ACCELERATING UNBINNED LIKELIHOOD COMPUTATION IN JUNO WITH GPU PARALLELIZATION

A. Serafini [a, b, ✉], B. Jelmini [b, c]
on behalf of the JUNO collaboration
[a] Department of Physics and Astronomy, University of Padua and INFN, 35131 Padua, Italy
[b] National Institute for Nuclear Physics (INFN), Padua Section, 35131 Padua, Italy
[c] Department of Information Engineering, University of Padua and INFN, 35131 Padua, Italy
✉ andrea.serafini@pd.infn.it

BIG DETECTORS FOR RARE PHENOMENA

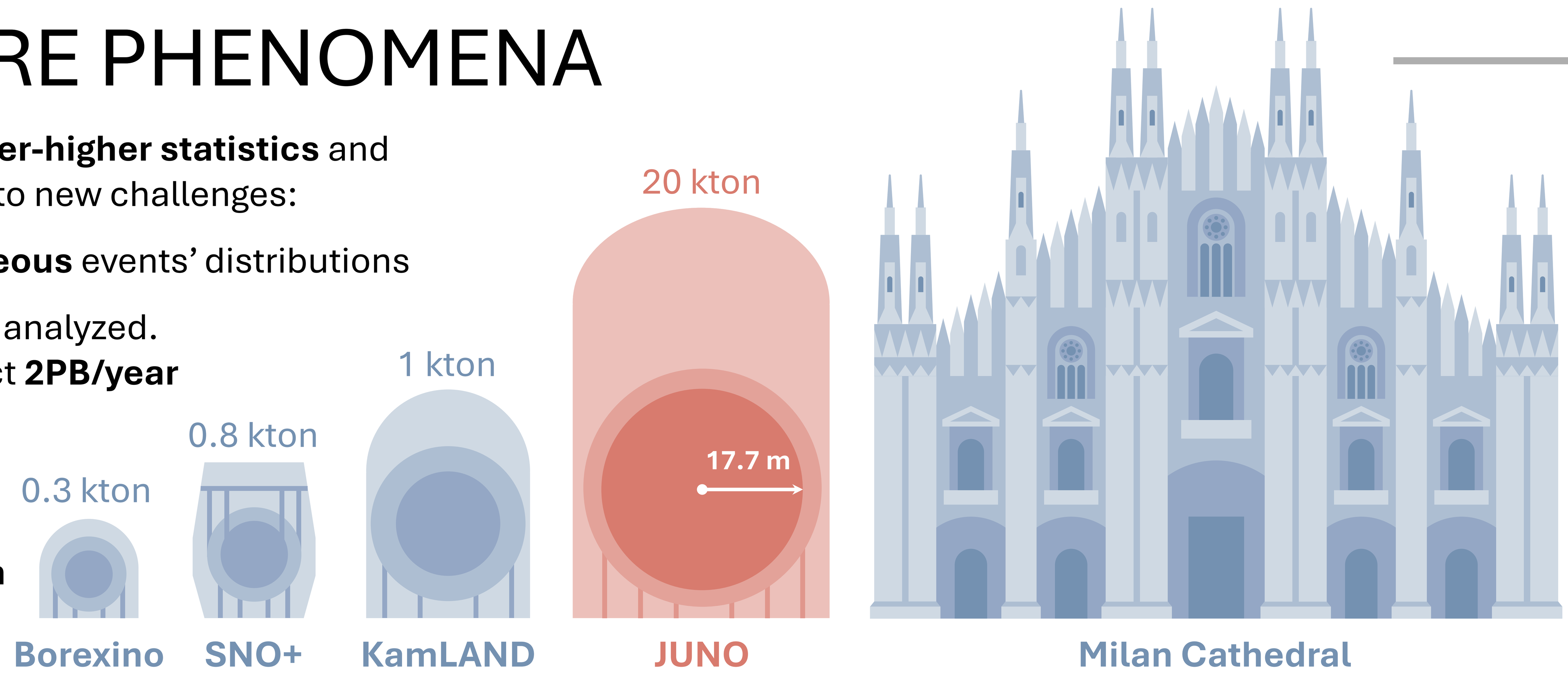
The search for **ever-rarer phenomena** and the need for **ever-higher statistics** and **precision** led to the construction of **larger detectors** and to new challenges:

1. **Detectors** with **complex responses** and **inhomogeneous** events' distributions
2. **Increasing number of events** that must be efficiently analyzed. JUNO, used here as a case study, is expected to collect **2PB/year**

DOI 10.1016/j.ppnp.2021.103927

New **flexible and efficient** approaches are needed to analyze this next generation of neutrino physics data.

