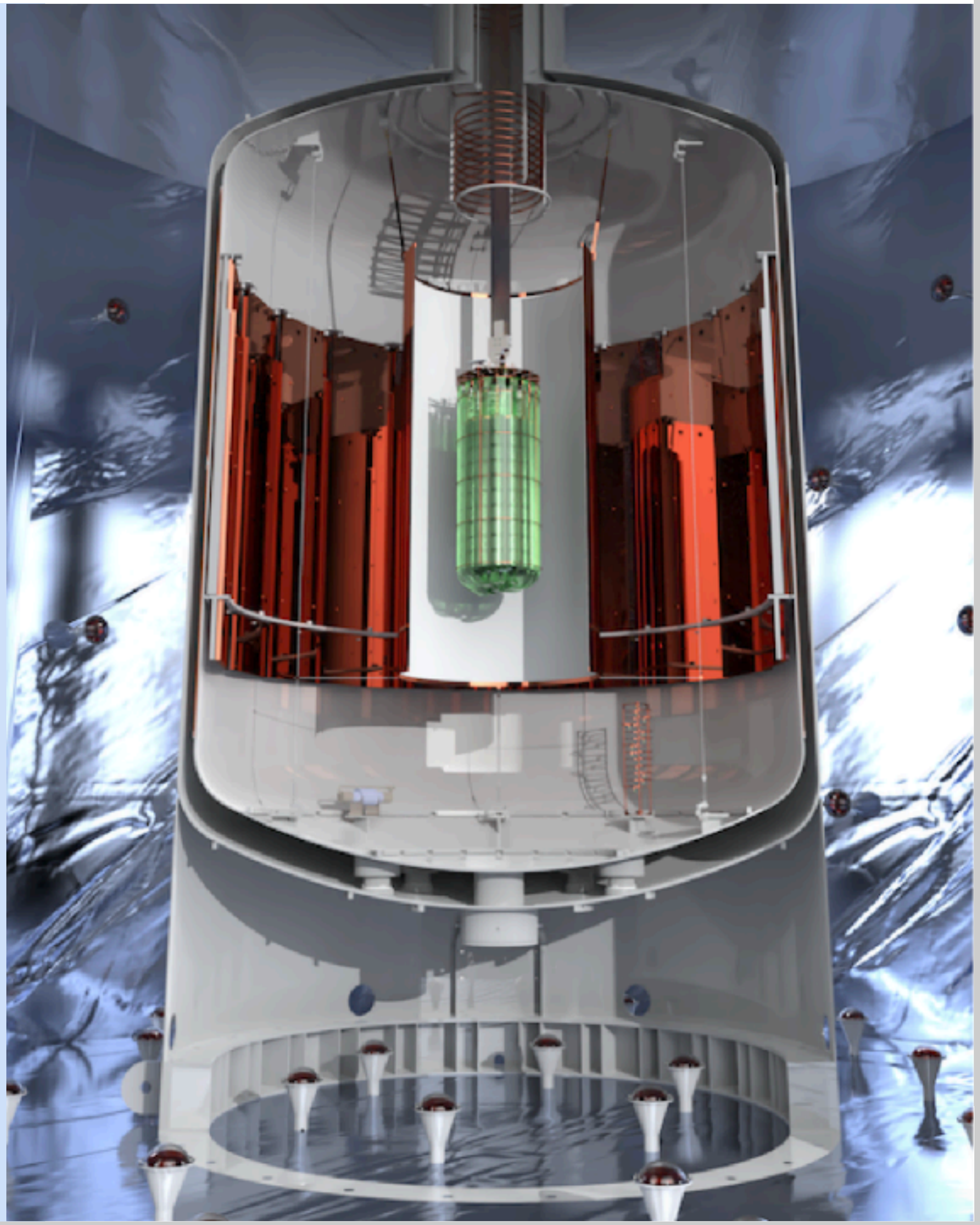
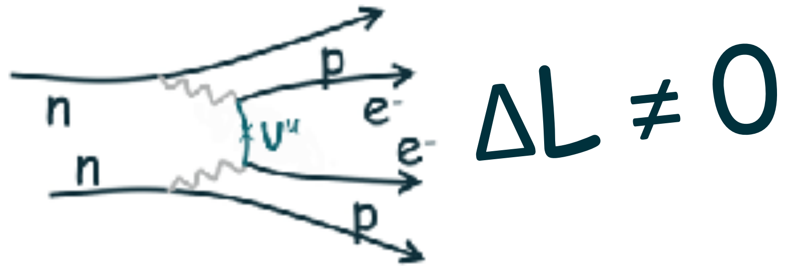
# RESuM and Beyond: Surrogate Modeling for Physics Detector Design



**Ann-Kathrin Schuetz (LBNL)**
Lawrence Berkeley National Laboratory

NSD Staff Meeting - February 4, 2025

Large Enriched Germanium Experiment for Neutrinoless ββ Decay






## 0νββ decay - Experimental sensitivity

$$T^{0\nu}_{1/2} \propto \varepsilon \cdot a \cdot \sqrt{\frac{M \cdot t}{BI \cdot \Delta E}}$$

Background index

**LEGEND-1000 background goal:**

$<10^{-5}$ cts/keV/kg/yr

$^{77(m)}$Ge background at LNGS:

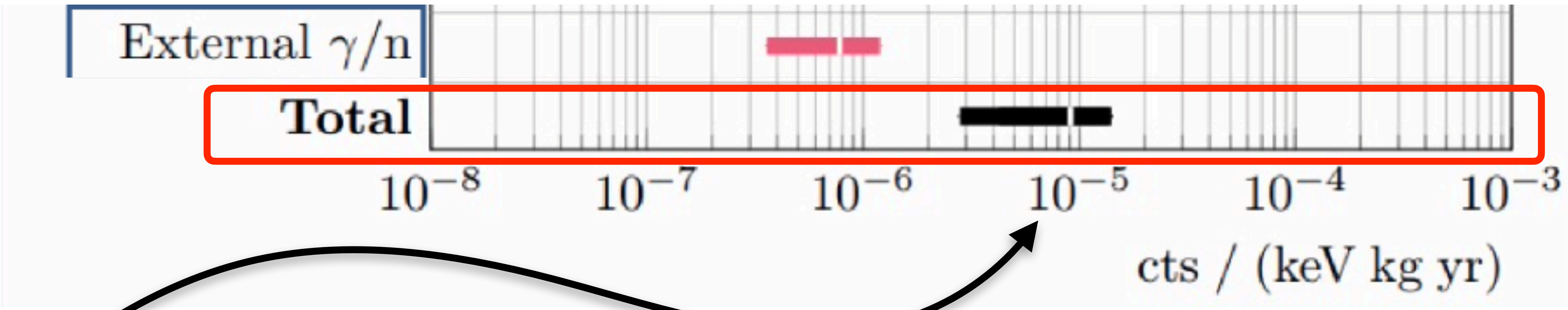
$>10^{-5}$ cts/keV/kg/yr

arXiv:2107.11462

BUT: many opportunities to reduce and actively suppress this background

arXiv:1802.05040

Why do we have to reduce the cosmogenic background at LNGS?

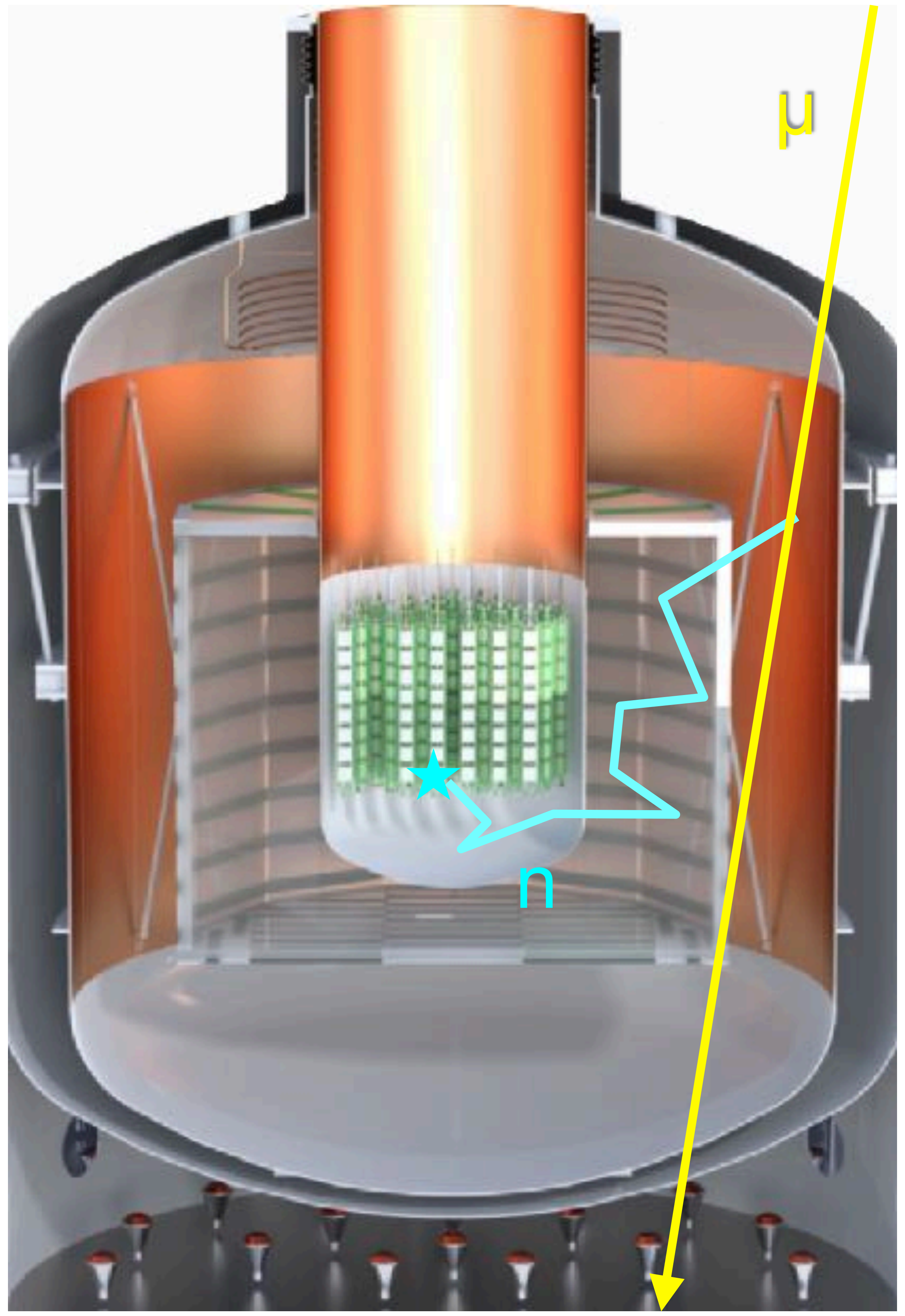



| Location | Depth [km.w.e.] | $^{77(m)}$Ge background contribution (w/o new cuts*) [cts/keV/kg/yr] |
|---|---|---|
| SNOLab (Reference Site) | 6 | $4.2\times10^{-7}$ [0] |
| LNGS (Alternative Site) | 3.5 | $2.7\times10^{-5}$ |

* standard background rejection are applied which strongly supresses $^{77}$Ge

# Cosmogenic background reduction

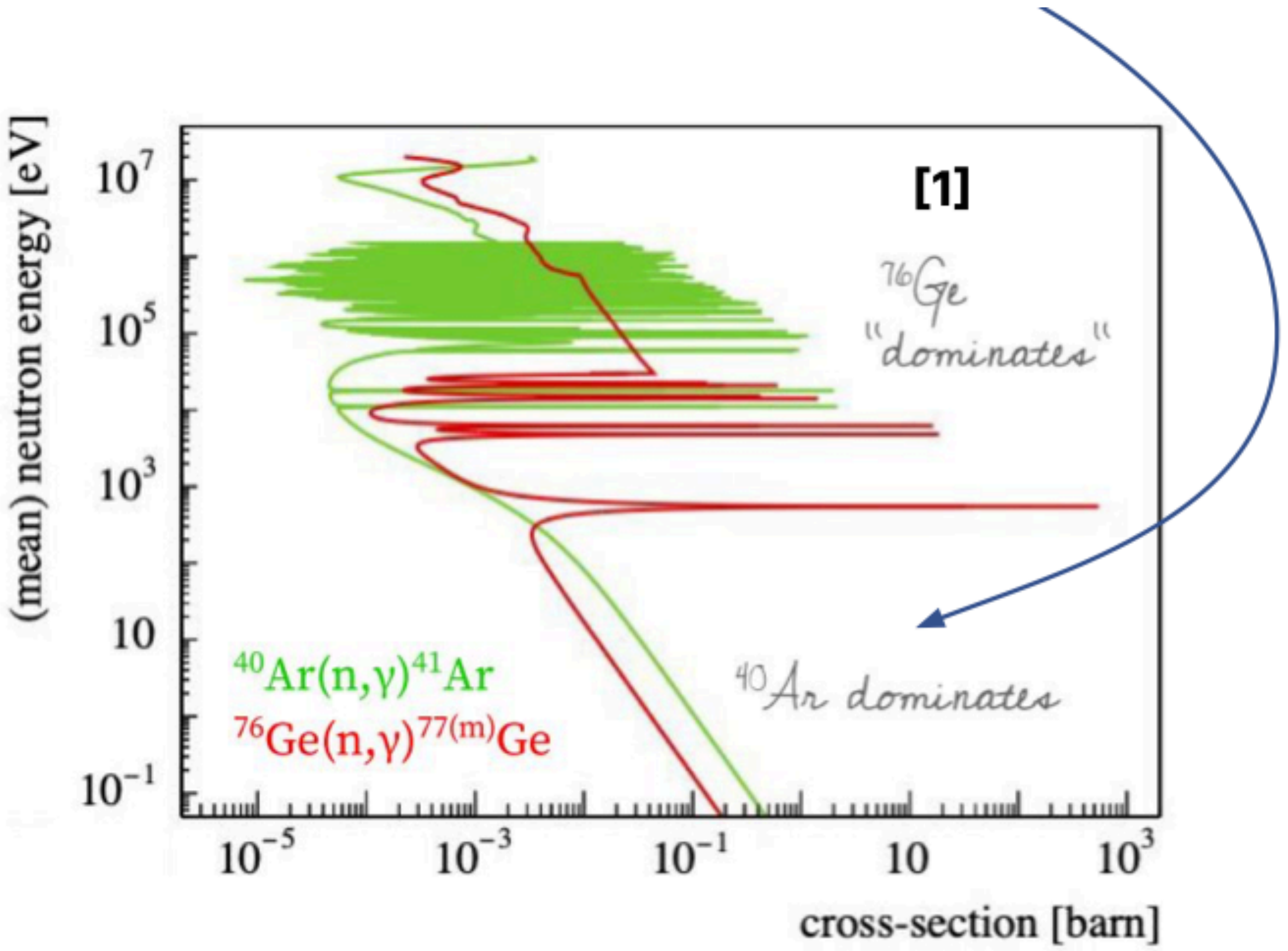
## What options are there to reduce the impact of cosmogenic background?

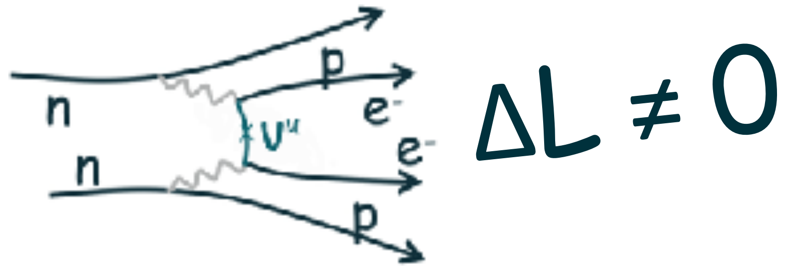



1. Reduce the muon flux → increase overburden.
2. Reduce the neutron flux around the detectors.
3. Tag the $^{77(m)}$Ge production and apply a delayed coincidence cut.

## Reduce the neutron flux around the detectors - *Idea*:

add neutron moderators to slow neutrons down and increase their likelihood to be captured by LAr instead of $^{76}$Ge.

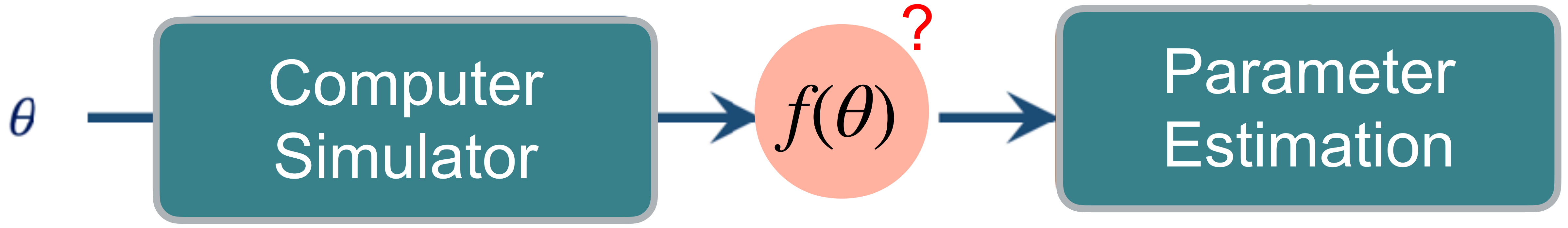


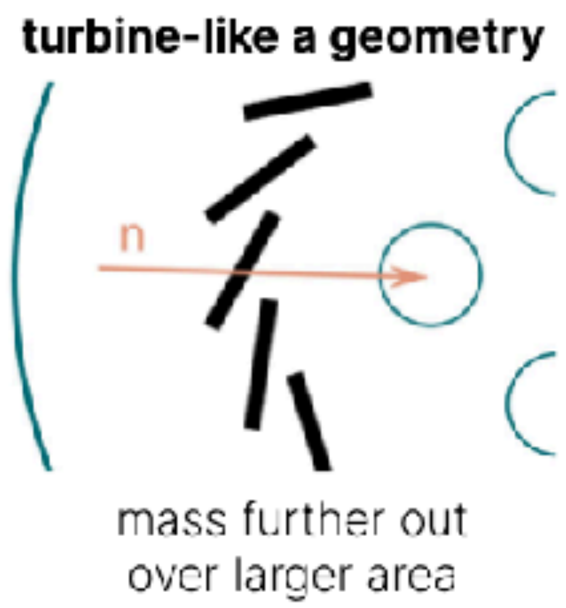



# Optimal Design Parameters

## How to find the optimal design parameter?

**Run a few simulations at different parameters**



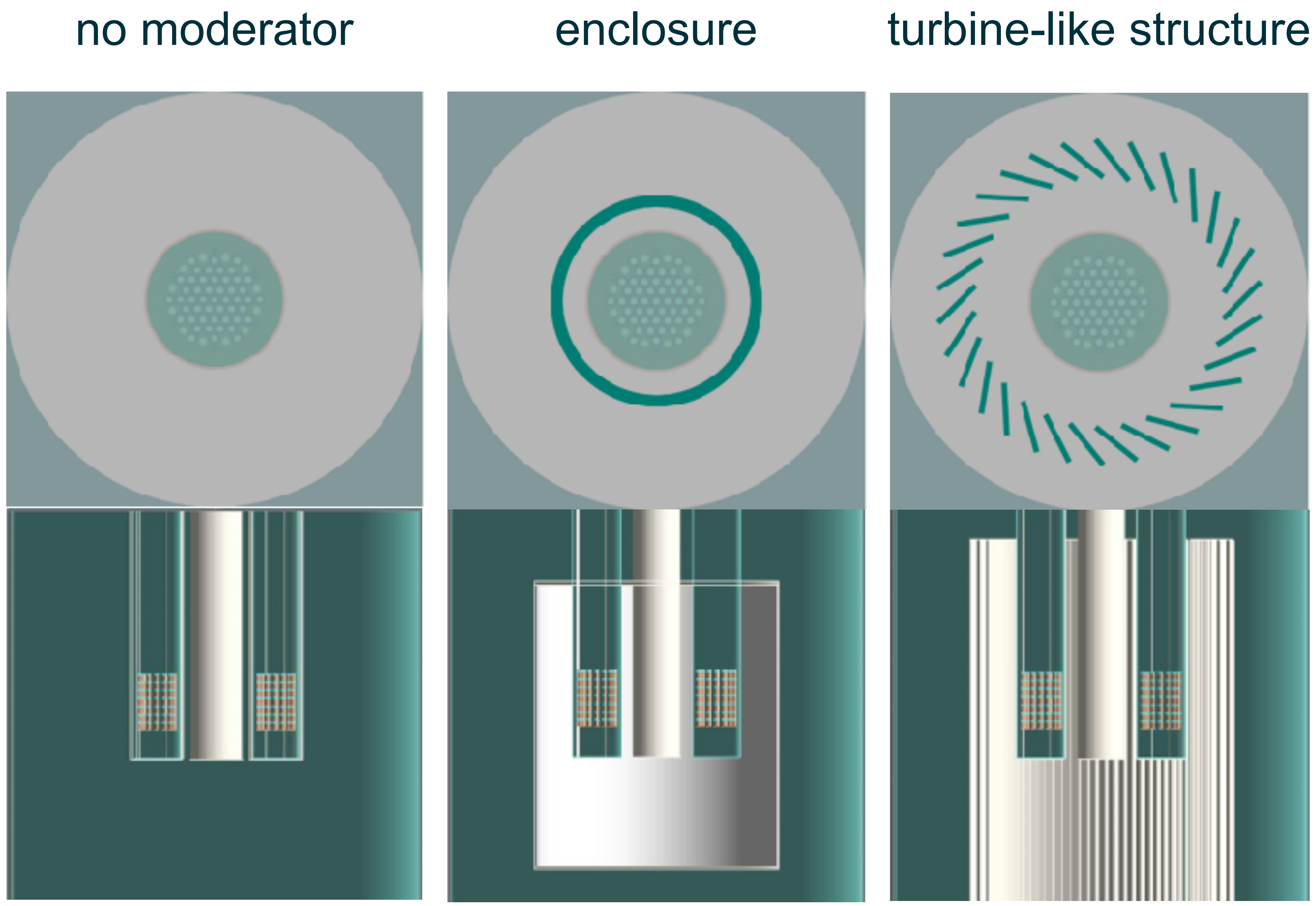




- MC studies using a custom simulation module[3] based on LEGEND-1000 and GERDA setup[3] implementation
- Solid neutron moderator design: enclosing tube or turbine-like structure
- 5 design parameters: Radius r, n Panels, Thickness d, Length L and Angle $\theta$

