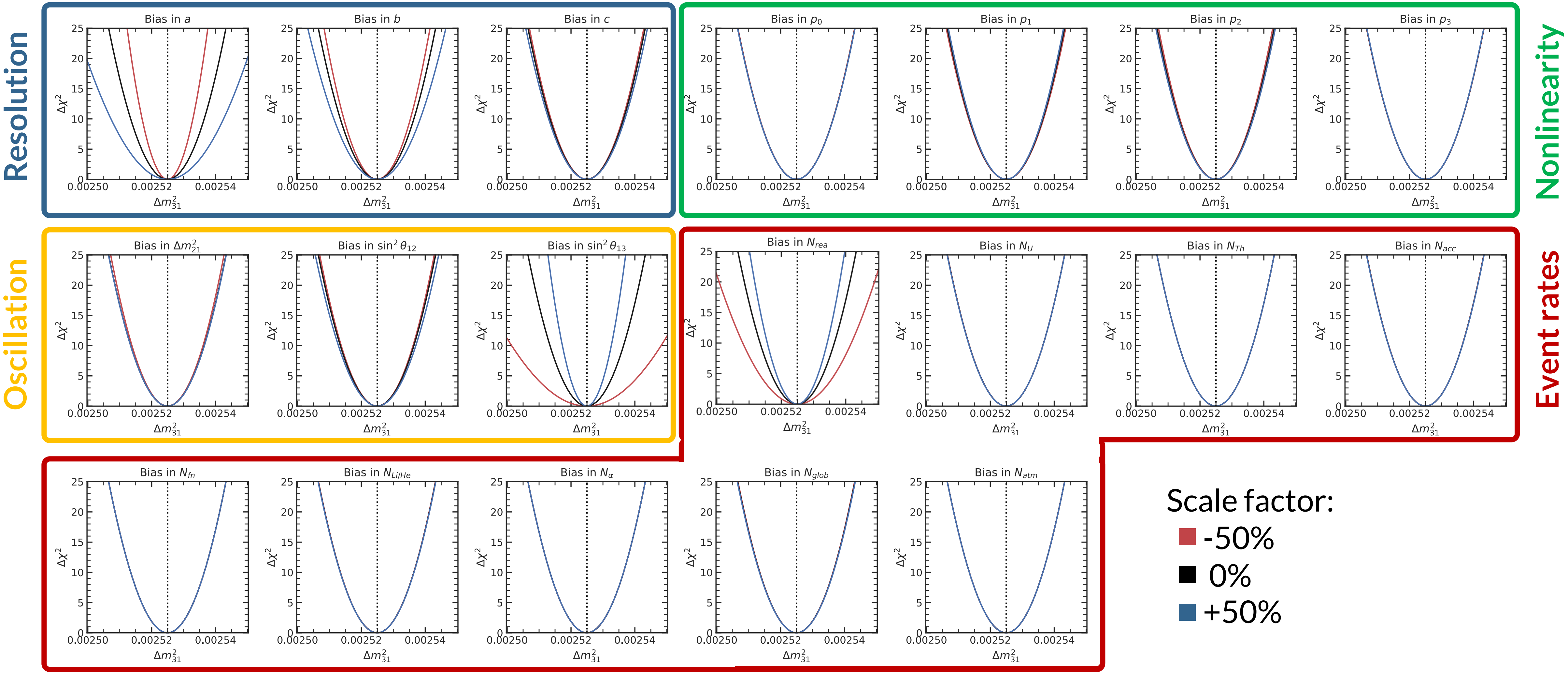
Multi-threading on **GPU** can achieve **performances x100**