Unbinned likelihood and its **computational optimization** can be a **practical solution** to these challenges.



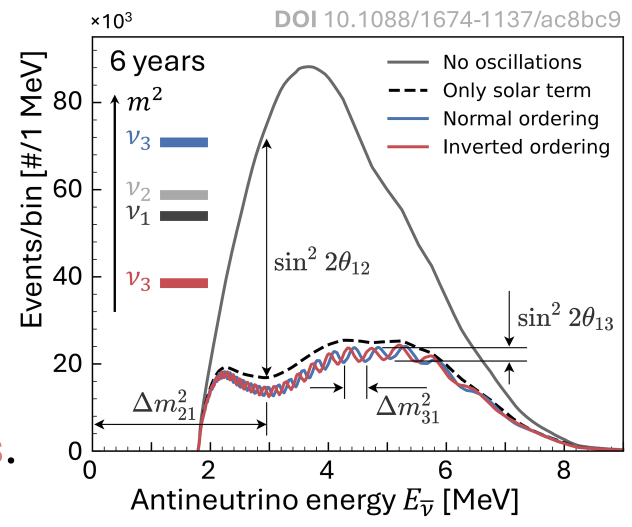
DETECTOR RESPONSE

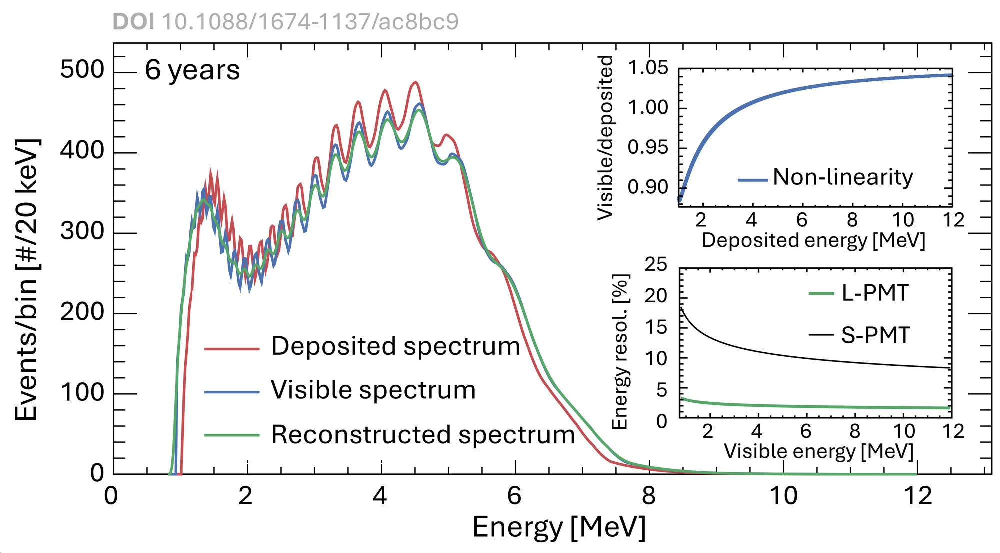
JUNO uses $\bar{\nu}_e$ emitted by 8 reactor cores at a 52.5 km distance to access neutrino **oscillation properties**.

The detection occurs via Inverse Beta Decay (IBD):

$$\bar{\nu}_e + p \to e^+ + n \qquad E_{dep} \simeq E_{\bar{\nu}} - 0.782\ MeV$$

The energy deposited by the e^+ provides an indirect link to the incoming $\bar{\nu}_e$ energy and four oscillation parameters.





Effects **impacting** the **reconstructed energy** must be **modelled** to accurately and precisely estimate the parameters.

