Searching for Rare Astrophysical Events with Rare AI

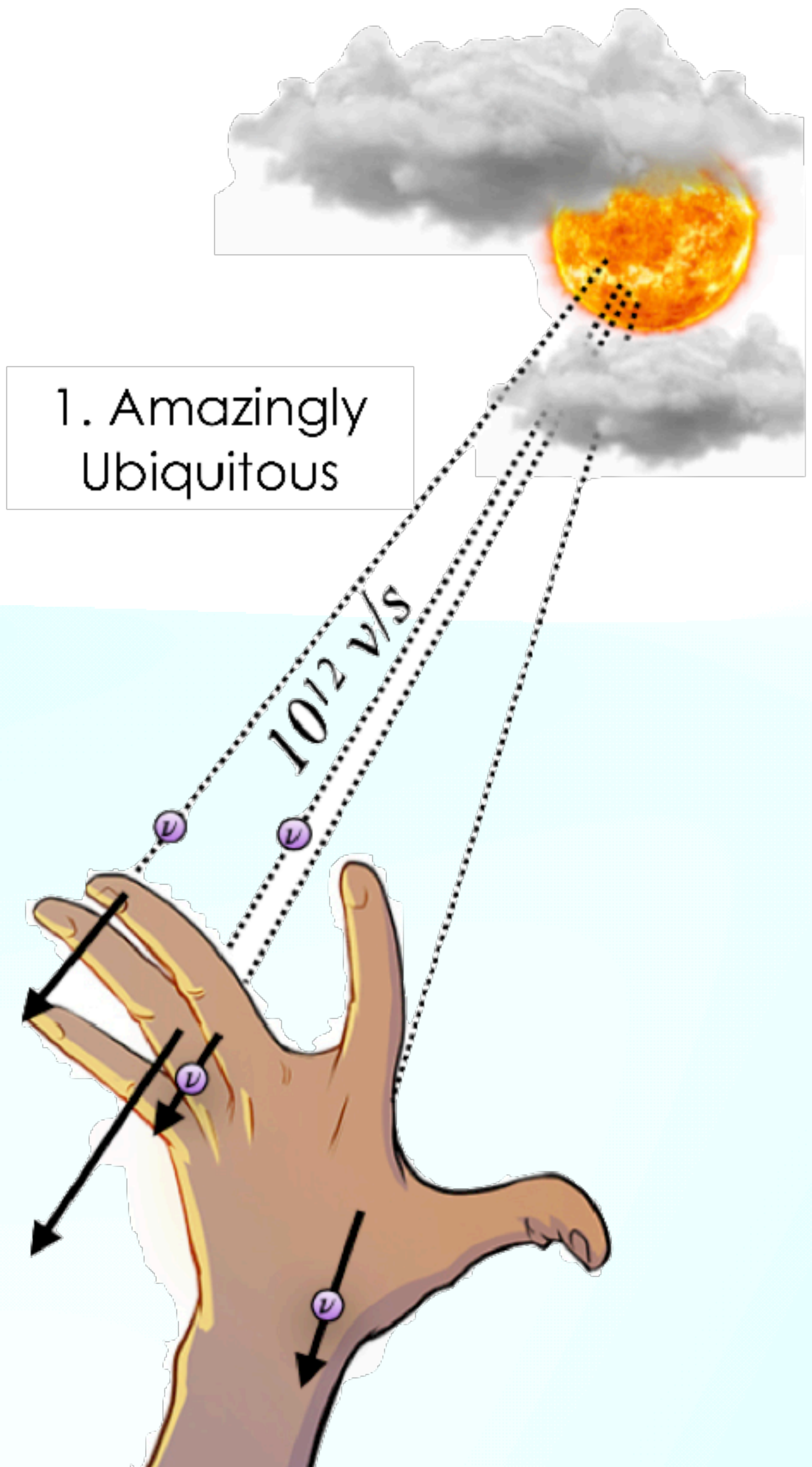
Aobo Li
Halıcıoğlu Data Science Institute
Department of Physics
UC San Diego

Workshop on Machine Learning for Analysis of High-Energy Cosmic Particles 06/19/2024

UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

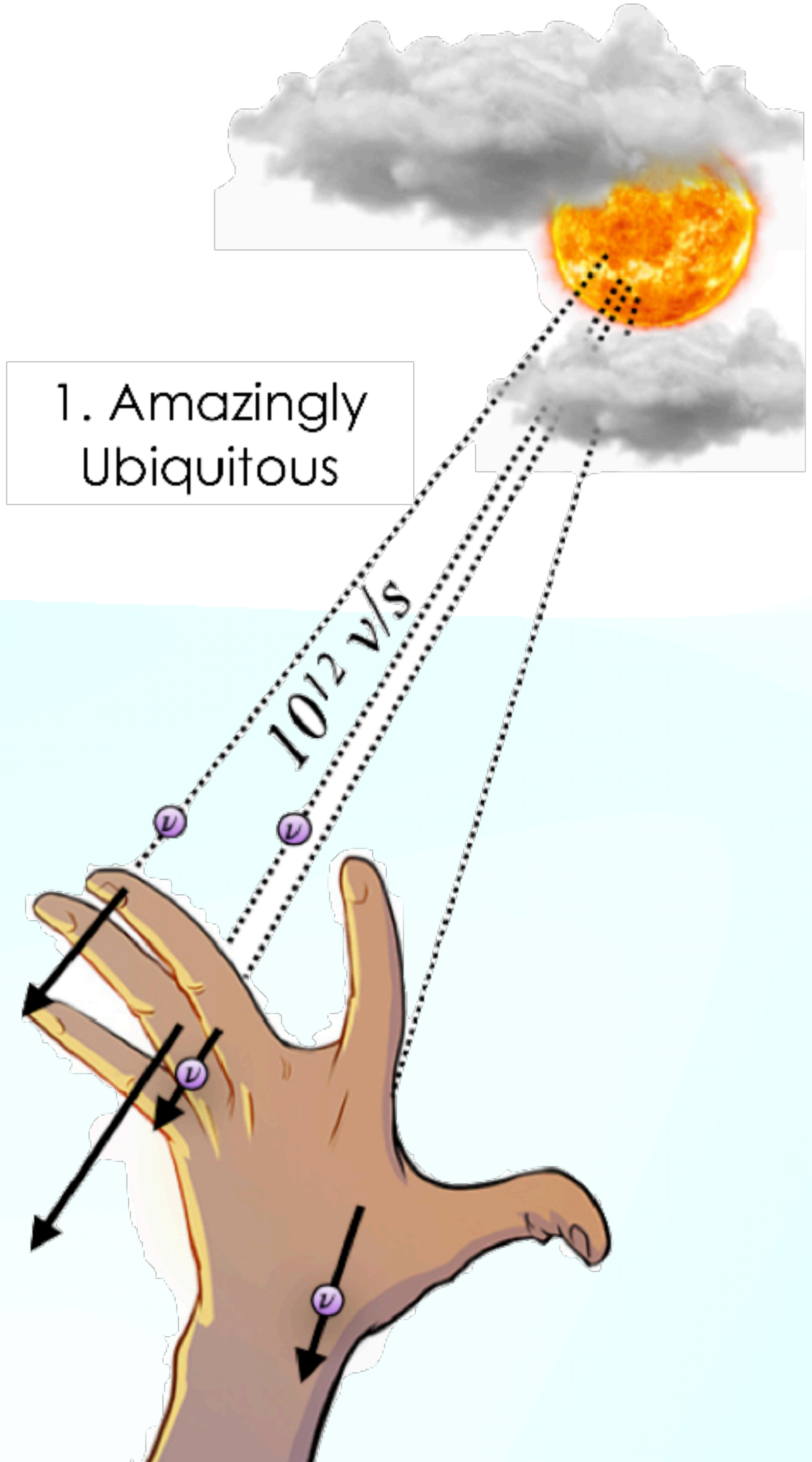
UC San Diego
PHYSICS

Naturally Occurring Rare Events



Naturally Occurring Rare Events

1. Amazingly Ubiquitous



Rare Event Search in 1950s

The Cowan-Reine Neutrino Experiment

First detection of neutrino (via inverse beta decay):

$$\bar{\nu}_e + p \rightarrow n + e^+$$

Extremely low cross section, but unique signature:

- $e^+ + e^- \rightarrow 2\gamma$
- Neutron capture γ

Nobel Prize of 1995

Rare Event Search in 2025

Rare Event Search in 2025

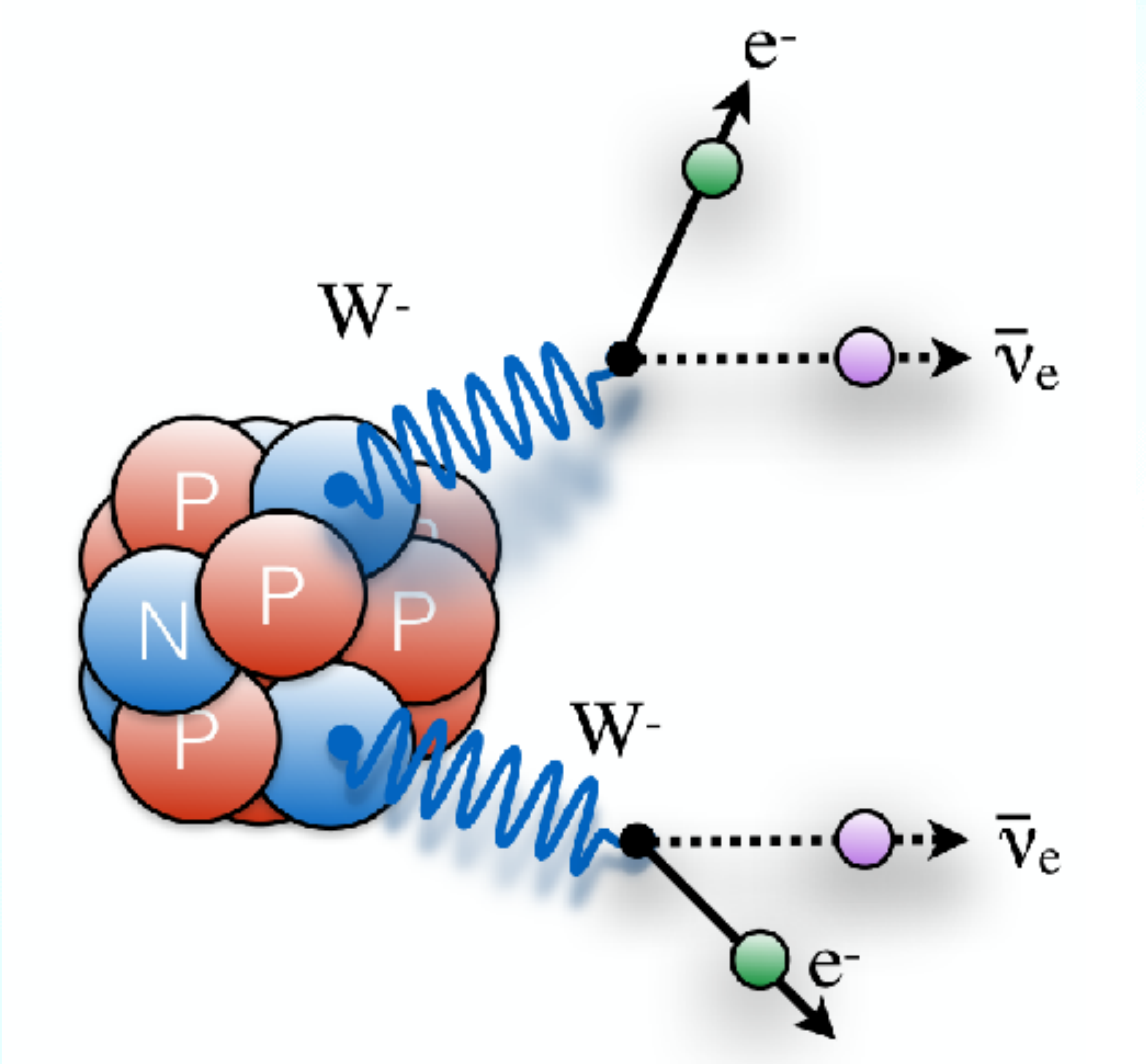
Double Beta Decay ($2\upsilon\beta\beta$)

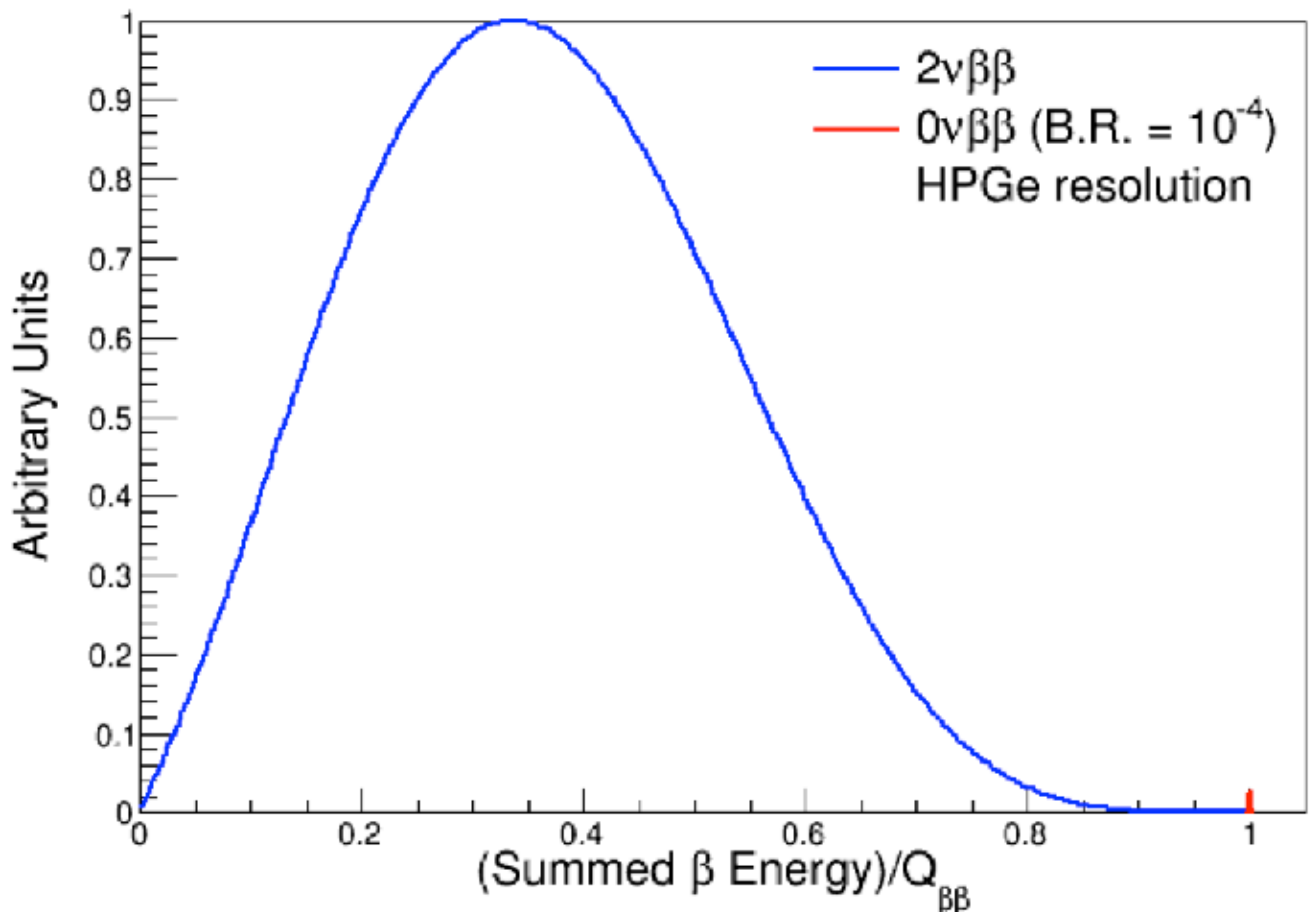
First proposed by Maria Goeppert Mayer in 1935

First detection by Elliott, Hahn, Moe, in 1987

Decay half-life $T_{\frac{1}{2}} \sim 10^{14} - 10^{24} yrs$

Much longer than the age of universe!





Rare Event Search in 2025

Double Beta Decay ($2\upsilon\beta\beta$)

First proposed by Maria Goeppert Mayer in 1935

First detection by Elliott, Hahn, Moe, in 1987

Decay half-life $T_{\frac{1}{2}} \sim 10^{14} - 10^{24} yrs$

Much longer than the age of universe!

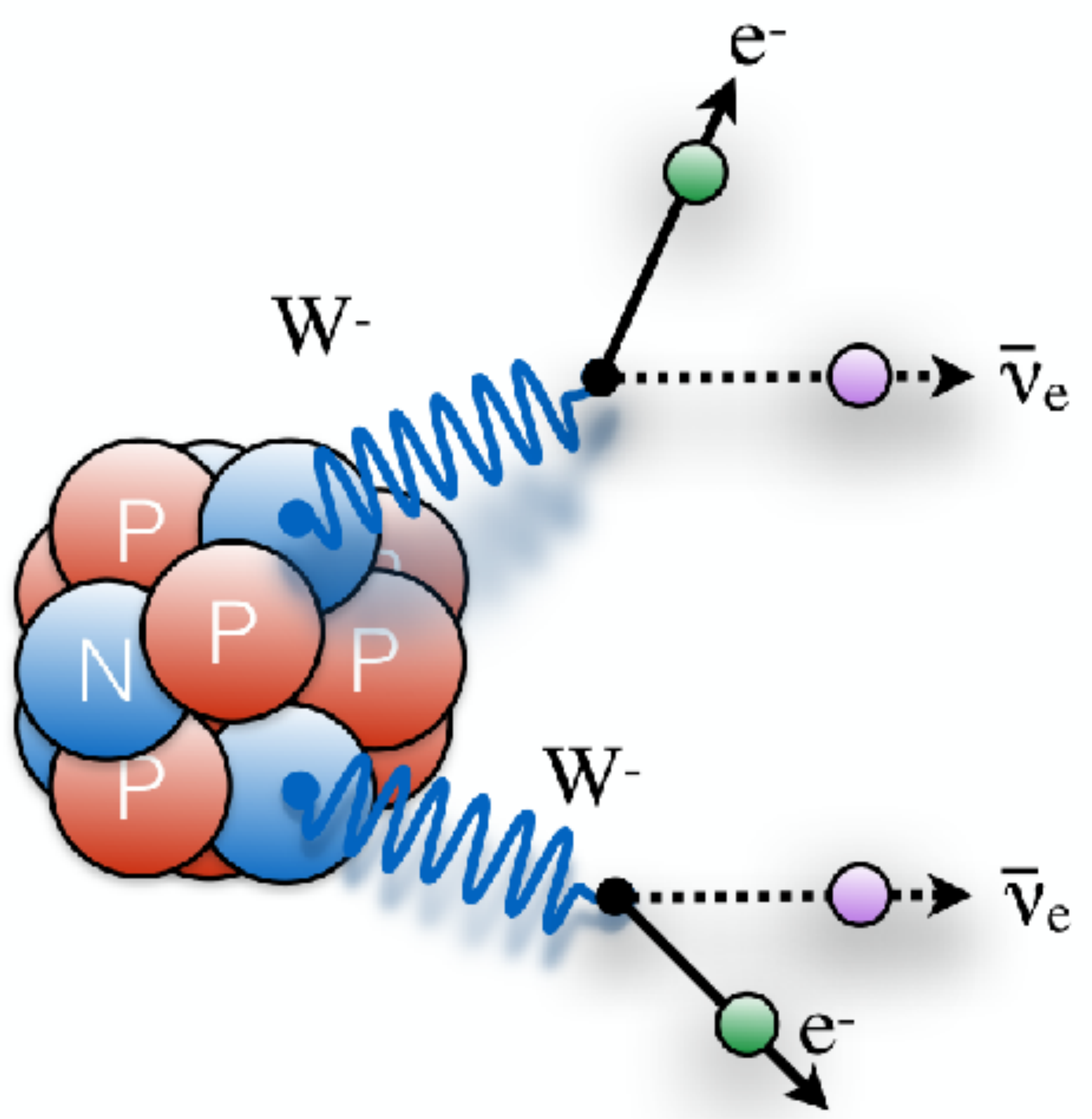
Neutrinoless Double-Beta Decay ($0\upsilon\beta\beta$)

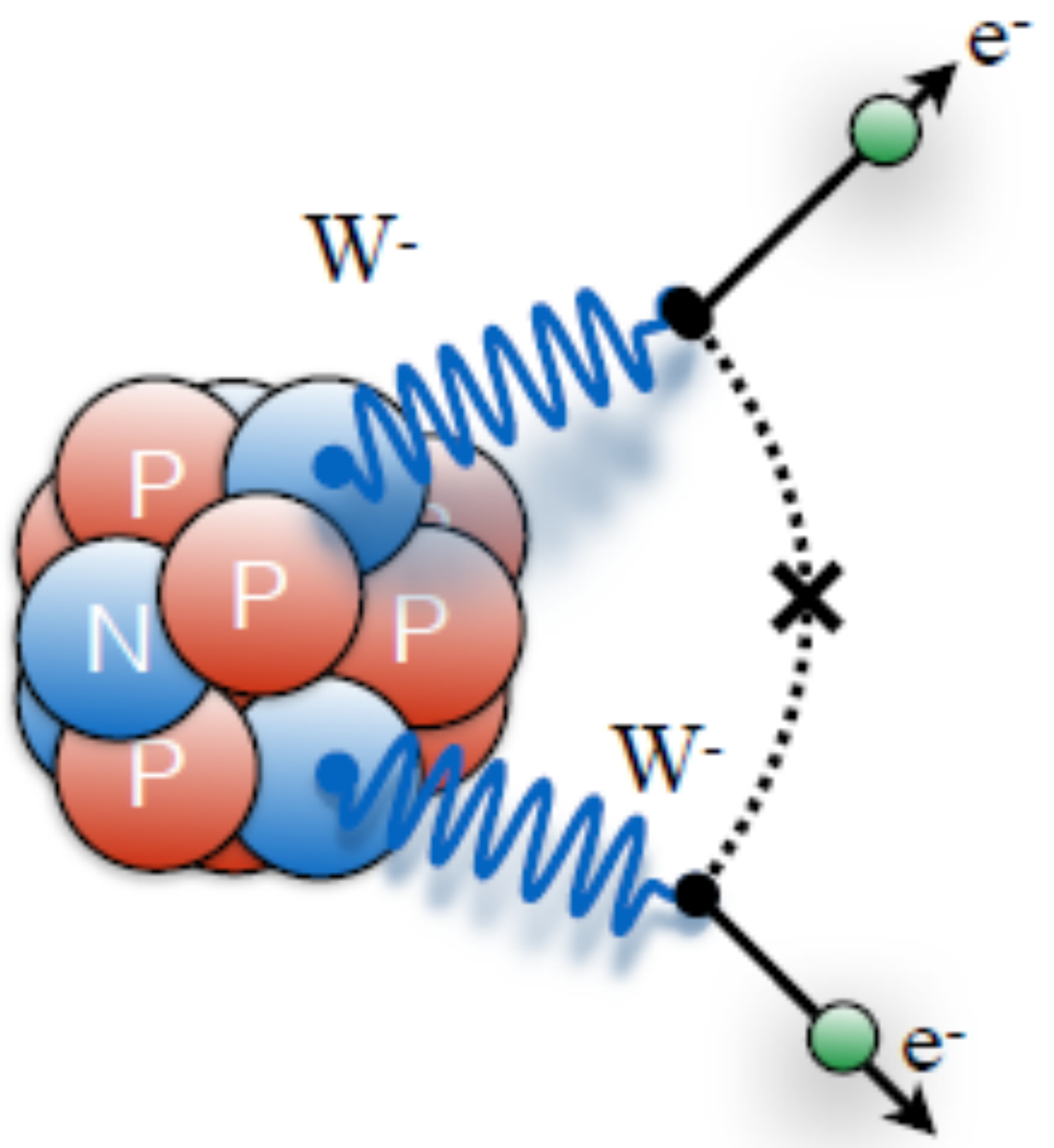
ΔL = 2 lepton number violation process

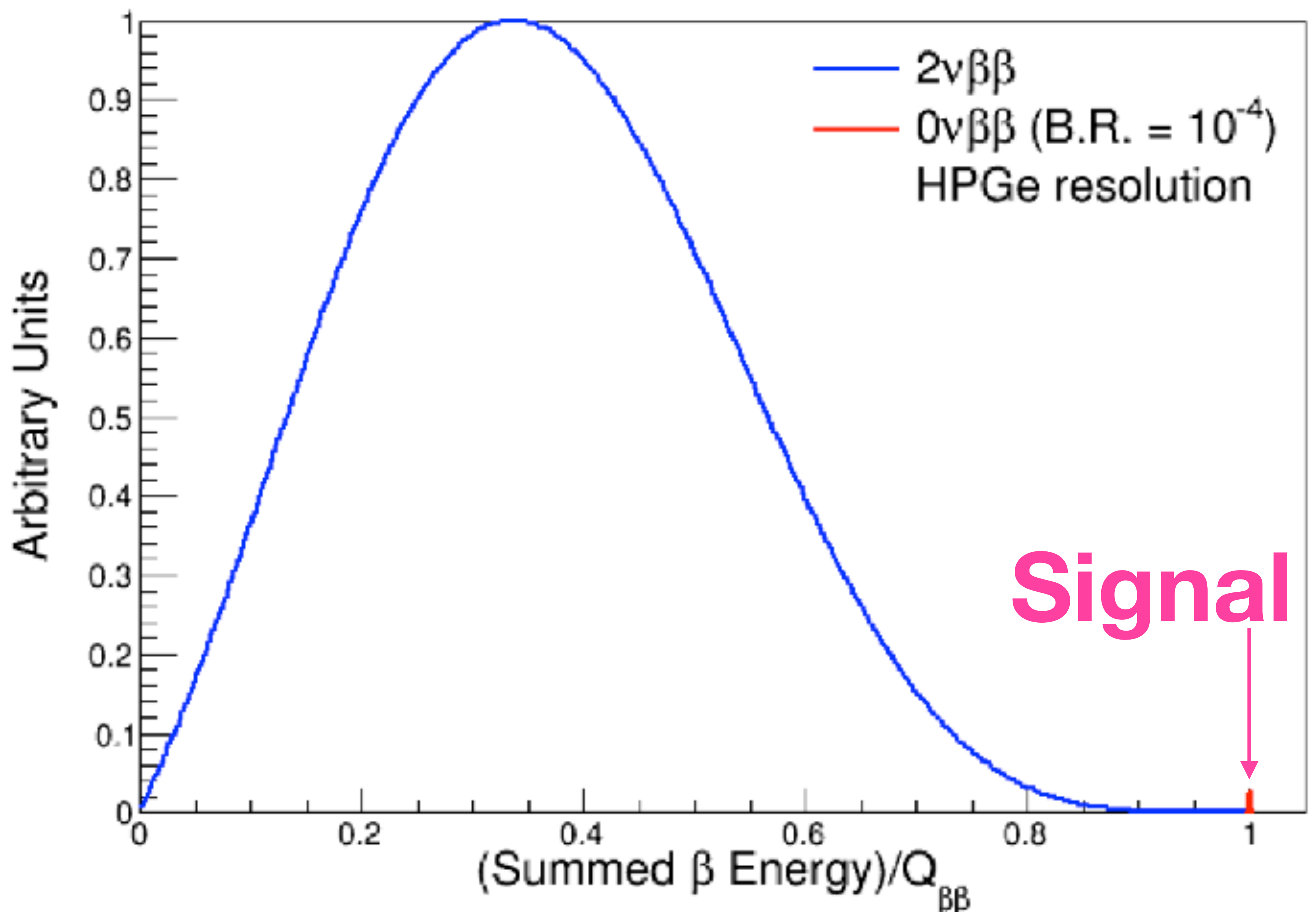
Explain the **matter-antimatter asymmetry** in our universe