– Less time for computations
+ More time for analyses ☺



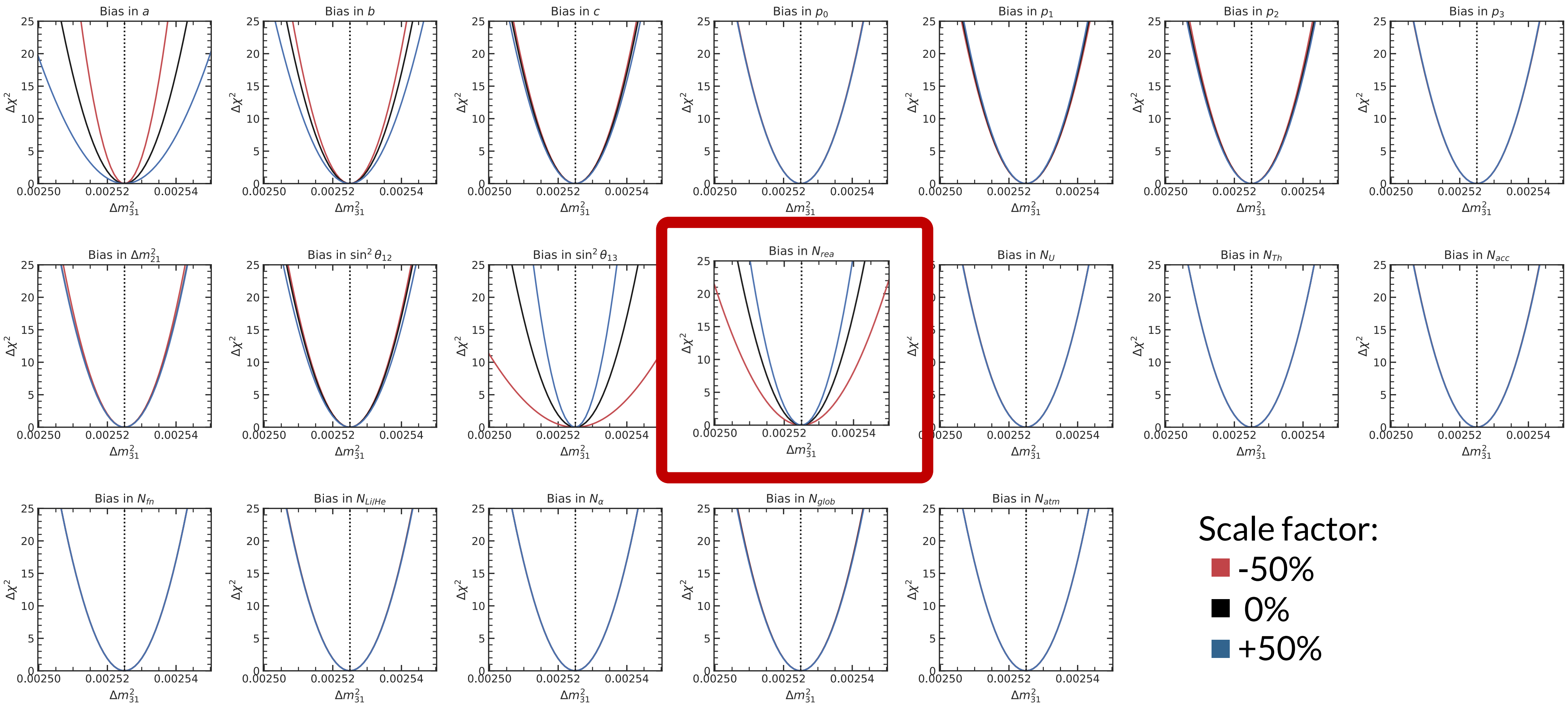
A case study: $\boldsymbol{\Delta m^2_{31}}$ - precision

"What if parameters are different from what we expect? (but we still manage to measure them)" → **precision loss**