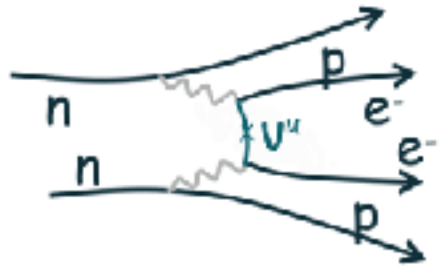
➡ **High-dimensional parameter spaces**

➡ **High computational cost of Geant4 MC simulations (~200 CPUh)**

➡ **Traditional methods like grid searches are impractical**




**Starting point: 4 high fidelity simulation data points only!!**




$\Delta L \neq 0$

# Introduction: Rare Event Design Problem

very small trigger probability $t(\theta, \phi)$, rare event

parameter to optimize

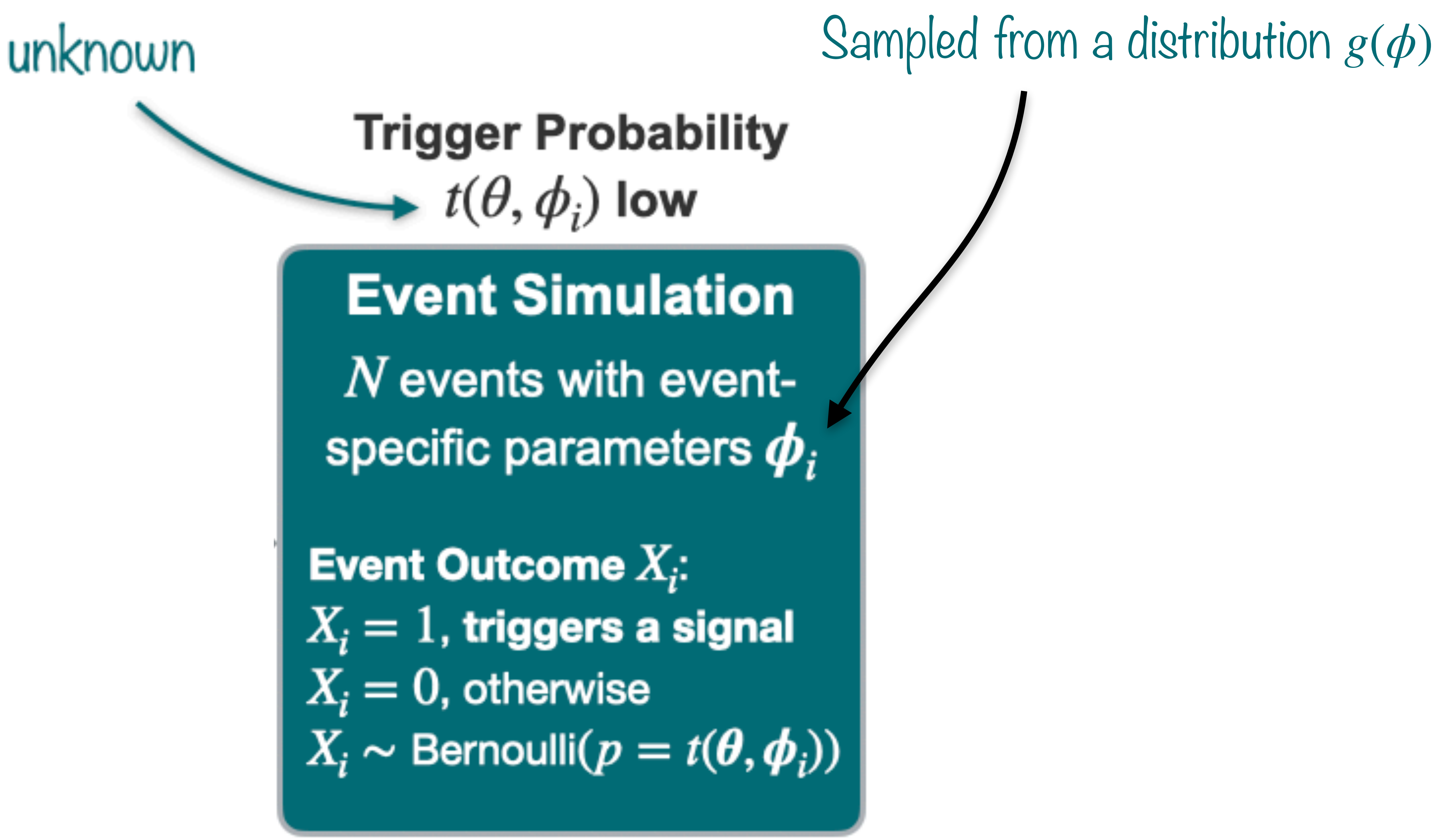
Goal: Optimize $\theta$ such that trigger rate $y$ is minimal under Rare Event Assumption (REA) and expensive simulation

**Design parameter $\theta$**

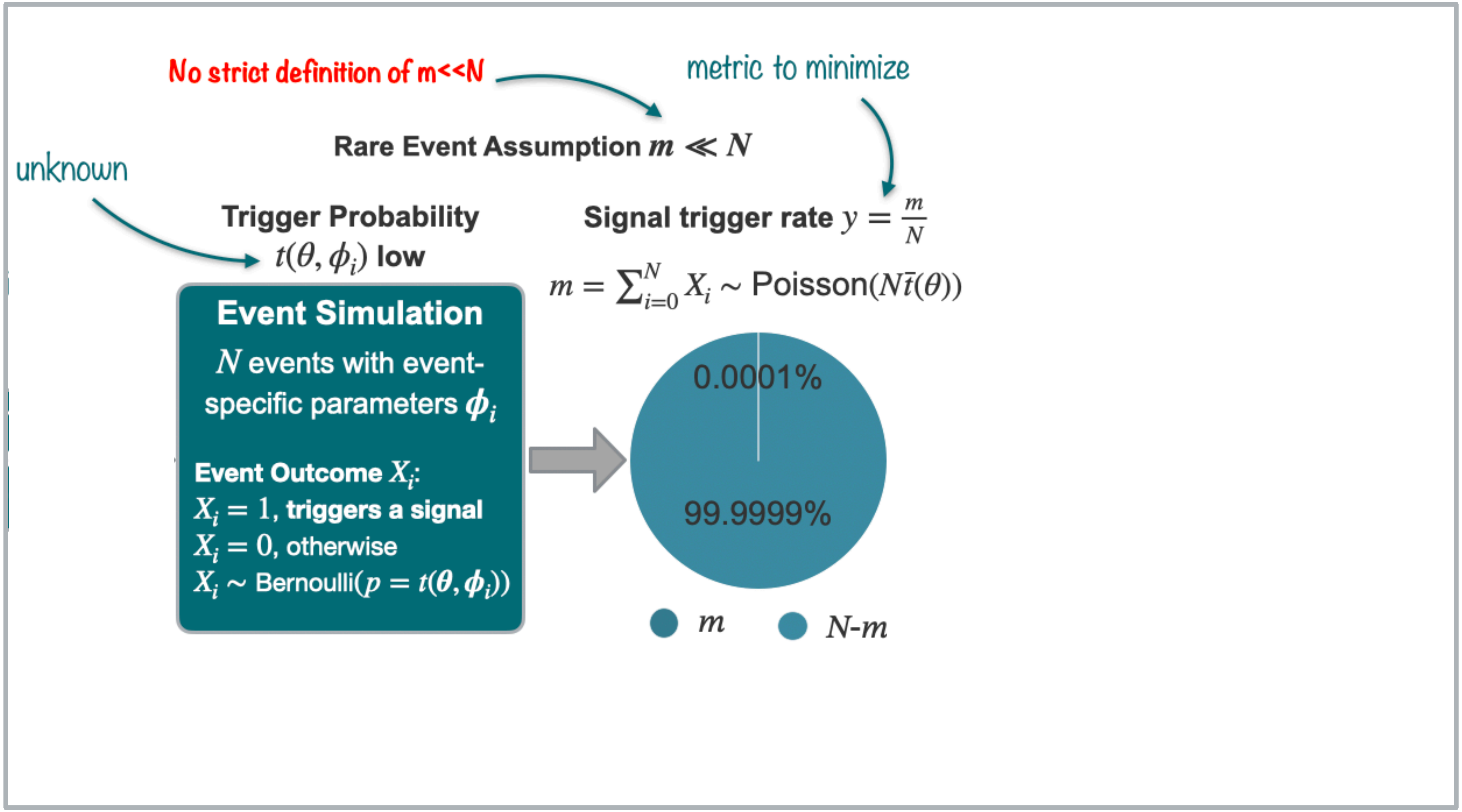
**HF sim**
$\theta, \phi_i$

expensive (200 CPUh per run)

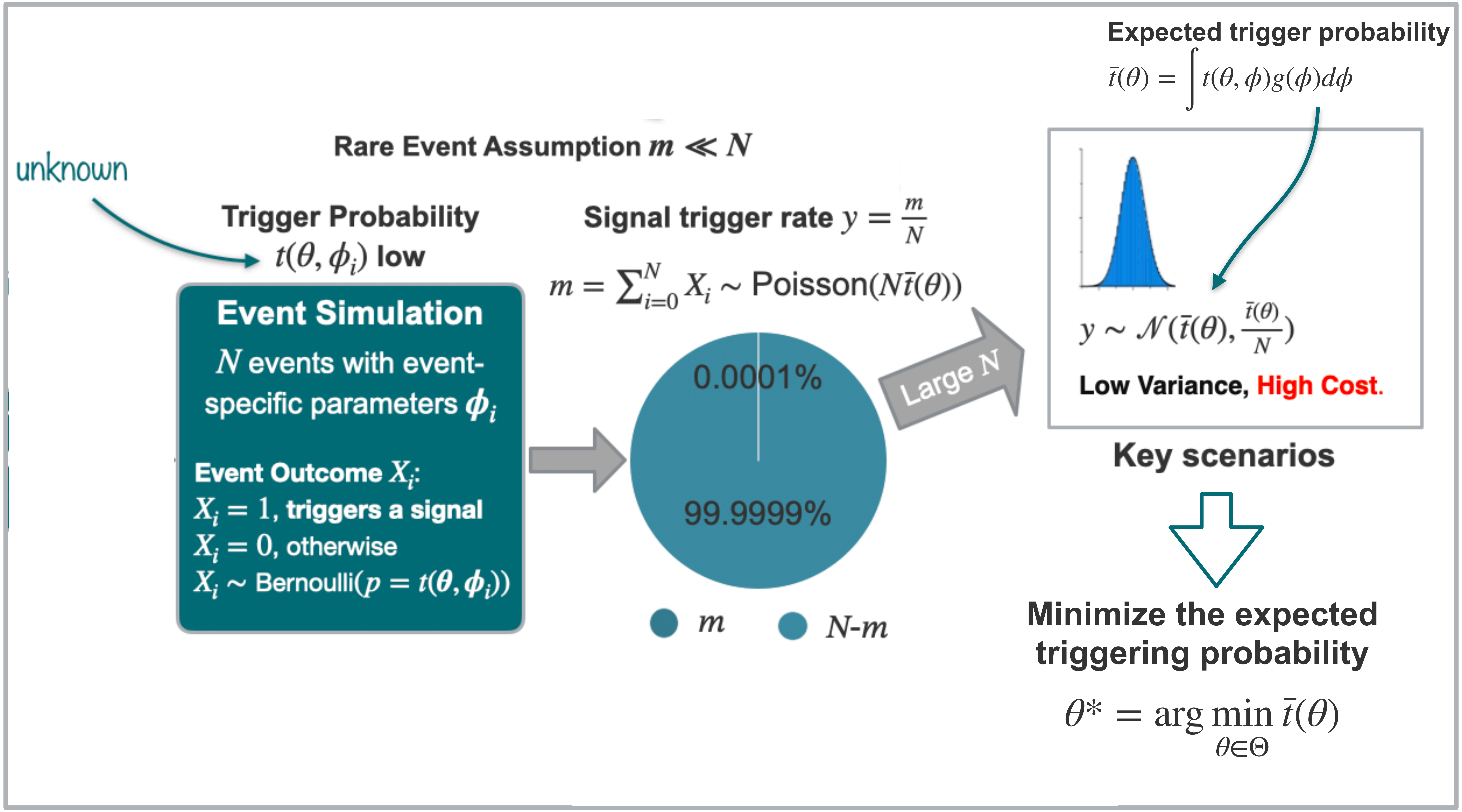
# Rare Event Problem



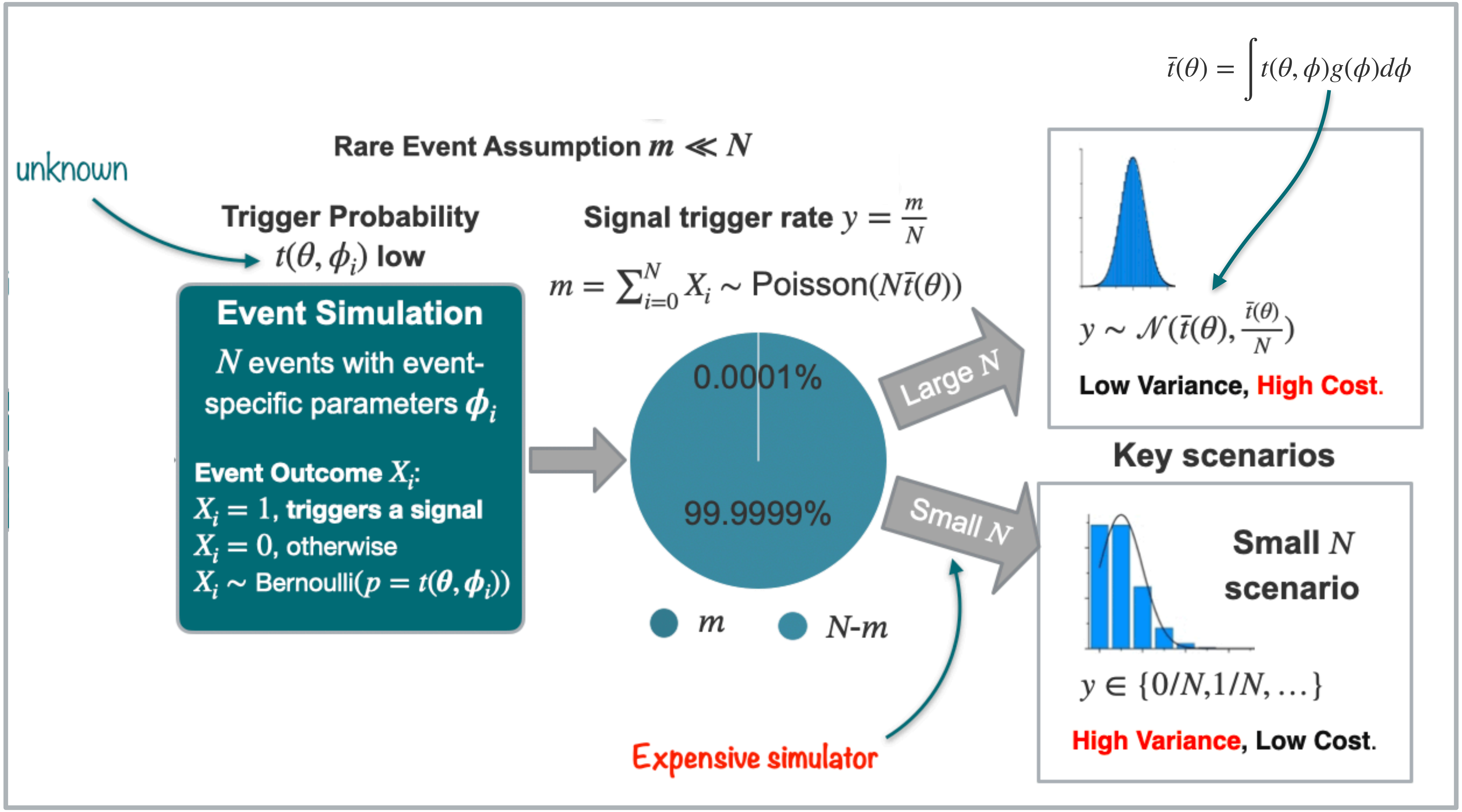
# Rare Event Problem

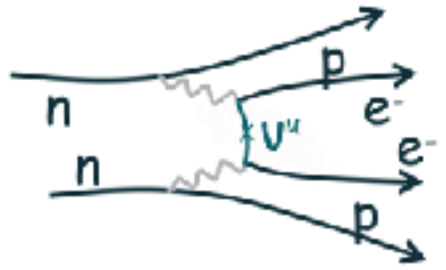


# Rare Event Problem



# Rare Event Problem



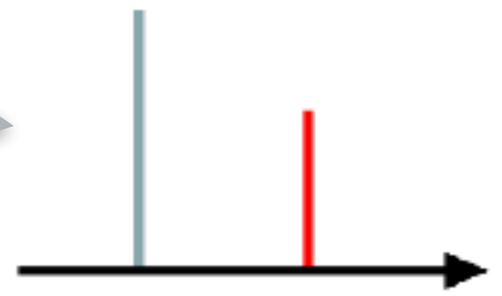



$\Delta L \neq 0$

## How to surrogate our data?

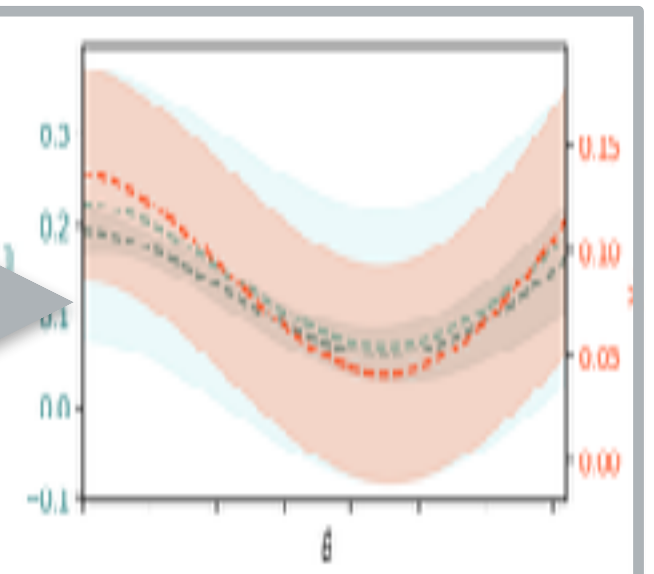
simulation

**discret**



$$f(x) = \begin{cases} 1, & \text{n captured,} \\ 0, & \text{n not captured} \end{cases}$$