Changes our fundamental understanding of particle physics

Has not been observed at $T_{\frac{1}{2}} > 10^{26} yrs$



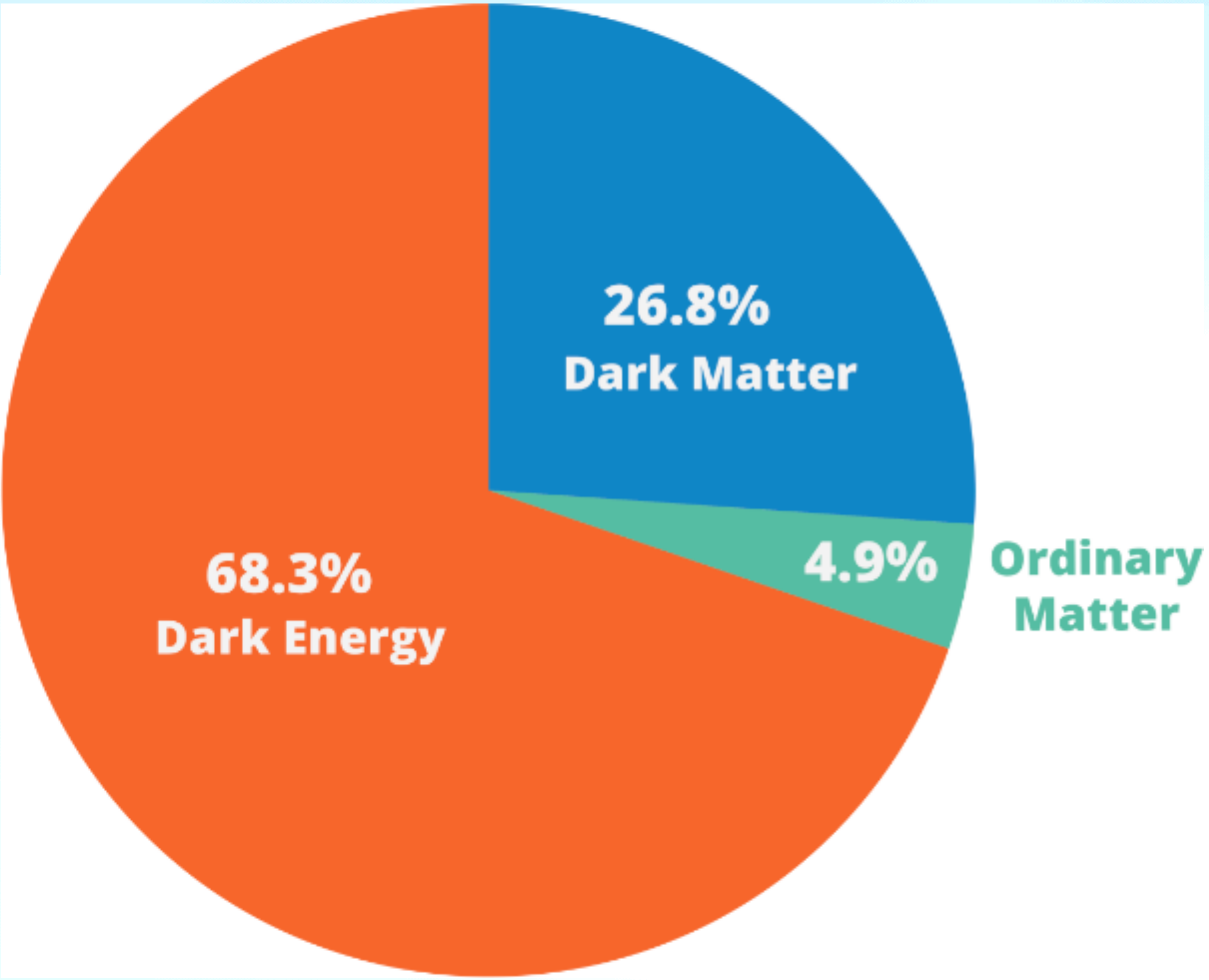


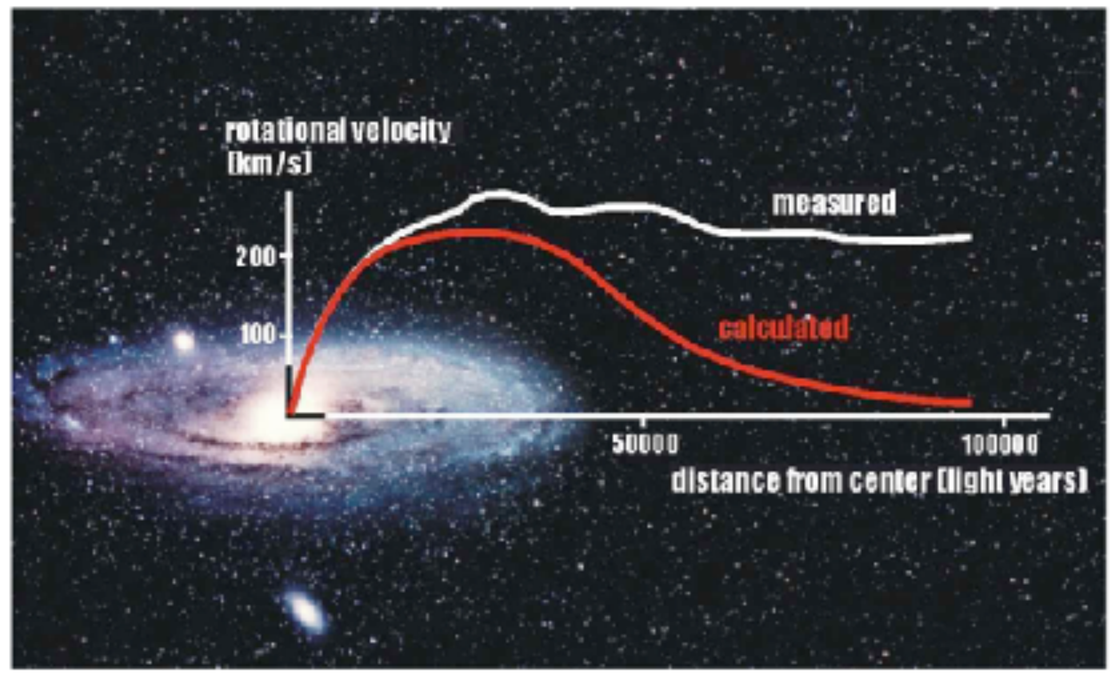


Rare Event Search in 2025

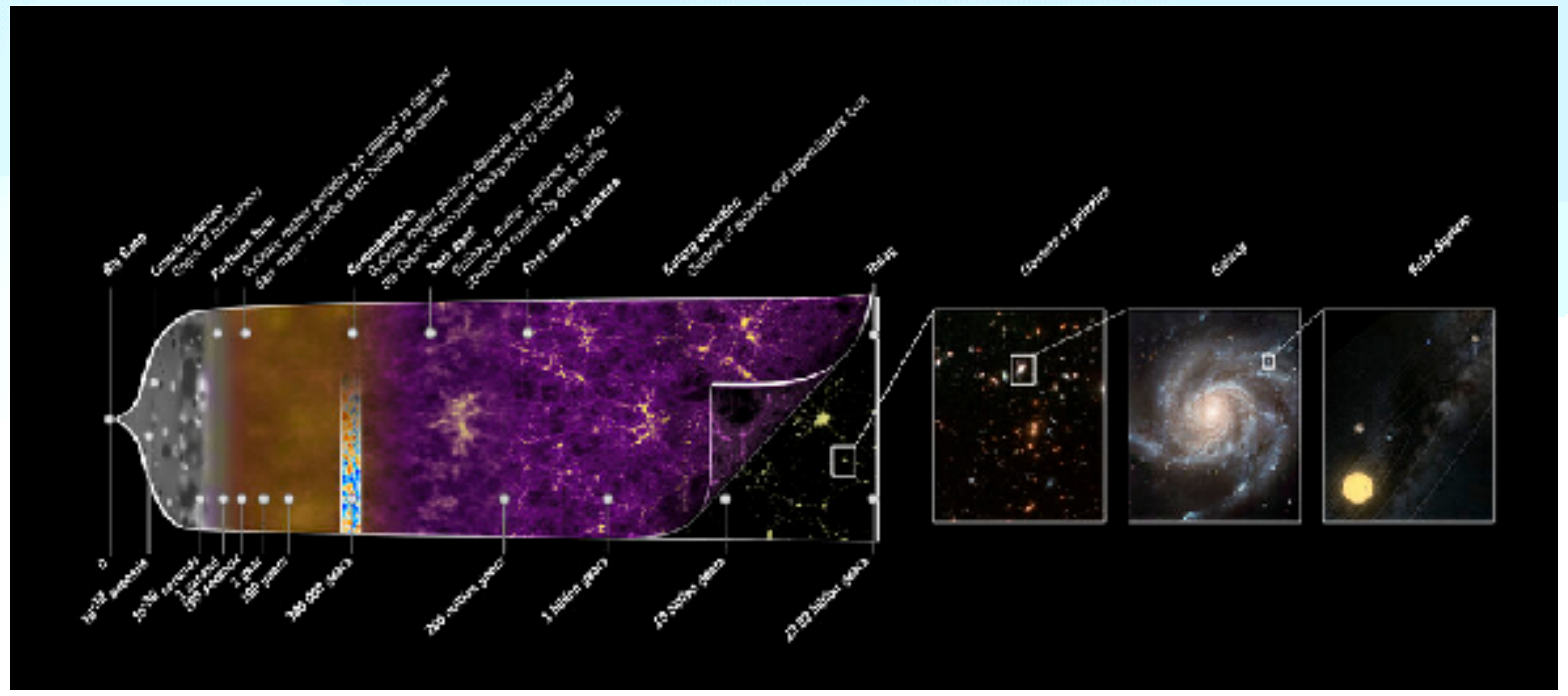
Dark Matter

The evidence for the existence of dark matter has been plenty



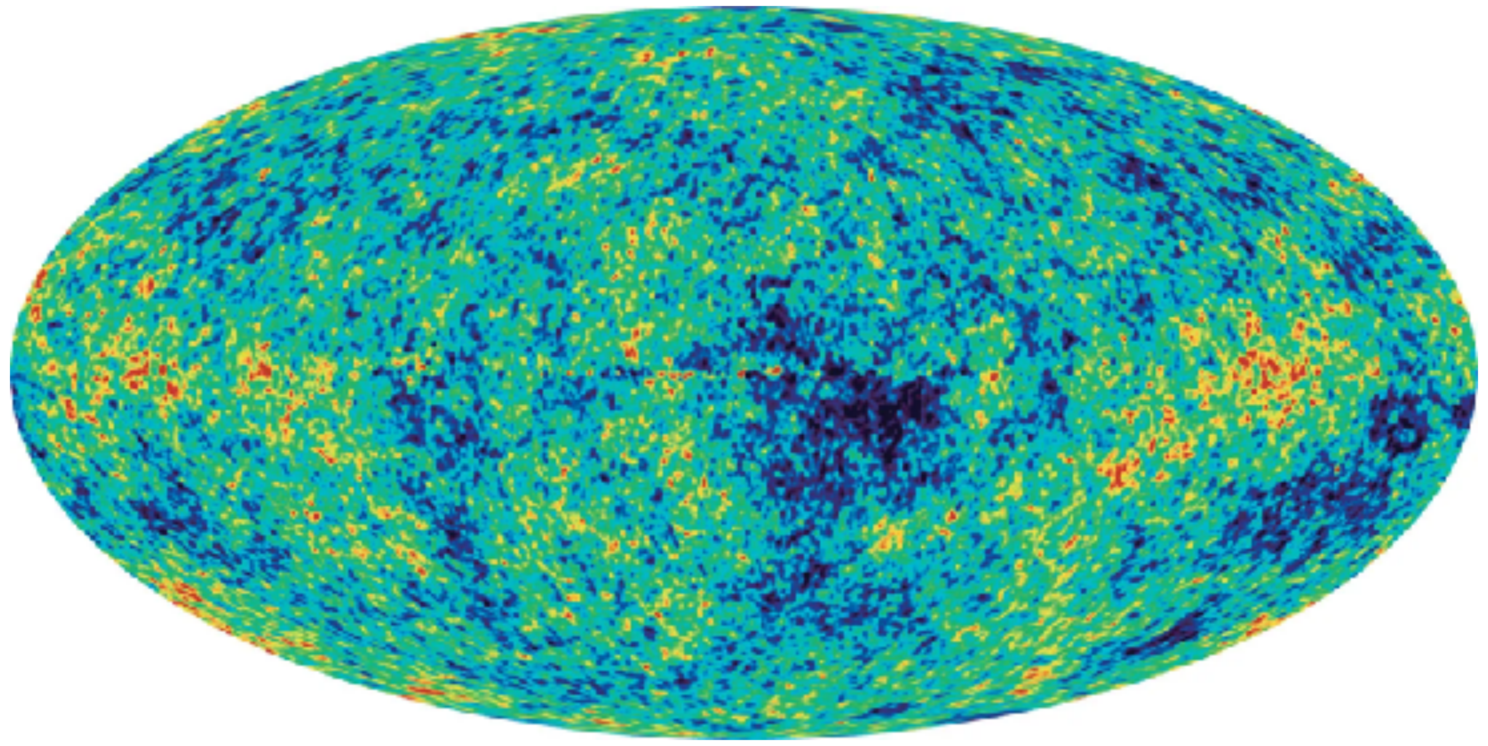


Galaxy Rotation Curve



Large Scale Structure Formation

Gravitational Lens



Cosmic Microwave Background

Rare Event Search in 2023

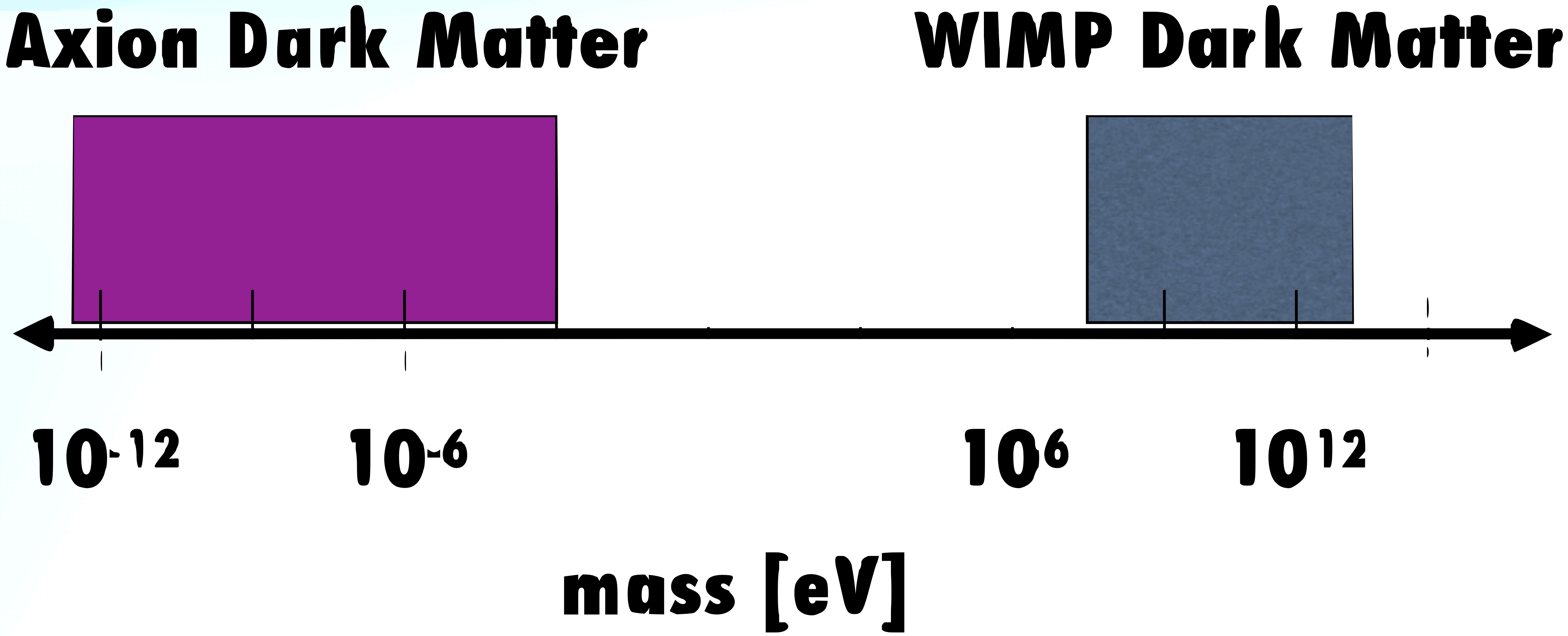
Dark Matter

The evidence for the existence of dark matter has been plenty

Many DM candidates have been proposed (WIMP, Axion, etc.)

None has been observed.

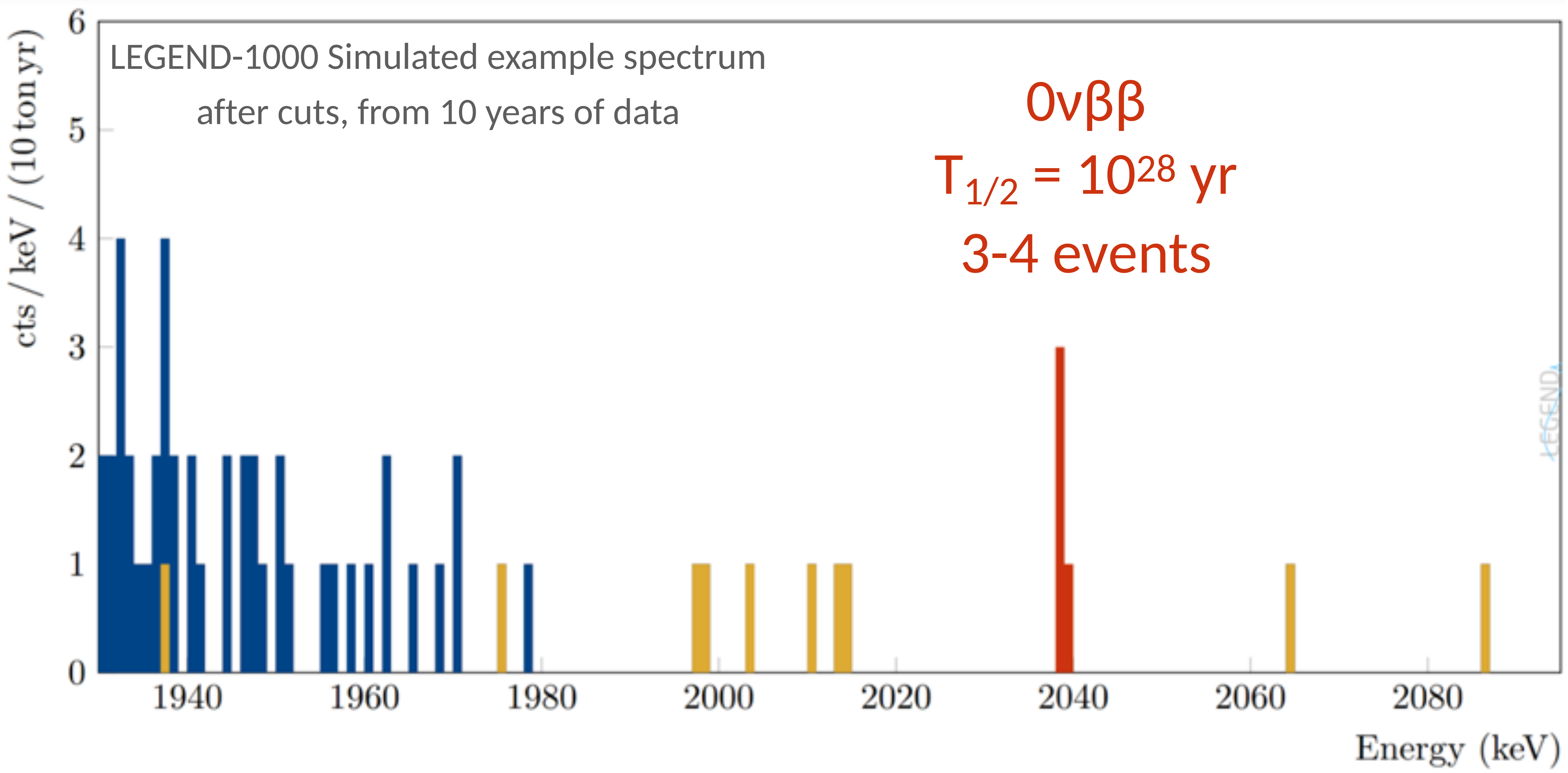
> *Dark Matter can feel like a zoo.*
>
> *—Prof. Lindley Winslow*



What Makes Rare Event Search Hard?

It is extremely rare! Using $0\upsilon\beta\beta$ as an example …

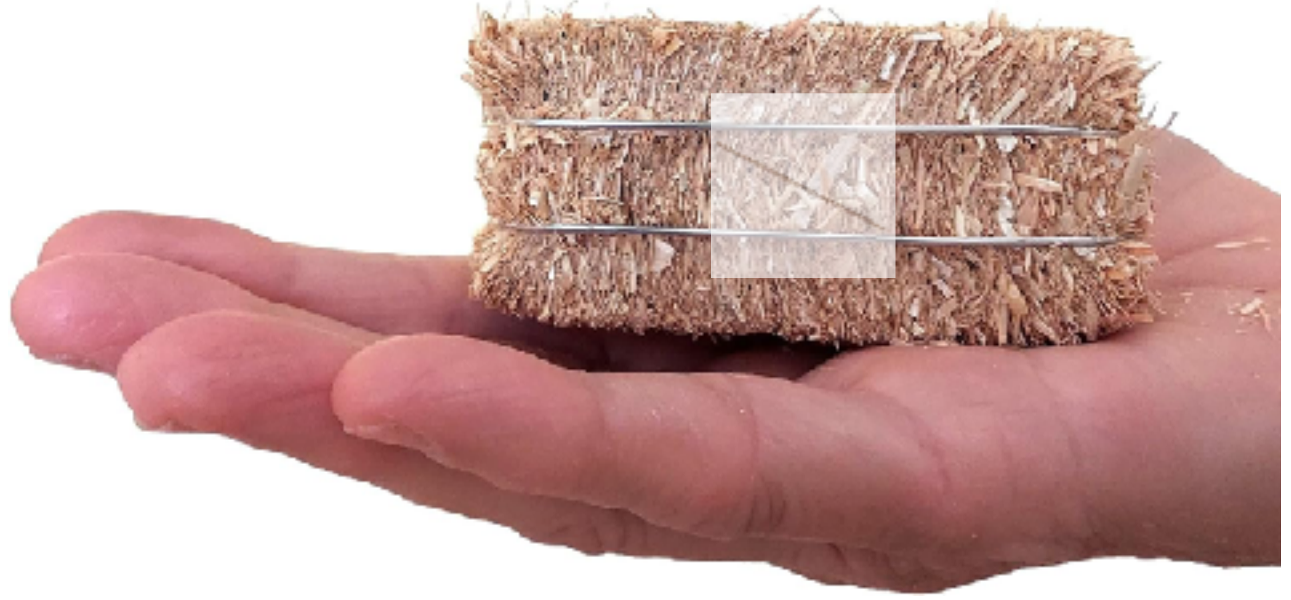
- We have not seen $0\upsilon\beta\beta$ at half life of $T_{\frac{1}{2}} > 10^{26} yrs$
- Next-generation experiments typically aims at $T_{\frac{1}{2}} > 10^{28} yrs$ ($\times$100 improvement)
- Correspond to **3-4 event** after **10 years** of data taking



What Makes Rare Event Search Hard?

- **1 event** every **2.5-3.3 year**, we need ultra-sensitive detector to capture every event
- As our detector gets more sensitive, we also collect lots of background events that are not 0νββ/WIMP DM

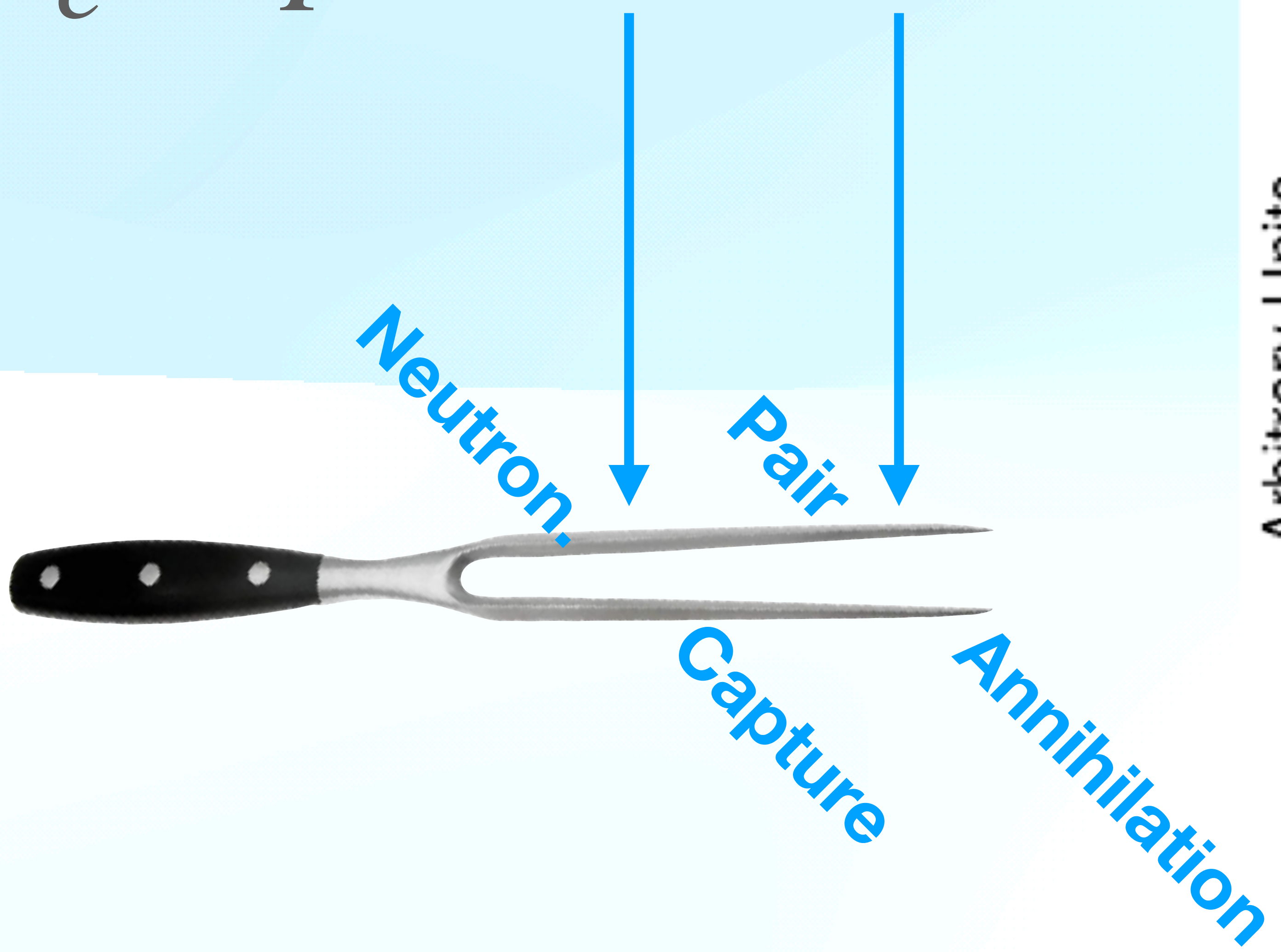
Search for needle in a haystack



What Makes Rare Event Search Hard?