Considered **detector response** includes:

- **Energy transfer** in the **IBD** reaction
- Liquid scintillator **non-linearity**
- Detector **energy resolution**

UNBINNED LIKELIHOOD

Unbinned likelihood can easily incorporate events' features like **energy**, **position** and **time**, for a more flexible analysis tuned on an **event-by-event basis**:

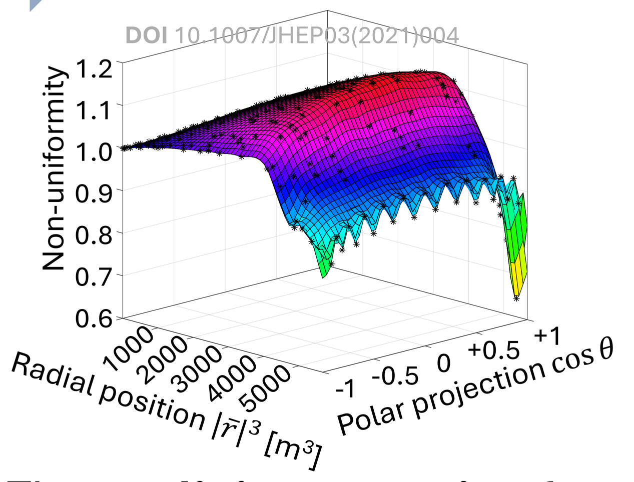
$$\vec{\phi} = (\Delta m^2_{21}, \theta_{12}, \Delta m^2_{31}, \theta_{13})$$

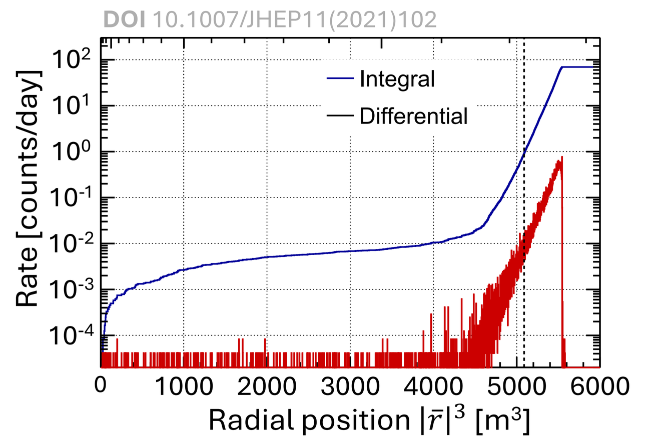
$$P(data \mid \vec{\phi}, \vec{\eta}) = \frac{\mu(\vec{\phi},\vec{\eta})^N e^{-\mu(\vec{\phi},\vec{\eta})}}{N!} \prod_{i=1}^{N} p(E_i', \vec{r}_i', t_i' \mid \vec{\phi}, \vec{\eta})$$

Model parameters (physics + response); Predicted events; Observed events; Product over events; Model PDF; Event features; Likelihood; Poisson normalization; Unbinned term

It is possible to characterize in $p(E, \vec{r}, t \mid \vec{\phi})$ the **space-** and **time-dependent detector response** and the distributions of signal and background events. As an example:

Detector light response (changing in space) — Radiogenic background (increasing with radius)



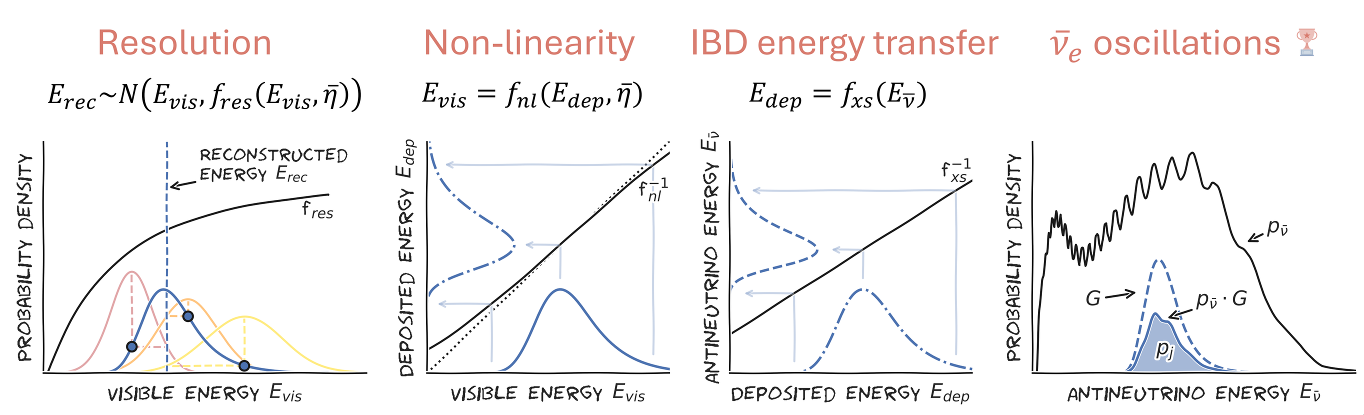


The **explicit expression** for evaluating the $\bar{\nu}_e$ model PDF convoluted with the detector response for a given event is:

$$p(E_i' \mid \vec{\phi}, \vec{\eta}) = \int p_{\bar{\nu}}(E_{\bar{\nu}} \mid \vec{\phi}) \cdot G(f_{nl}(f_{xs}(E_{\bar{\nu}}), \vec{\eta}), E_{rec}, \vec{\eta}) \cdot \frac{dE_{vis}}{dE_{dep}} \frac{dE_{dep}}{dE_{\bar{\nu}}} \cdot dE_{\bar{\nu}} \Big|_{E_{rec}=E_i'}$$

Detector response convolution; IBD energy transfer; Resolution; Response parameters; Evaluation for event i; Antineutrino spectrum; Non-linearity; Oscillation parameters; Jacobian

Which can be schematically illustrated as:

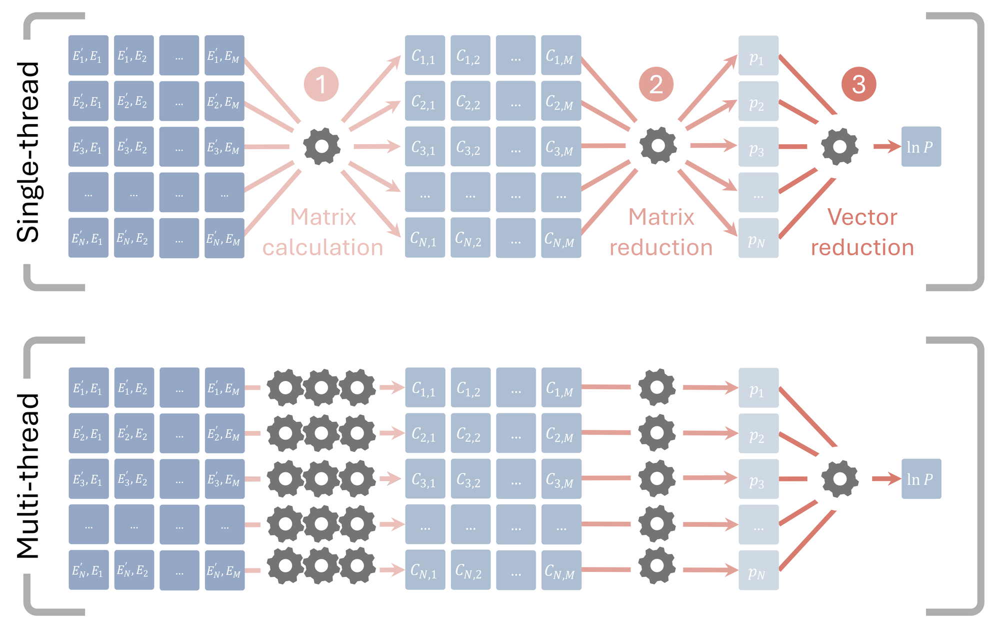


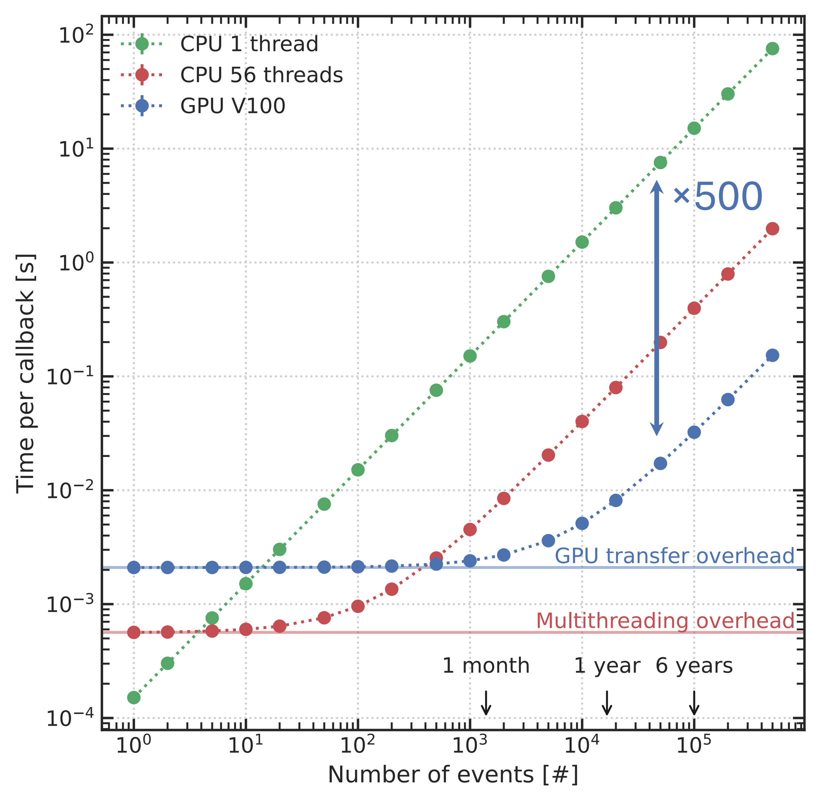
GPU PARALLELIZATION

The **likelihood** is approached through its **logarithm**. **Integrals** are computed as finite **sums**. The calculation simplifies to a form easily **parallelizable** on **CPU** and **GPU**:

$$P \to \ln P \qquad \ln \Pi \to \Sigma \ln \qquad \int dE \to \Sigma \Delta E$$

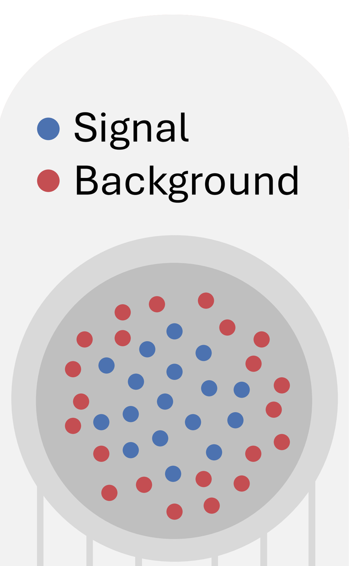
$$\ln P(data \mid \vec{\phi}, \vec{\eta}) \sim \sum_{i=1}^{N} \ln \sum_{k=1}^{M} C_{i,k}(\vec{\phi}, \vec{\eta})$$