Gaussian Process Regression

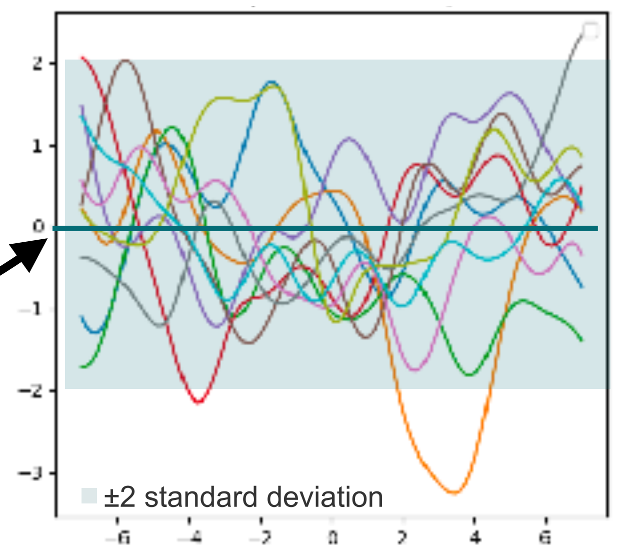


***A Gaussian process is a probability distribution over possible functions that fit a set of points.***

$$f(\theta) \sim GP(m(\theta), k(\theta, \theta'))$$

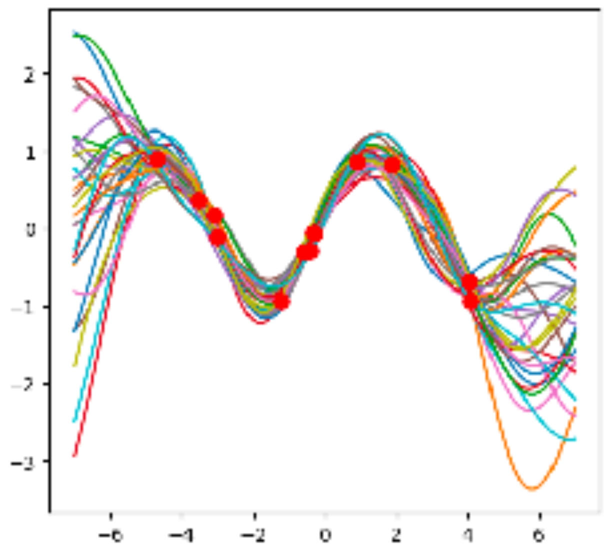
mean of posterior

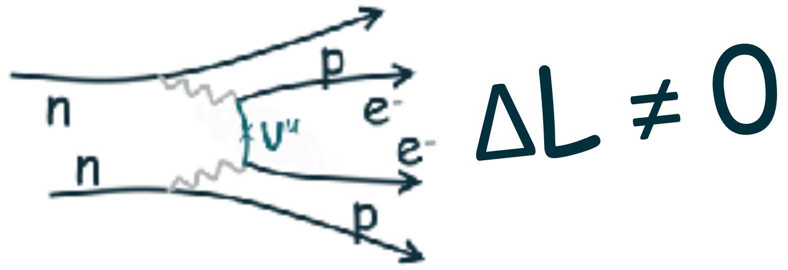
**GP prior**




data

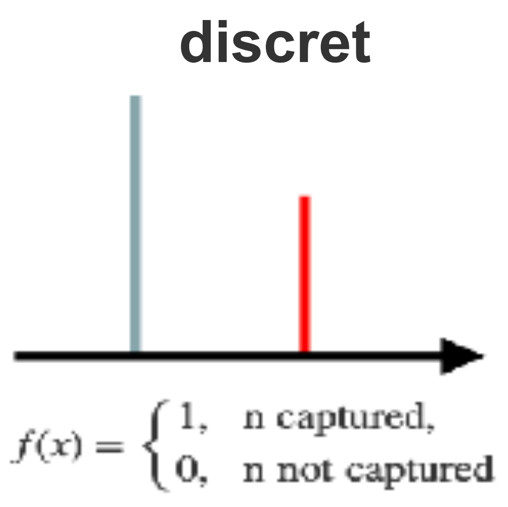
**GP posterior**



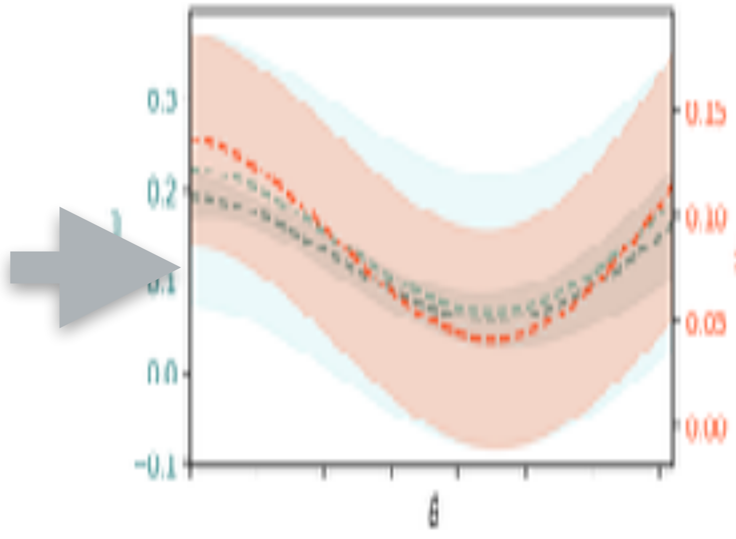





## How to  reduces the computational burden when generating a larger training dataset?



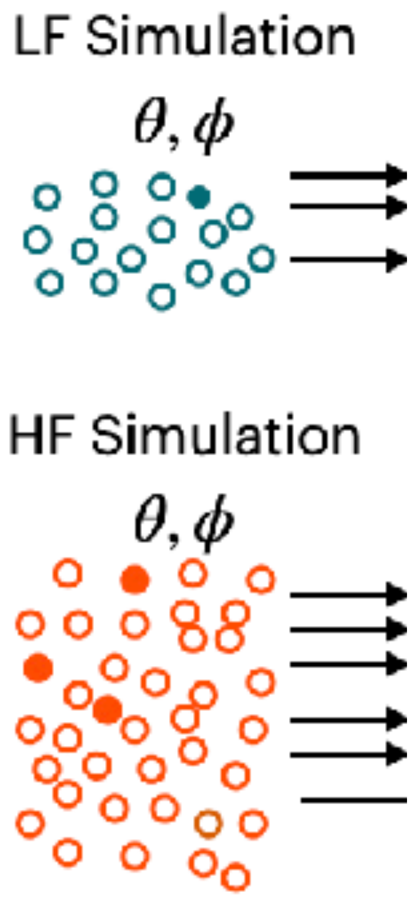

Gaussian Process Regression with Multi-Fidelity (MF) approach



➡ LF used to explore the design space

➡ LF provide prior information for HF

➡ retaining critical information from the more accurate HF

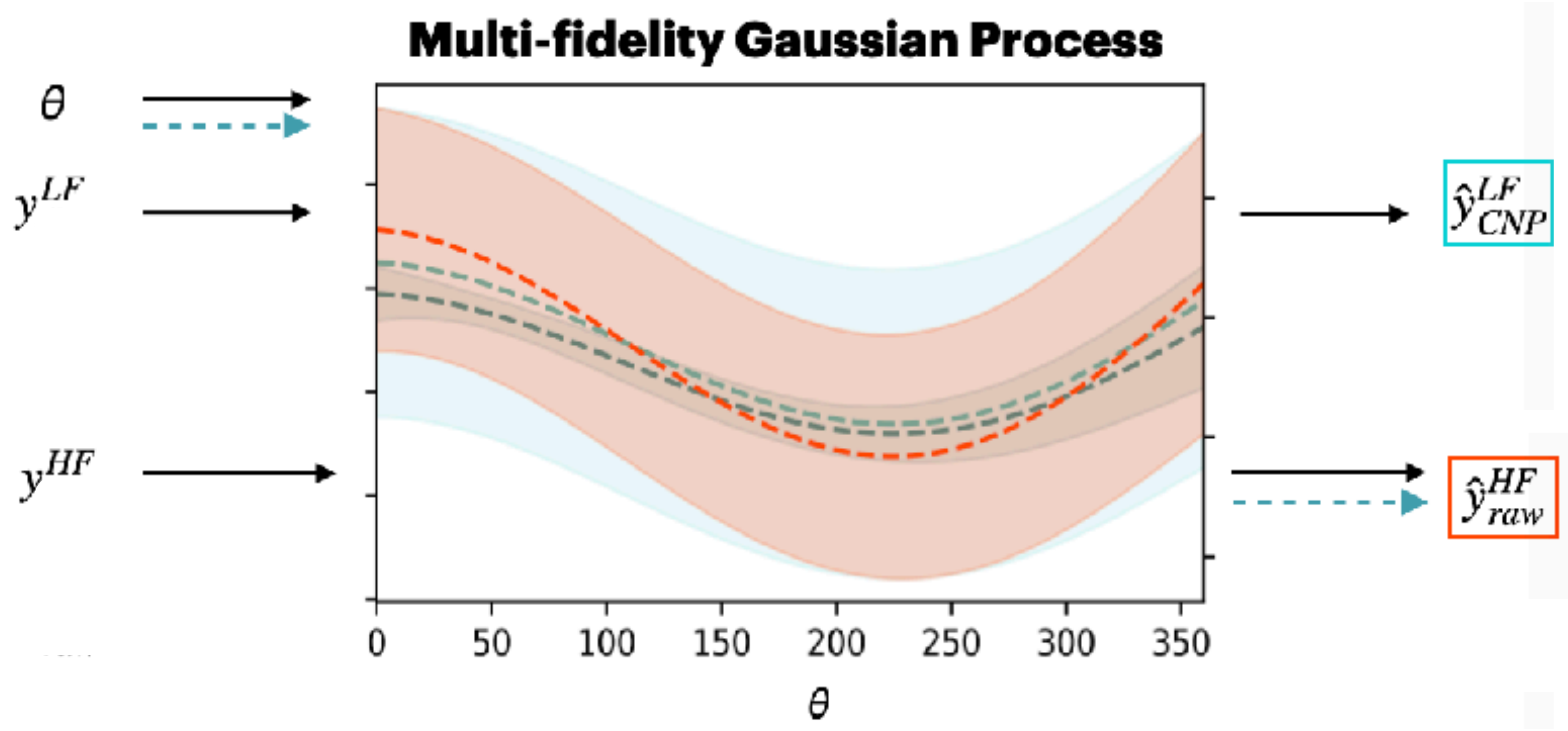


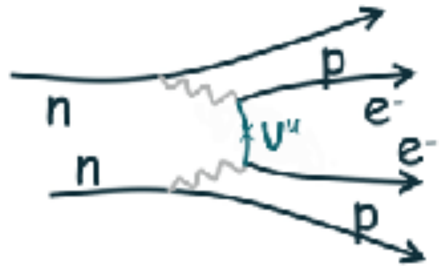

- **combine fast low-fidelity (LF) simulations with costly high-fidelity (HF) simulations**
- each level **share some basic features** and include **most important features**
- Fidelities ranked hierarchically by accuracy (h=0,…,m)
- Use “co-kriging” model with **GP**:

$$\eta_h(x) = \rho_{h-1}\eta_{h-1}(x) + \delta_h(x)$$

*discrepancy term modeled by **GP***

*correlation to lower fidelity (**GP**)*






$\Delta L \neq 0$

# Multi-Fidelity Approach

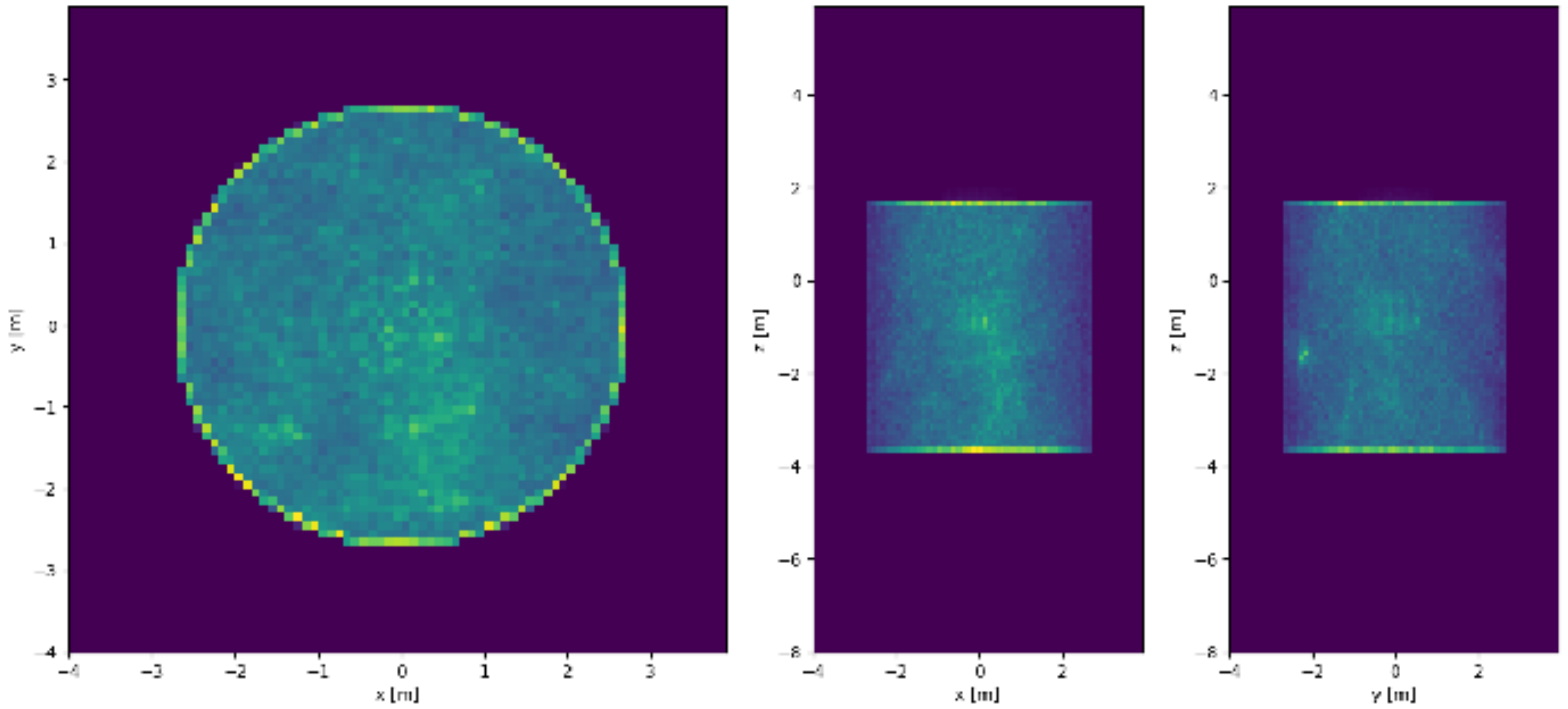
## Event specific distribution

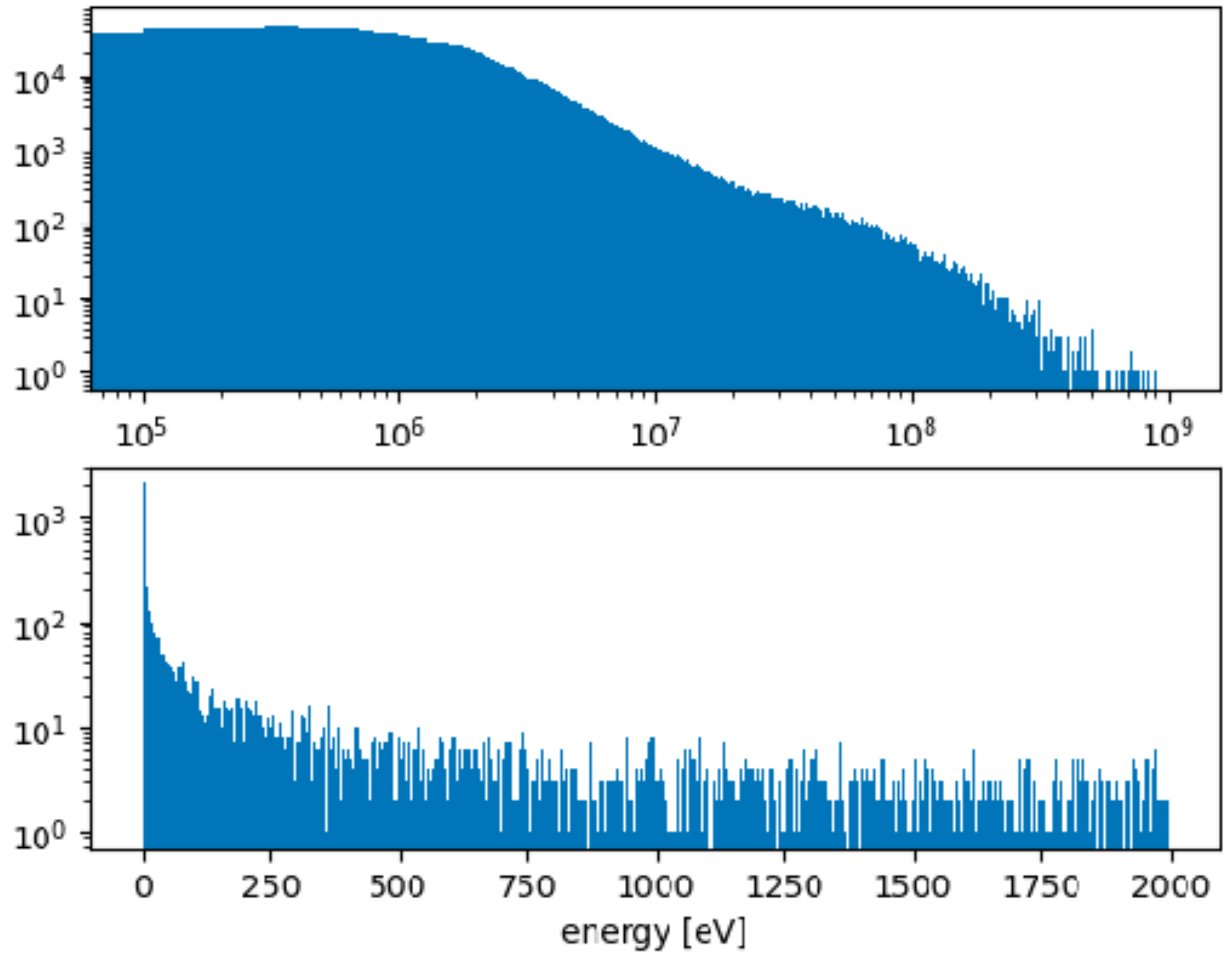
$$g(\phi) = g_\mu(\phi_\mu)\, d\phi_\mu \cdot g_n^{\text{ini}}(\phi_n^{\text{ini}} \mid \phi_\mu) \cdot g_n^{\text{fin}}(\phi_n^{\text{fin}} \mid \phi_n^{\text{ini}})$$