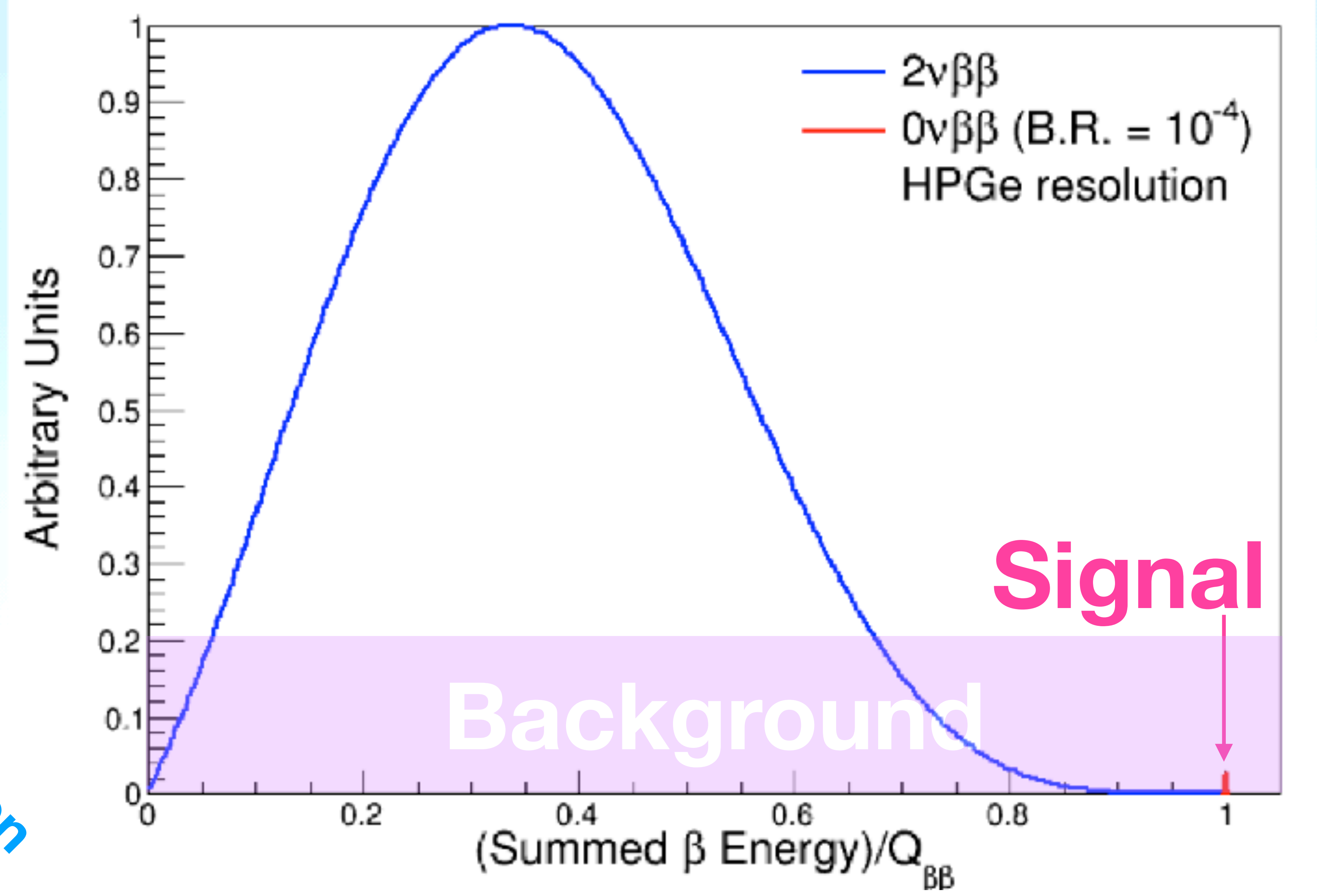
The Cowan-Reine Exp.

$$\bar{\nu}_e + p \rightarrow n + e^+$$

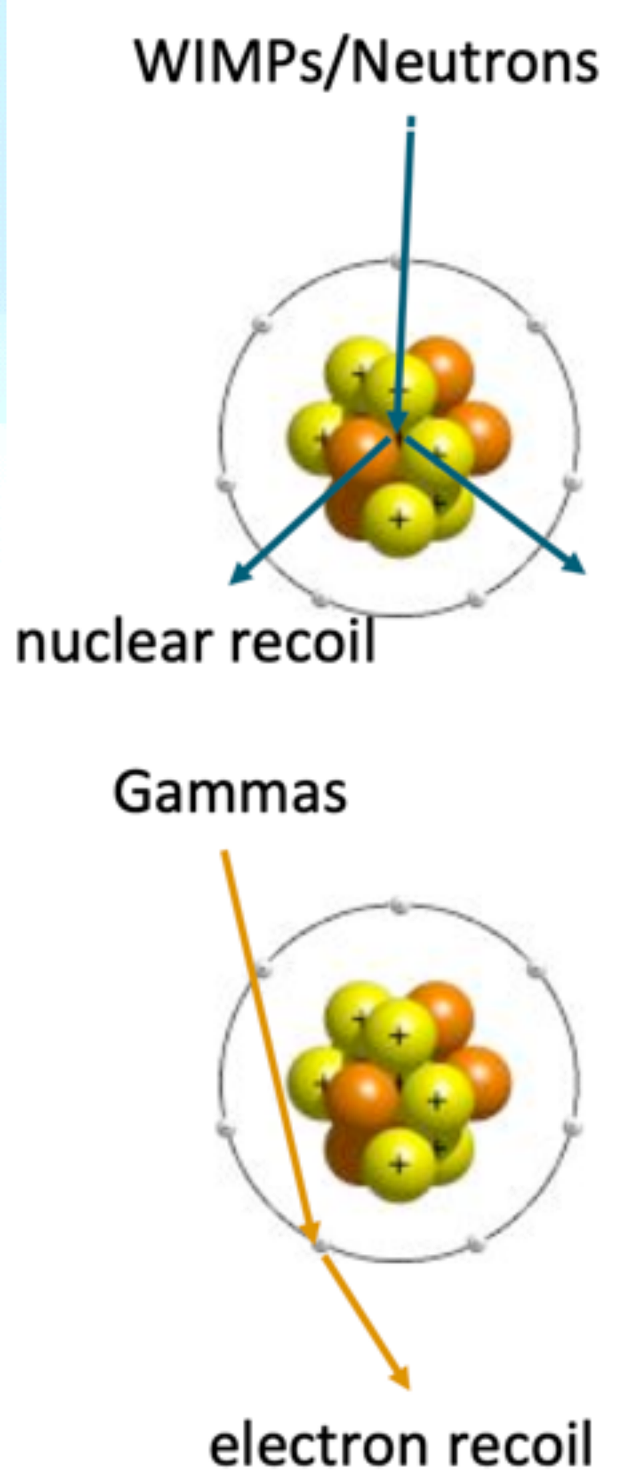


Nearly background-free

$0\upsilon\beta\beta$



WIMP Dark Matter

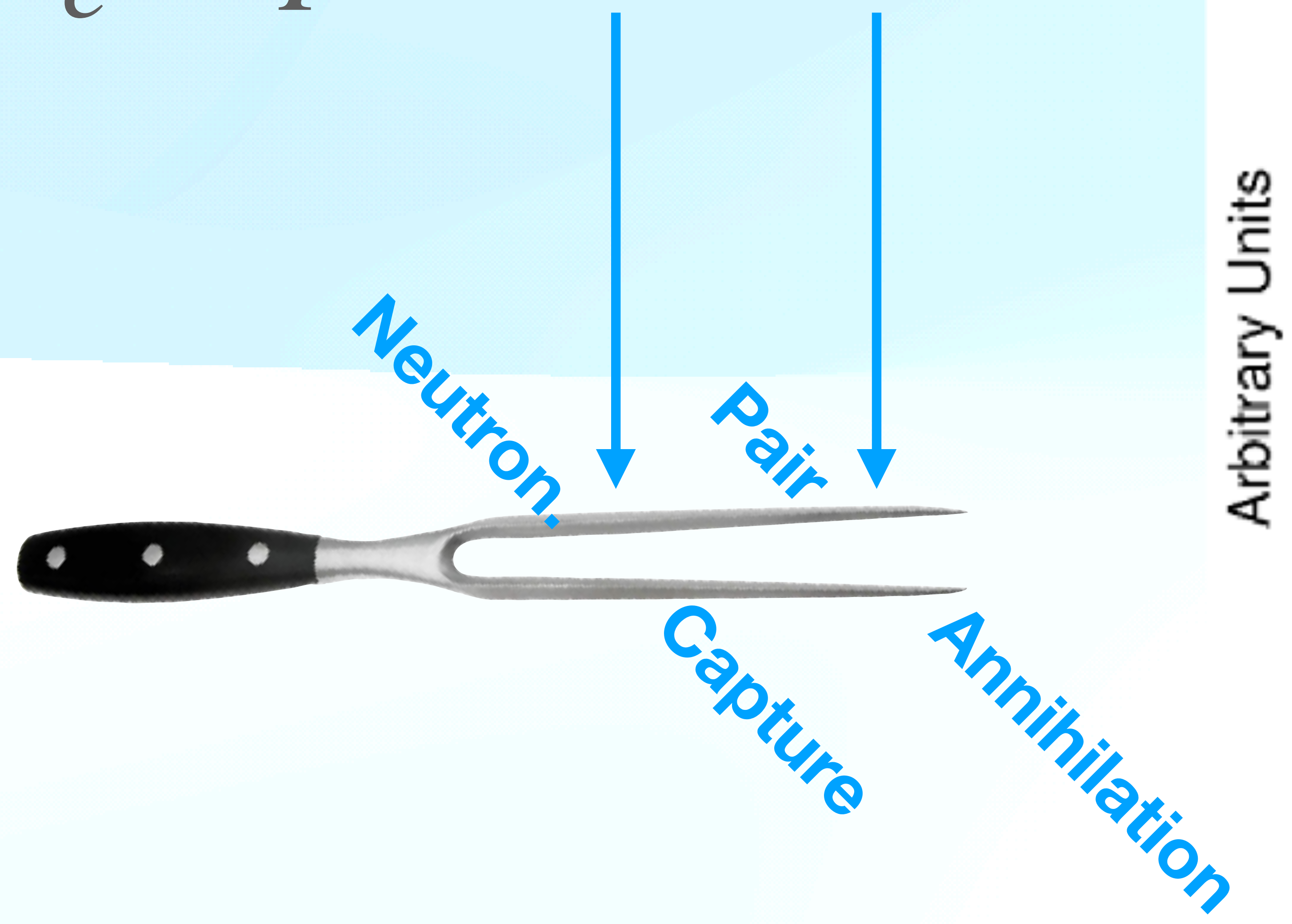


Naturally radioactive and cosmic ray background

What Makes Rare Event Search Hard?

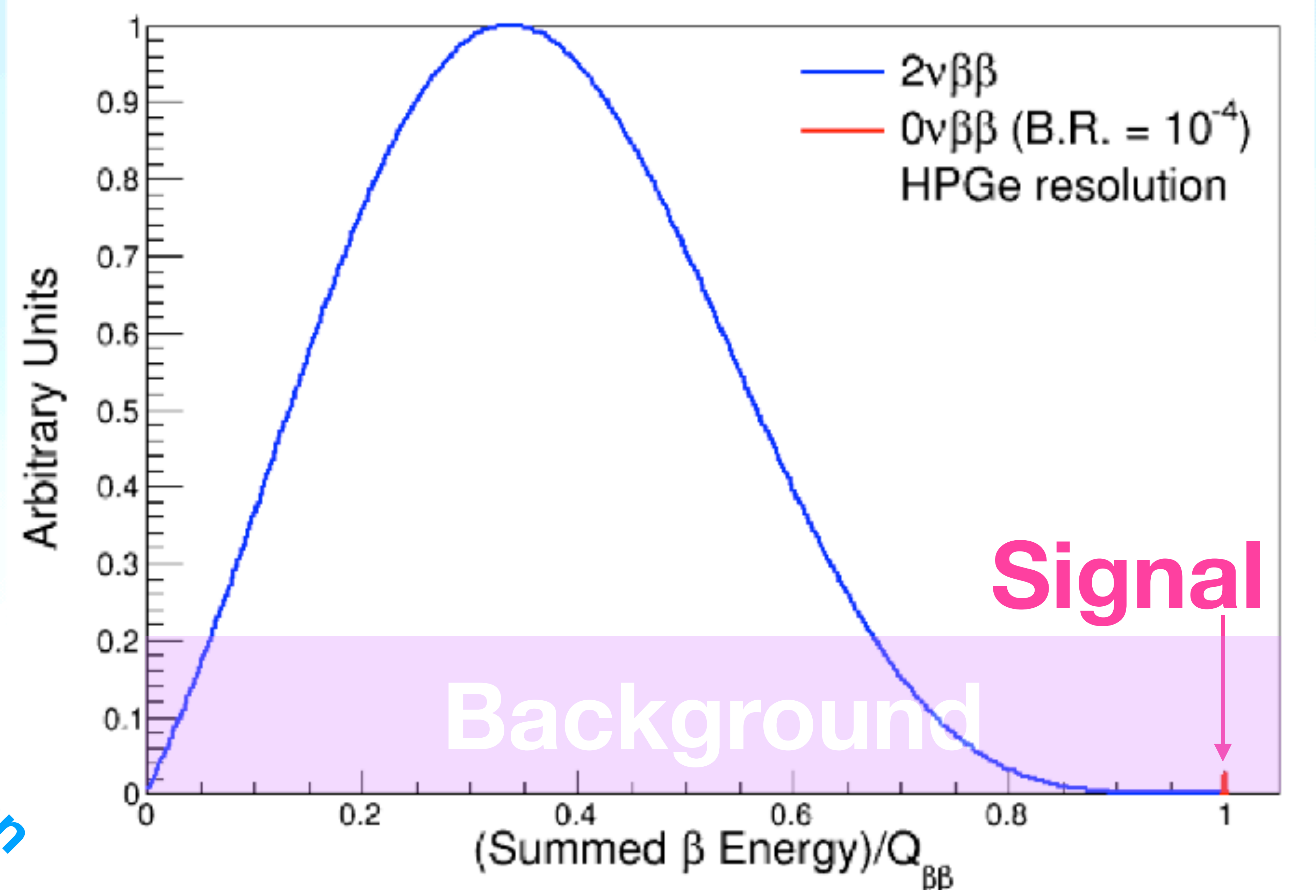
The Cowan-Reine Exp.

$$\bar{\nu}_e + p \rightarrow n + e^+$$

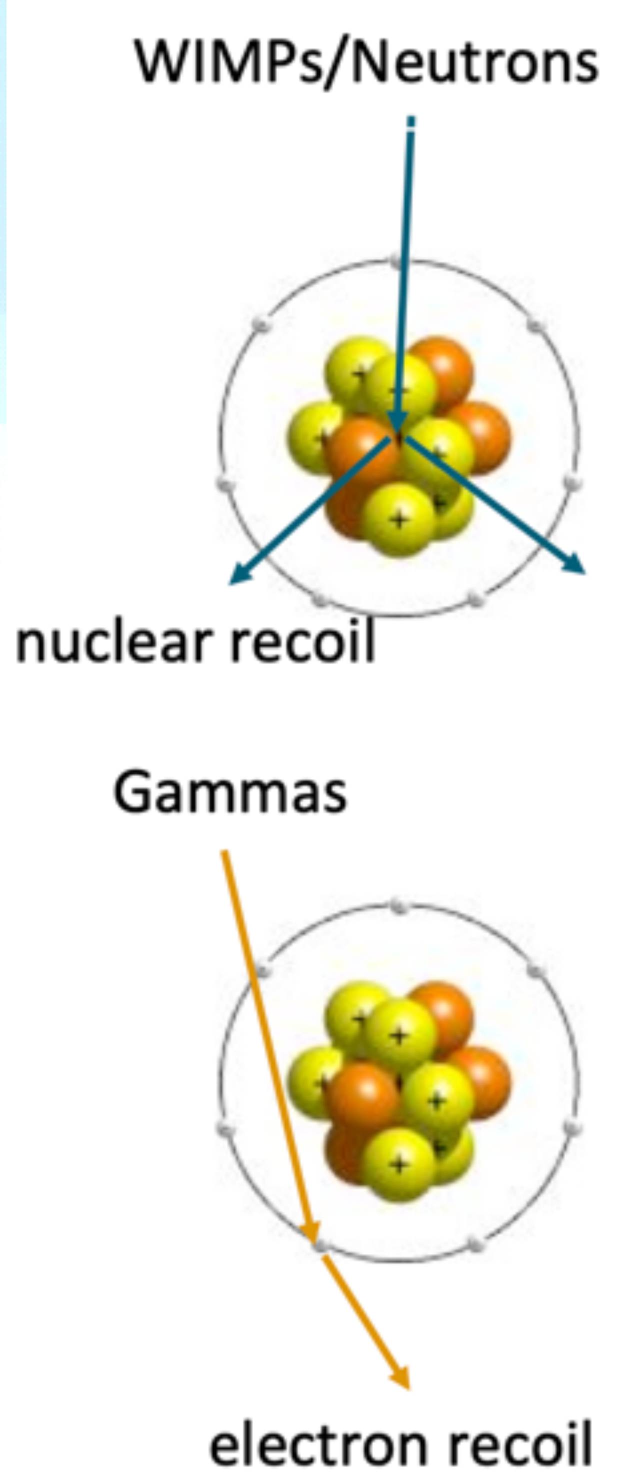


Nearly background-free

$0\upsilon\beta\beta$



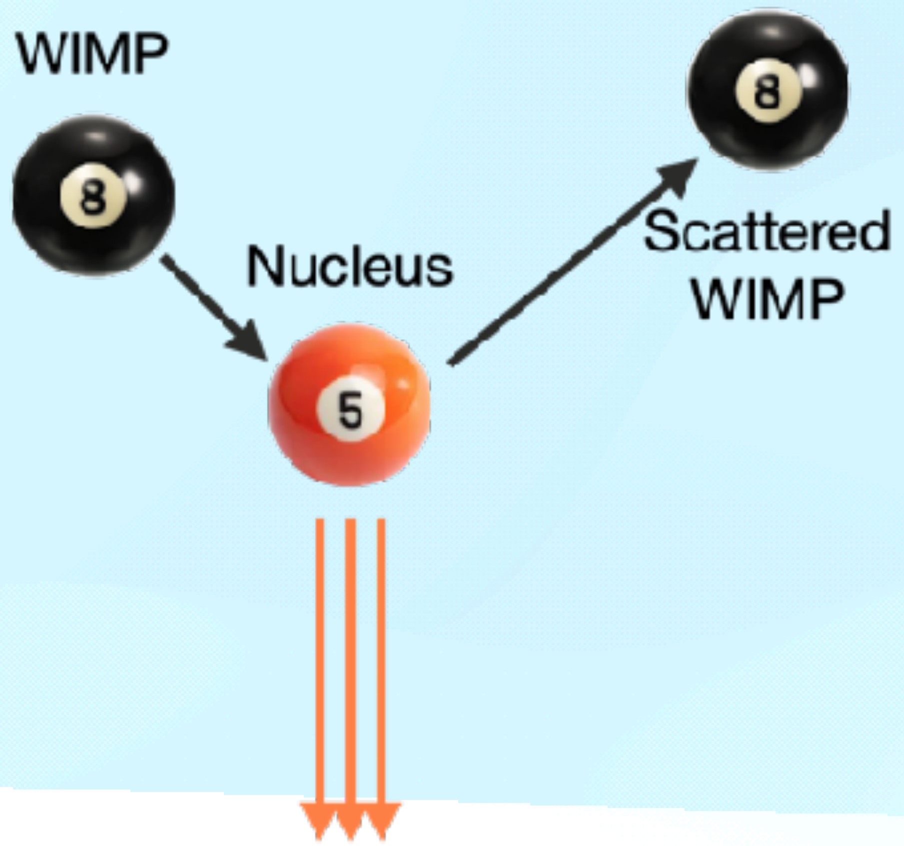
WIMP Dark Matter

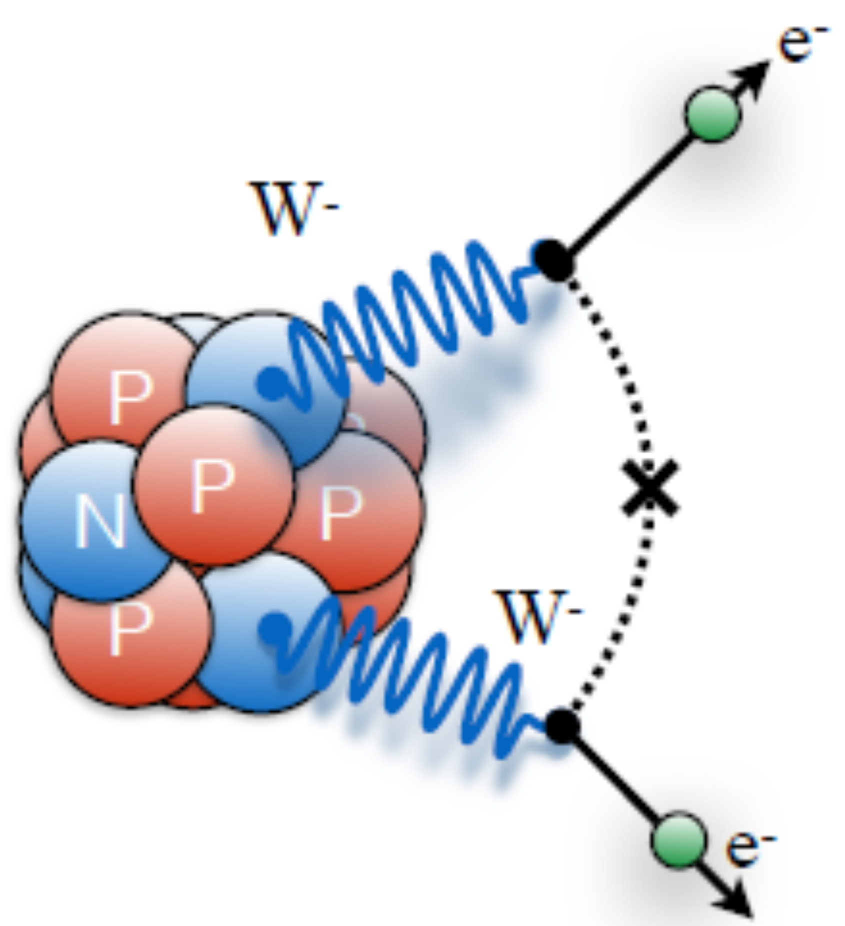


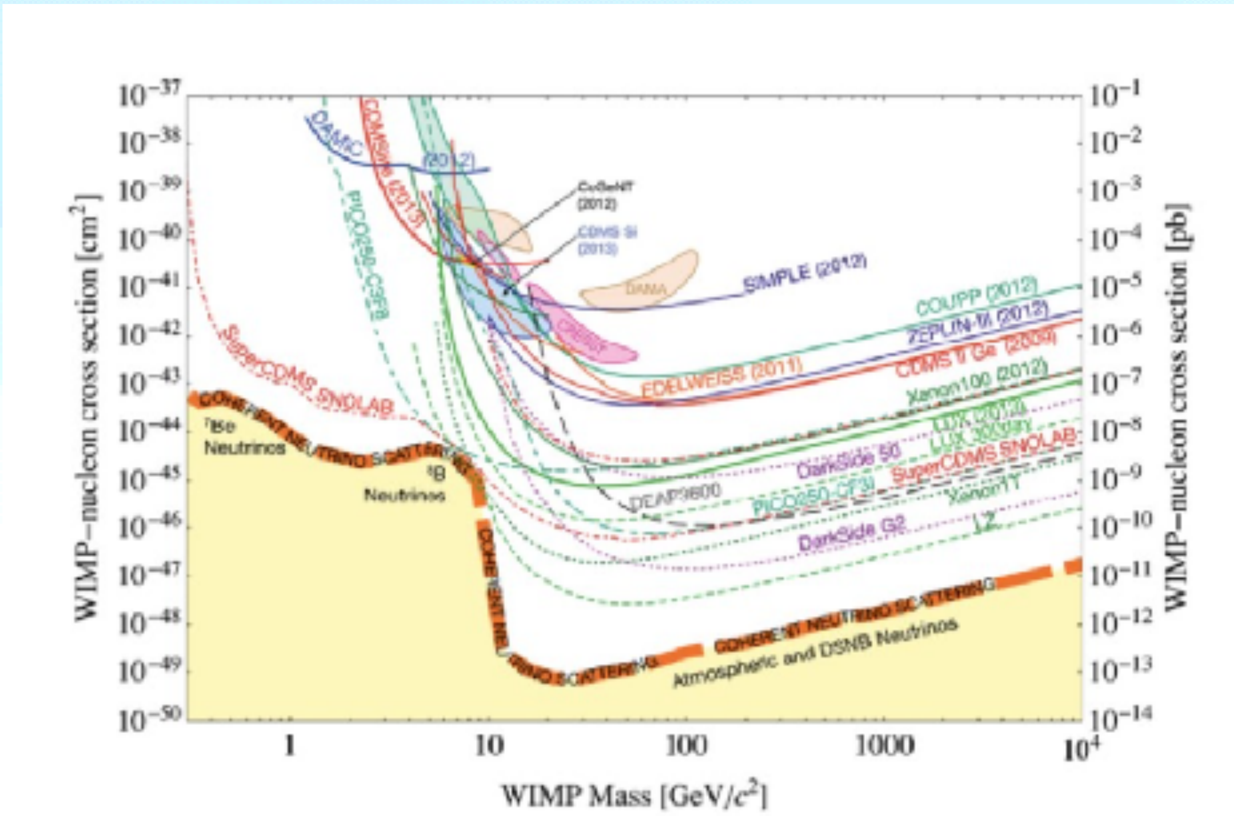
Naturally radioactive and cosmic ray background

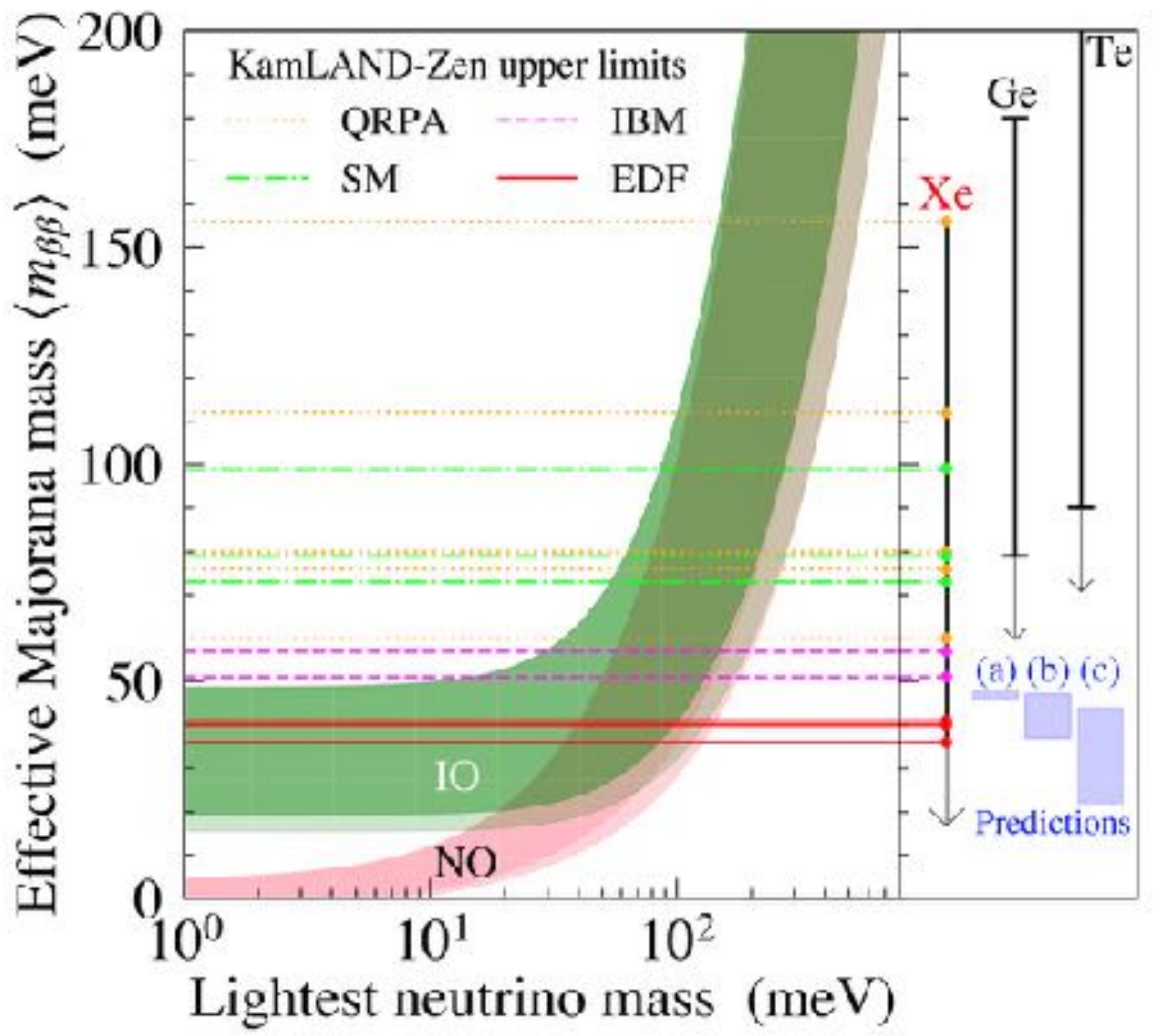
Control background is of unparalleled importance in rare event search experiment!