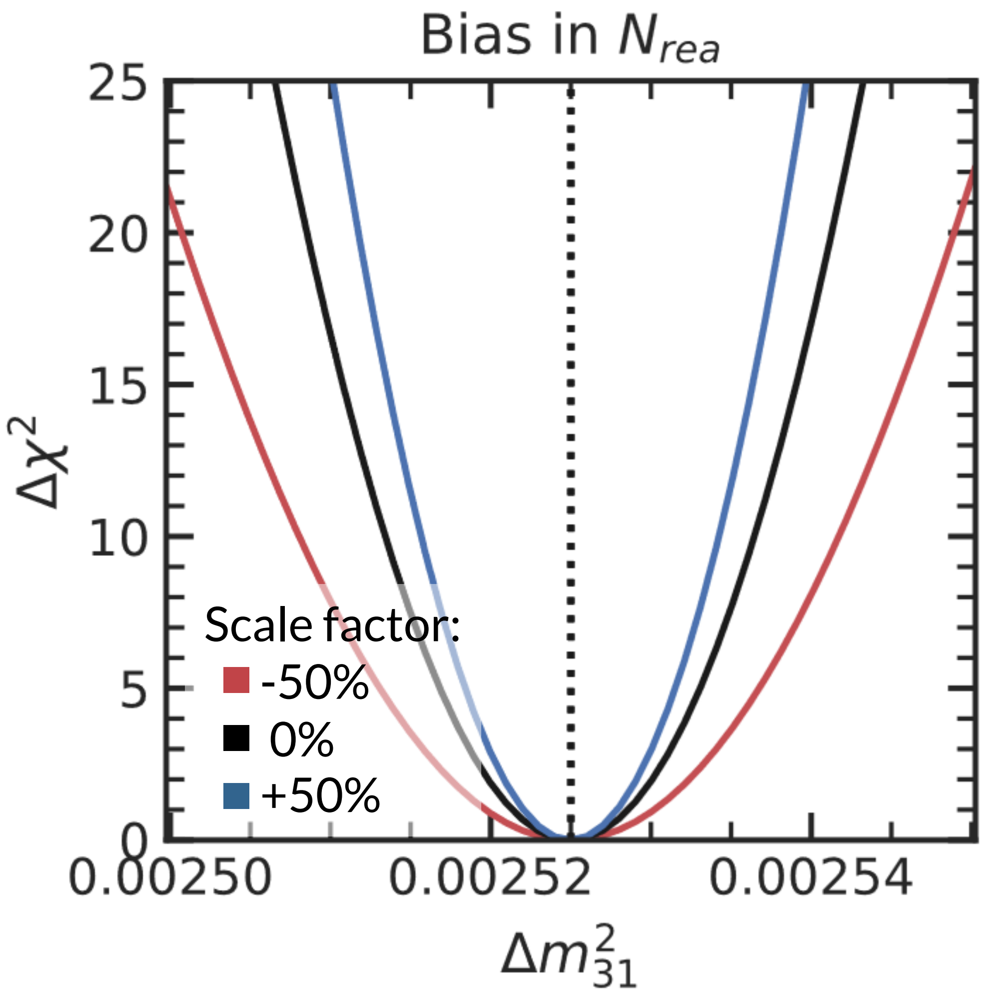
A case study: $\boldsymbol{\Delta m^2_{31}}$ - precision

“What if parameters are different from what we expect? (but we still manage to measure them)” → **precision loss**