muon — in — neutron — out

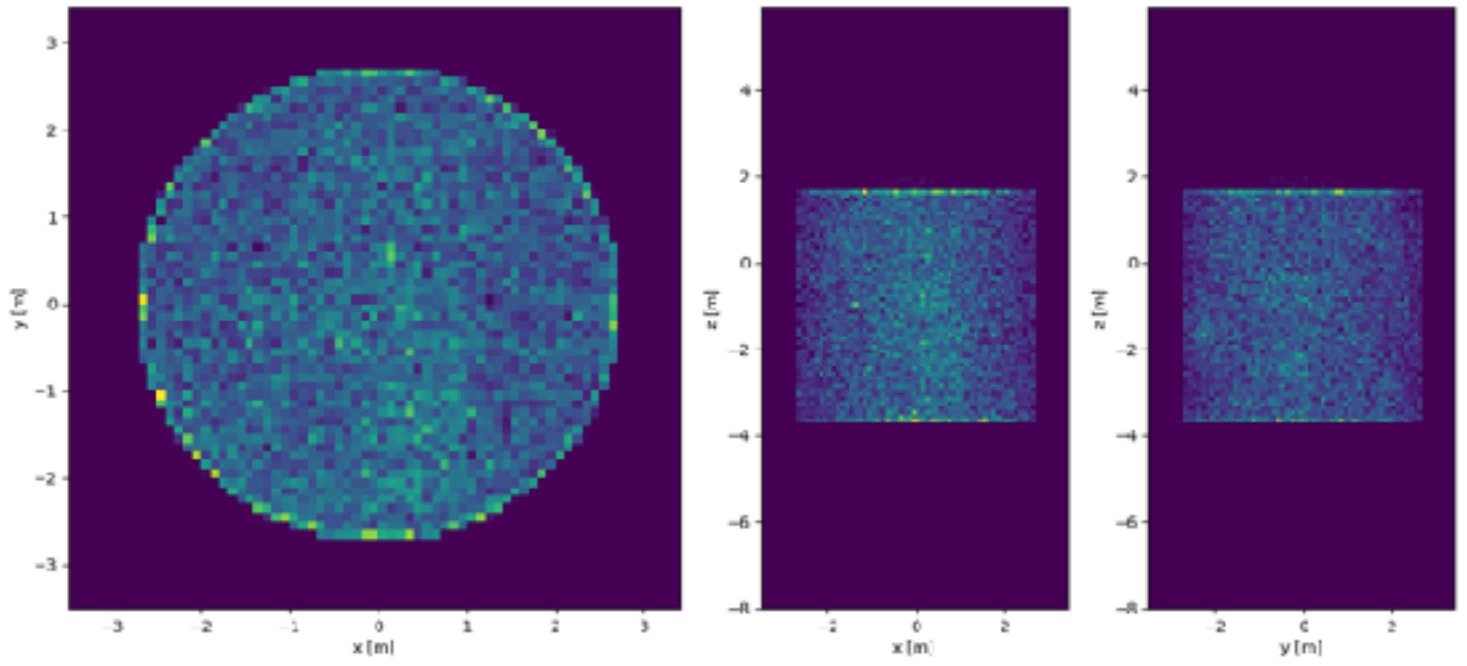
LF

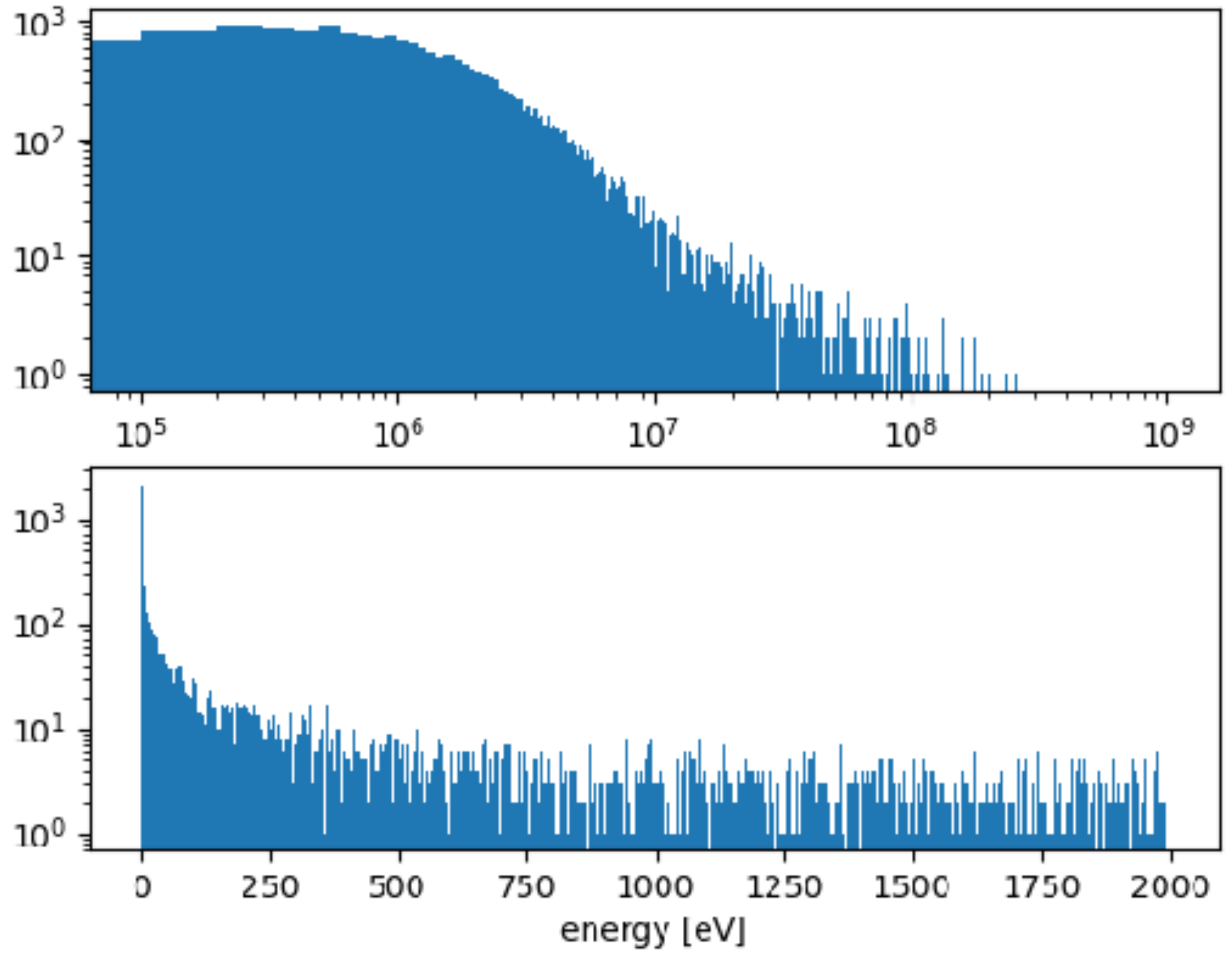
LF input



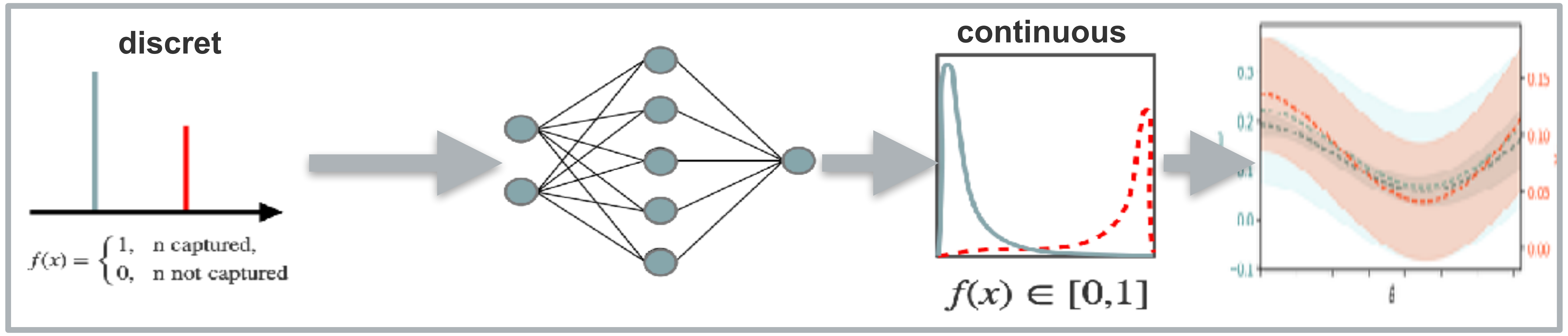

draw distributions with random neutron starting points from high fidelity simulation



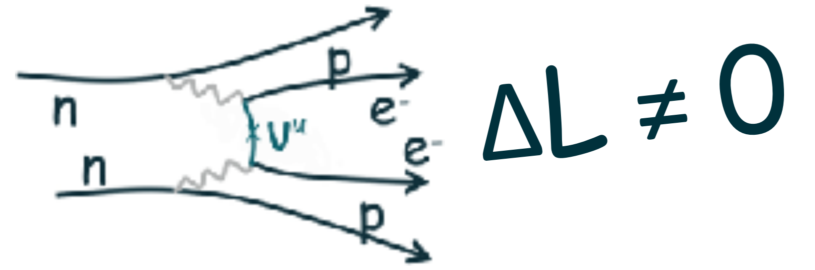



| | HF | LF |
|---|---|---|
| **Primary particle** | Muon | Neutron |
| **CPUh per neutron** | $1.5 \cdot 10^{-4}$ ($2 \cdot 10^{-5}$ per muon) | $3 \cdot 10^{-6}$ |
| **Full detector geometry** | ✓ | ✓ |
| **Full neutron physics lists** | ✓ | ✓ |
| **Timing to primary and production info** | ✓ | ✗ |

## How to mitigate statistical noise?

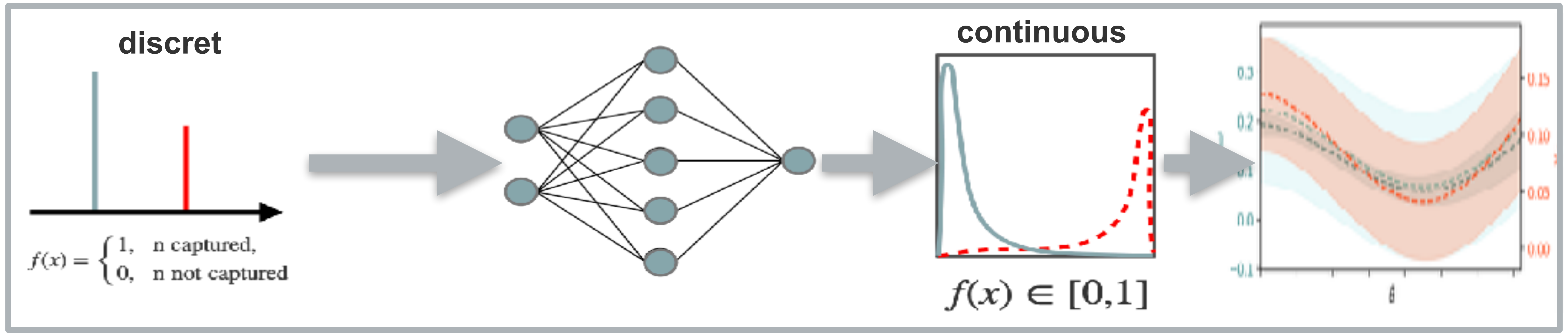



→ Convert the discrete nature of $X_i$ into a continuous score $\beta_i$ , approximating the triggering probability.

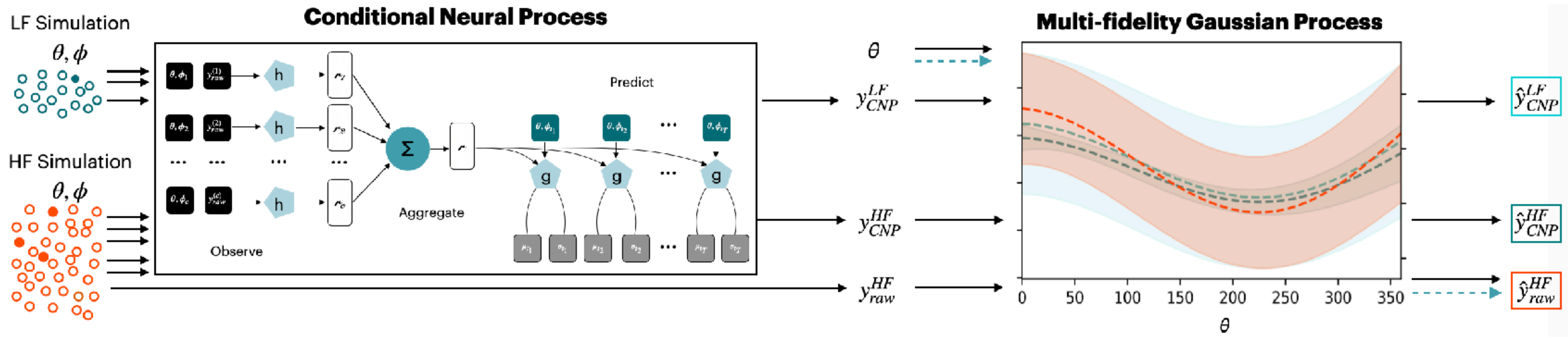


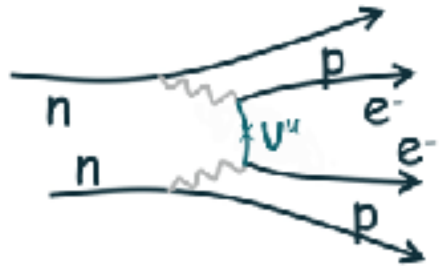



# How to mitigate statistical noise?

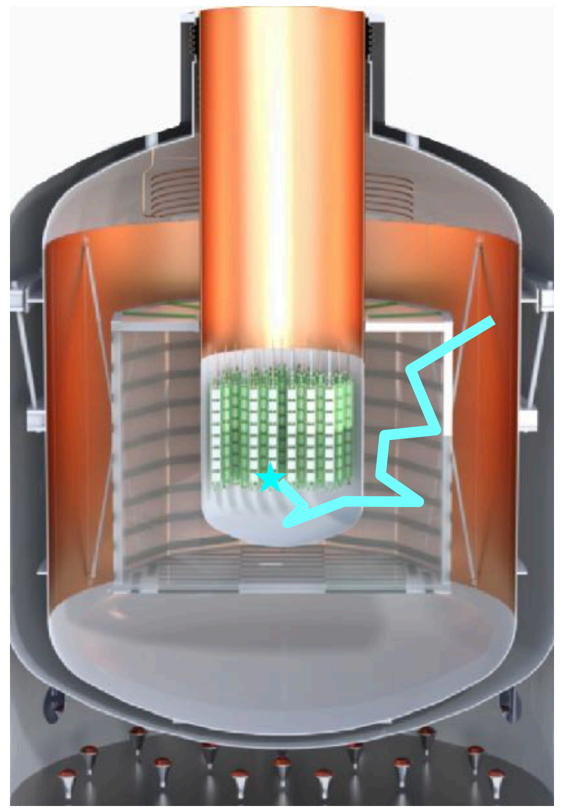



Converts the discrete nature of $X_i$ into a continuous score $\beta_i$ , approximating the triggering probability.



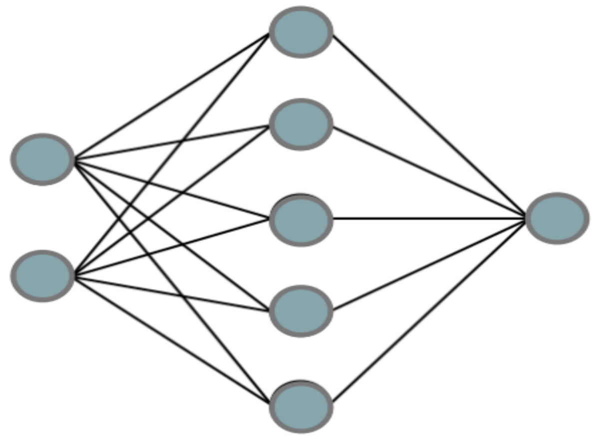



# $\Delta L \neq 0$

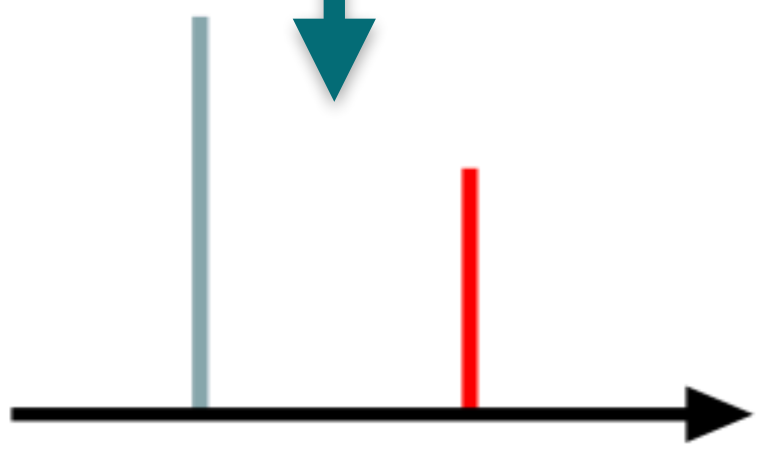


Run Geant4 LF simulations for different moderator configurations

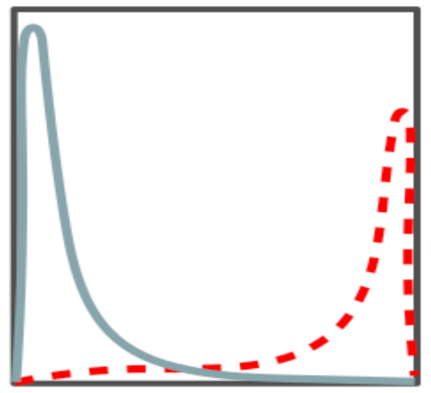
Count number of neutrons being captured given the configuration



$$f(x) = \begin{cases} 1, & \text{n captured,} \\ 0, & \text{n not captured} \end{cases}$$