The Rare Event Search Pipeline

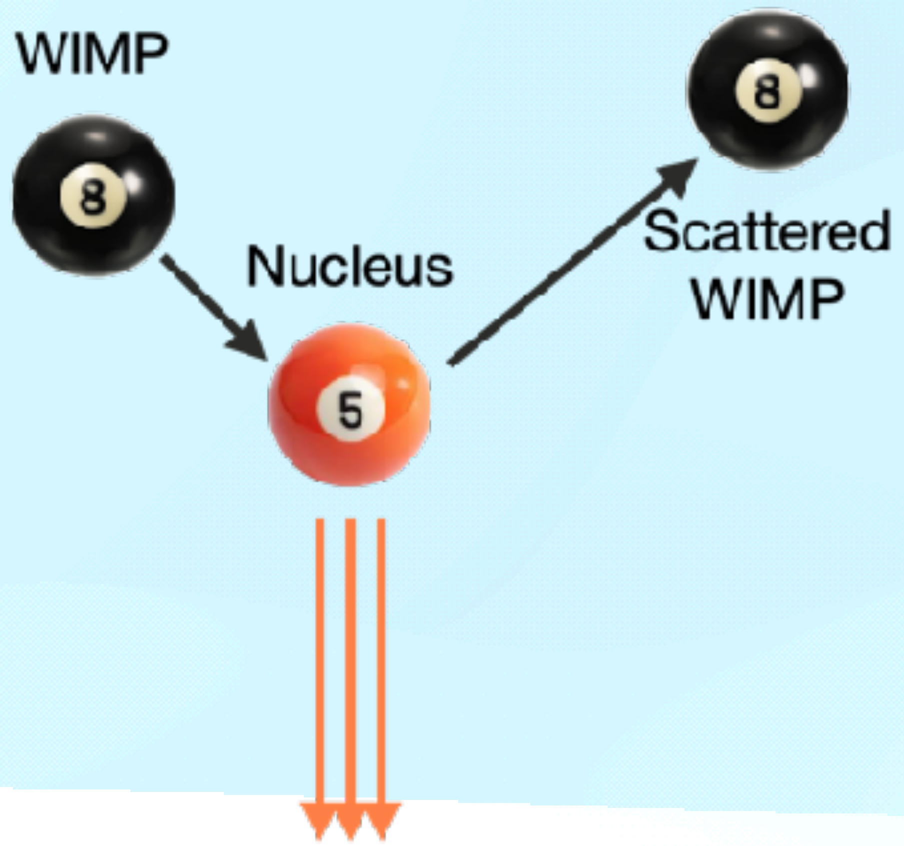


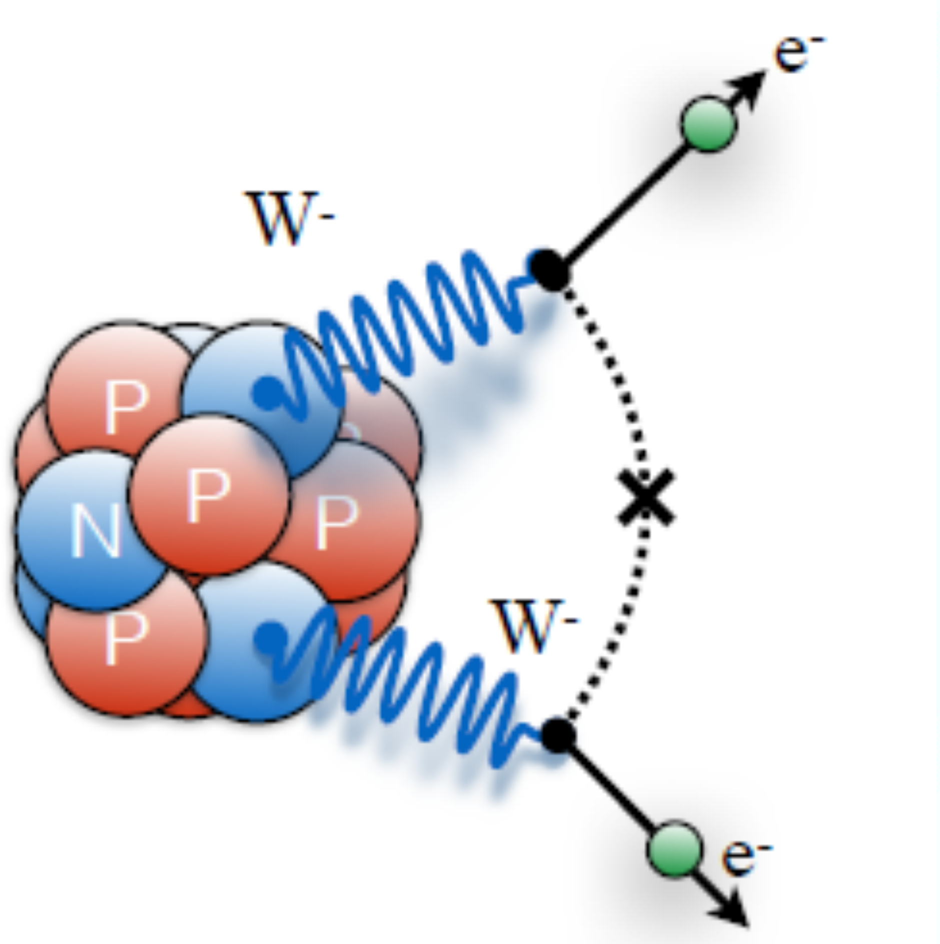






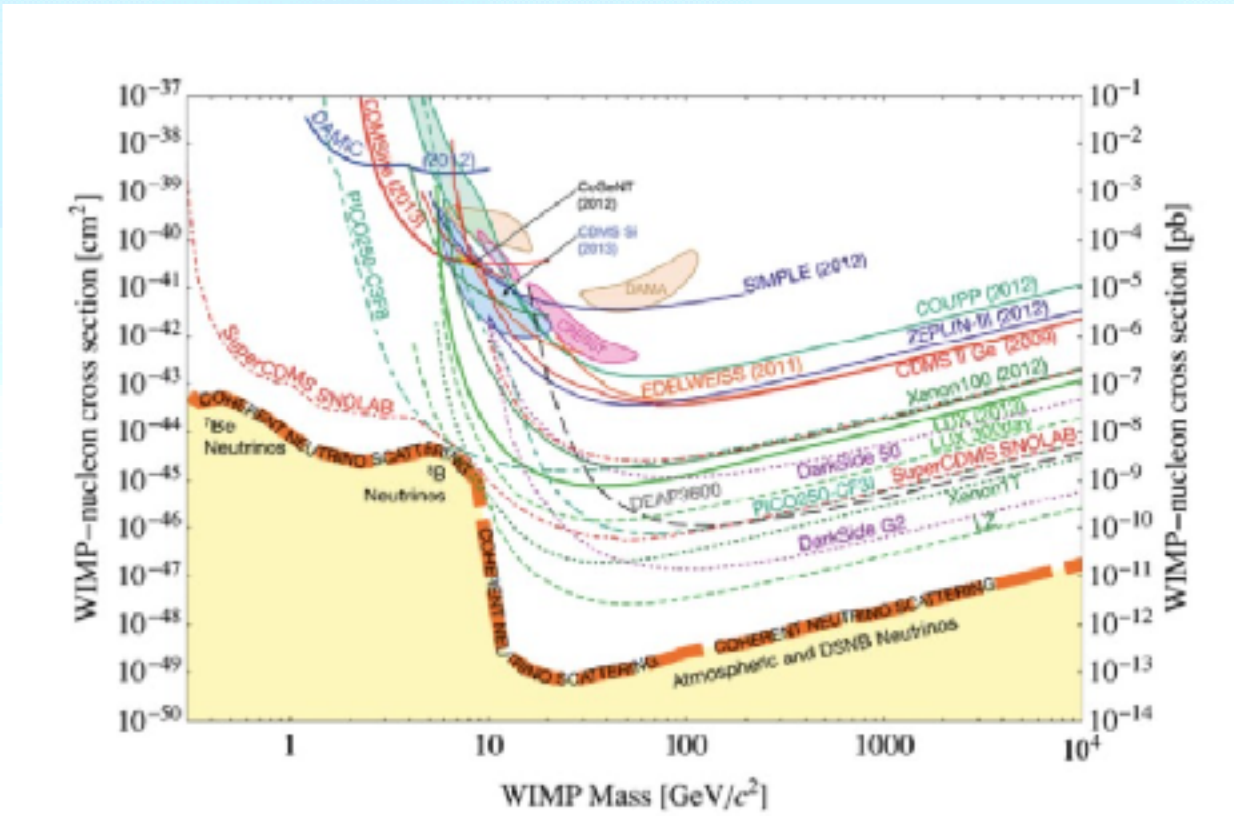
The Rare Event Search Pipeline

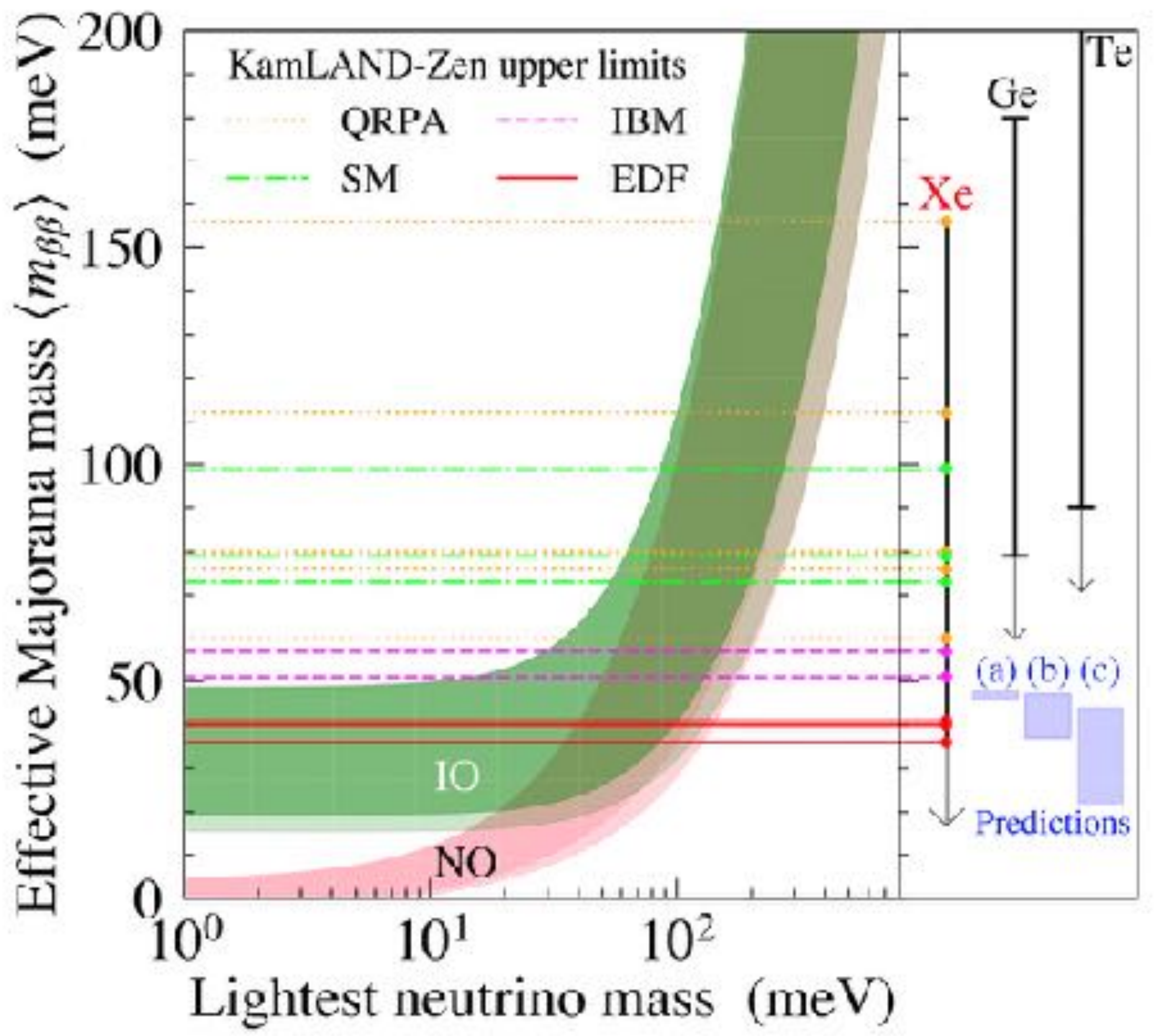




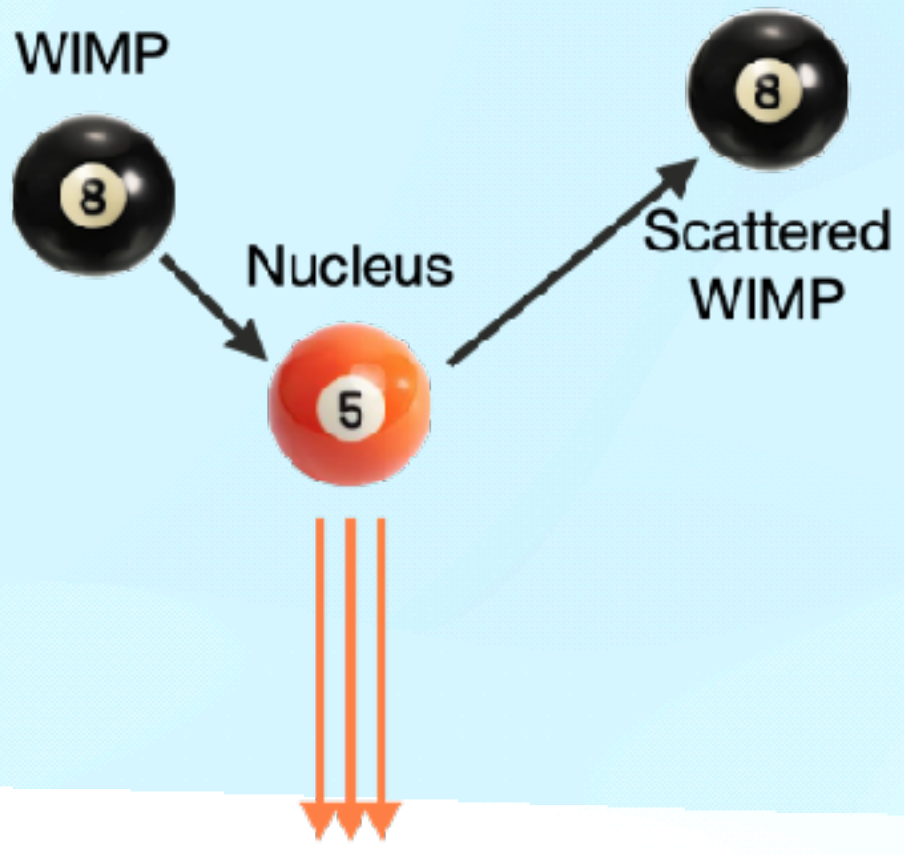
Radiation Detector

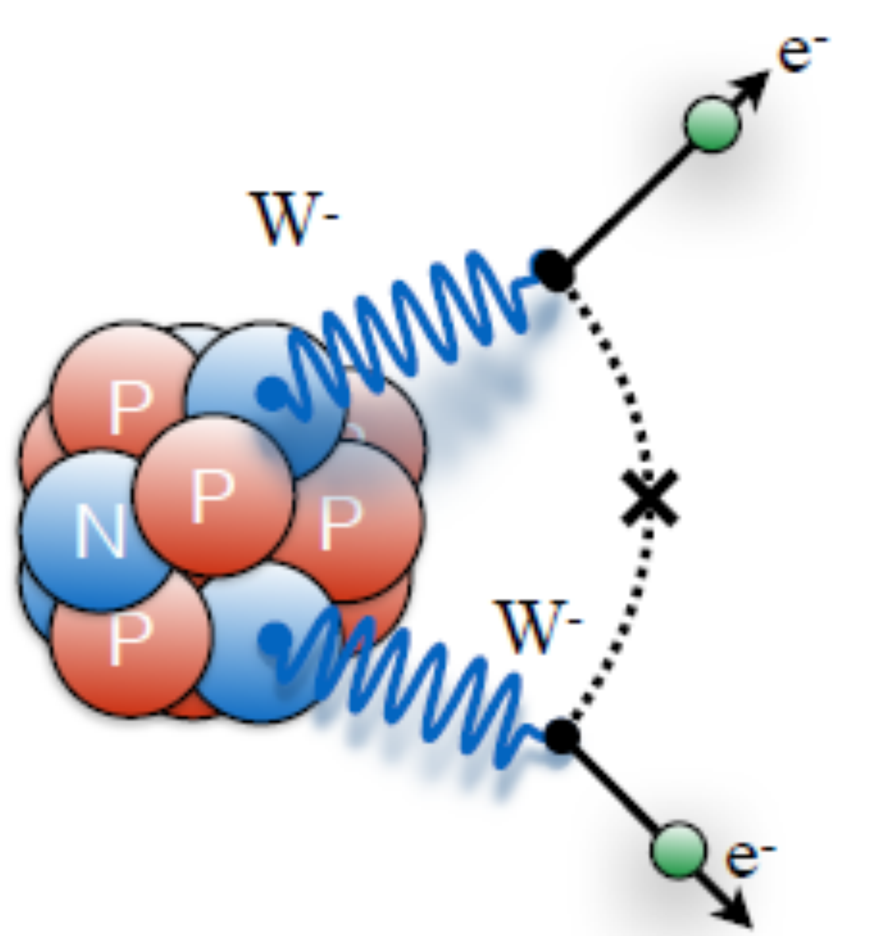
The "magnifying glass" that help finding the needle





The Rare Event Search Pipeline



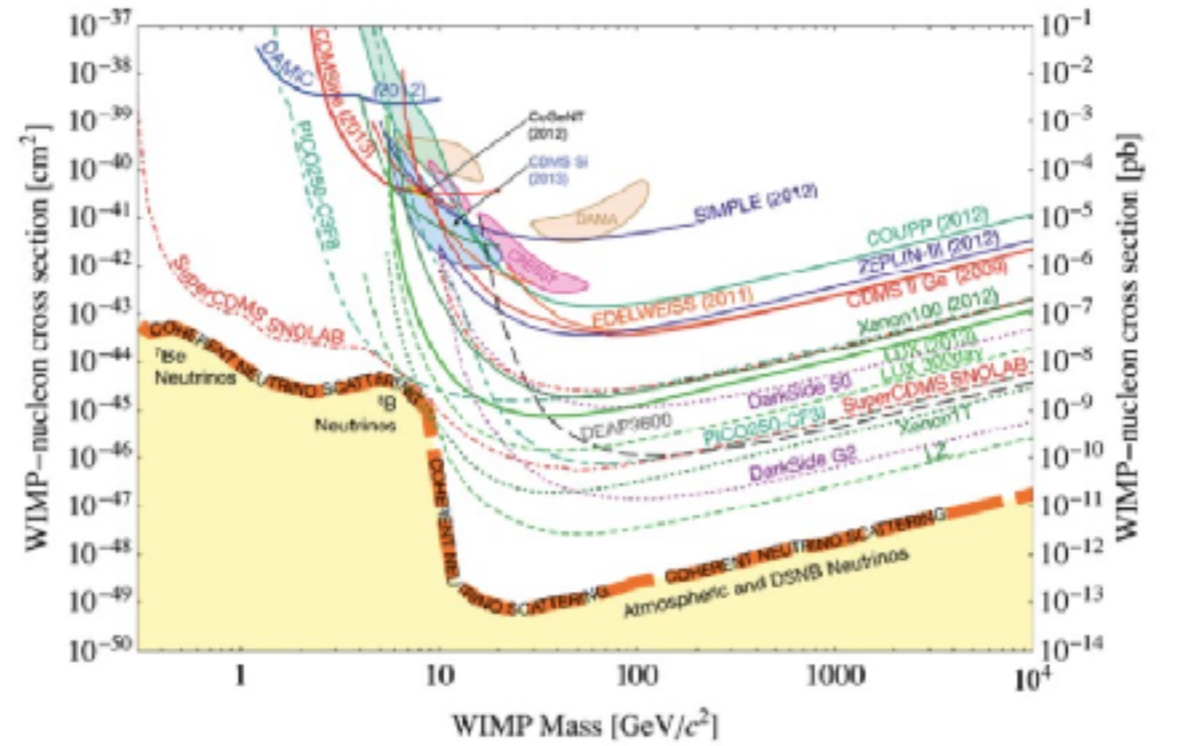


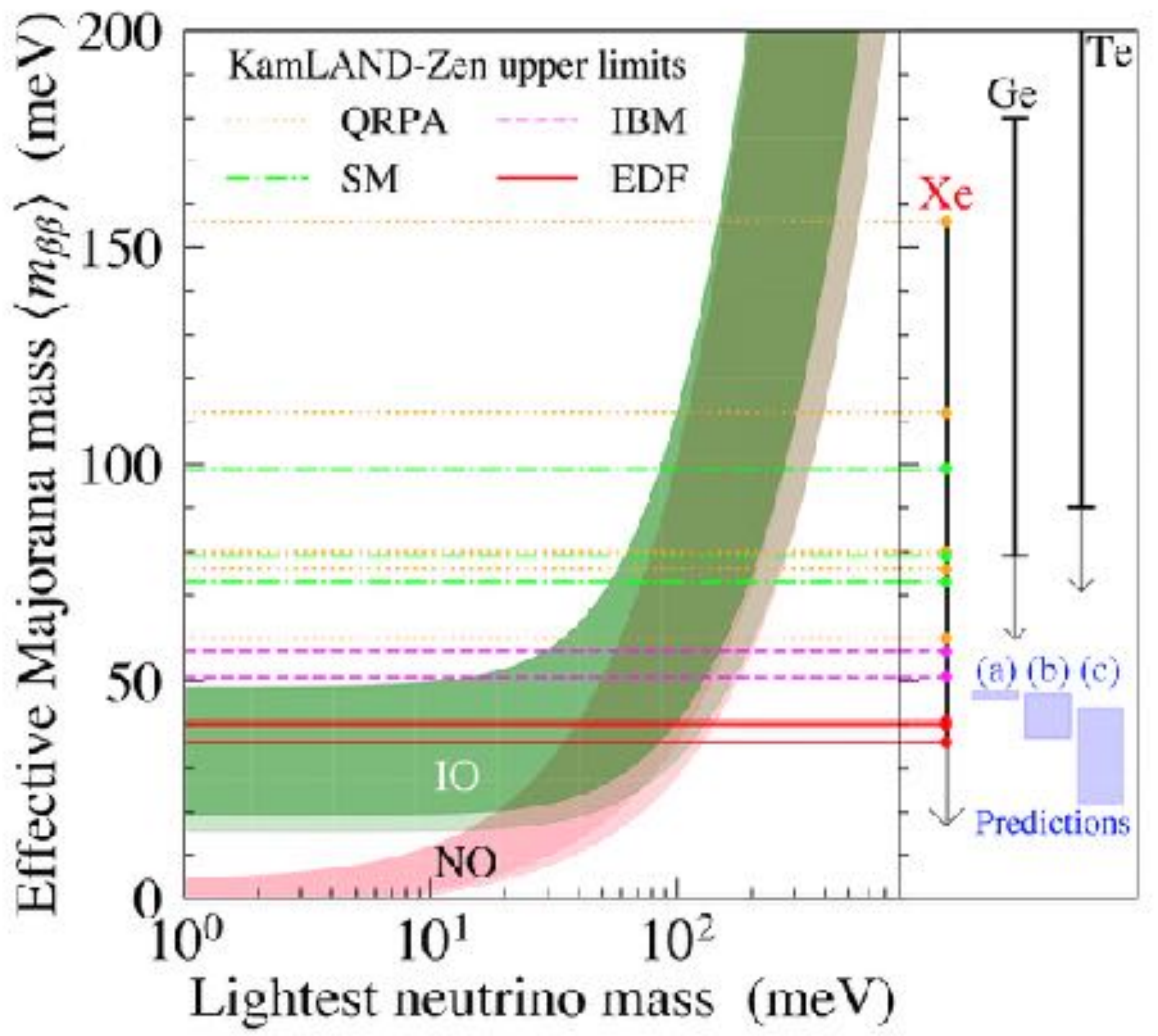
Radiation Detector

The "magnifying glass" that help finding the needle

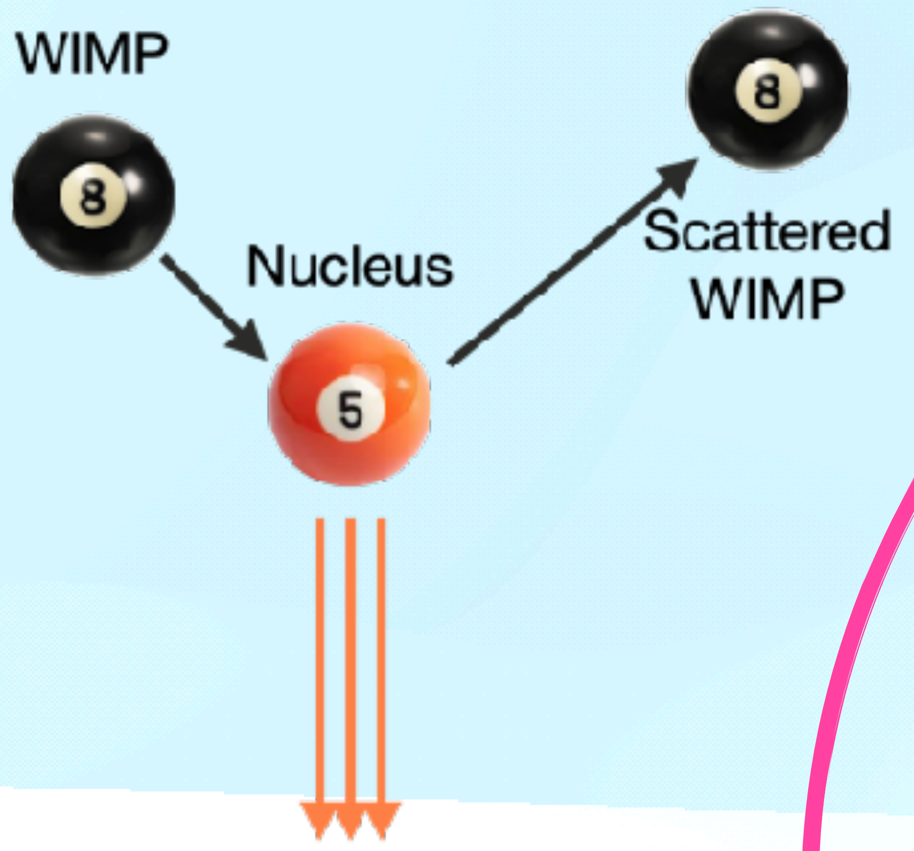
AI/ML

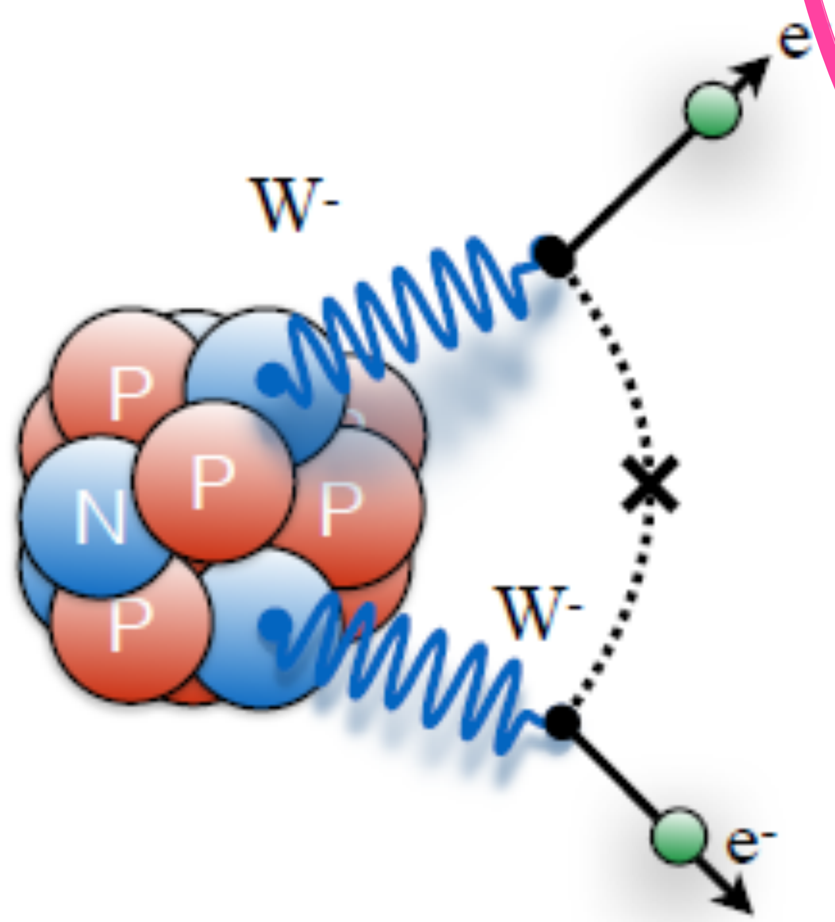
The "forklift" that help removing the haystack