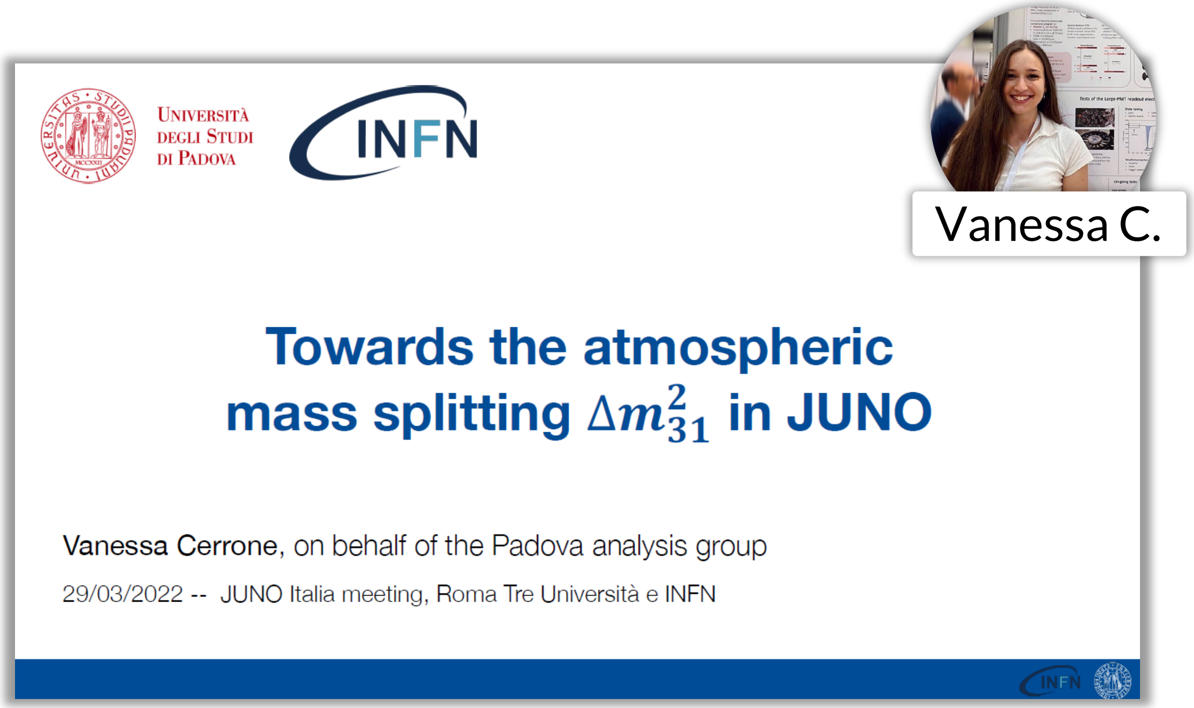


Scale factor:
- -50%
- 0%
- +50%

A case study: Δm^2_{31} - precision

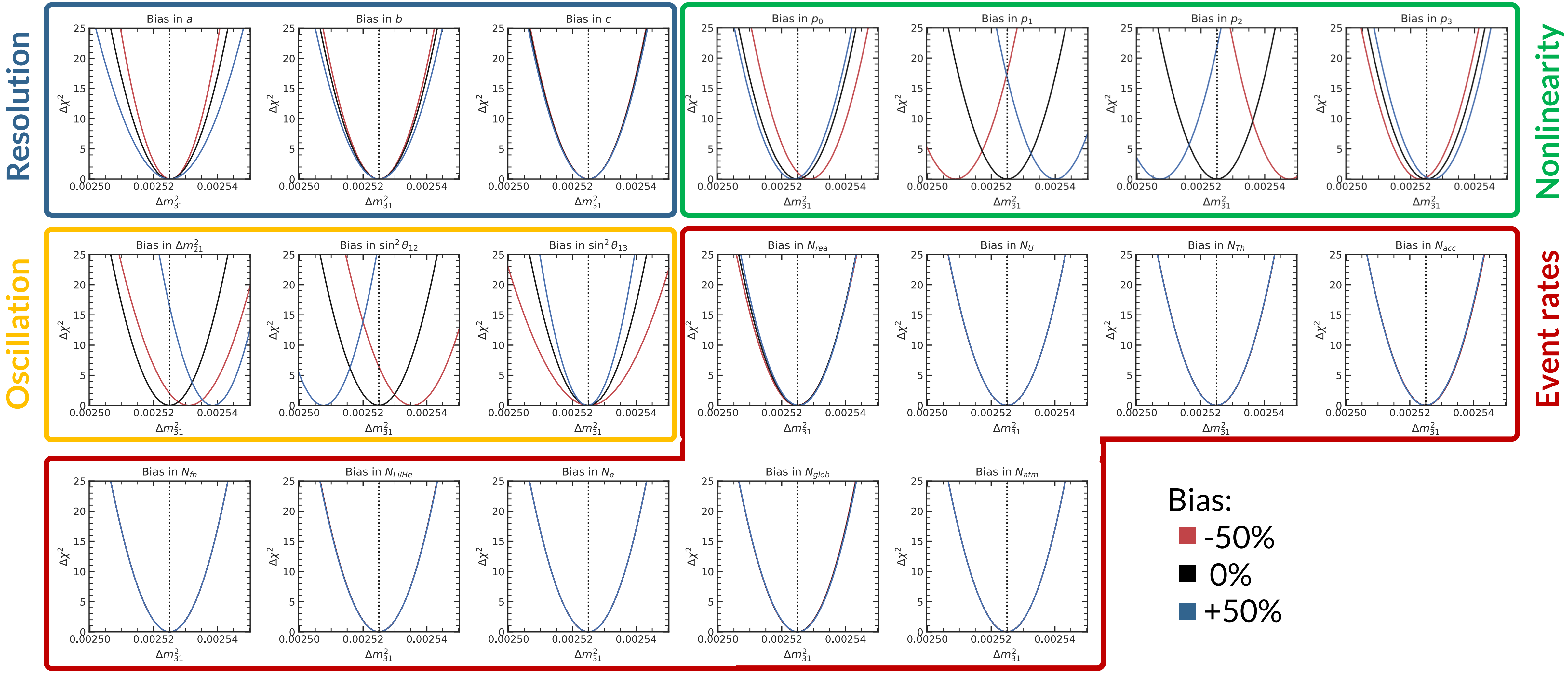


Important to maximize number of **selected events:**