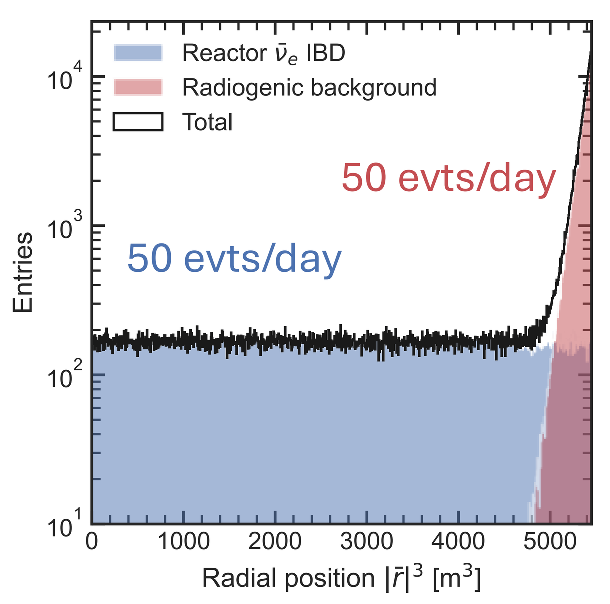




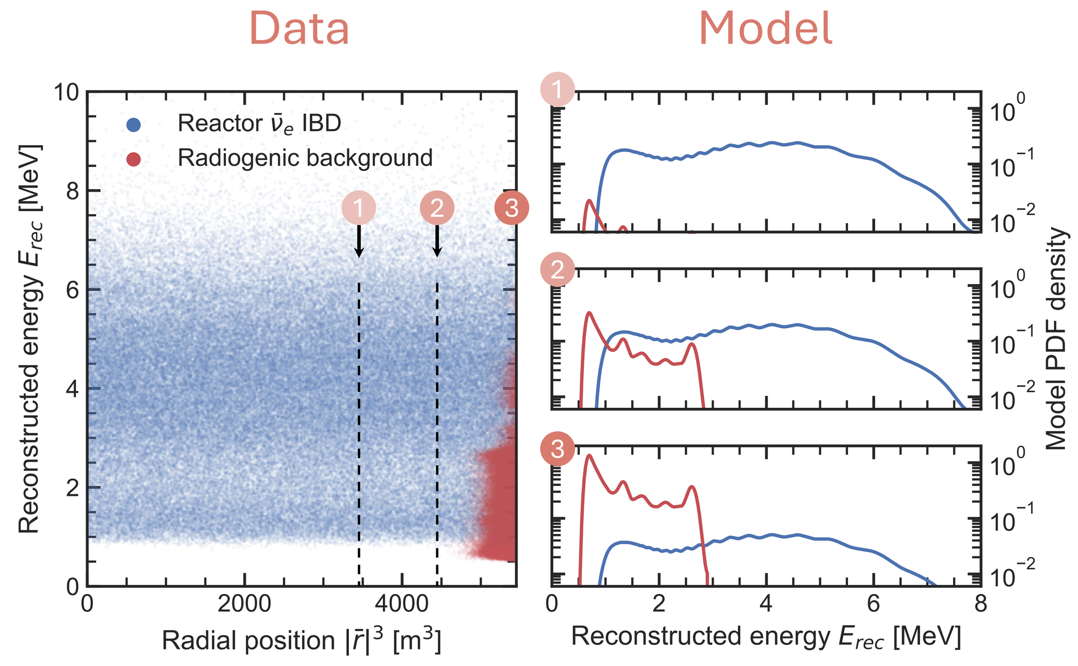
A CASE STUDY *

1. A **toy experiment** is designed considering a **homogeneously** distributed **signal** and a **background** exponentially **increasing** with radius.

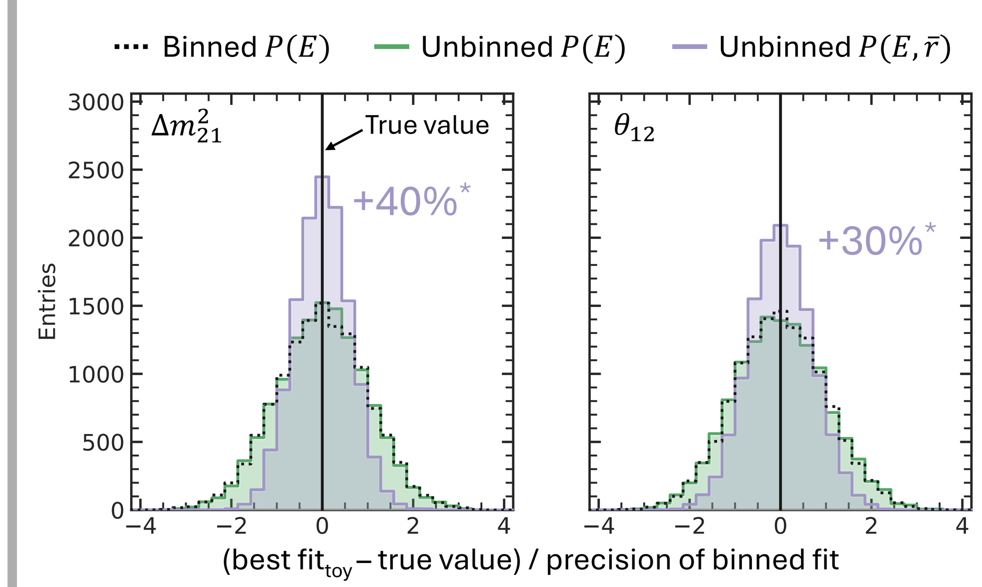




2. **Likelihood** is **modeled including** *a-priori* knowledge about signal and background **expected distributions**.



3. **Results**: 10k **fits** of **10k** toy experiments



Binned $P(E)$ **same performance** of unbinned $P(E)$.
Exploiting all event features → **better performance**.

*The inputs and results used are not intended to be representative of JUNO and its sensitivity. The case study is intended to demonstrate the potential benefits of the developed methodology. The rates, spectral shapes and distributions of the events are also not representative of JUNO.