The Rare Event Search Pipeline



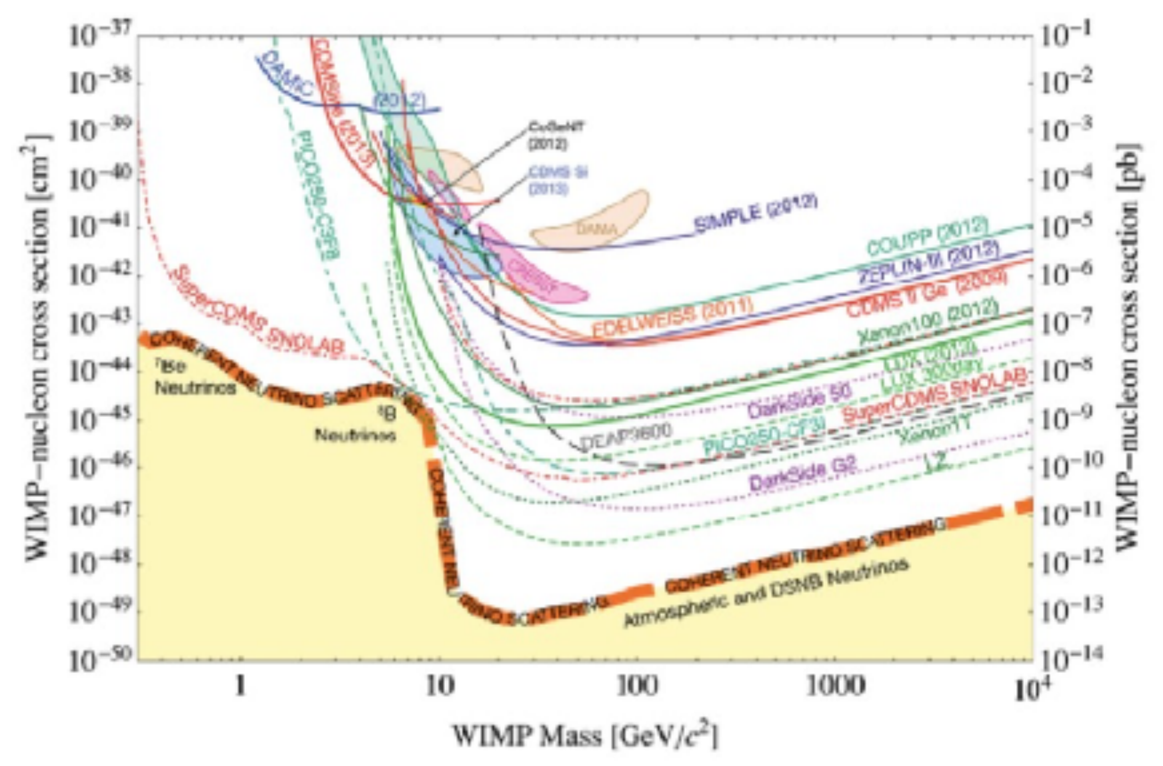


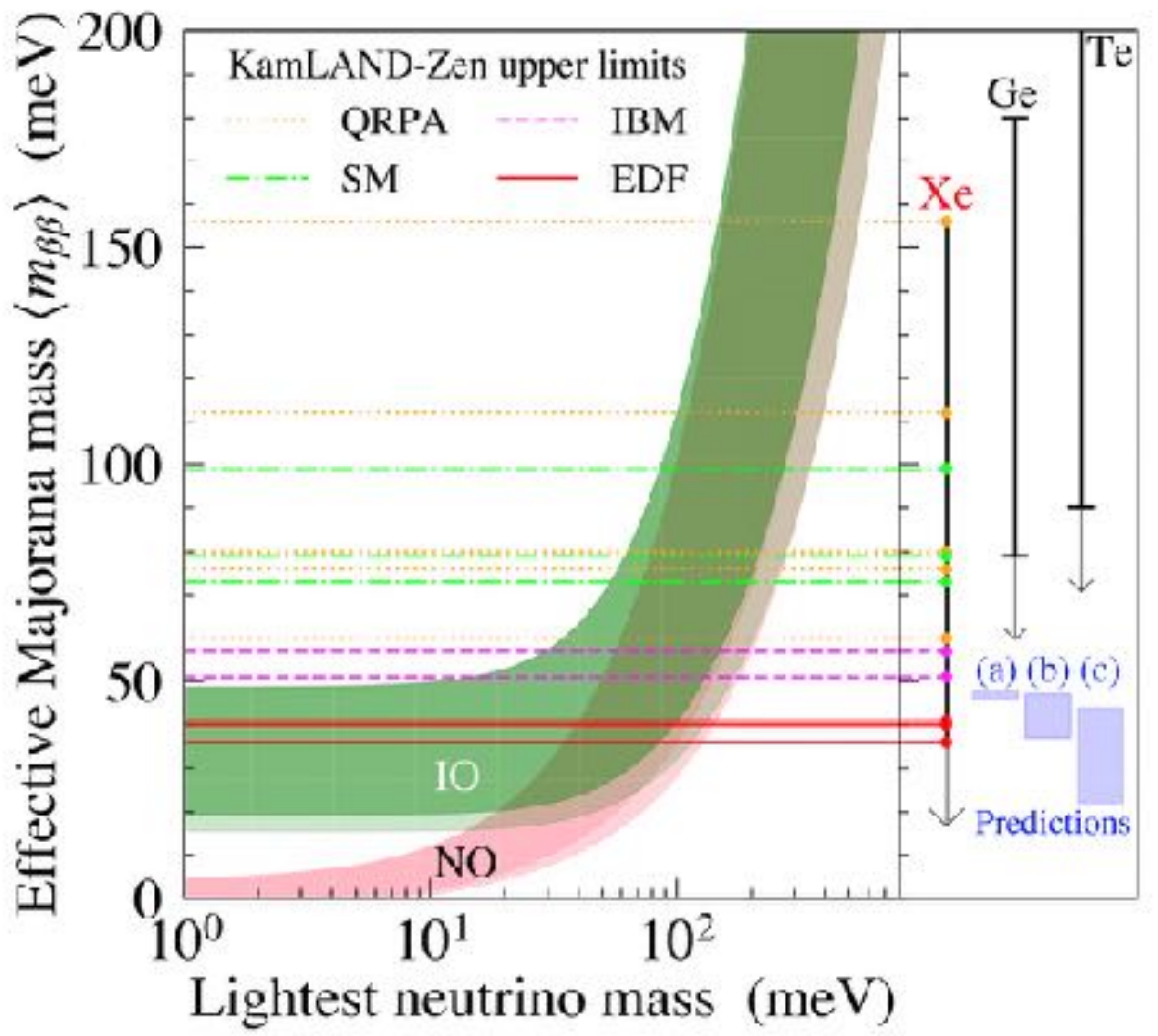
Radiation Detector

The "magnifying glass" that help finding the needle

AI/ML

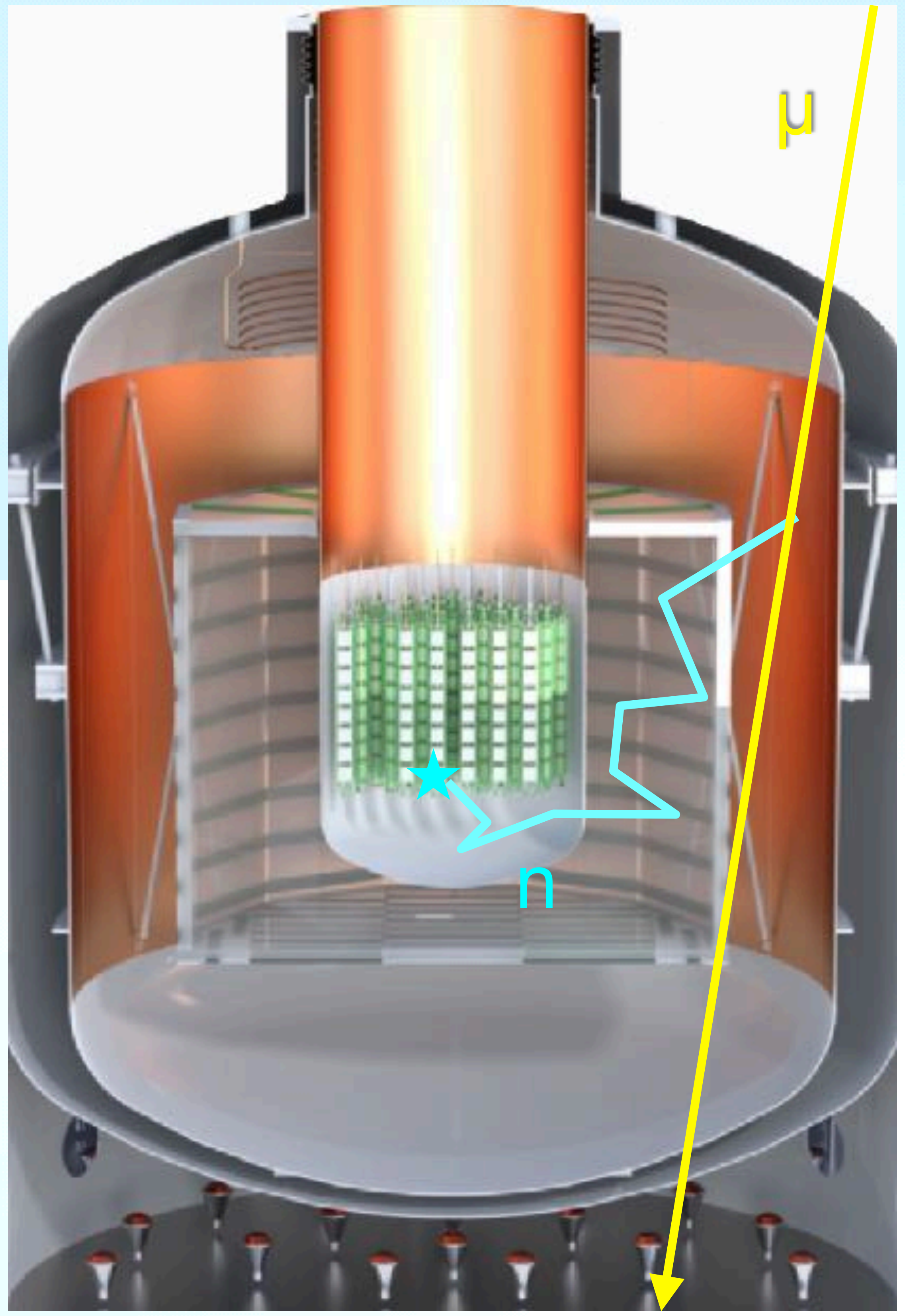
The "forklift" that help removing the haystack





Cosmogenic Background in LEGEND

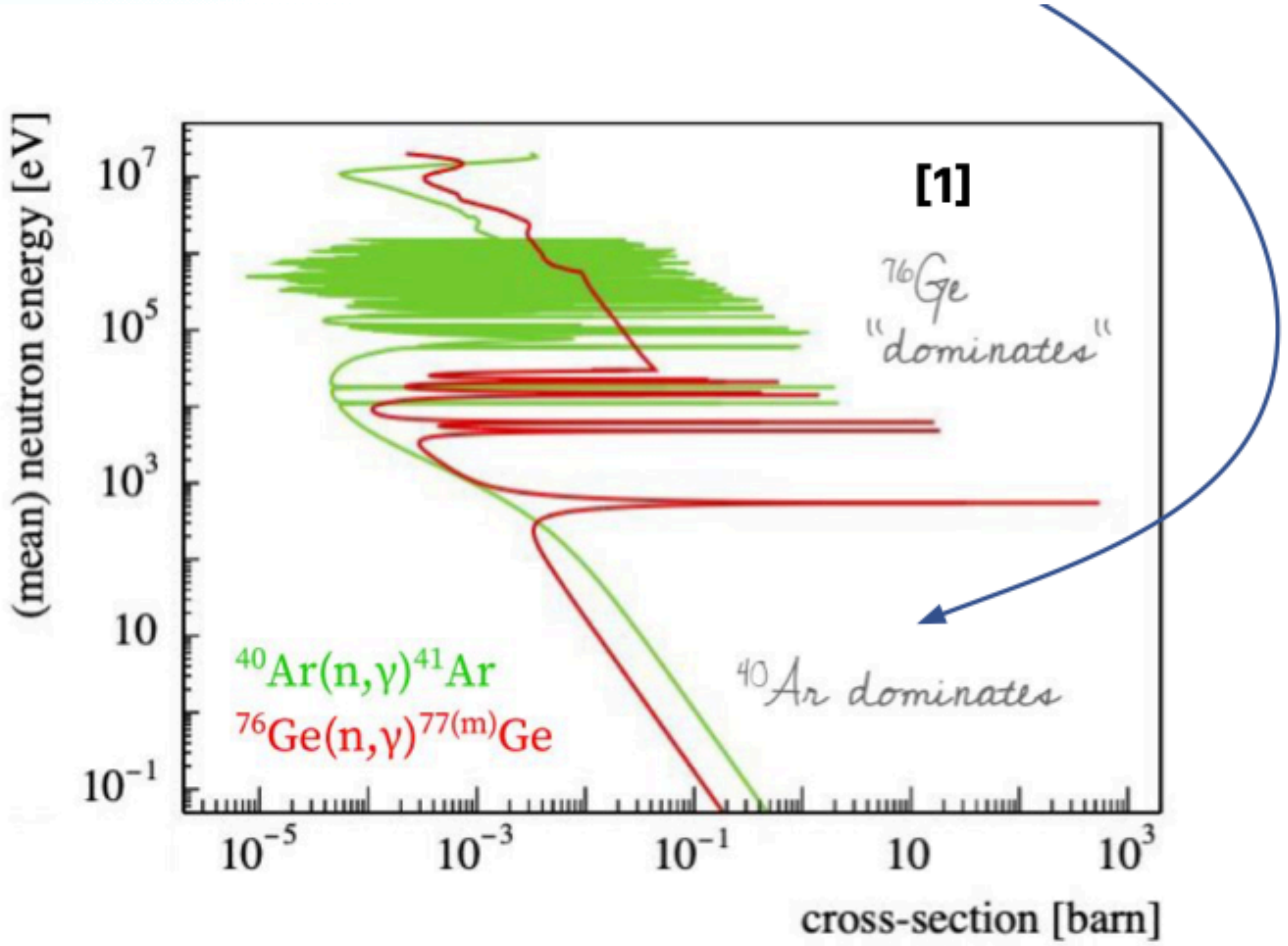
A flagship HPGe experiment searching for Neutrinoless Double-Beta Decay



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*:

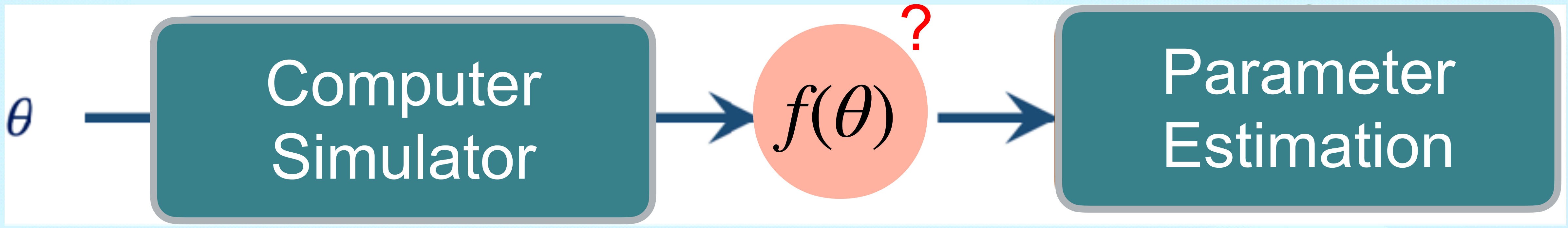
add neutron moderators to slow neutrons down and increase their likelihood to be captured by LAr instead of ^{76}Ge.

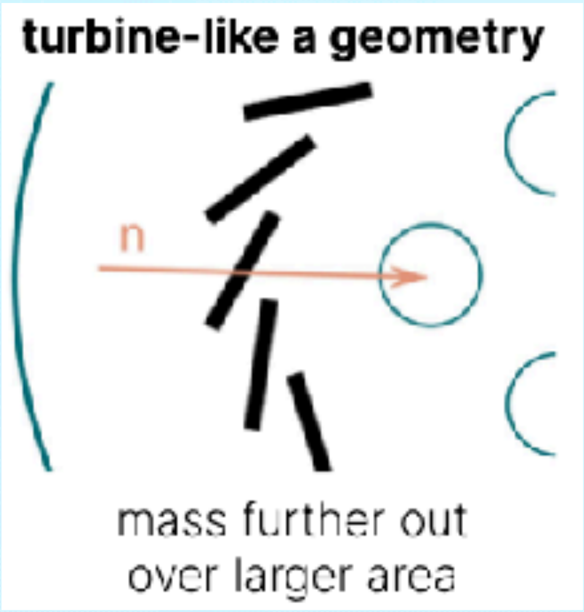


LEGEND Neutron Moderator

How to find the optimal design parameter?

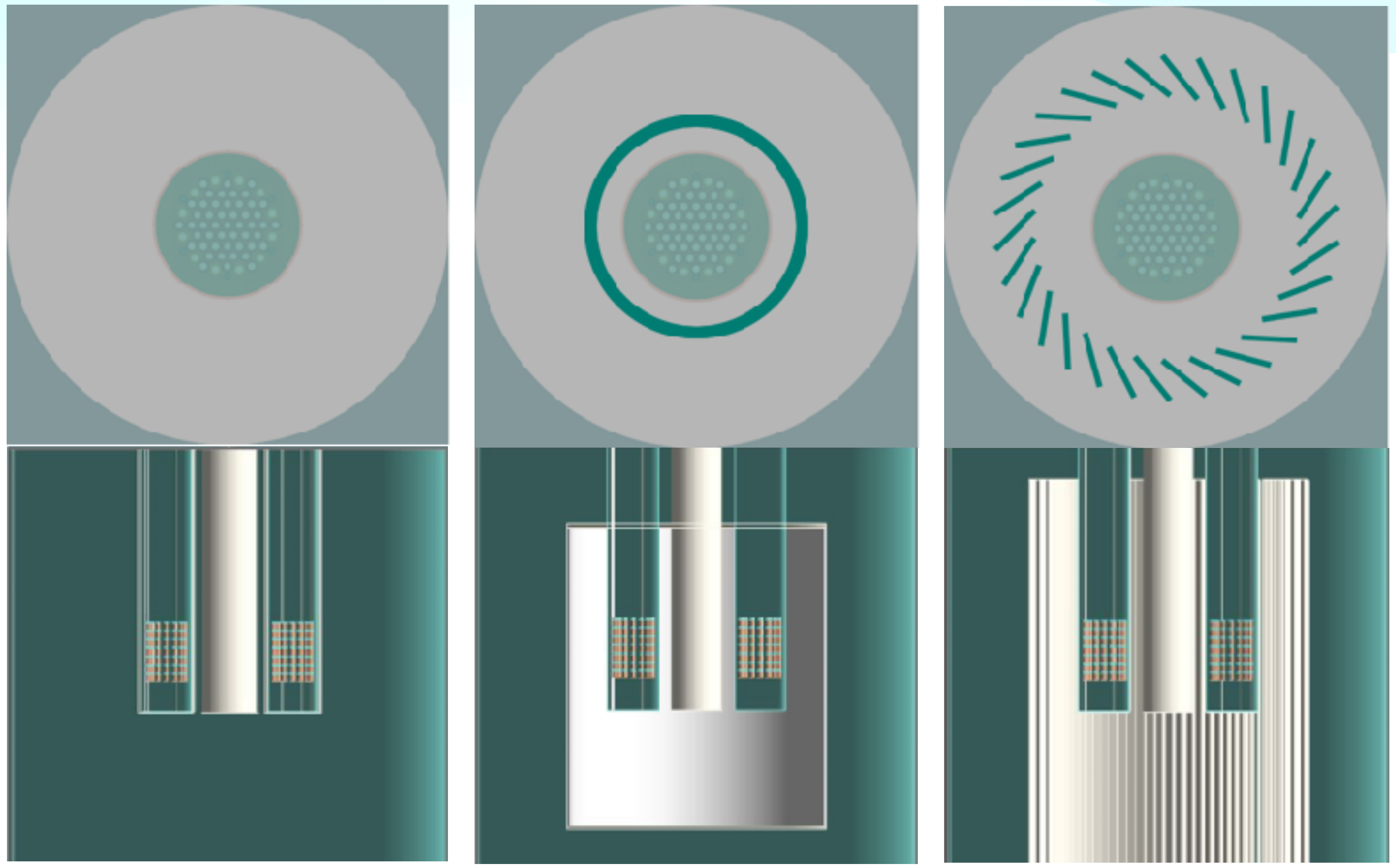
Run a few simulations at different parameters





- Solid neutron moderator design: enclosing tube or turbine-like structure
- 5 design parameters: Radius r, n Panels, Thickness d, Length L and Angle θ

no moderator enclosure turbine-like structure

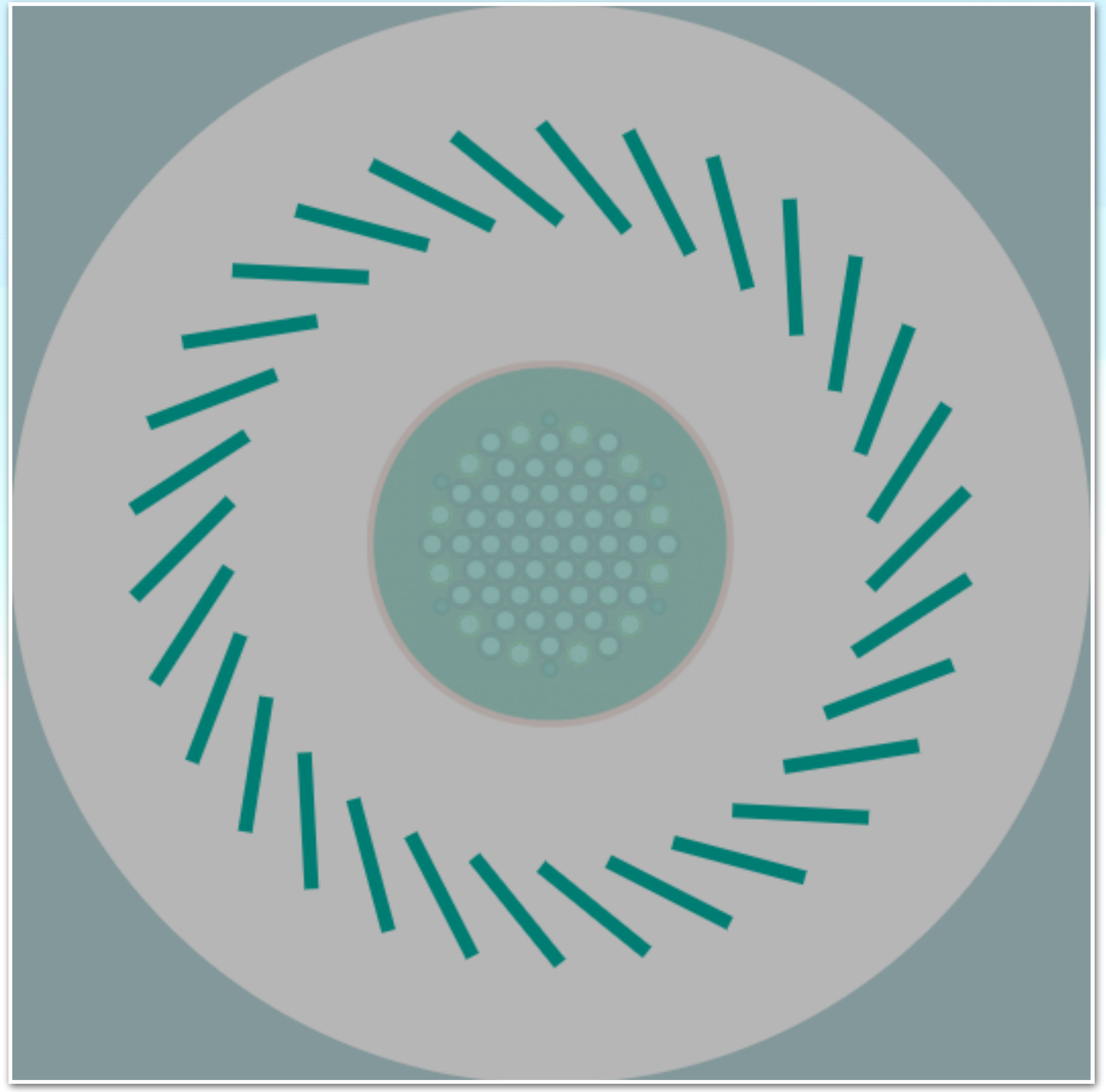


- ➡ **High-dimensional parameter spaces**
- ➡ **High computational cost of Geant4 MC simulations (~200 CPUh)**
- ➡ **Traditional methods like grid searches are impractical**

Why is Our Simulation So Expensive?

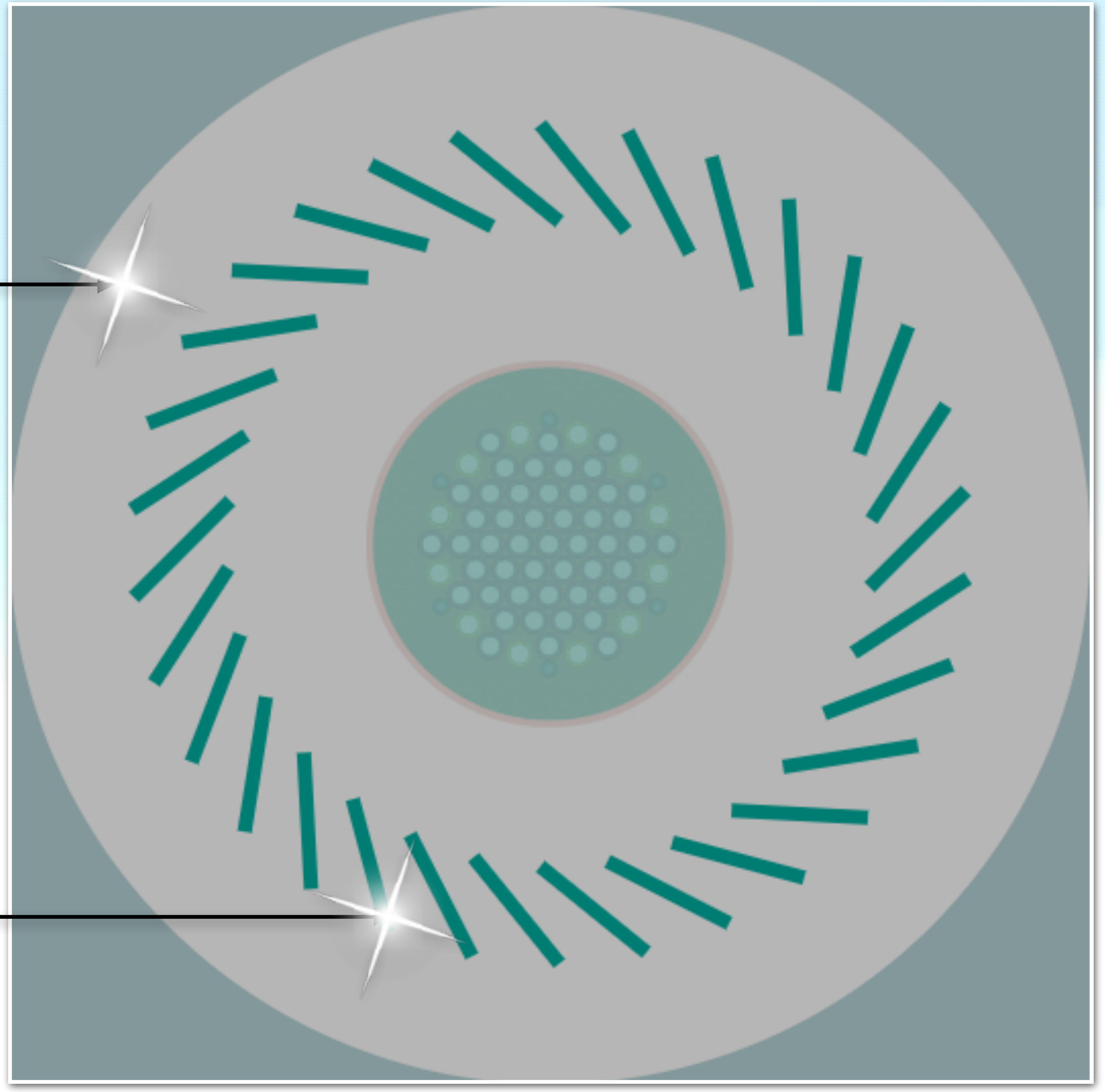
Why is Our Simulation So Expensive?