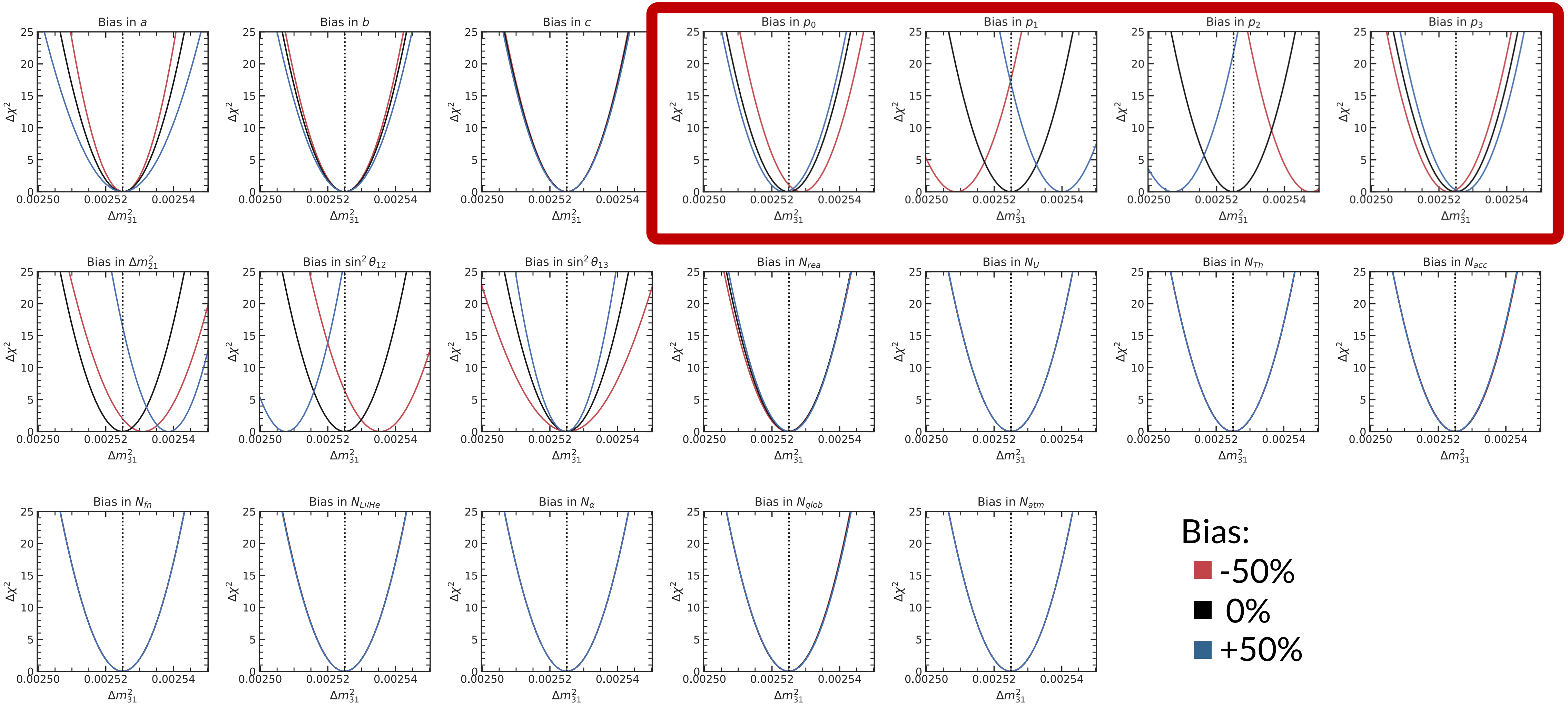
A case study: $\boldsymbol{\Delta m^2_{31}}$ - accuracy

"What if we fail in fitting/estimating one or more parameter?" → **possible bias**, accuracy loss