**continuous**



$f(x) \in [0,1]$

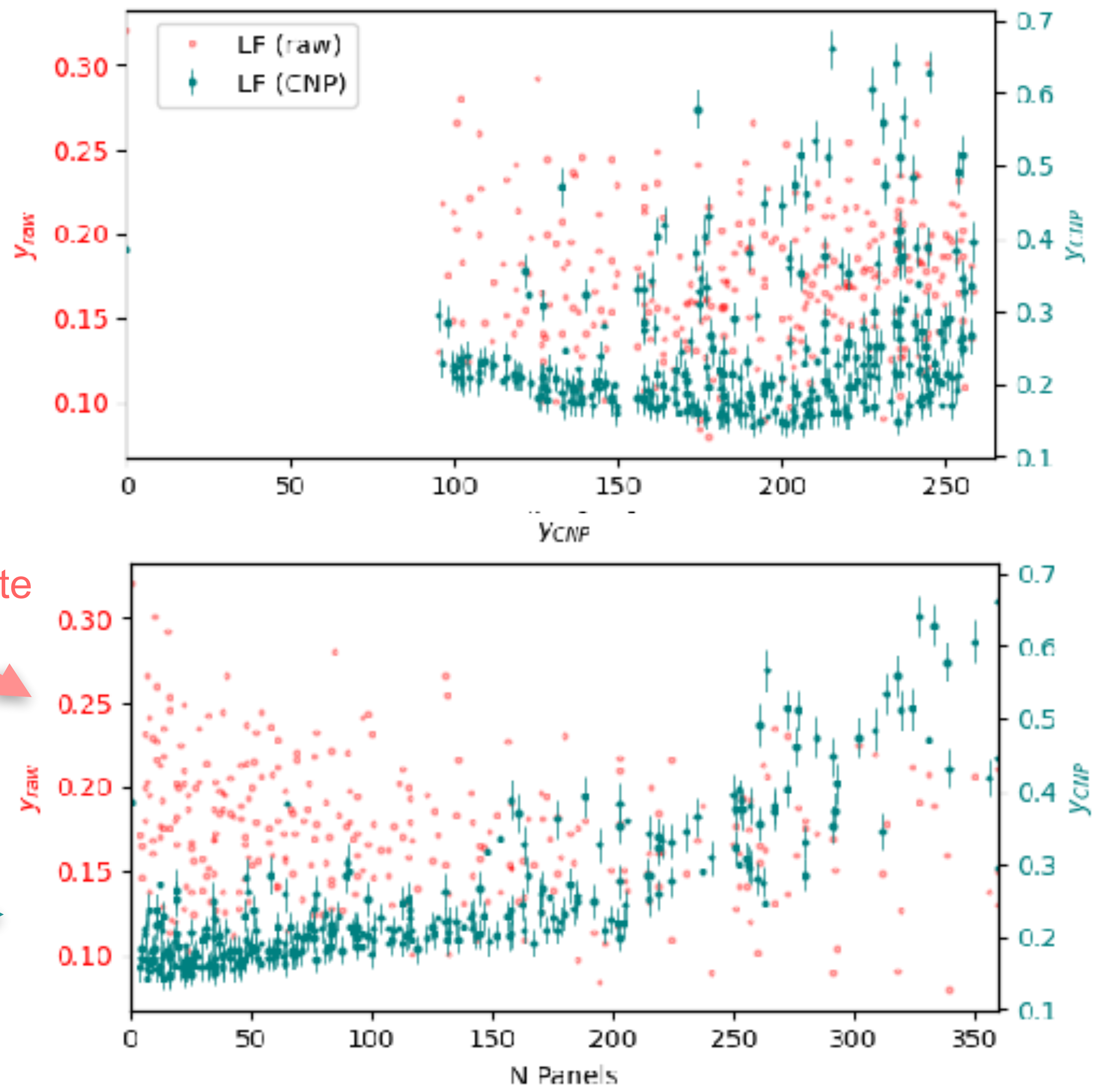
discrete

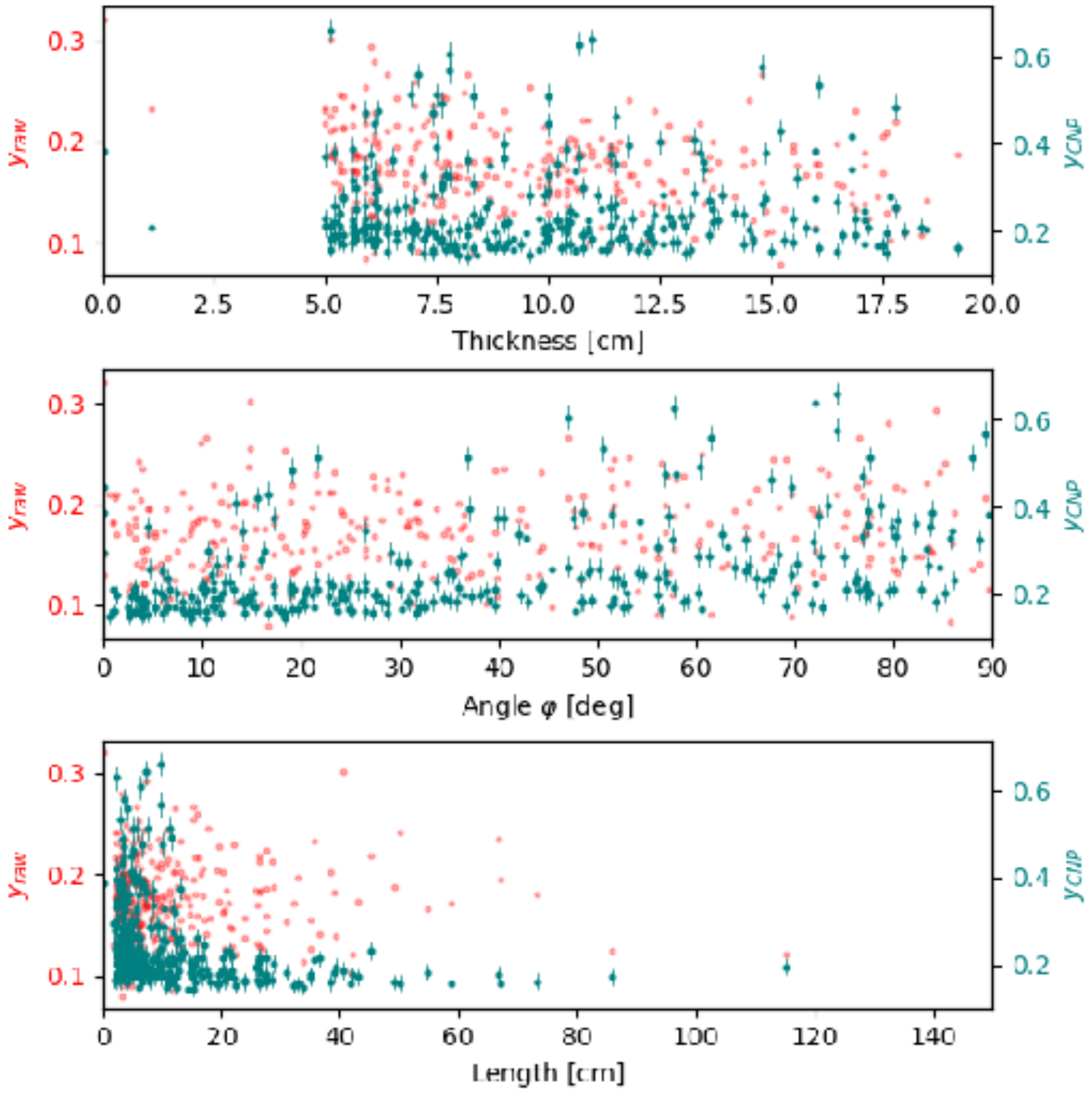
continuous

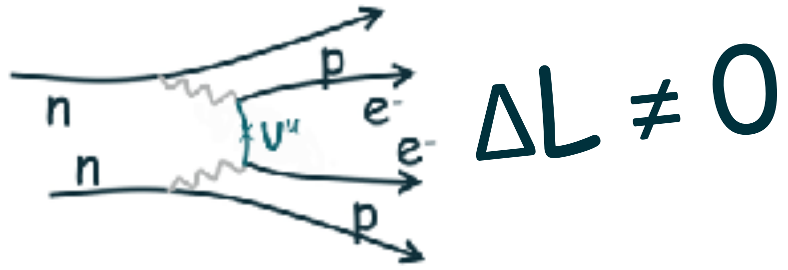
Neutron capture is shown as **1-dim projection**

→ **Strong fluctuations in the LF training data before**

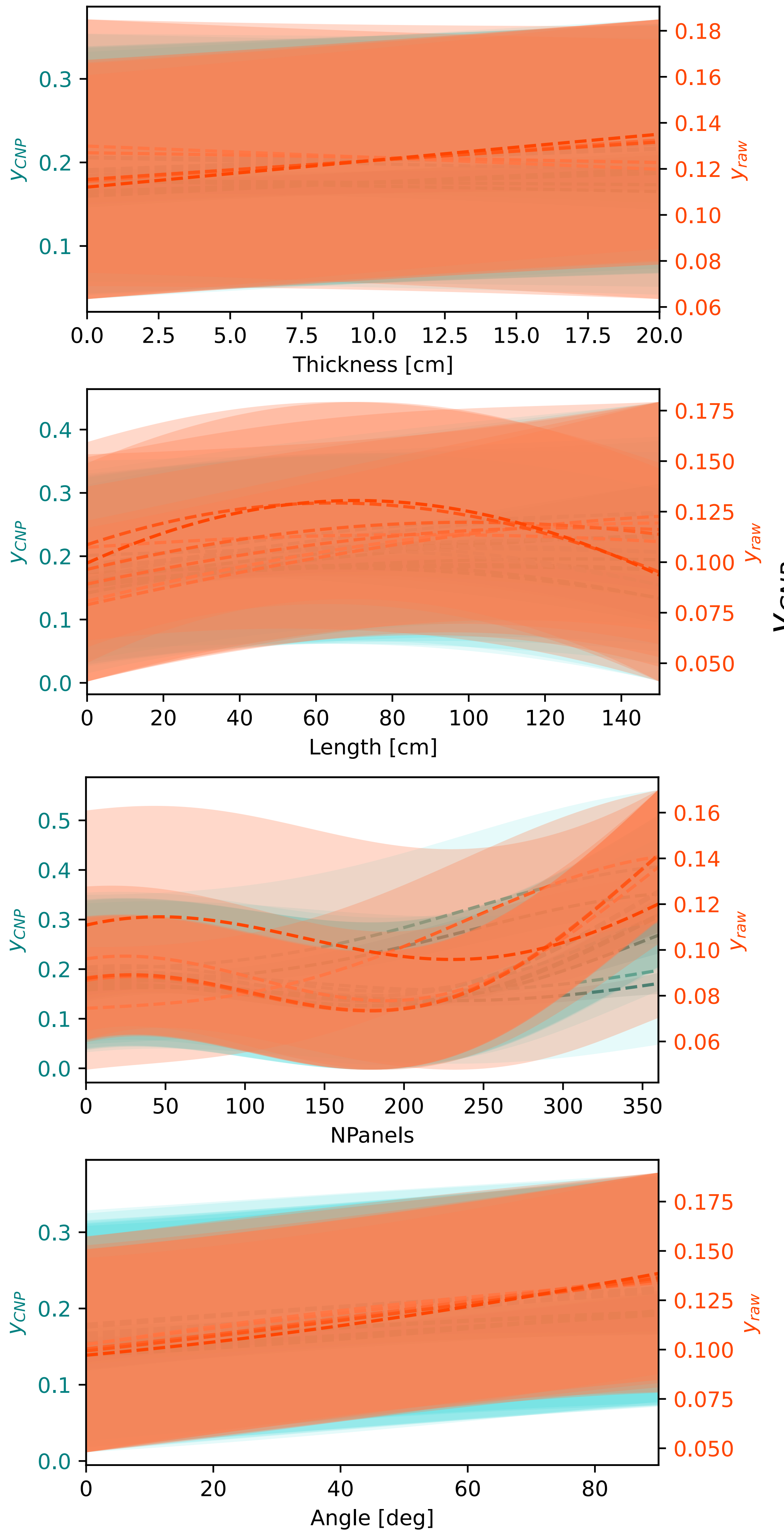
→ **Smoothening with CNP network**



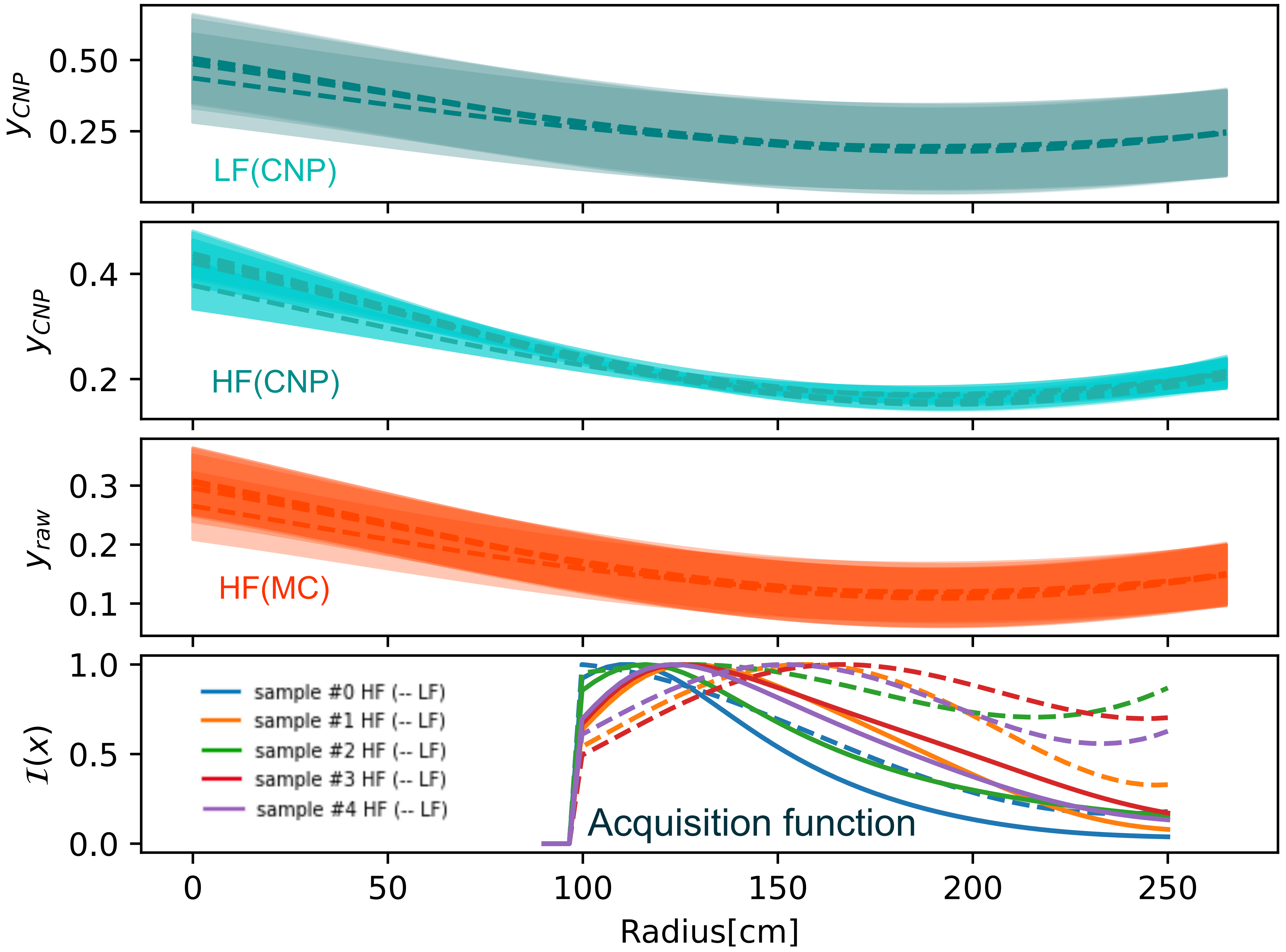


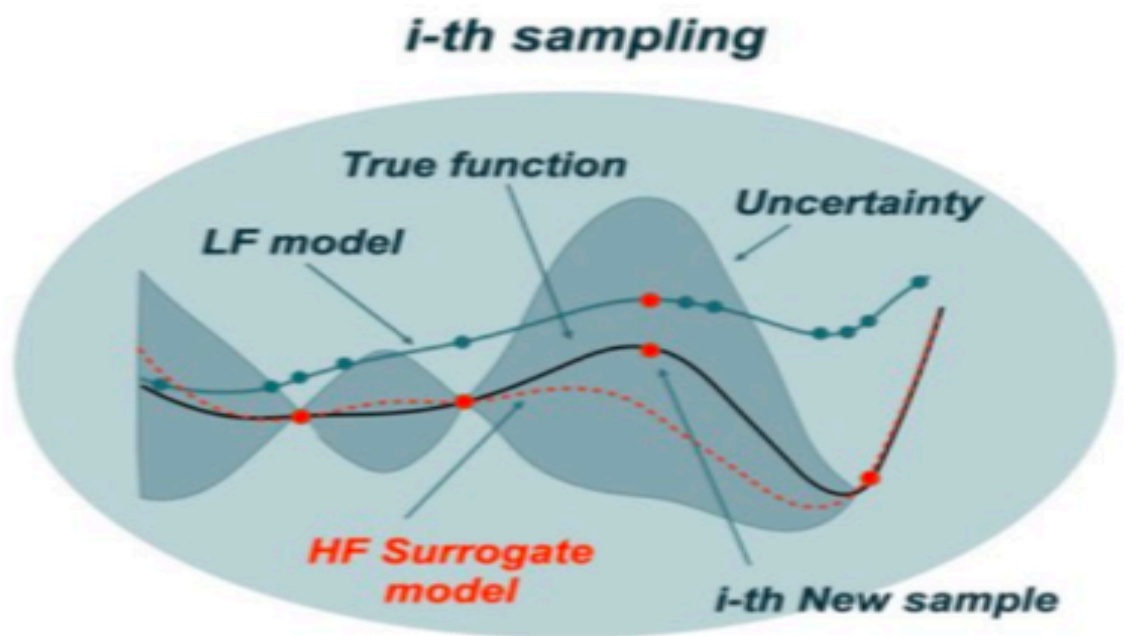



# RESuM - Results



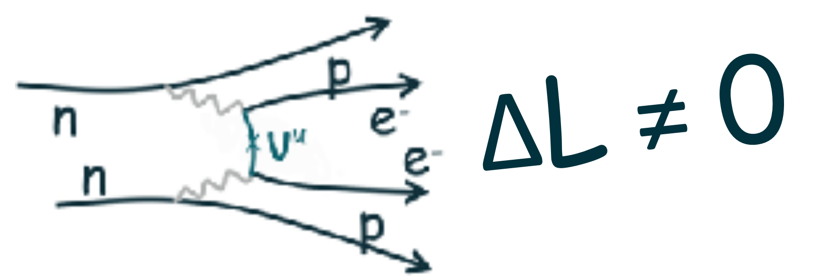

- Modeling of 5 dim space (r, t, $\theta$, n, L) with 3 fidelities (HF(MC), HF(CNP) and LF(CNP))
- model evolution shown as projection on r, t, n, $\theta$ and L at a random point in space
- Acquisition function: Integrated variance reduction with parameter constraints



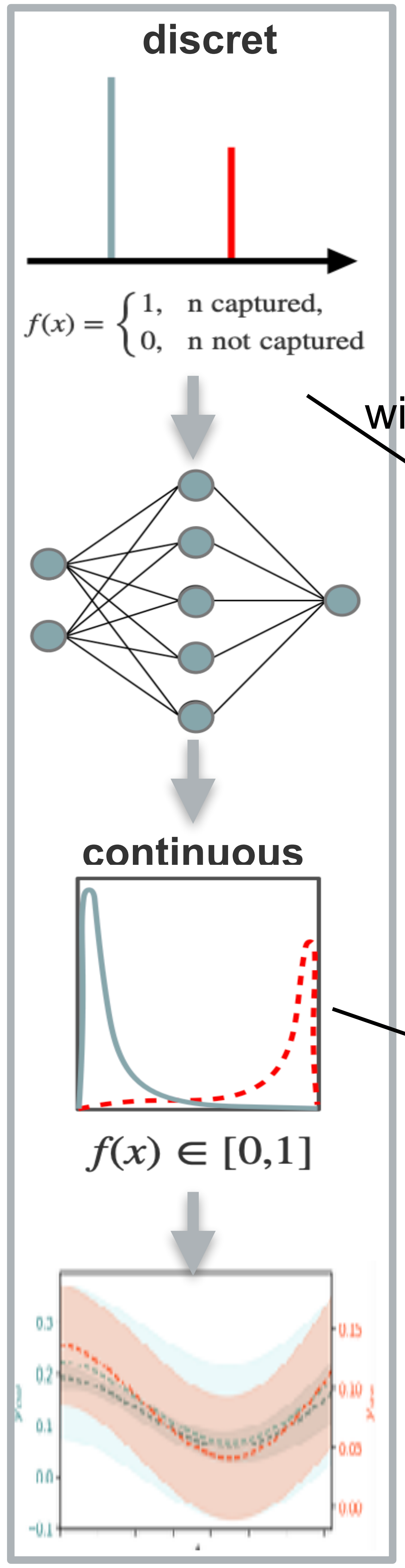

## Result & Conclusion

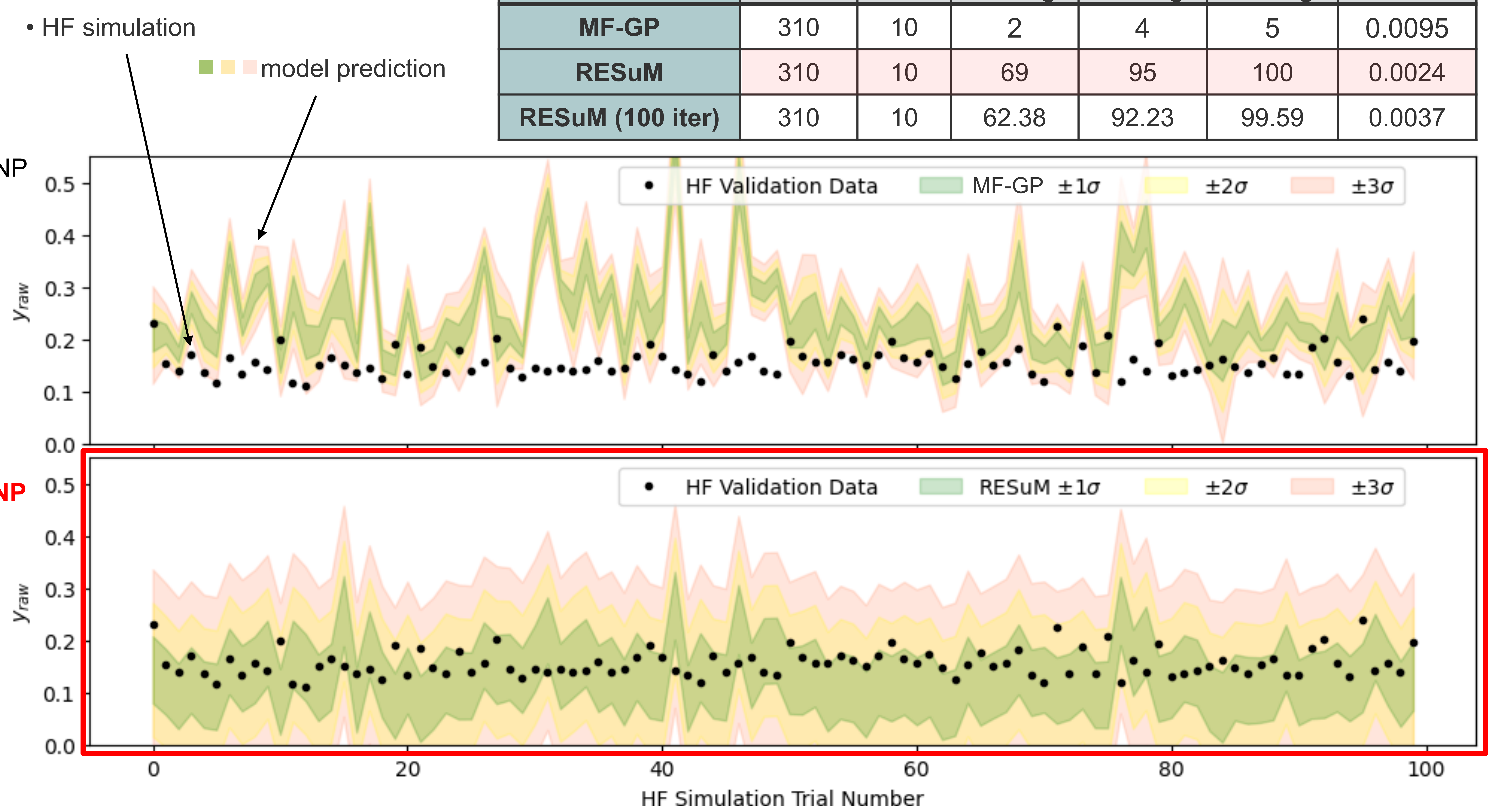
- **Impact:** Achieved a **66.5% reduction in neutron background** with **uncertainty predictions**
- **Efficiency**: Used **only 3.3% of the computational resources** compared to traditional method.