Given design parameter θ, we have …



Why is Our Simulation So Expensive?

Given design parameter θ, we have …

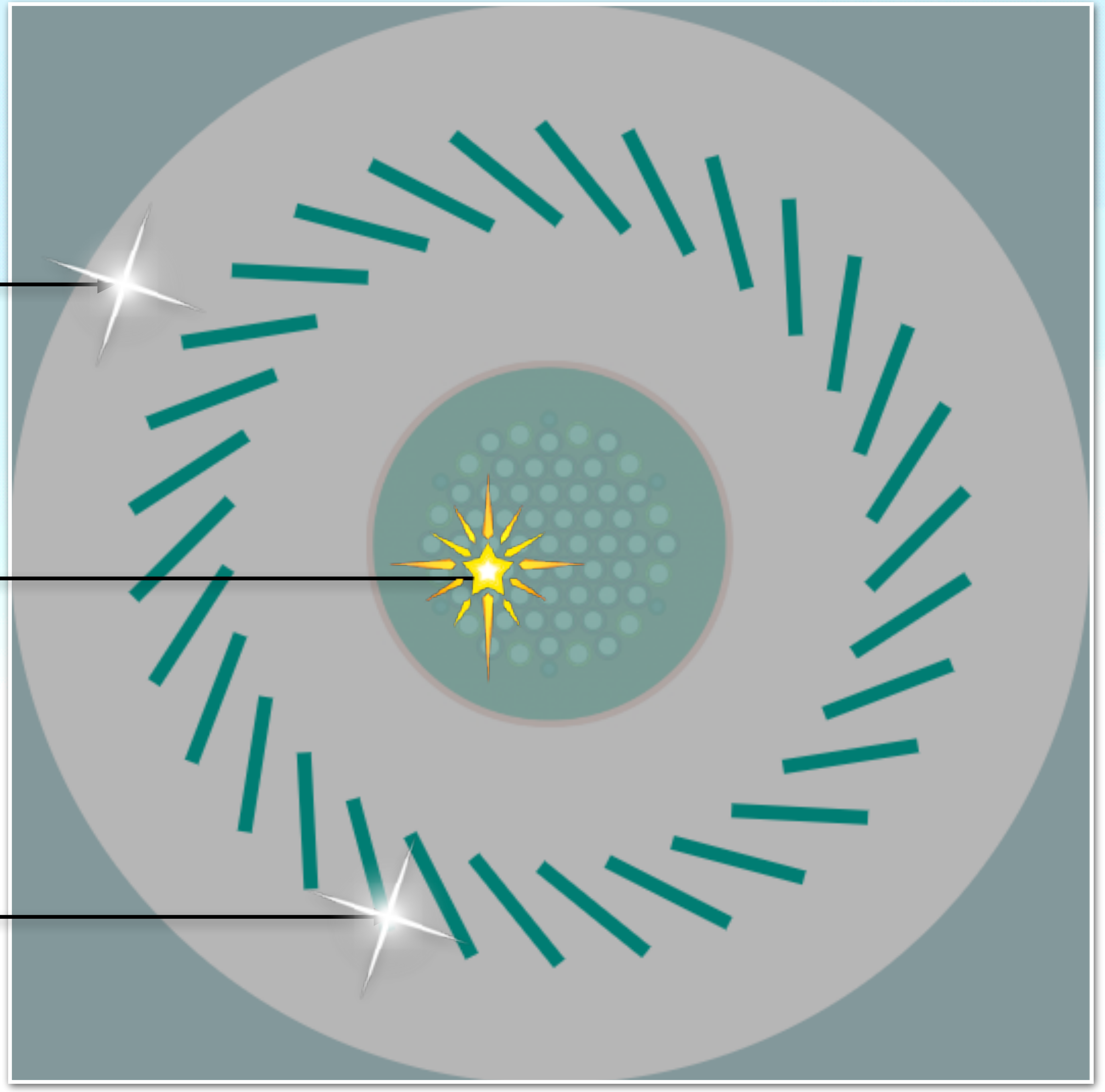


99.99% Neutron that deposit energies elsewhere

99.99% Neutron that are absorbed/slowed by neutron moderator

Why is Our Simulation So Expensive?

Given design parameter θ, we have …



99.99% Neutron that deposit energies elsewhere

0.01% "Lucky" neutron that enters the detector

99.99% Neutron that are absorbed/slowed by neutron moderator

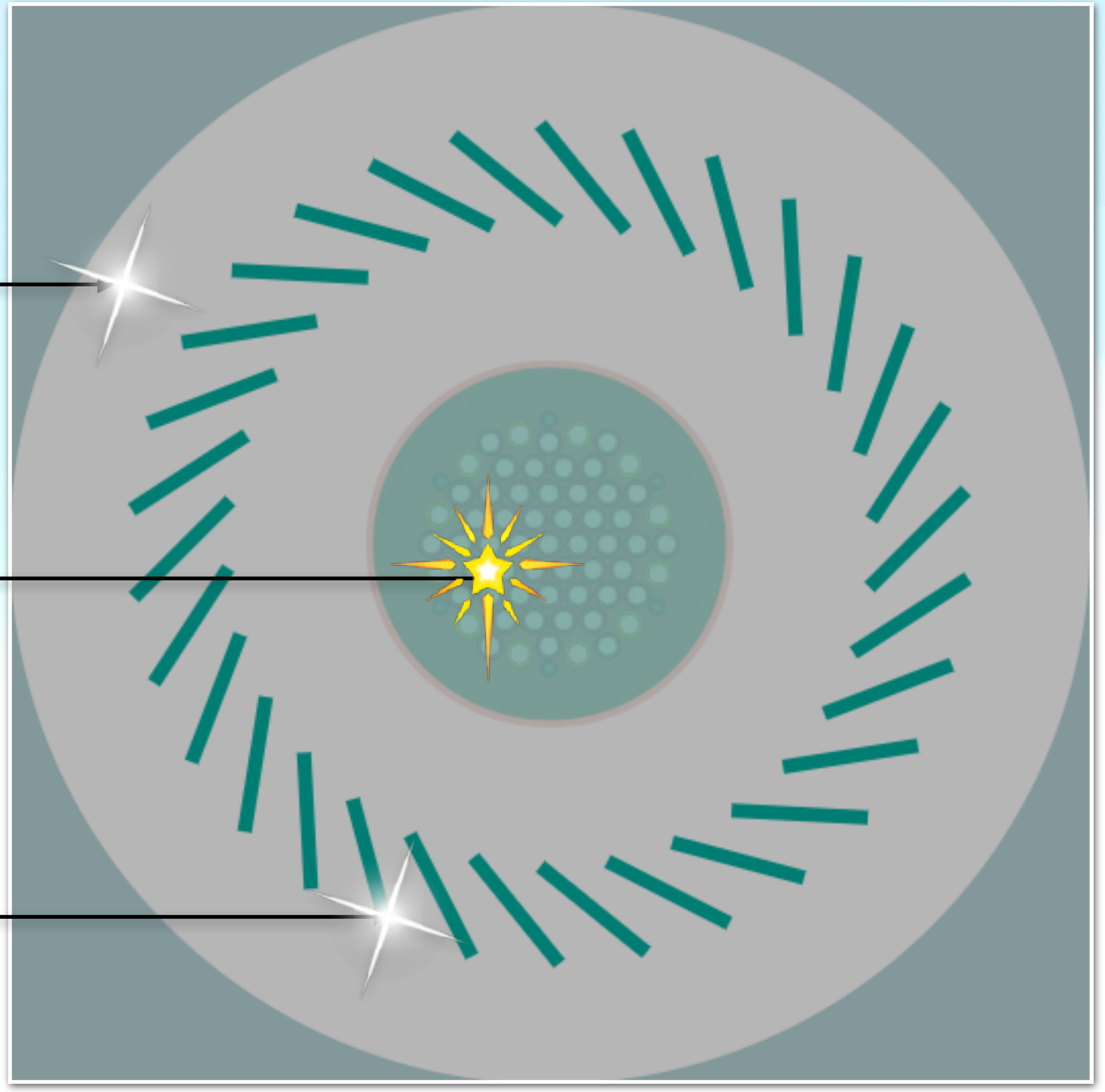
Why is Our Simulation So Expensive?

Given design parameter θ, we have …

99.99% Neutron that deposit energies elsewhere

0.01% "Lucky" neutron that enters the detector

99.99% Neutron that are absorbed/slowed by neutron moderator



Design Metric

$$y = \frac{\#\star}{\#\star} = \frac{m}{N}$$

y is intrinsically very small!

The Rare Event Design Problem

Given design parameter θ:

Event Simulation

Simulate **N** event, each with event-specific parameters ϕ_i (Neutron energy, position etc.)

Each event can be considered as a **Bernoulli RV:**

- $X_i = 1$ if triggered a signal
- $X_i = 0$ otherwise
- $X_i \sim \text{Bernoulli}(p = t(\theta, \phi_i))$

Underlying trigger probability

The Rare Event Design Problem

High-Fidelity (HF) Simulation

Given design parameter θ:

Event Simulation

Simulate **N** event, each with event-specific parameters ϕ_i (Neutron energy, position etc.)

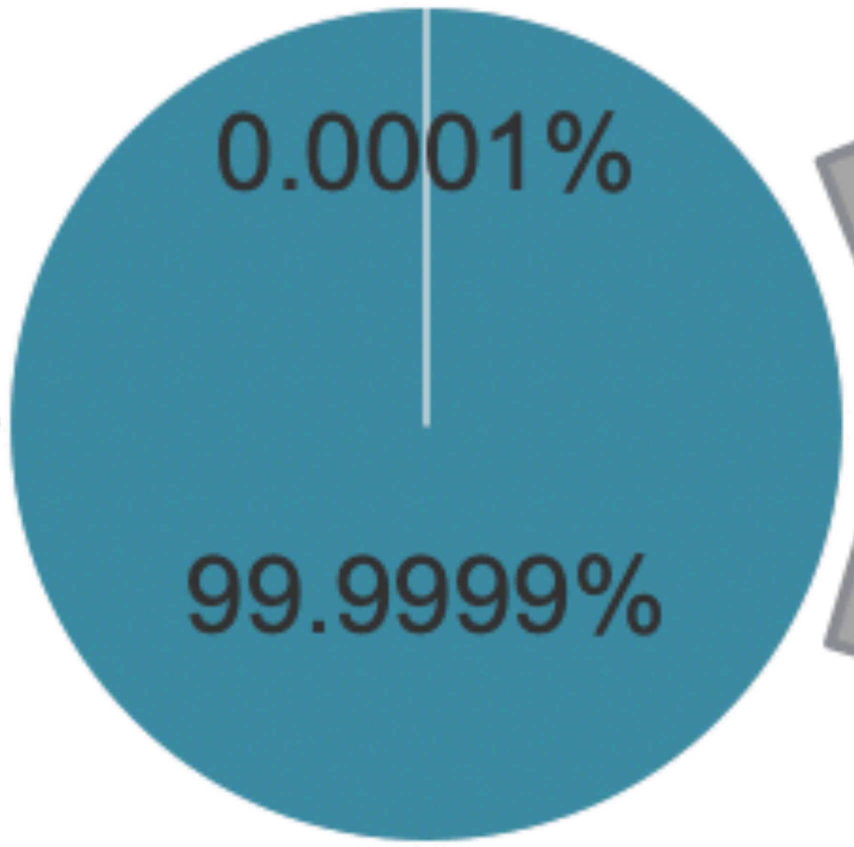
Each event can be considered as a **Bernoulli RV:**

- $X_i = 1$ if triggered a signal
- $X_i = 0$ otherwise
- $X_i \sim \text{Bernoulli}(p = t(\theta, \phi_i))$

Underlying trigger probability

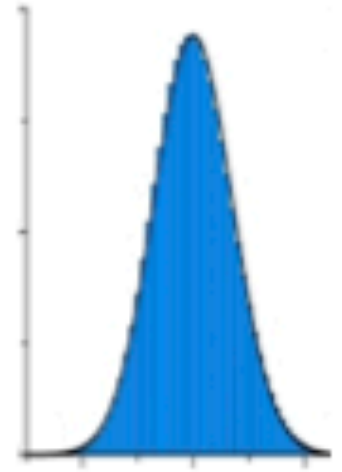
Signal trigger rate $y = \frac{m}{N}$

$m = \sum_{i=0}^{N} X_i \sim \text{Poisson}(N\bar{t}(\theta))$



m $N\text{-}m$

Large N



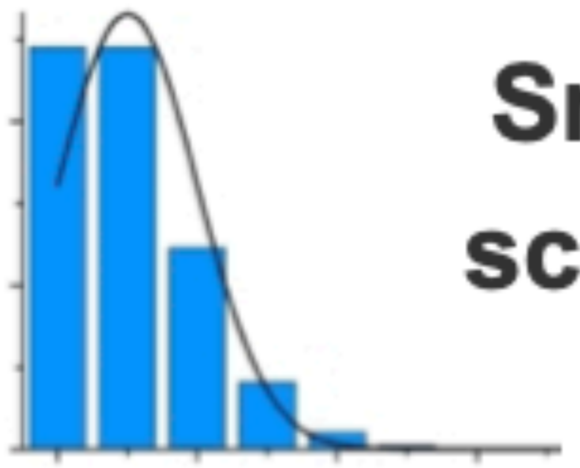
$\bar{t}(\theta) = \int t(\theta, \phi) g(\phi) d\phi$

$y \sim \mathcal{N}(\bar{t}(\theta), \frac{\bar{t}(\theta)}{N})$

Low Variance, High Cost.

Key scenarios

Small N



Small N scenario

$y \in \{0/N, 1/N, \ldots\}$

High Variance, Low Cost.

Expensive simulator

Low-Fidelity (LF) Simulation

The Rare Event Design Problem

Ultimate Goal

Emulate $\bar{t}(\theta)$, or function $f:\theta \rightarrow y$ with as small N as possible

Given design parameter θ:

Event Simulation

Simulate **N** event, each with event-specific parameters ϕ_i (Neutron energy, position etc.)

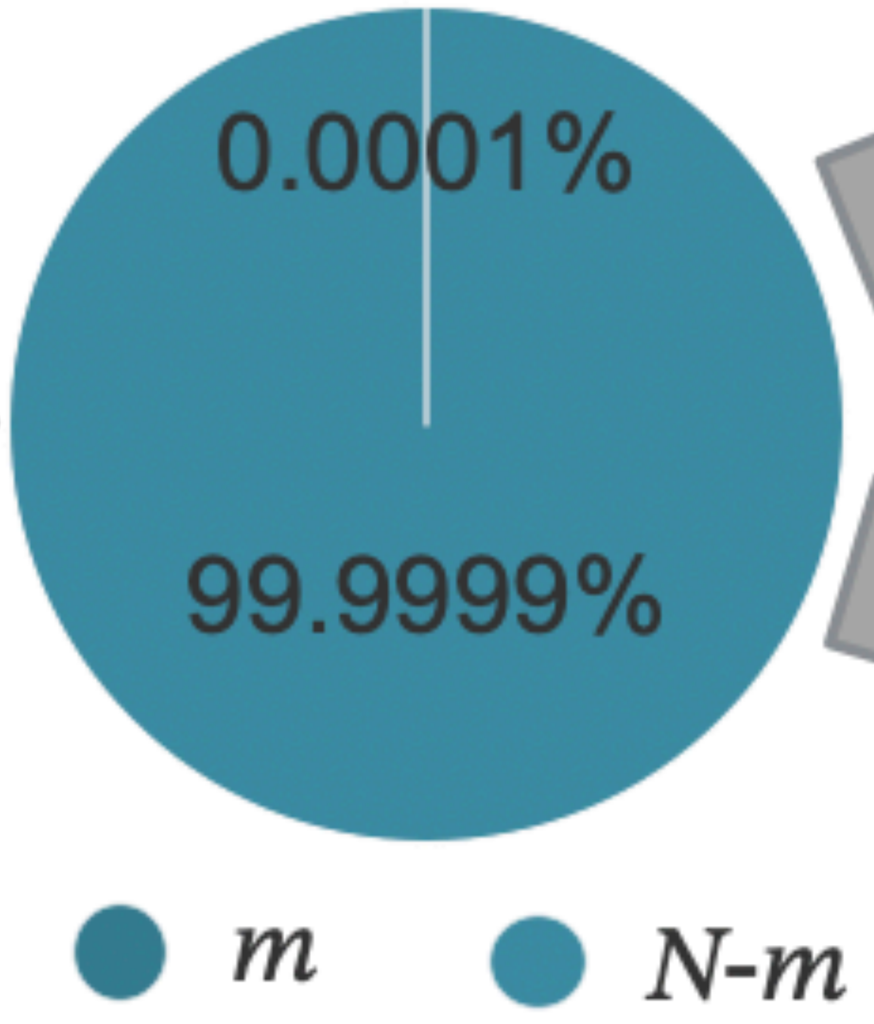
Each event can be considered as a **Bernoulli RV:**

- $X_i = 1$ if triggered a signal
- $X_i = 0$ otherwise
- $X_i \sim \text{Bernoulli}(p = t(\theta, \phi_i))$

Underlying trigger probability

Signal trigger rate $y = \frac{m}{N}$

$m = \sum_{i=0}^{N} X_i \sim \text{Poisson}(N\bar{t}(\theta))$



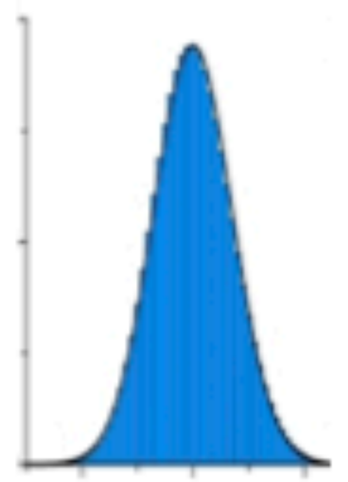
Large N

Small N

Expensive simulator

High-Fidelity (HF) Simulation

$\bar{t}(\theta) = \int t(\theta, \phi) g(\phi) d\phi$

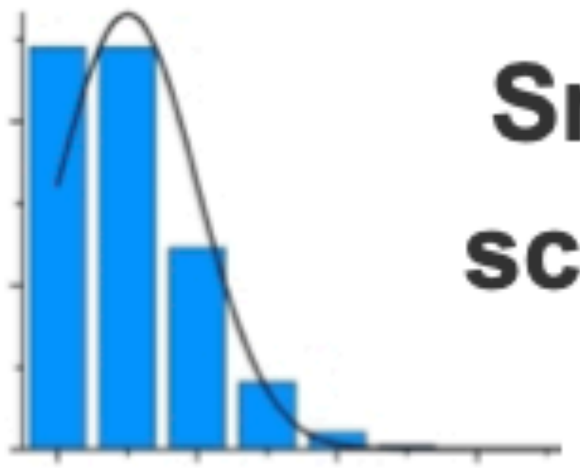


$y \sim \mathcal{N}(\bar{t}(\theta), \frac{\bar{t}(\theta)}{N})$

Low Variance, High Cost.

Key scenarios

Small N scenario



$y \in \{0/N, 1/N, \dots\}$

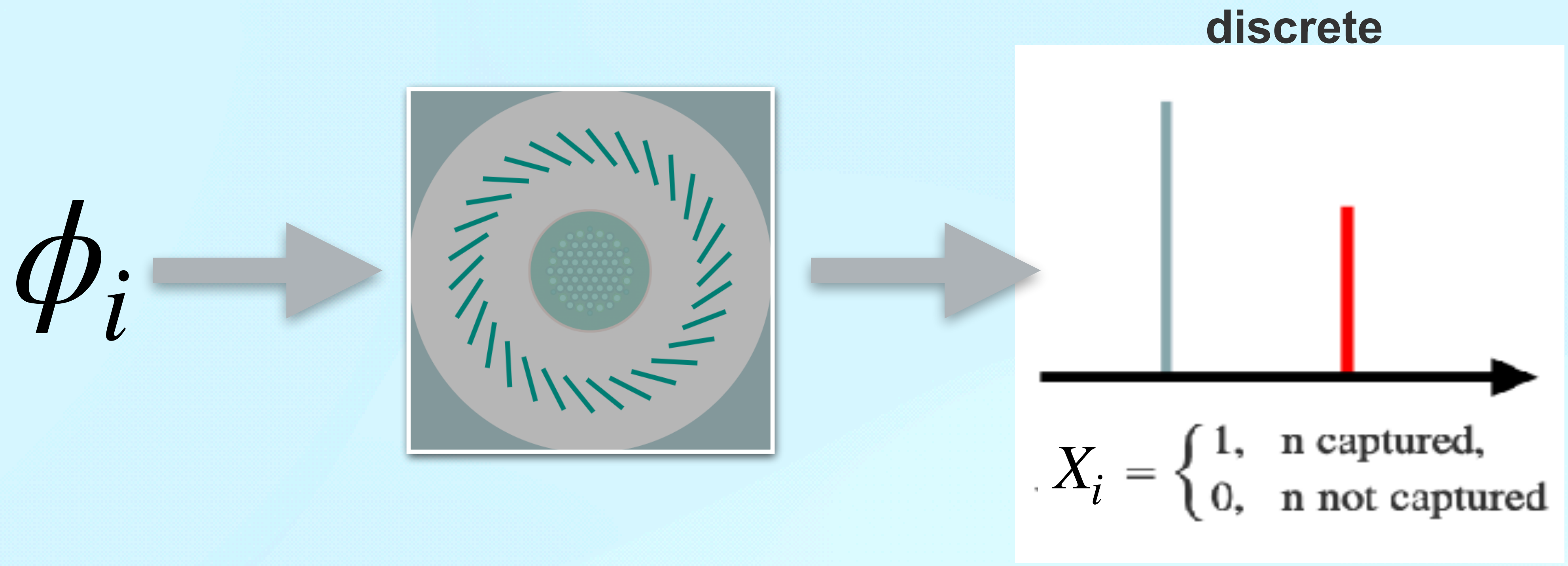
High Variance, Low Cost.

Low-Fidelity (LF) Simulation

RESuM: The Rare Event Surrogate Model

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

Key Insight 1: Incorporating Prior Information with Conditional Neural Process

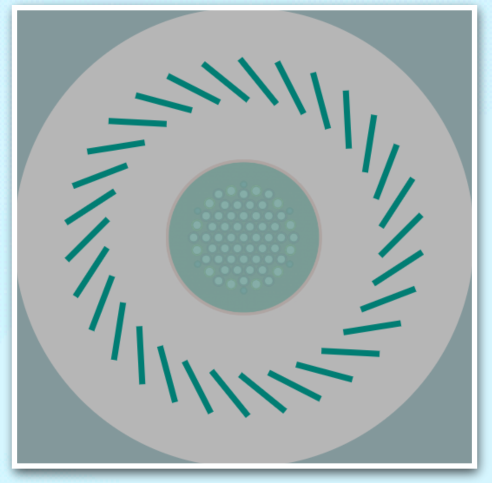


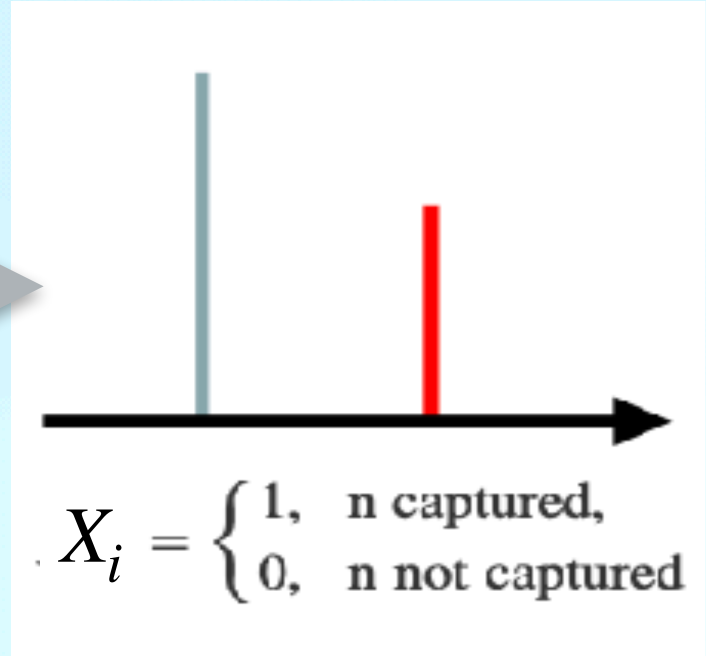
RESuM: The Rare Event Surrogate Model

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

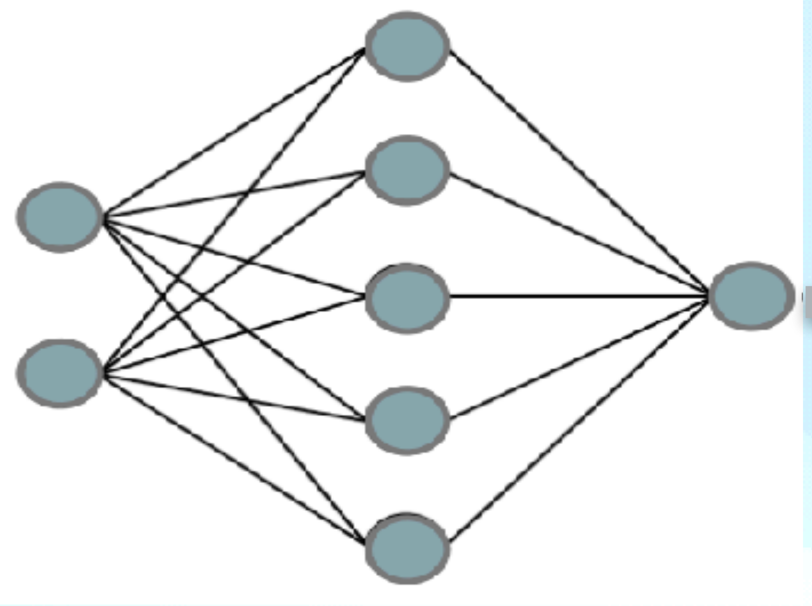
Key Insight 1: Incorporating Prior Information with <u>Conditional Neural Process</u>