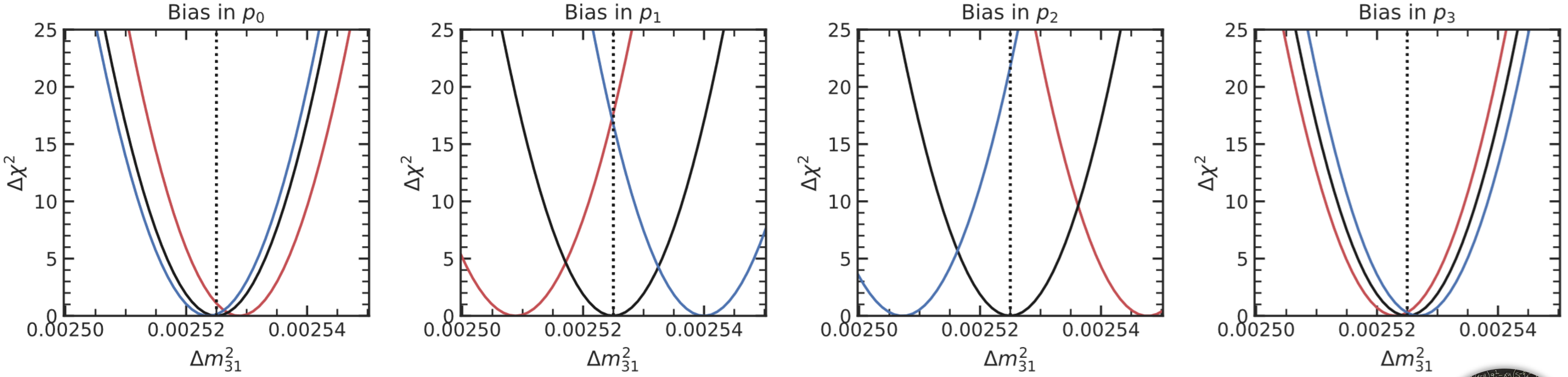


A case study: $\boldsymbol{\Delta m^2_{31}}$ - accuracy

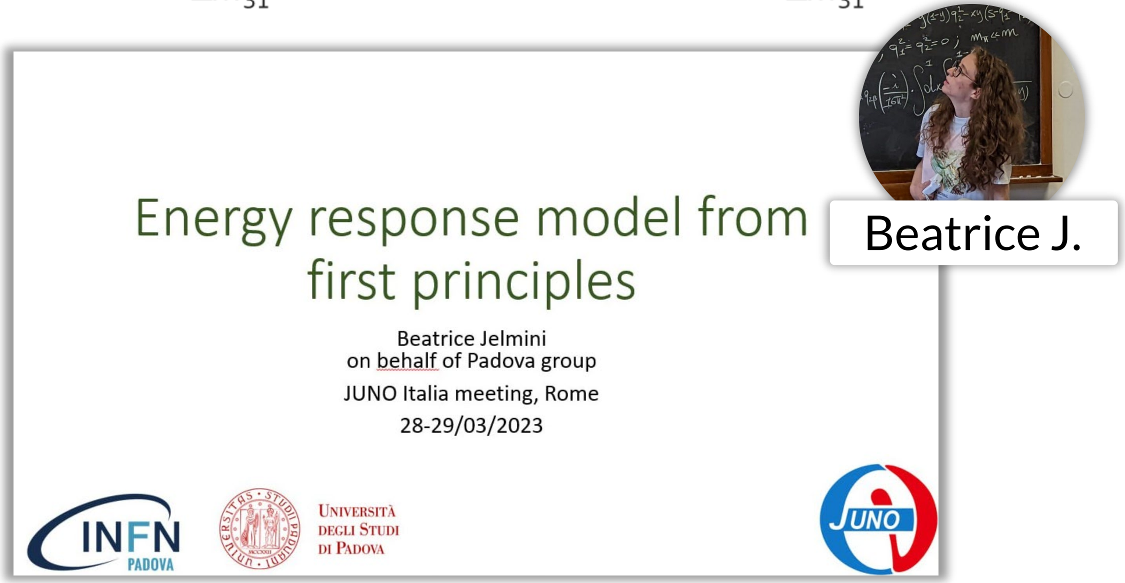
"What if we fail in fitting/estimating one or more parameter?" → **possible bias**, accuracy loss



A case study: Δm^2_{31} - accuracy



Important to understand the **energy scale** of JUNO and to characterize its **nonlinearity**



Final remarks

- **Unbinned likelihood** gives us **more freedom** in treating space/time-dependent effects:
  - **performs comparably to binned likelihood/χ^2** when no additional info is provided
    → we need to test performances when space/time effects are included

- In unbinned likelihood **computation** time **scales** linearly **with** the **number of events**
  - luckily the $\boldsymbol{N_{evts}} \times \boldsymbol{N_{bins}}$ computations required are almost fully independent
    → perfect for parallelization

- **GPUs** are a “no-brainer” when considering **parallelization**
  - **Multi-threading implemented** in the **ORSA** fitting framework for both **GPU** and CPU
    → **x100 reduction** in computation **time**
    (e.g., only 10 seconds for an Asimov binned fit with Minuit)