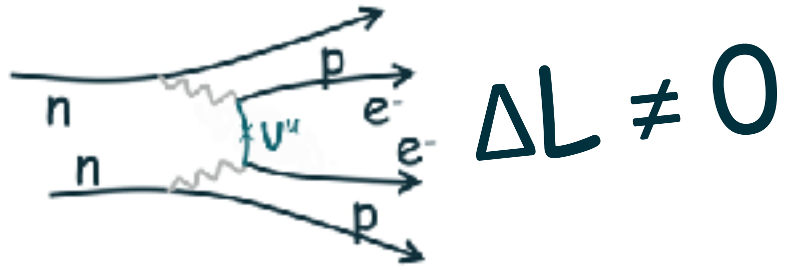



# Model Validation

| Model | LF | HF | 1σ coverage | 2σ coverage | 3σ coverage | MSE |
|---|---|---|---|---|---|---|
| MF-GP | 310 | 10 | 2 | 4 | 5 | 0.0095 |
| RESuM | 310 | 10 | 69 | 95 | 100 | 0.0024 |
| RESuM (100 iter) | 310 | 10 | 62.38 | 92.23 | 99.59 | 0.0037 |



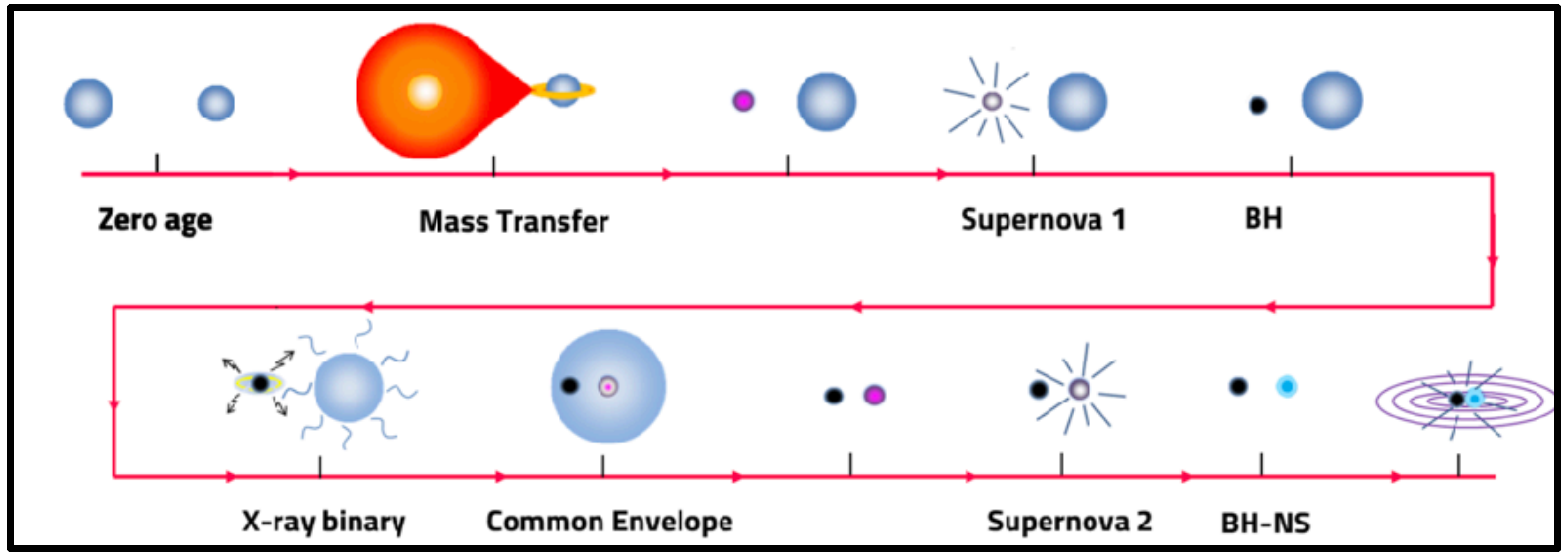






# Application I: Binary Black Hole Population Synthesis

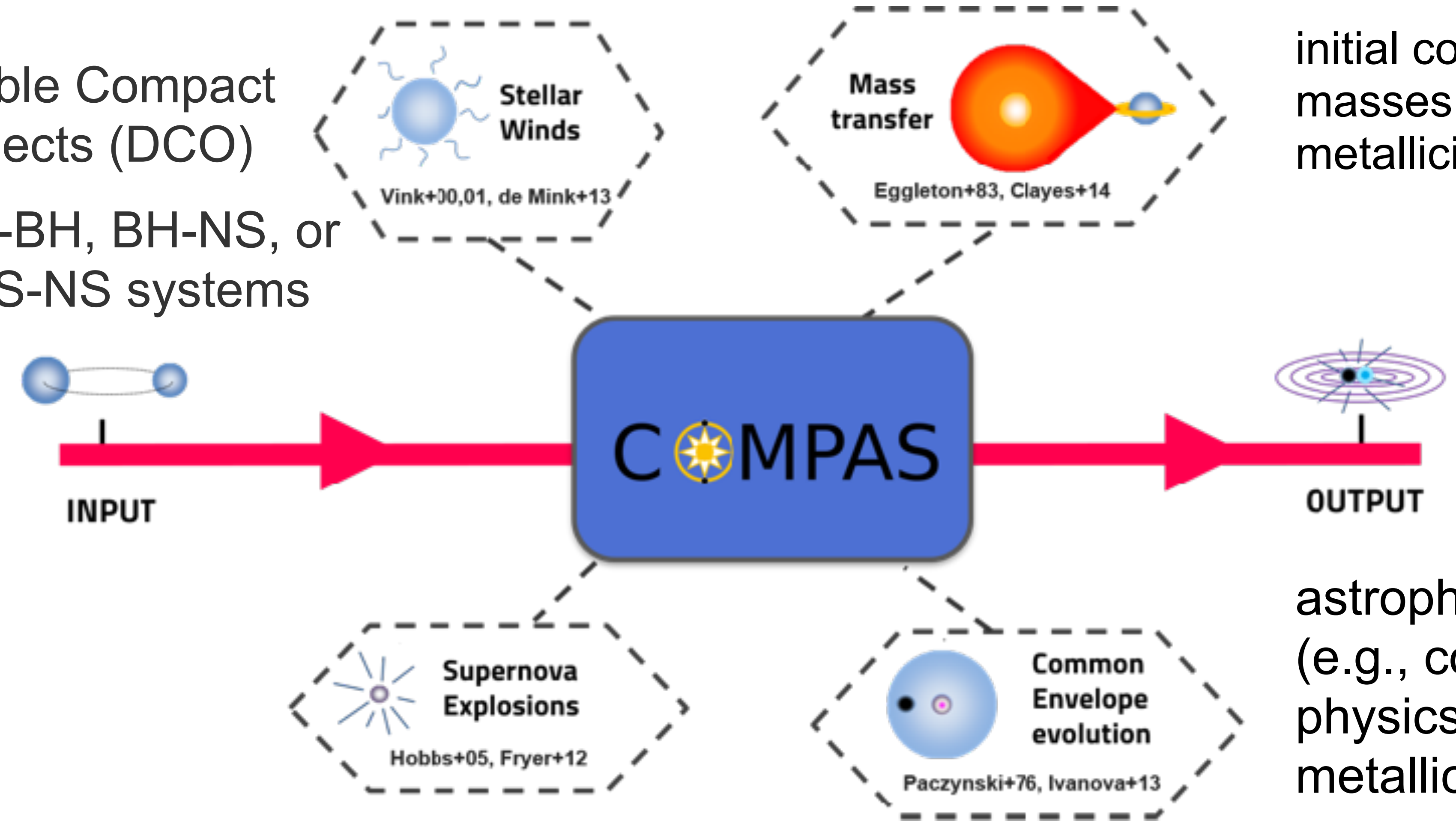
In Collaboration with **Prof. Floor Broeckgarden (UCSD)**

## Binary Black Hole Merger Simulation




Double Compact Objects (DCO)

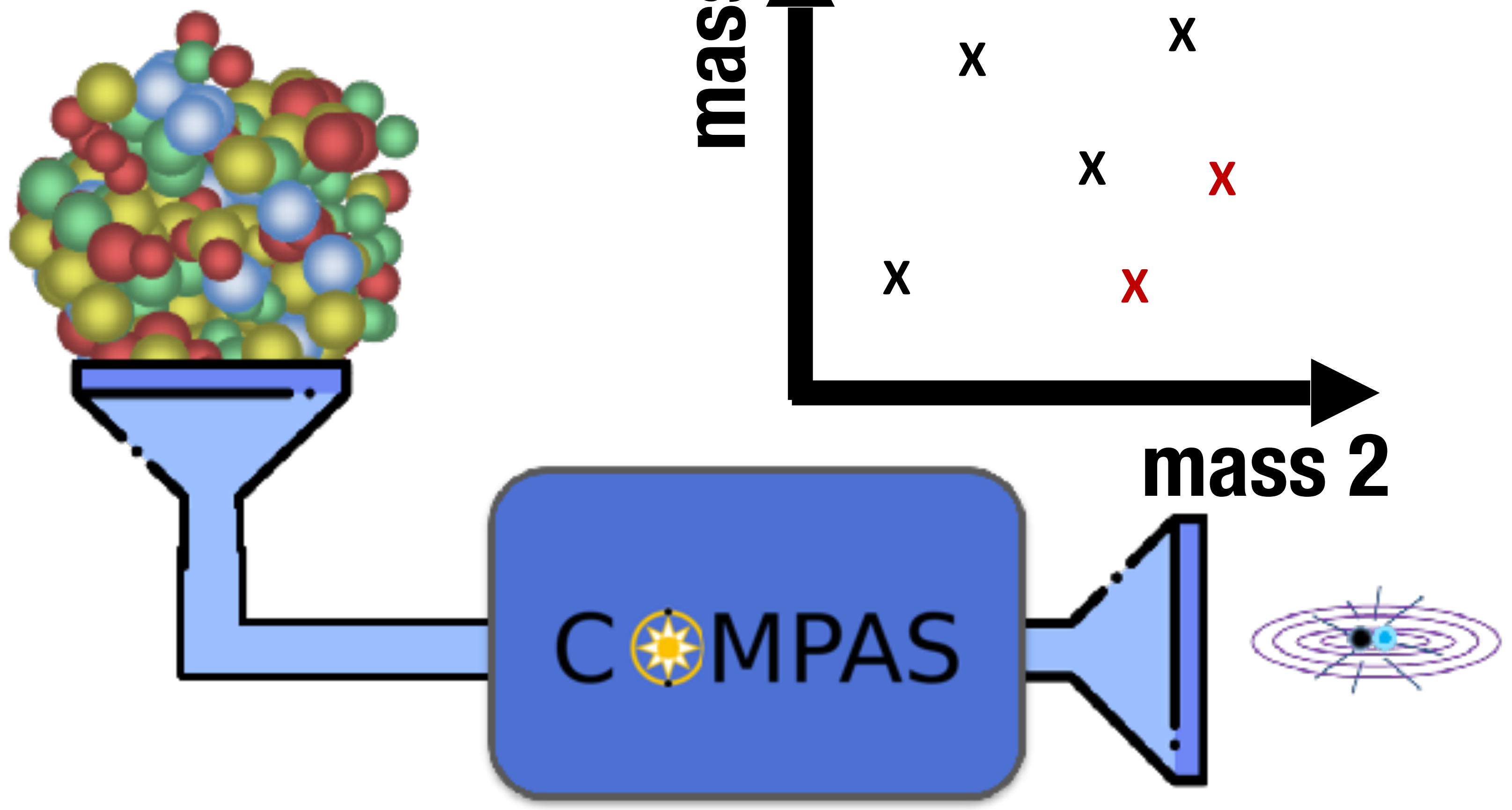
BH-BH, BH-NS, or NS-NS systems




initial conditions (e.g., masses, separations, metallicities)

astrophysical models (e.g., common-envelope physics, natal kicks, metallicity effects)

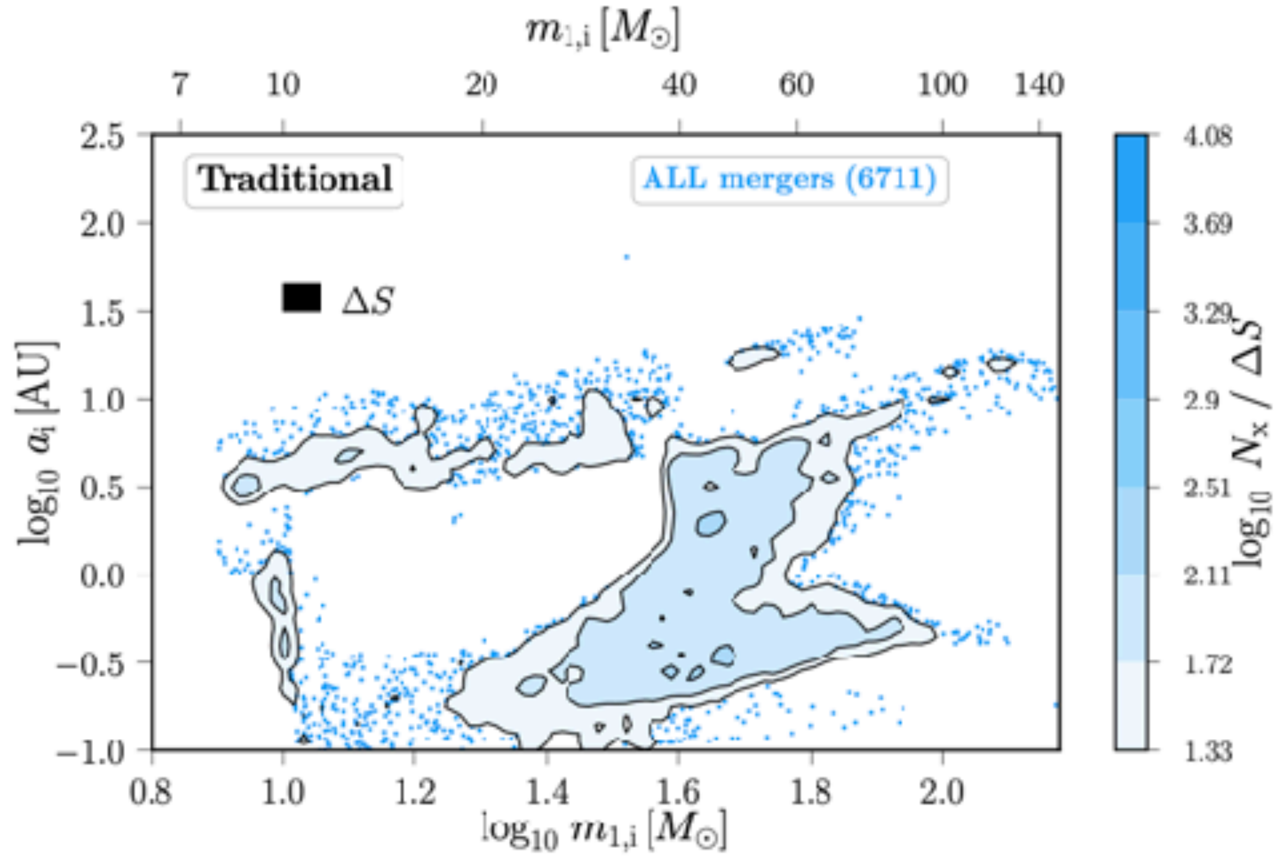
**1 000 000 binaries**




**only ~1 merger!**

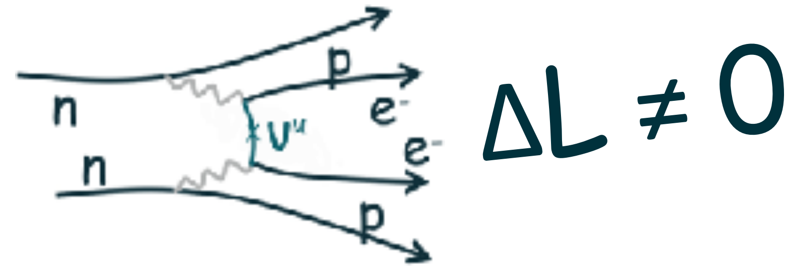
Leads to large Poisson (sampling) noise

- signal to background ratio 1:10⁶
- convergency only for N > 3·10⁸
- 180 CPUh
- high statistical uncertainties due to limited sample size
- most interesting gravitational wave sources occur in extreme tails of distr.



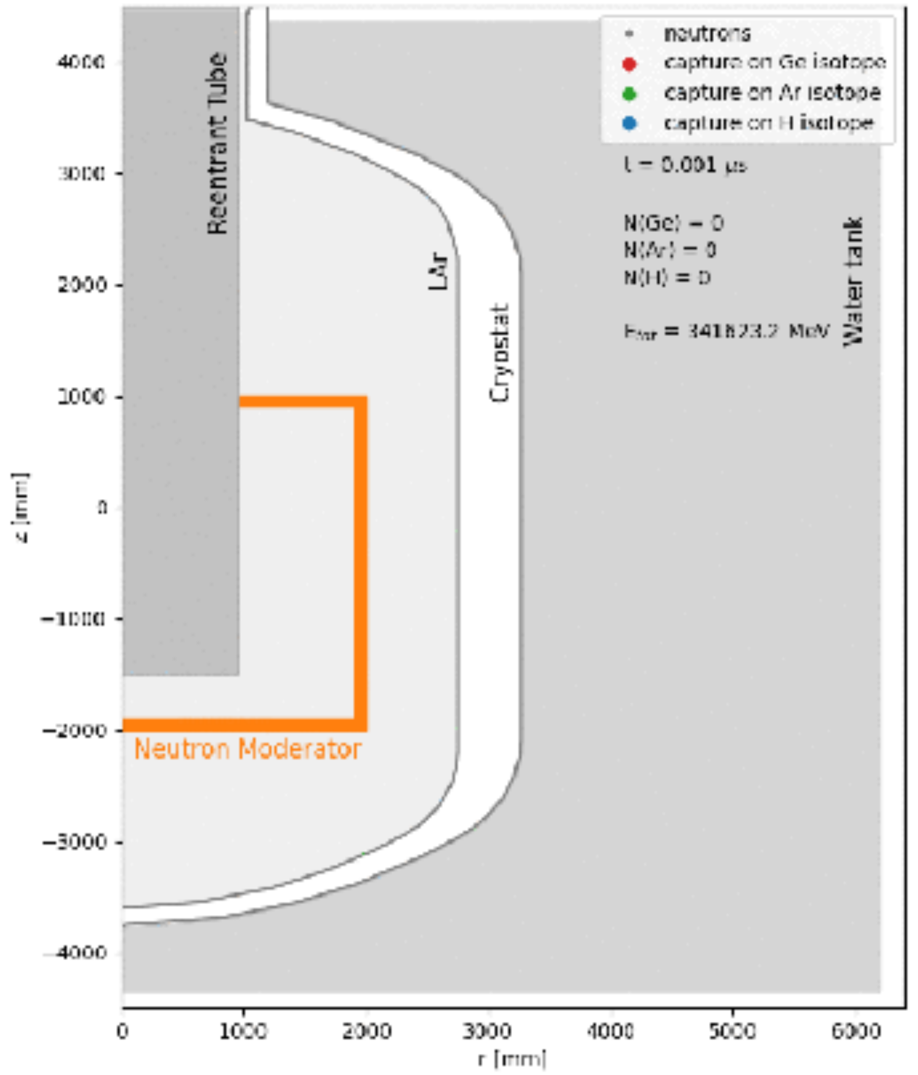

Goal: Higher resolution mapping of initial conditions leading to mergers.