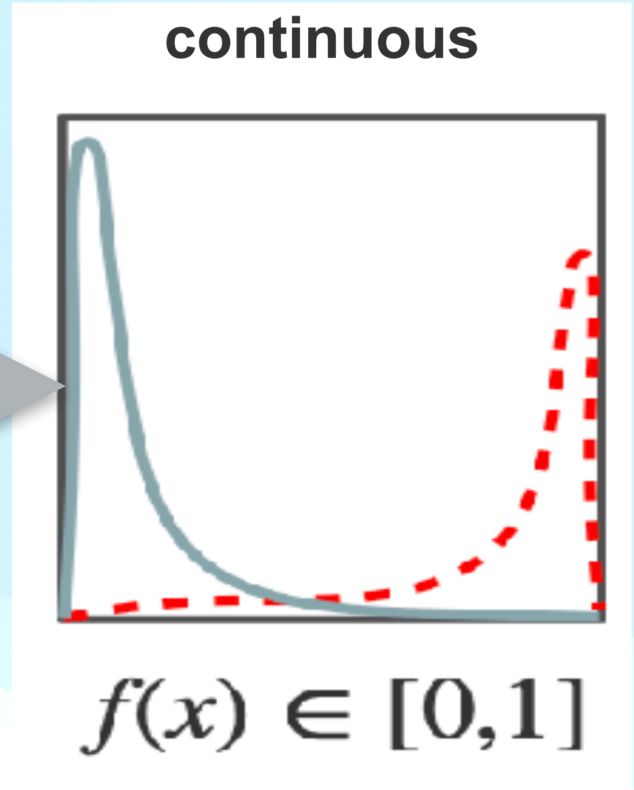
ϕ_i





ϕ_i



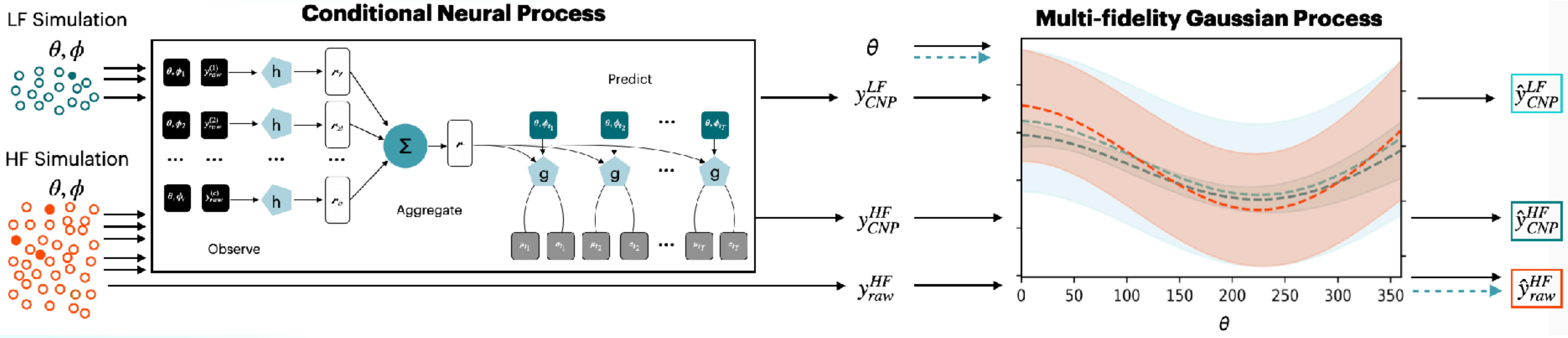


RESuM: The Rare Event Surrogate Model

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

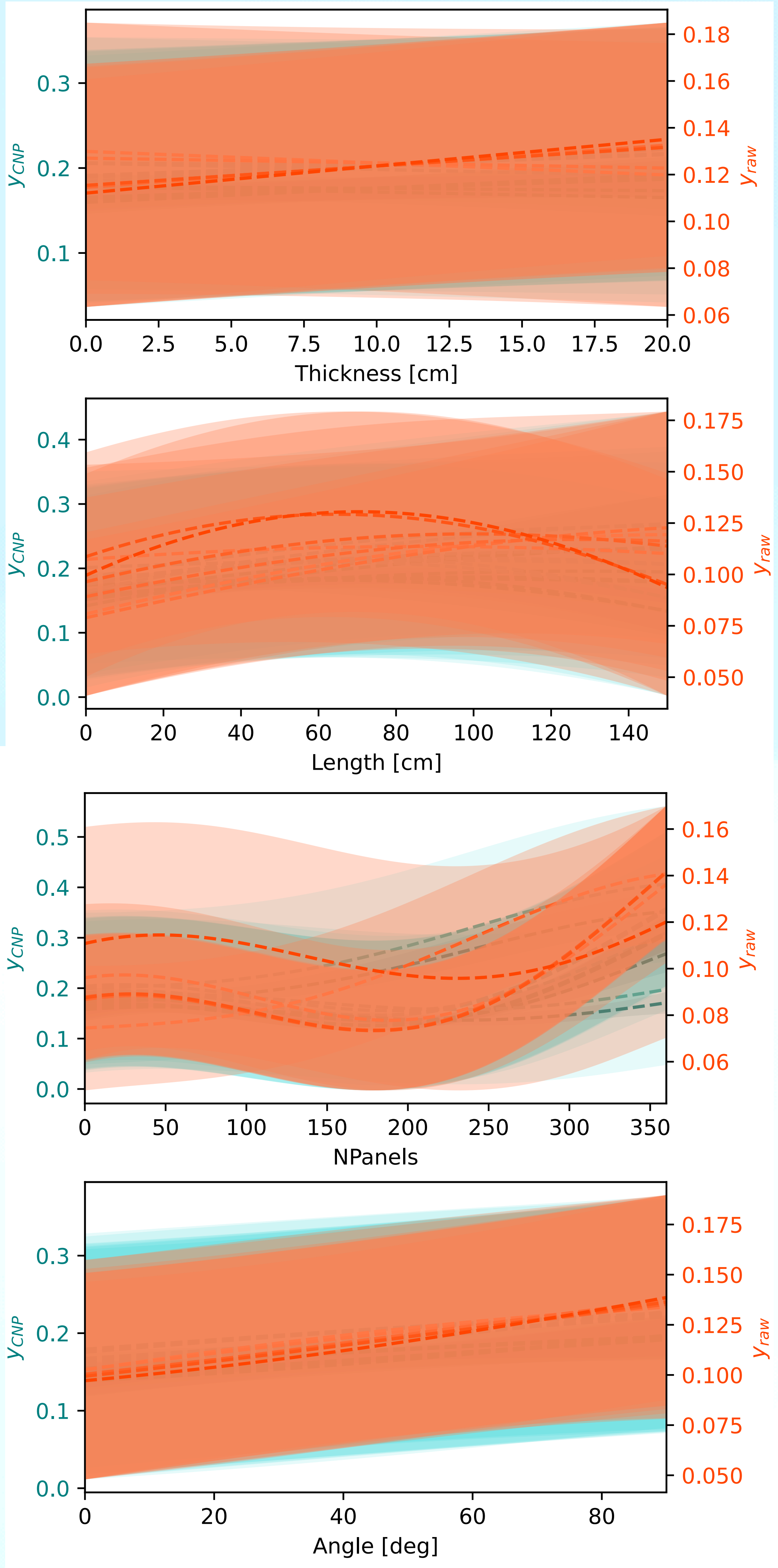
Key Insight 1: Incorporating Prior Information with Conditional Neural Process

Key Insight 2: Multi-Fidelity Gaussian Process

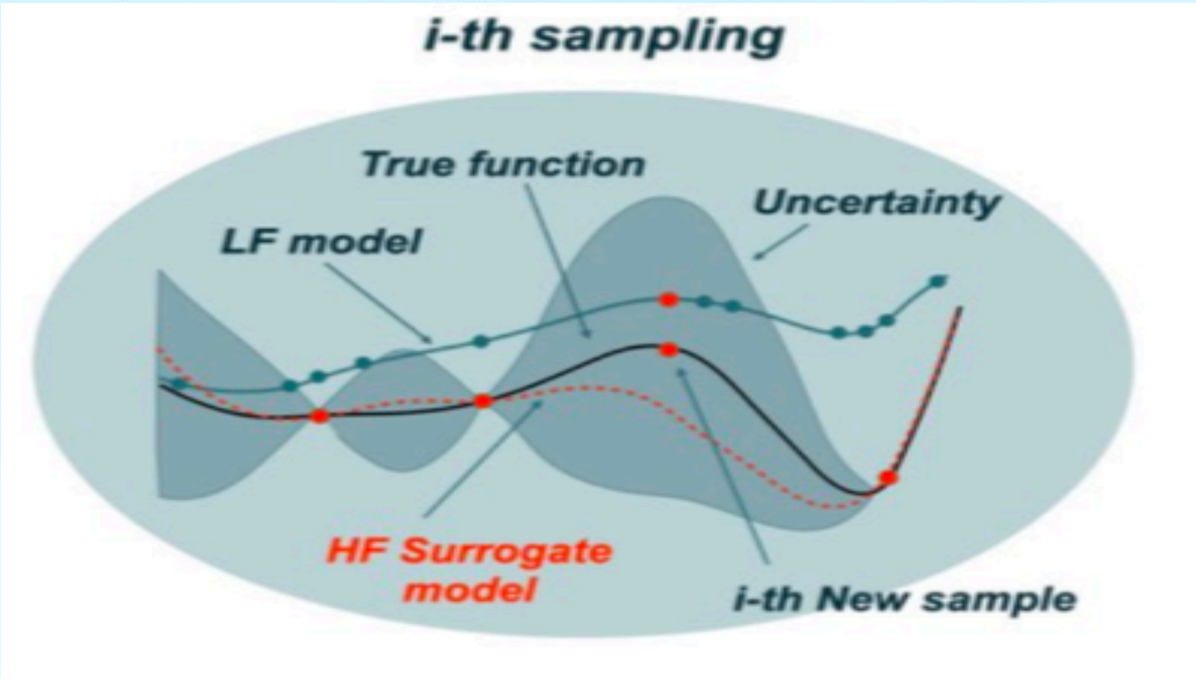


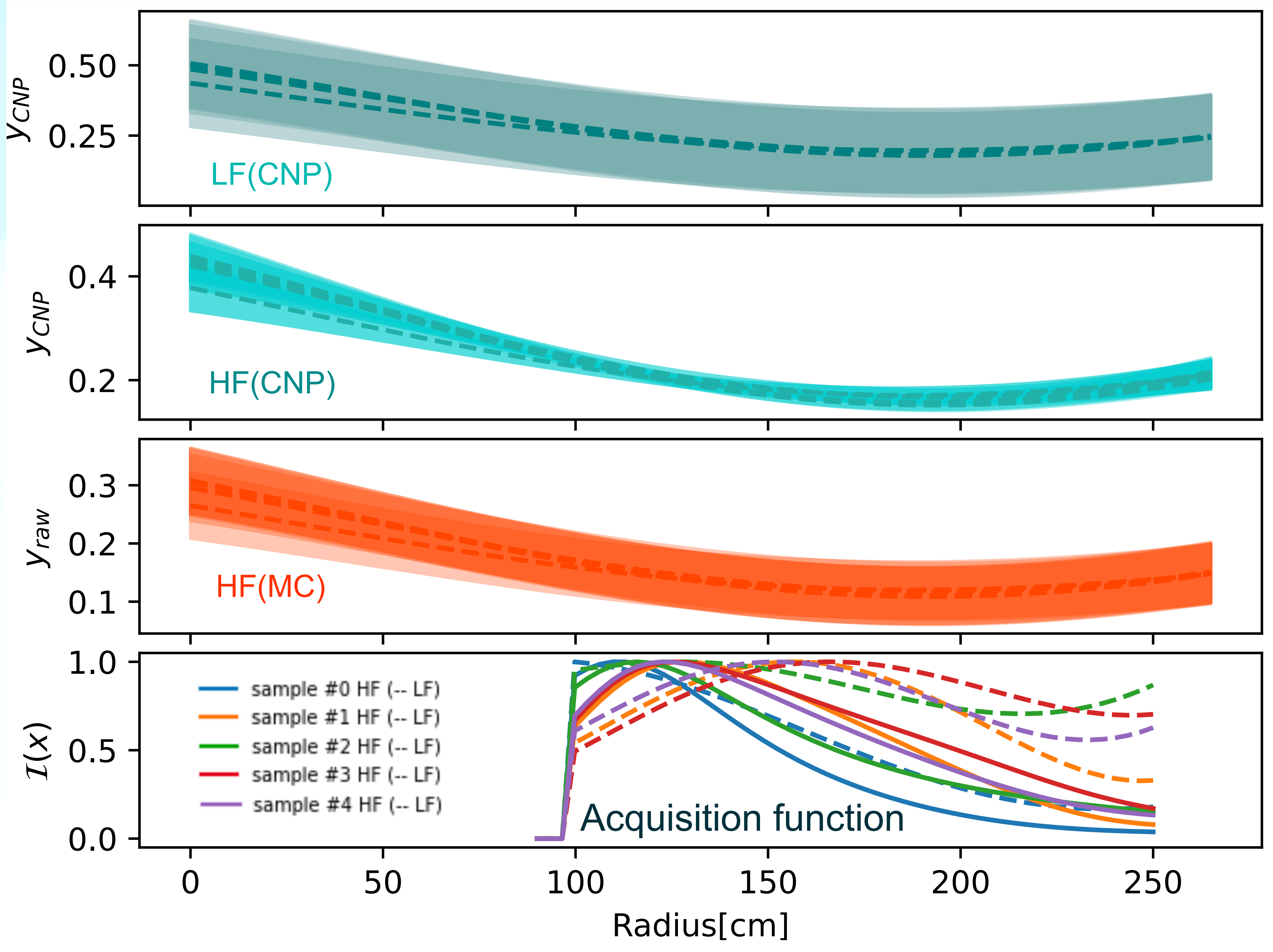
RESuM: The Rare Event Surrogate Model

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025



- Modeling of 5 dim space (r, t, θ, n, L) with 3 fidelities (HF(MC), HF(CNP) and LF(CNP))
- model evolution shown as projection on r, t, n, θ 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.

Benchmarking RESuM

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

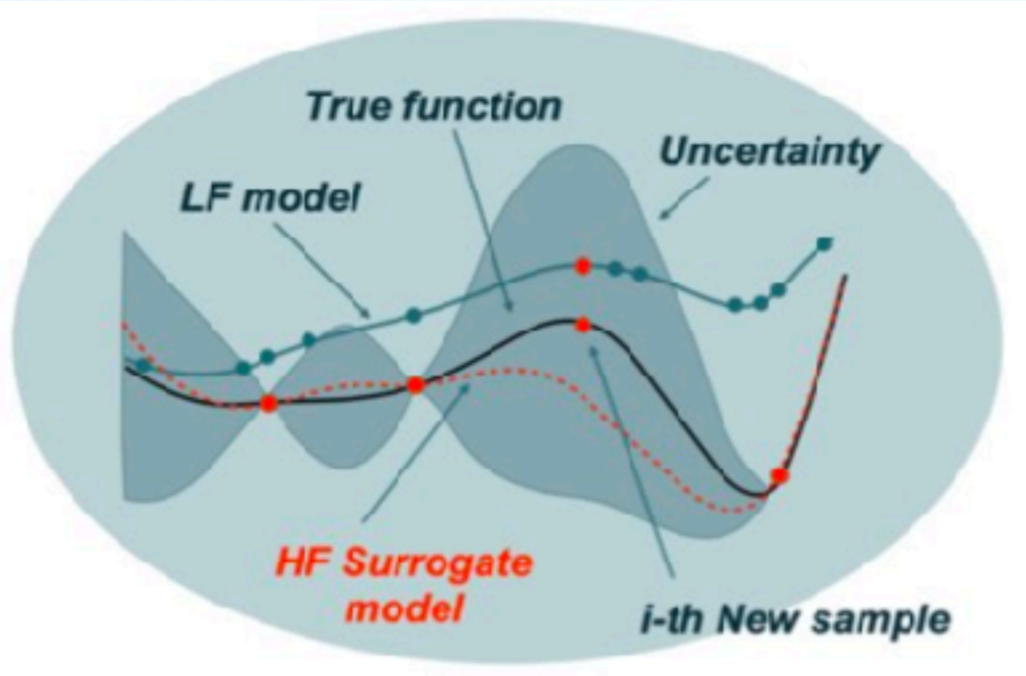
We test RESuM vs. Other model on 100 out-of-sample HF Simulation

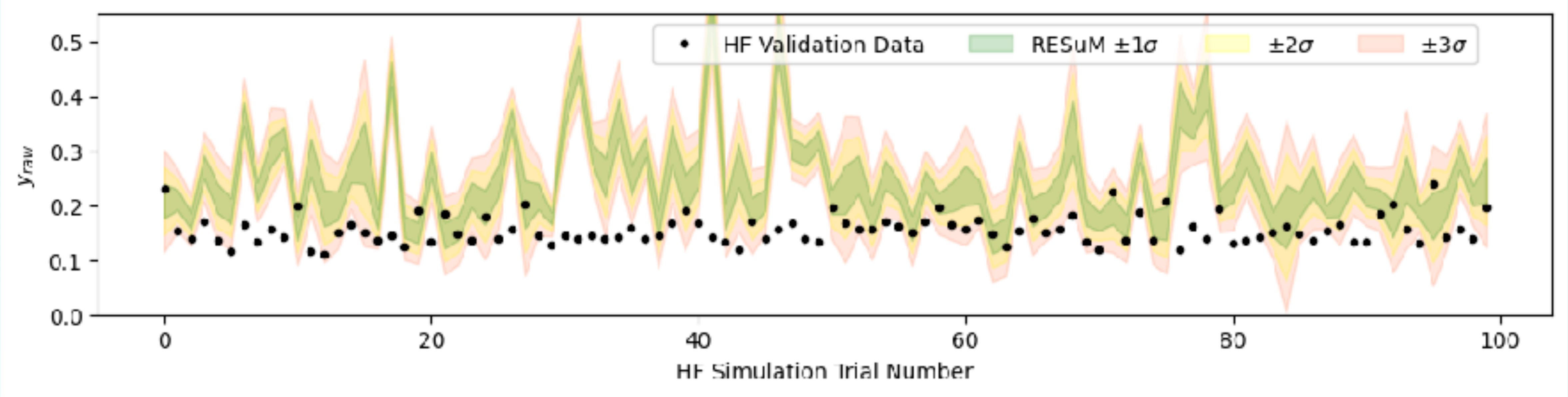
Benchmarking RESuM

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

We test RESuM vs. Other model on 100 out-of-sample HF Simulation

MFGP



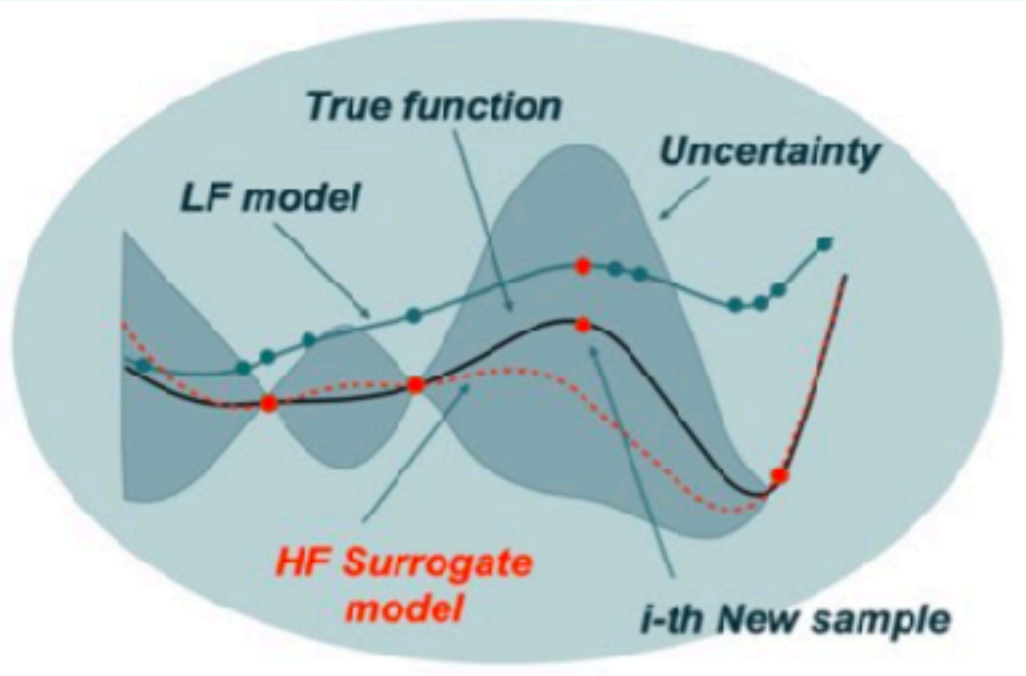


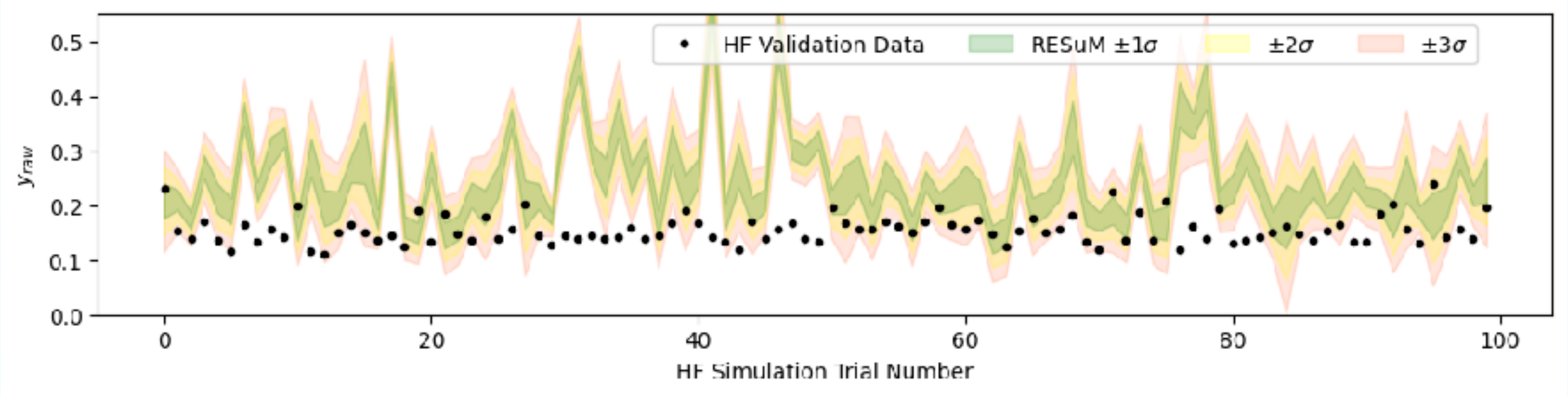
Benchmarking RESuM

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

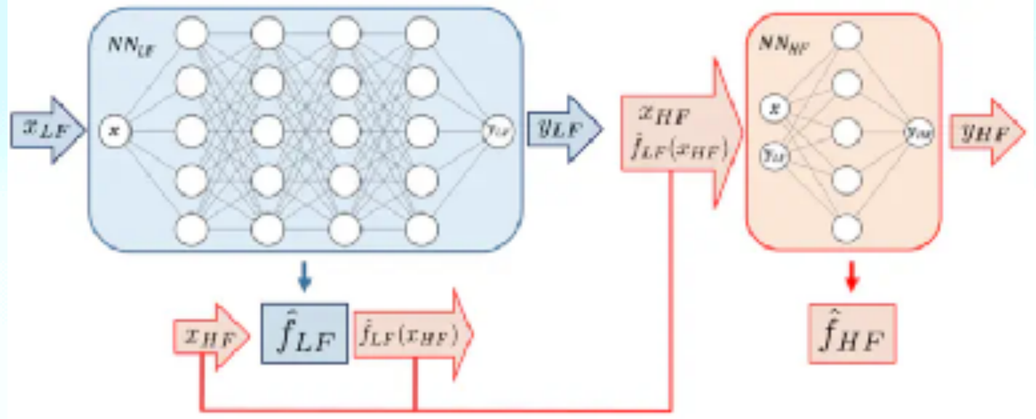
We test RESuM vs. Other model on 100 out-of-sample HF Simulation

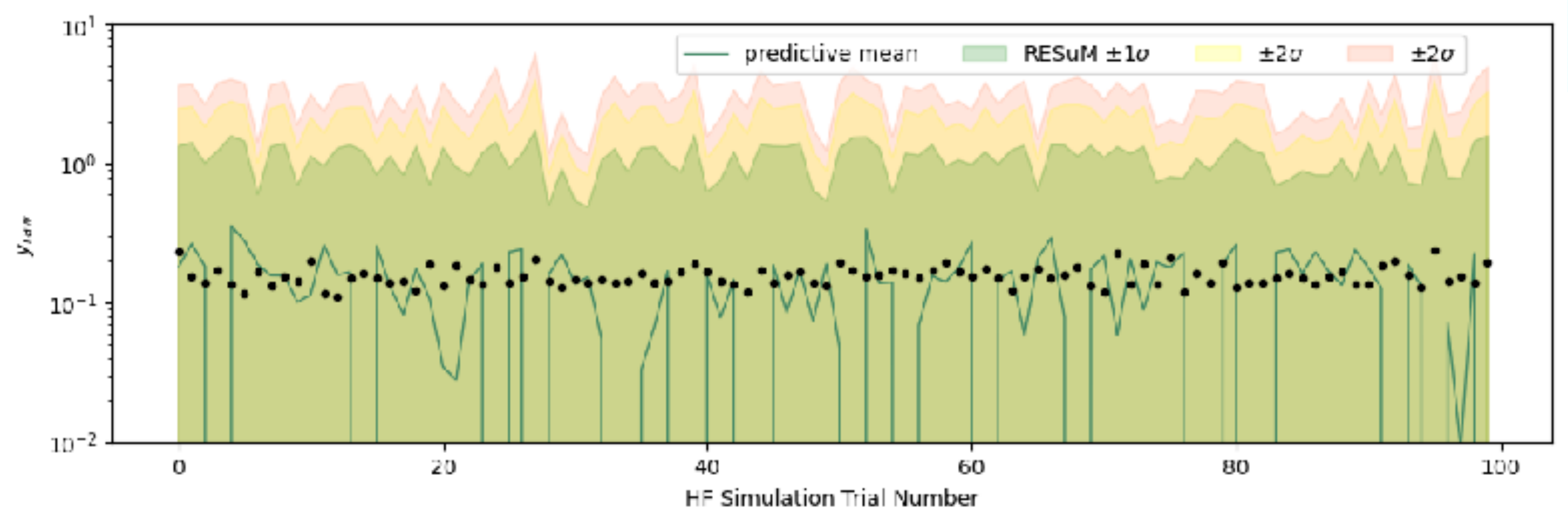
MFGP





MF-BNN



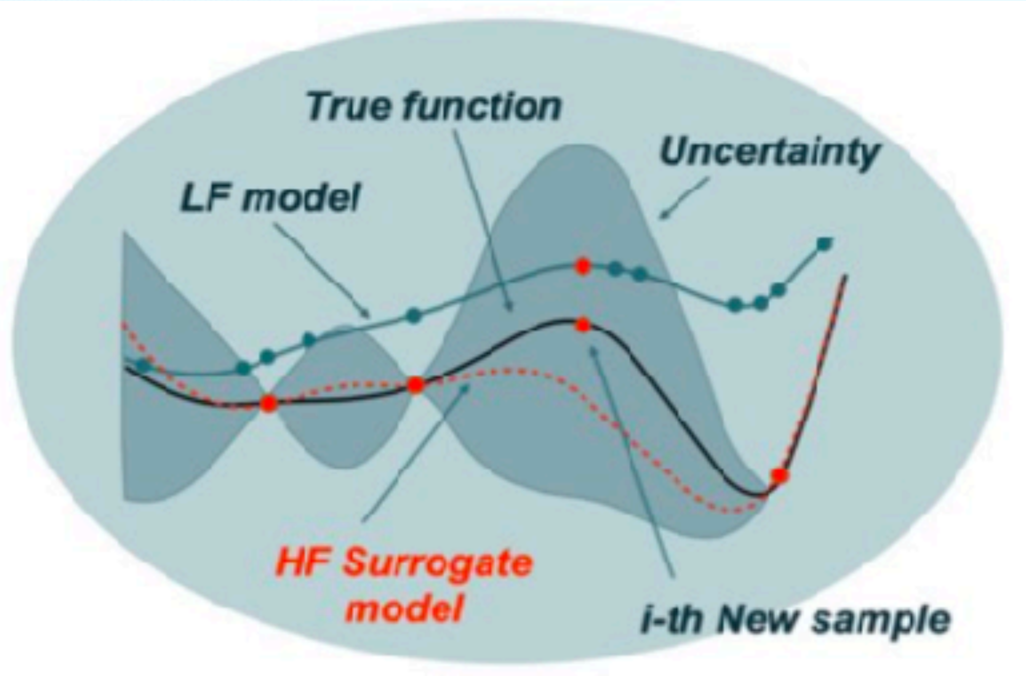


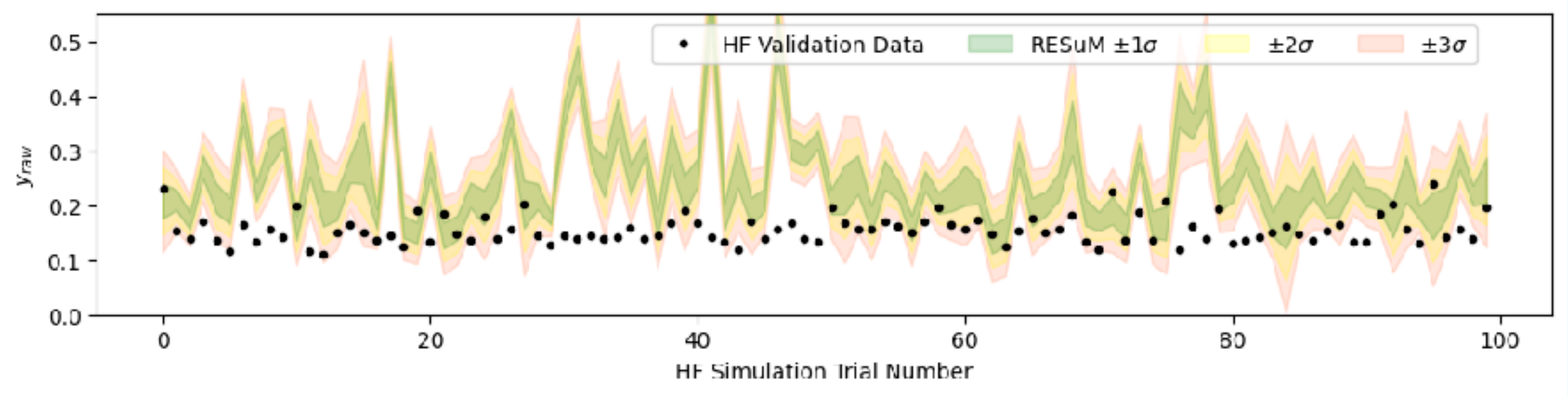
Benchmarking RESuM

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

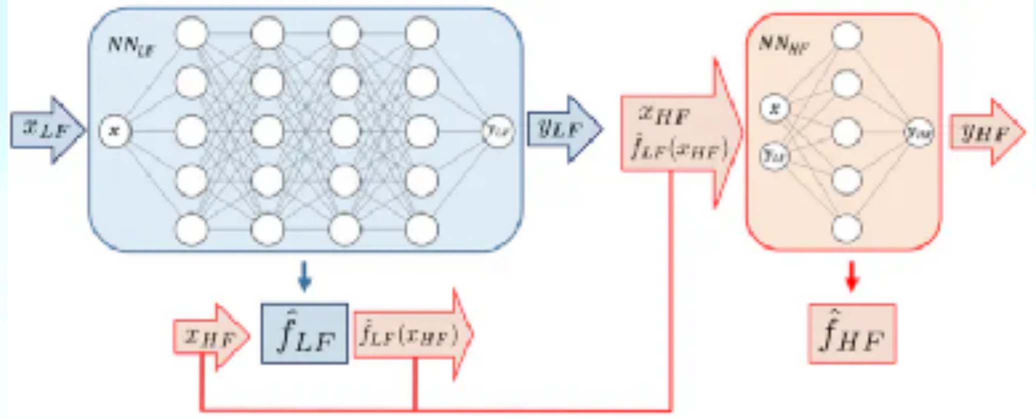
We test RESuM vs. Other model on 100 out-of-sample HF Simulation