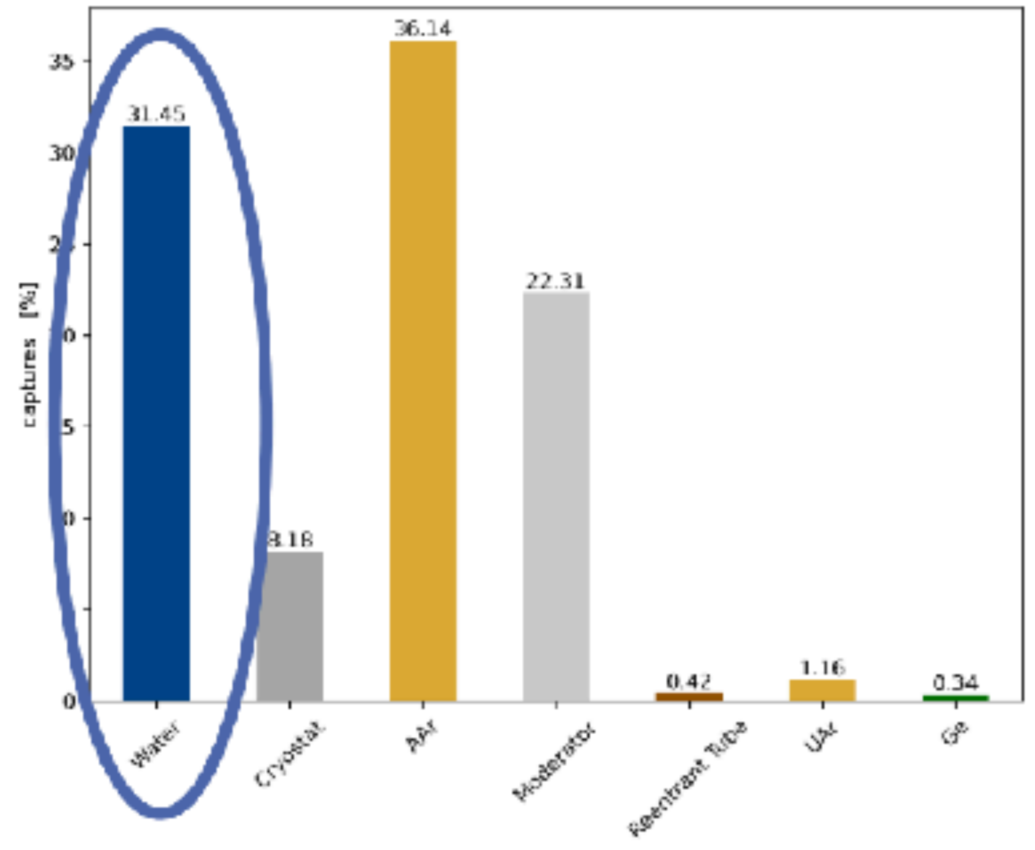
[Pictures from F. Broeckgarden]

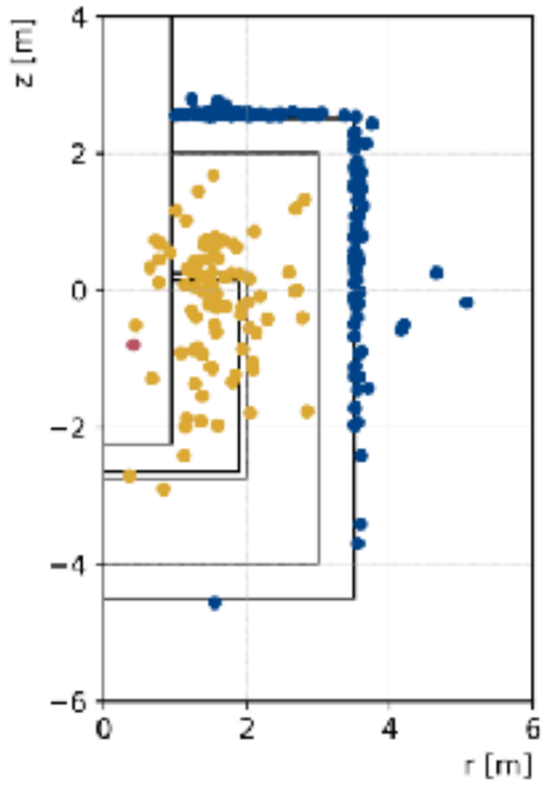



# Application II - Active Neutron Tagger

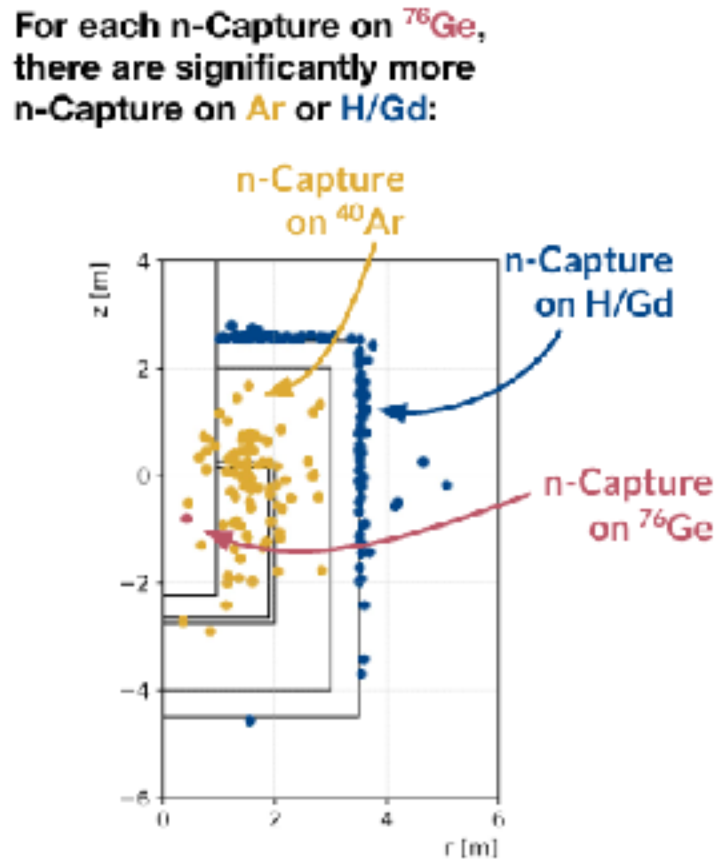
In Collaboration with **Prof. Josef Jochum (University of Tuebingen)**

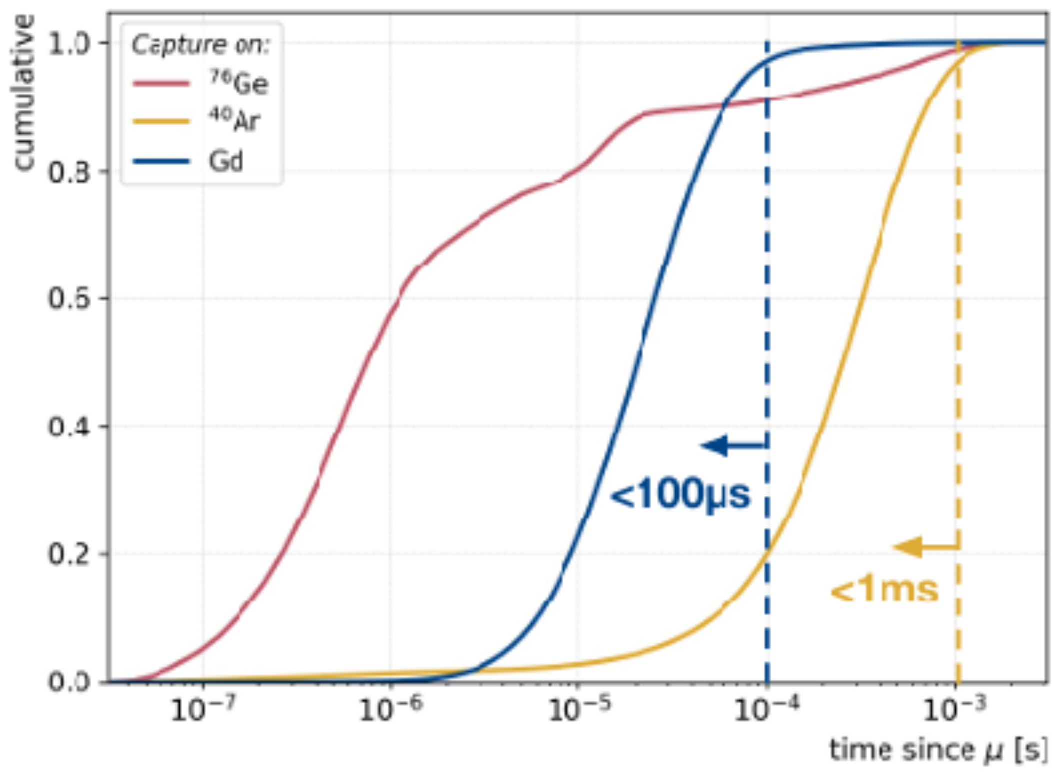


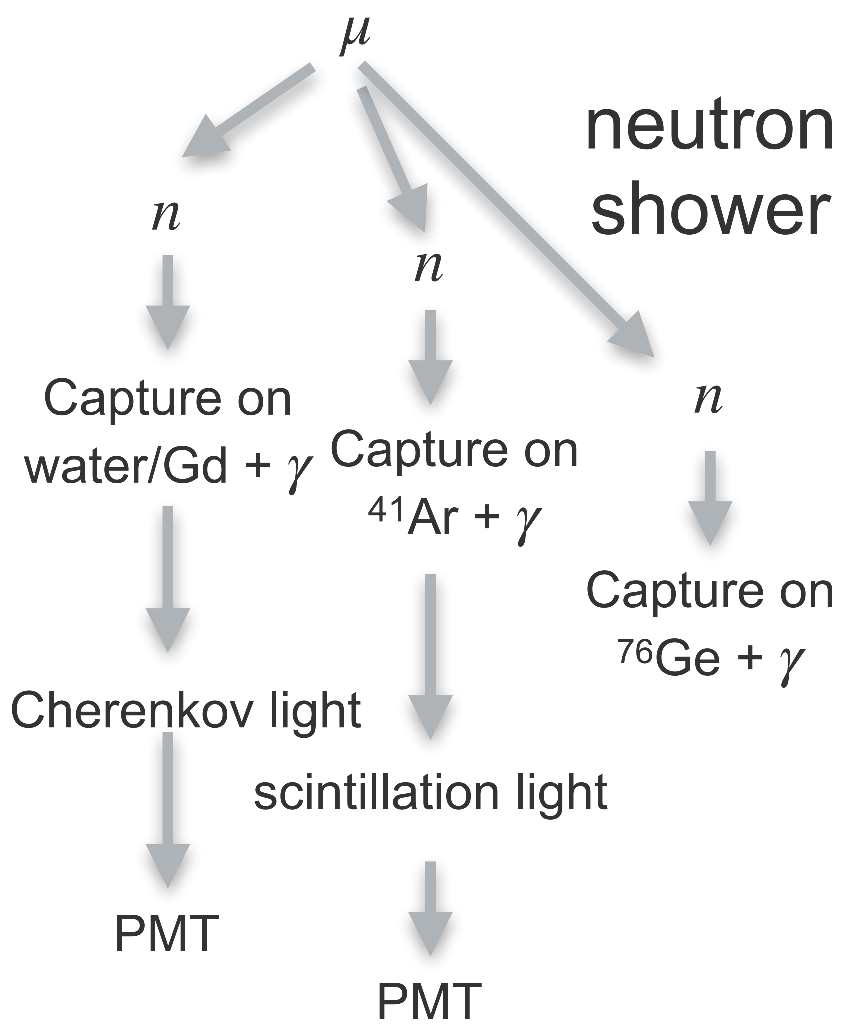




Cosmic-ray muons produce secondary neutron showers, some of which propagate into water or are created in coincidence within water



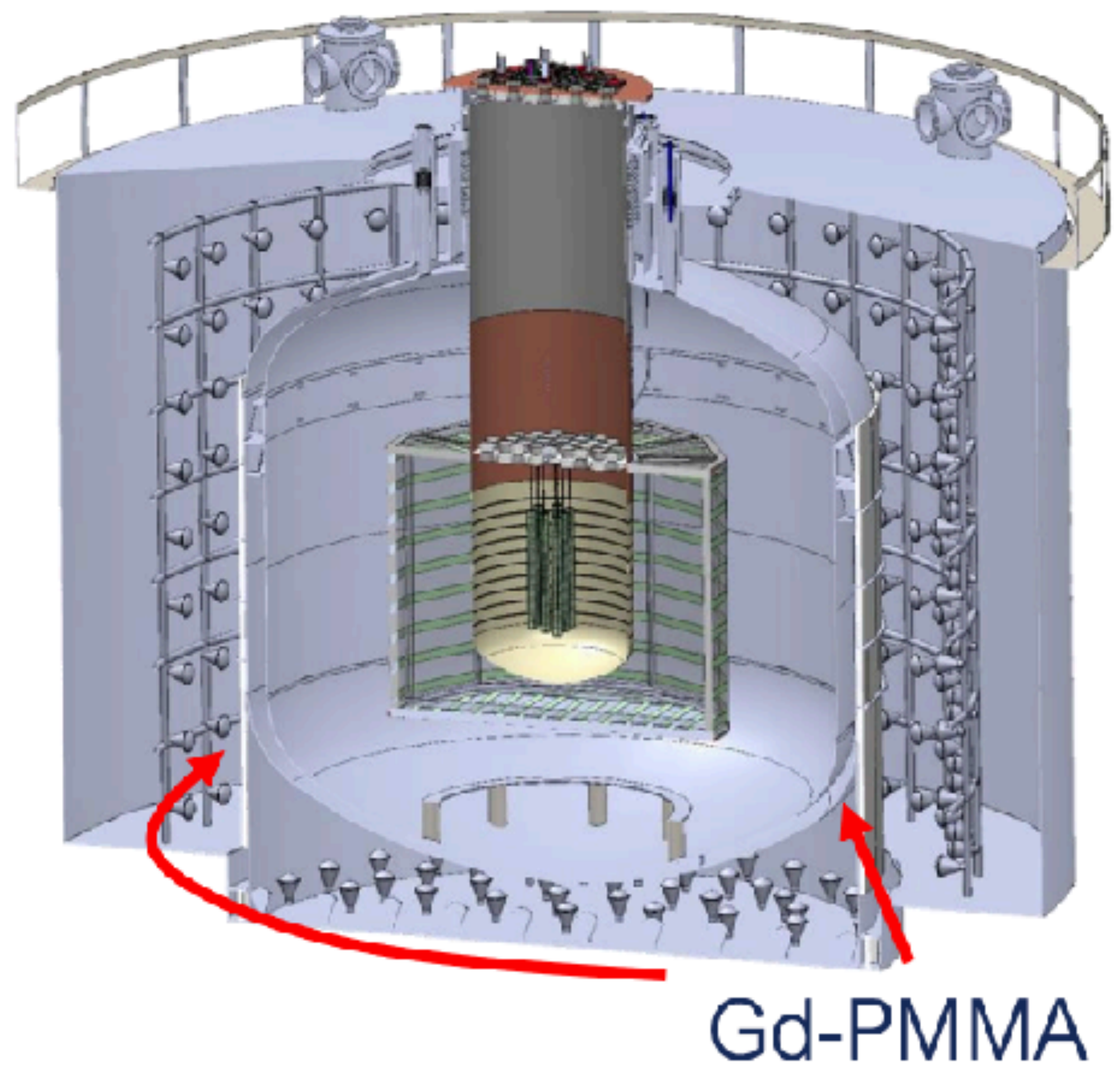



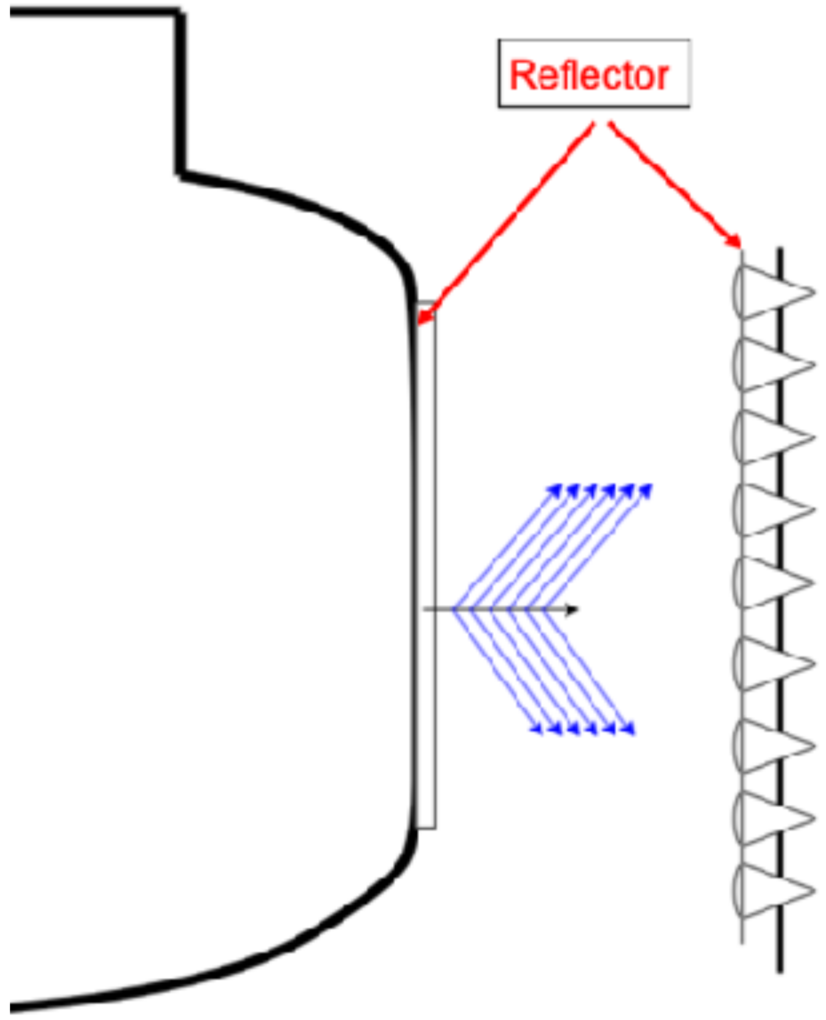




- Neutron capture by a nuclear (e.g. hydrogen, oxygen, dissolved gadolinium) with subsequent gamma emission
- The gamma rays from neutron capture Compton-scatter electrons, producing Cherenkov light that
- Detect Cherenkov light by an array of PMTs
- Veto signal to identify and reject neutron events

# Application II - Active Neutron Tagger

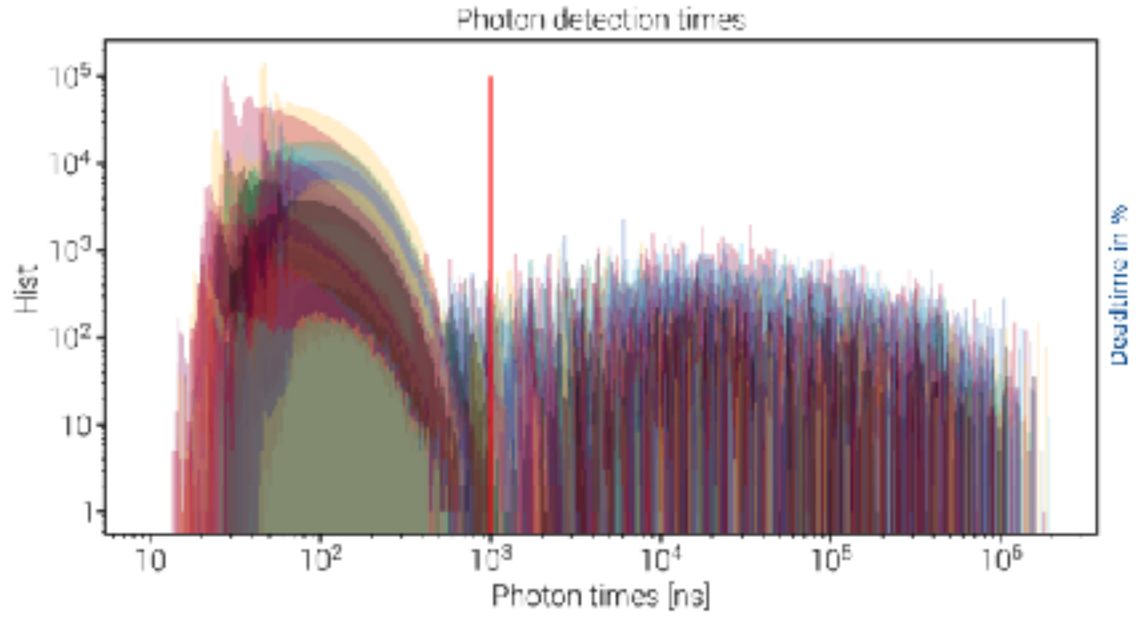
In Collaboration with **Prof. Josef Jochum (University of Tuebingen)**

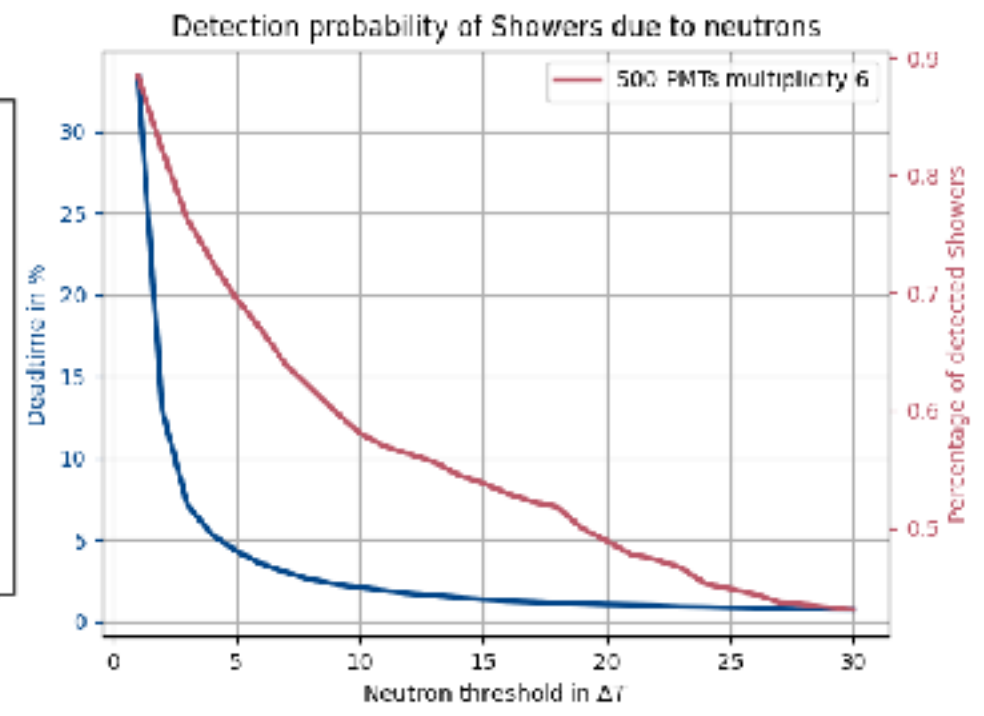



- How many PMTs? PMT distribution?
- what coverage ensures efficient light collection and spatial reconstruction?
- Timing, position, and energy information from PMTs?
- What is the probability/ rate of false positives?
- What is the efficiency vs. dead time tradeoff?

We want to use the approach of RESuM to help us answering these questions



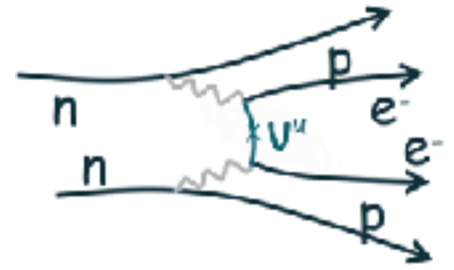




Neutron Tagger Design

- PMT distribution such that light yield maximal
- Minimal Dead Time (reduce false positives)

RESuM extensions:

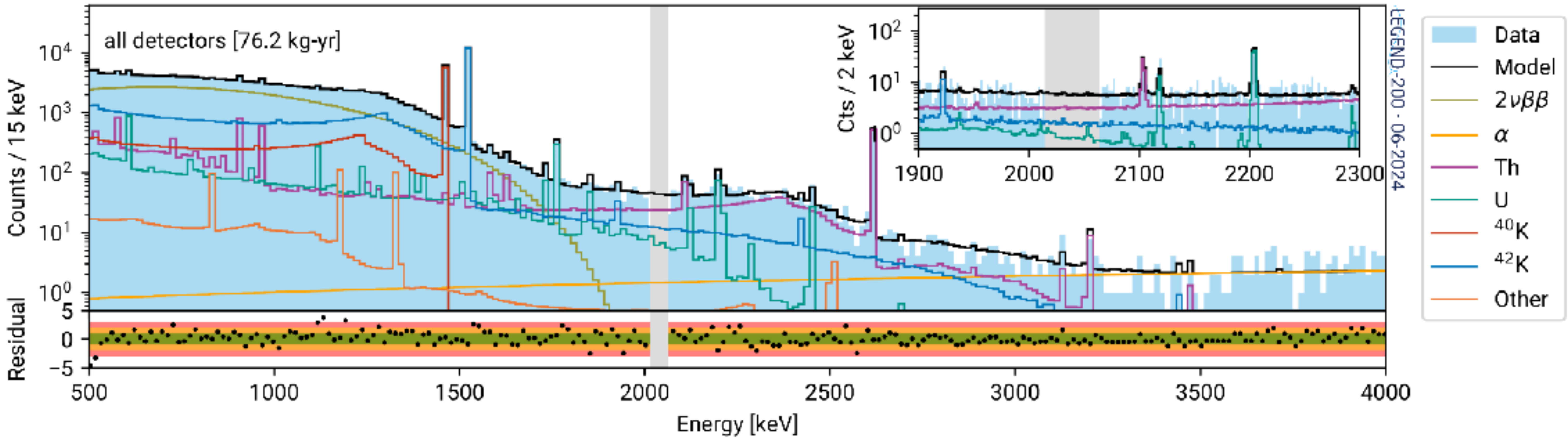
- Events are correlated → new network architecture needed
- Additional surrogate for optical properties
- Sequential multi-fidelity modeling
- multi-objective optimization




$\Delta L \neq 0$

# Application III: Spectral Decomposition + Anomaly detection

Binned likelihood fit to energy spectrum



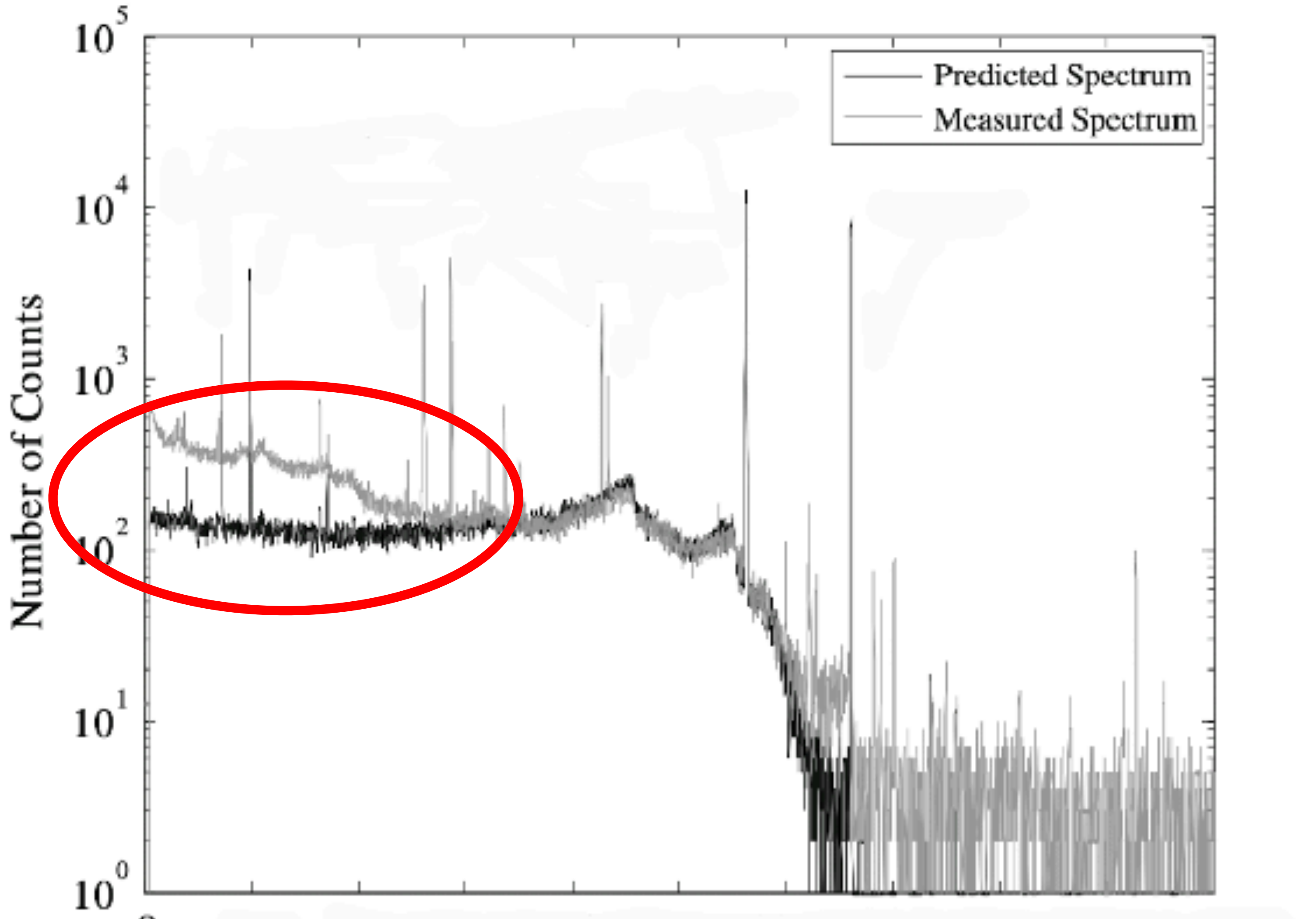

- Number of events in each bin is Poisson distributed
- Linear composition of individual contributions

The number of expected events in each bin:

$$\lambda_{d,i} = \sum_k \lambda_{d,i}^{(k)} \text{ with } \lambda_{d,i}^{(k)} = N_d^{(k)} \int_{\Delta E_i} \Phi_d^{(k)}(E)dE$$

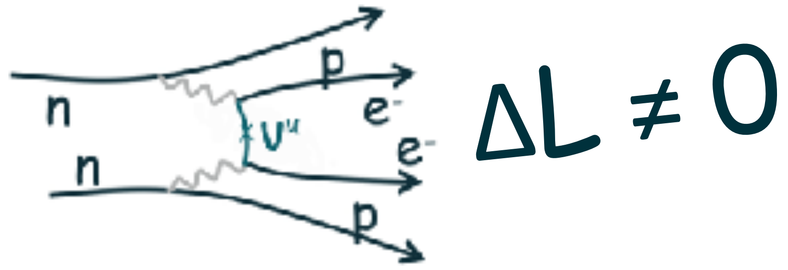
parameters of interest

Pdf = MC simulated spectrum




**Bottlenecks of widely used background modeling approach:**

- **Expensive Simulations:** Computational Cost & Scalability: Monte Carlo simulations are the standard but require massive computing resources
- **Source Degeneracy/ Ambiguity in Background Contribution:** Different isotopes and locations can produce similar energy spectra.
- **No Anomaly Detection/ Blind to Unexpected Backgrounds**: Traditional background modeling assumes **all contributions are known**. No mechanism to trace back outliers or systematic deviations in experimental data.



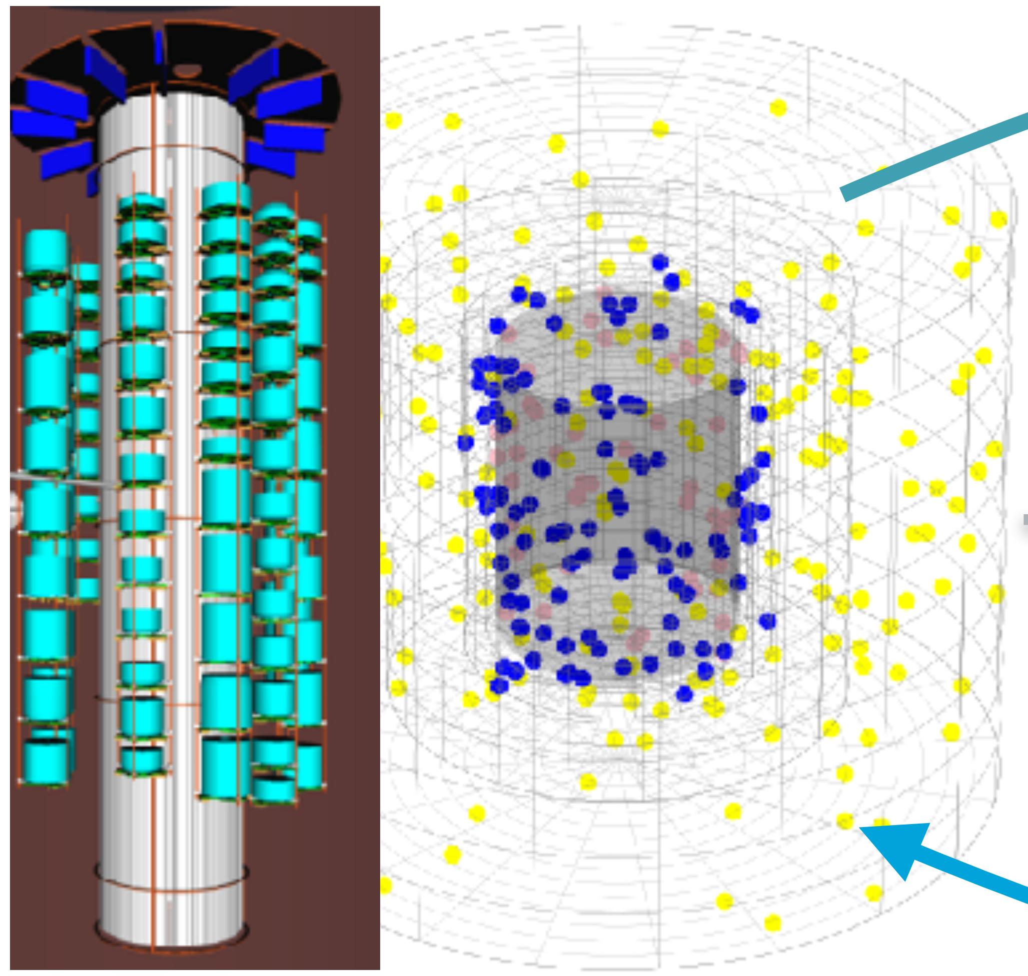

# Application III: Spectral Decomposition + Anomaly detection

New Methodology using RESuM concept

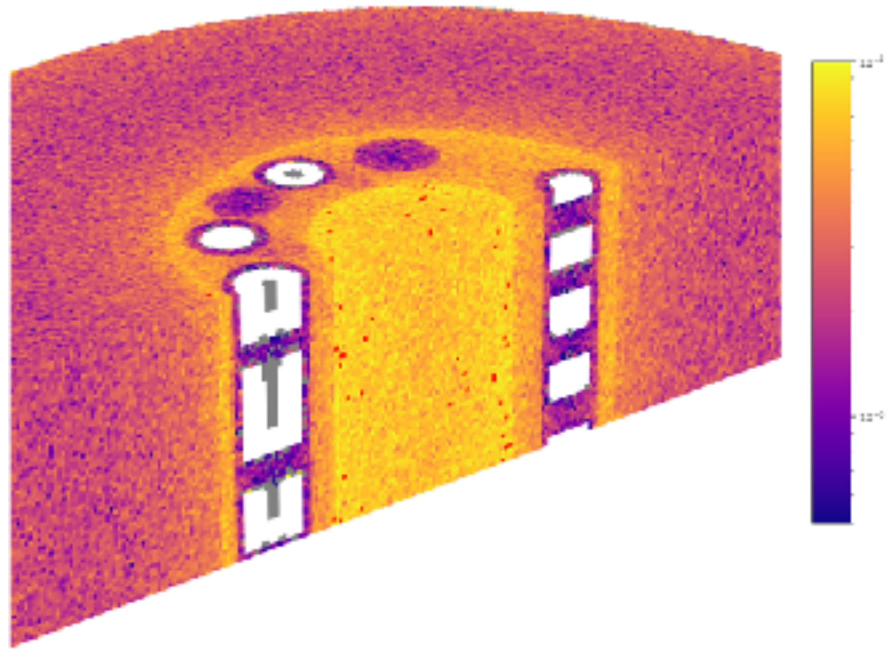
Multi-dimensional output to predict energy spectrum

Forward RESuM

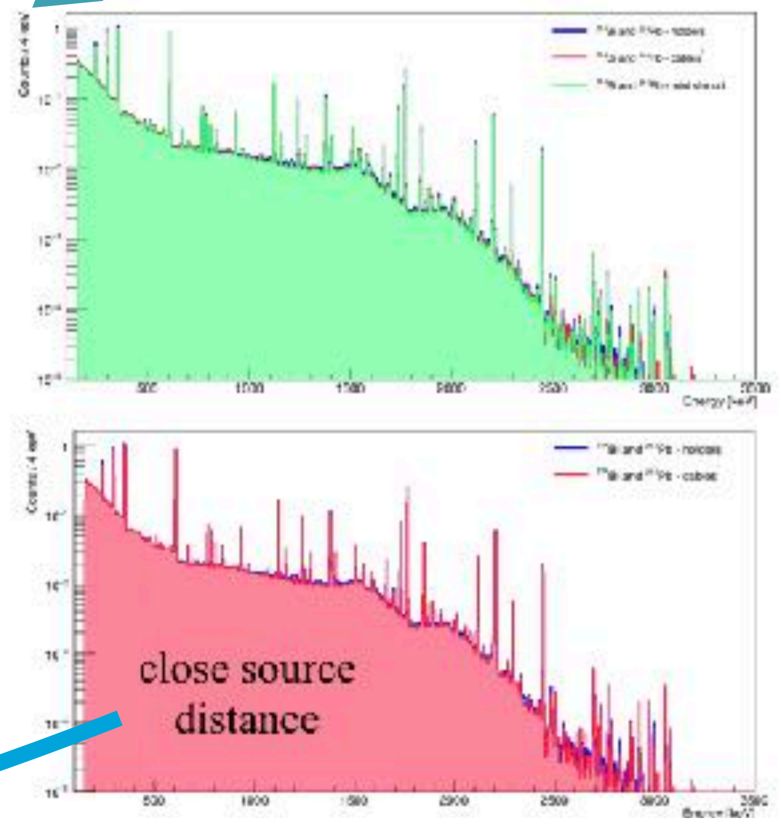
Background Modeling



Detection probability maps



Pdf.

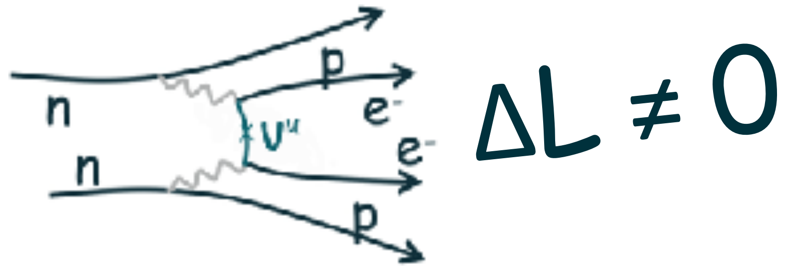



Anomaly detection

- Train network/model on energy-dependent shells for $\gamma, \beta$ and $\alpha$ particles

Inverse RESuM

Invertible network/model (VAE or Normalizing Flow)






Developed RESuM, a surrogate model optimized for rare event design (RED) problems in physics detector design

- Successfully optimized the neutron moderator design for the LEGEND experiment
- Reduced neutron-induced background by 66.5% while using only 3.3% of the computational resources required by traditional methods

- Achieved proper statistical coverage and robustness validated with independent simulations
- Incorporates Conditional Neural Processes (CNPs) for smoothing discrete design metrics
- Utilizes Multi-Fidelity Gaussian Processes (MFGPs) for efficient surrogate modeling
- Balances computational efficiency and accuracy with active learning strategies
- RED problems and the RESuM framework have potential applications (detector optimization, astronomy (e.g., BBH)

**Paper accepted at ICLR 2025!!**

A.Schuetz, A.W. Poon, A. Li arxiv:2410.03873