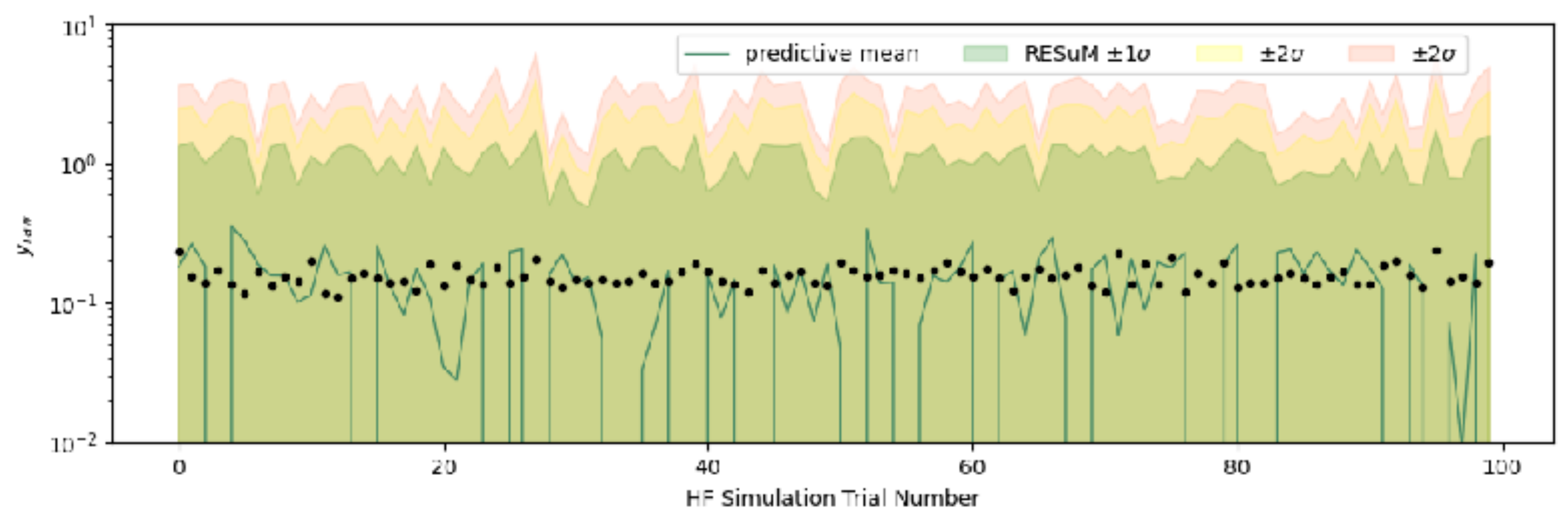
MFGP



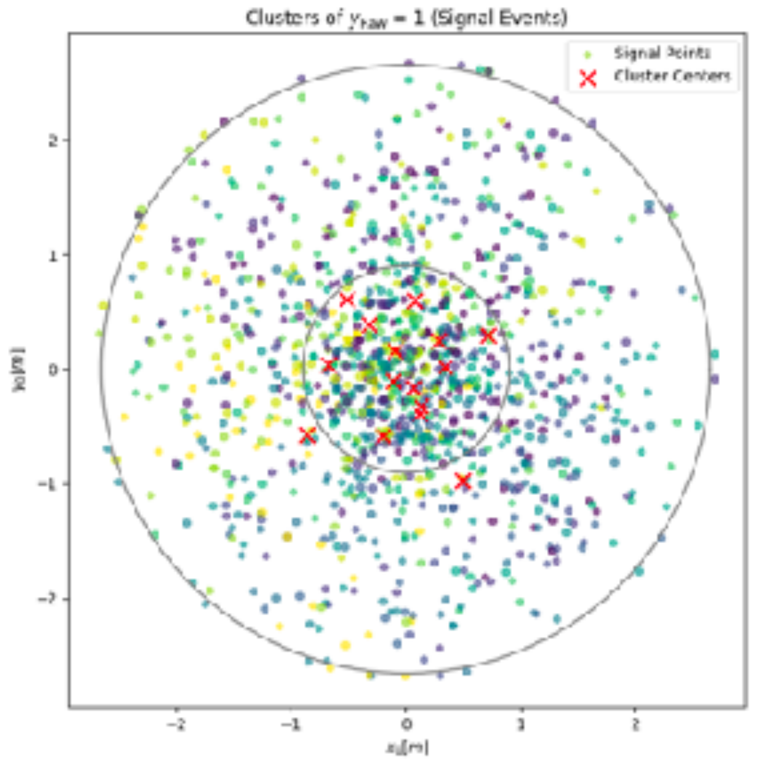


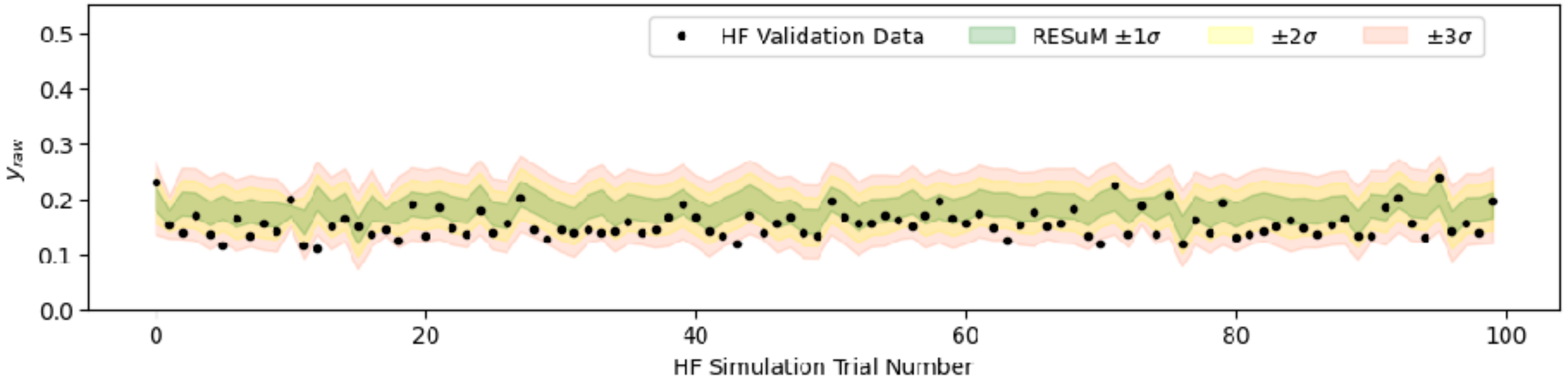
MF-BNN





MFGP with Adaptive Importance Sampling



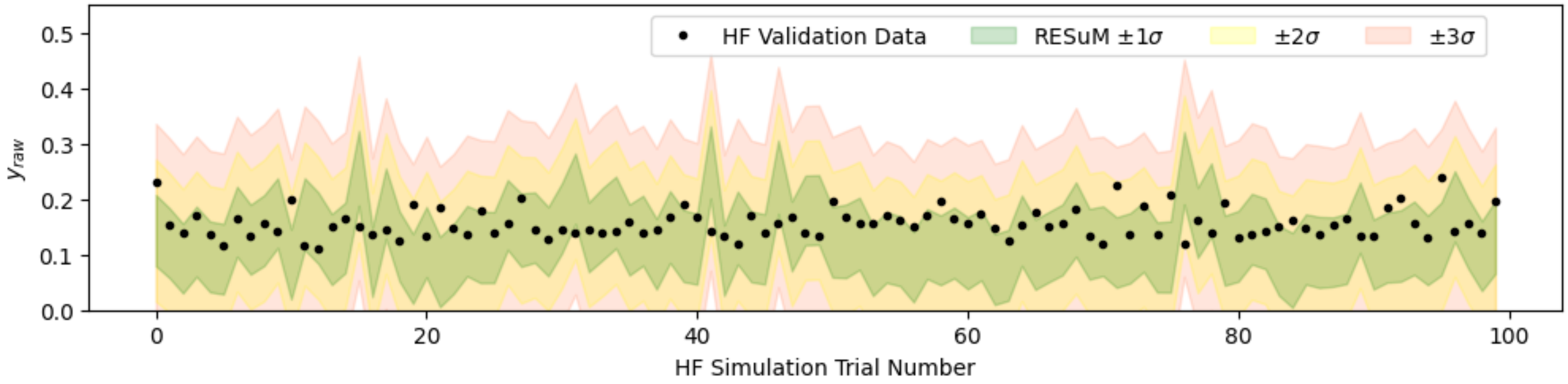


Benchmarking RESuM

A. Shuetz, A.W. Poon, A. Li,
arXiv:2410.03873
Accepted by ICLR 2025

We test RESuM vs. Other model on 100 out-of-sample HF Simulation

RESuM



Model	1σ Coverage	2σ Coverage	3σ Coverage	MSE
MFGP	2	4	5	0.0095
MF-BNN	100	100	100	0.471
AIS+MFGP	33	75	95	0.0012
RESuM	69	95	100	0.0024
RESuM (100 iter)	62.38	92.23	99.59	0.0037

MSE:

$(y-\hat{y})^2$

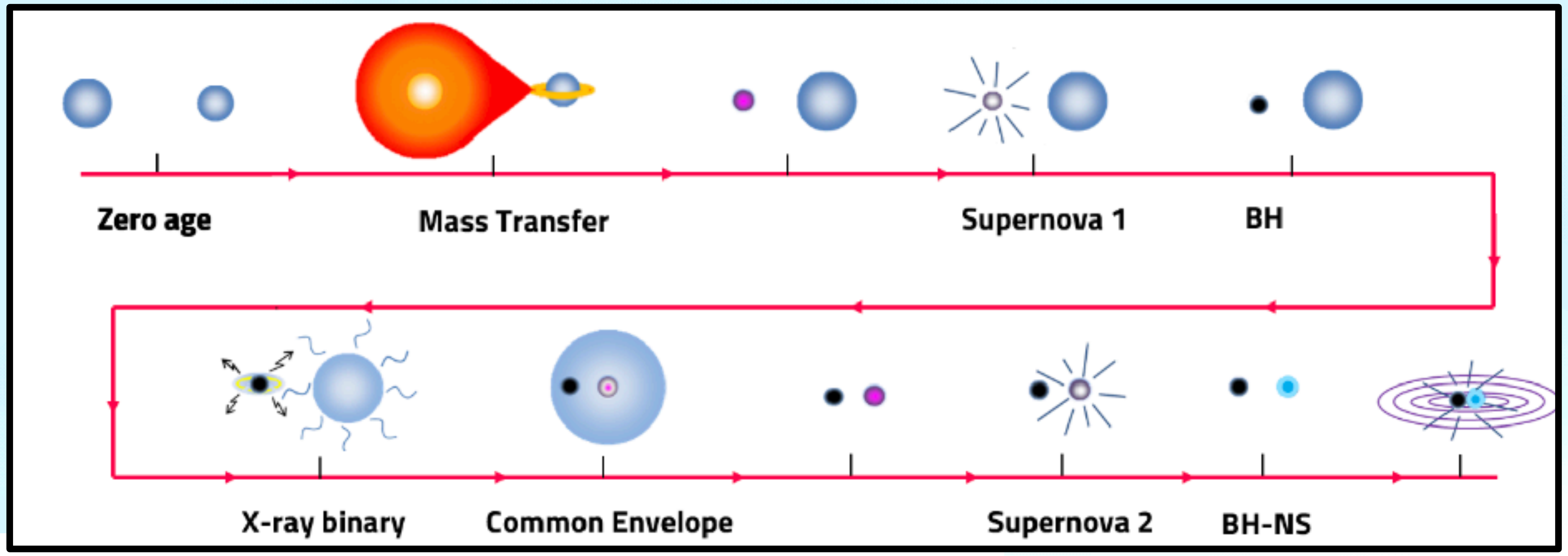
Coverage:

percentage of y falling in $\hat{y} \pm 1/2/3\sigma$

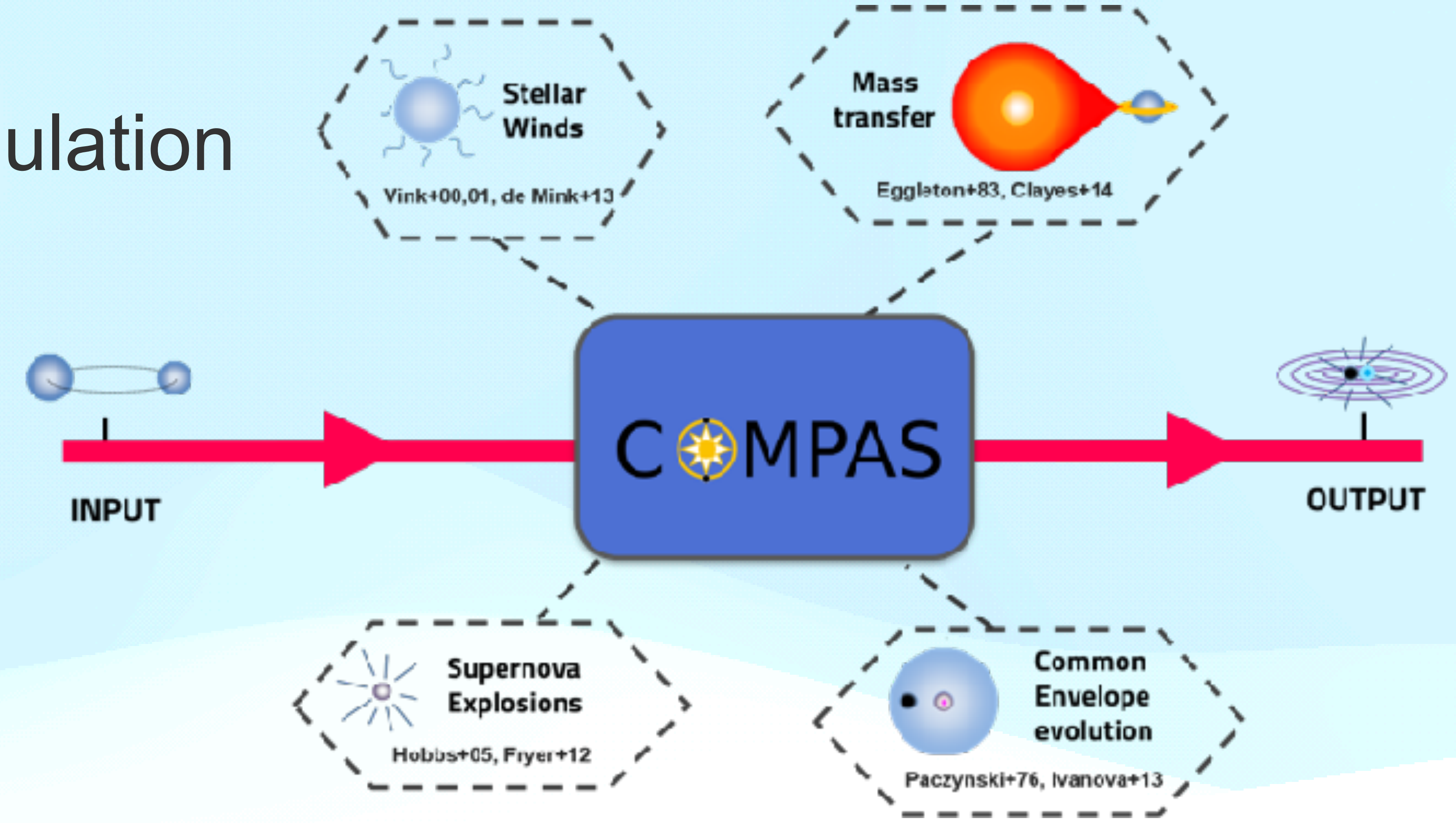
Application: Binary Black Hole Population Synthesis

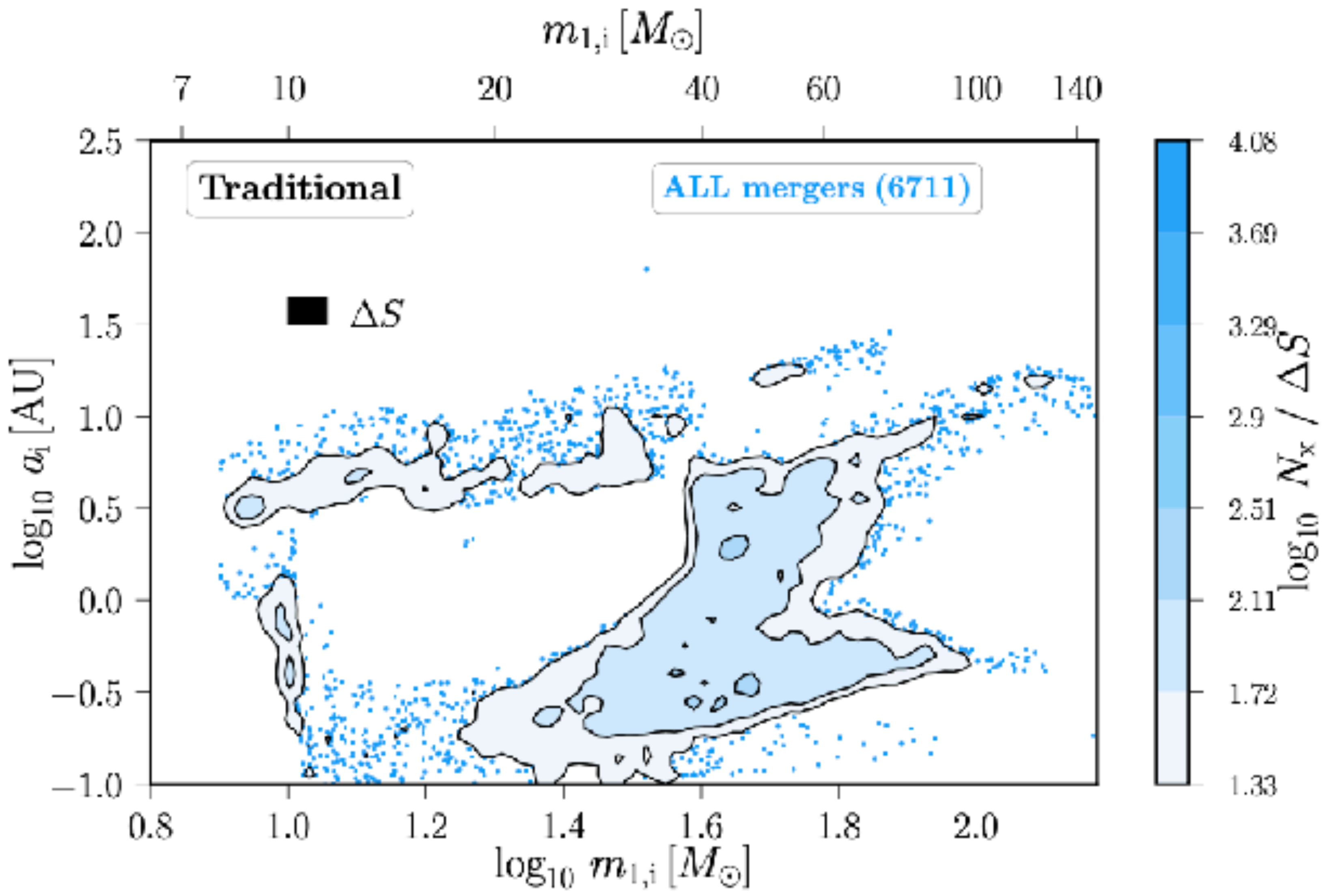
In Collaboration with Prof. Floor Broeckgarden (UCSD)

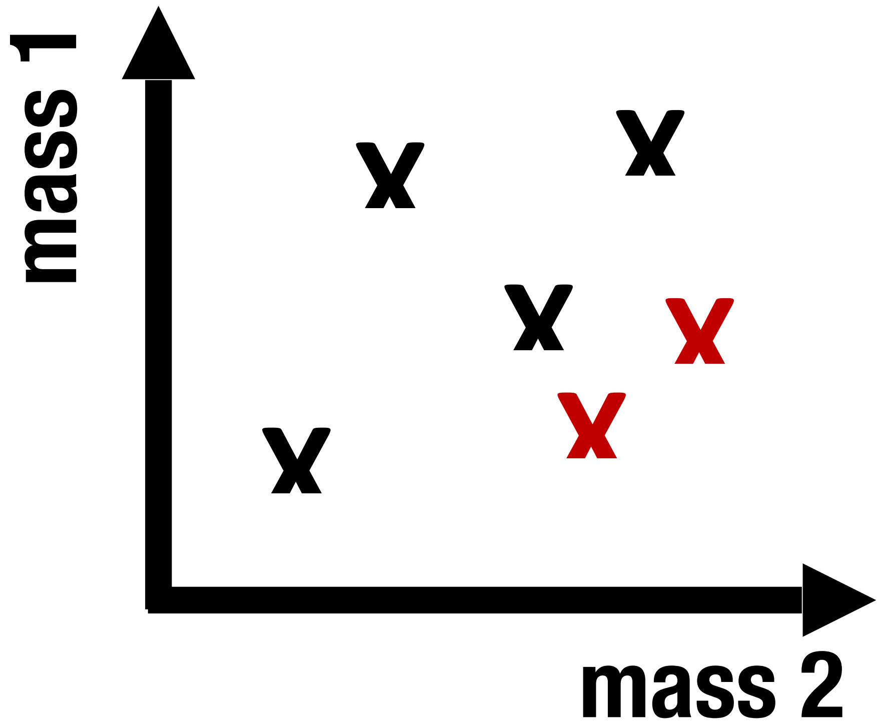
Binary Black Hole Merger

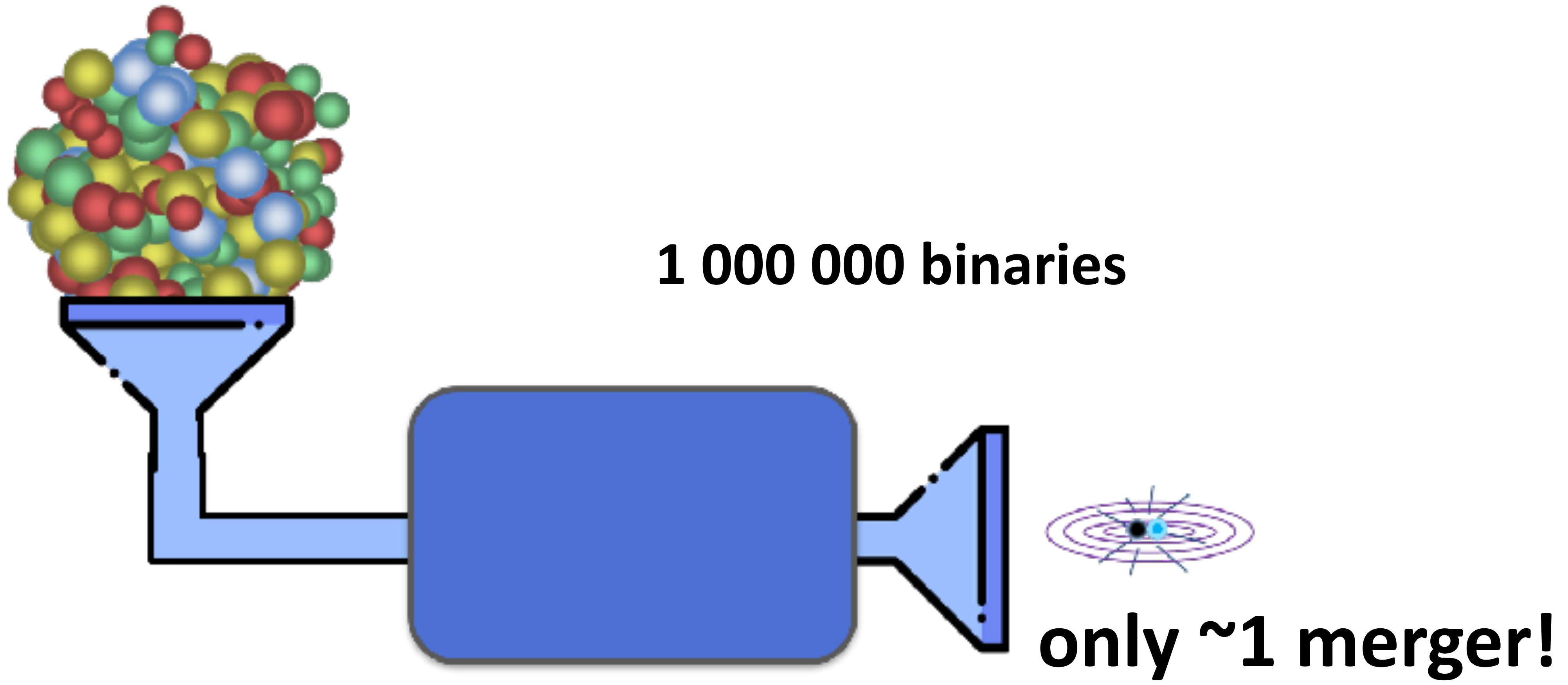


Simulation





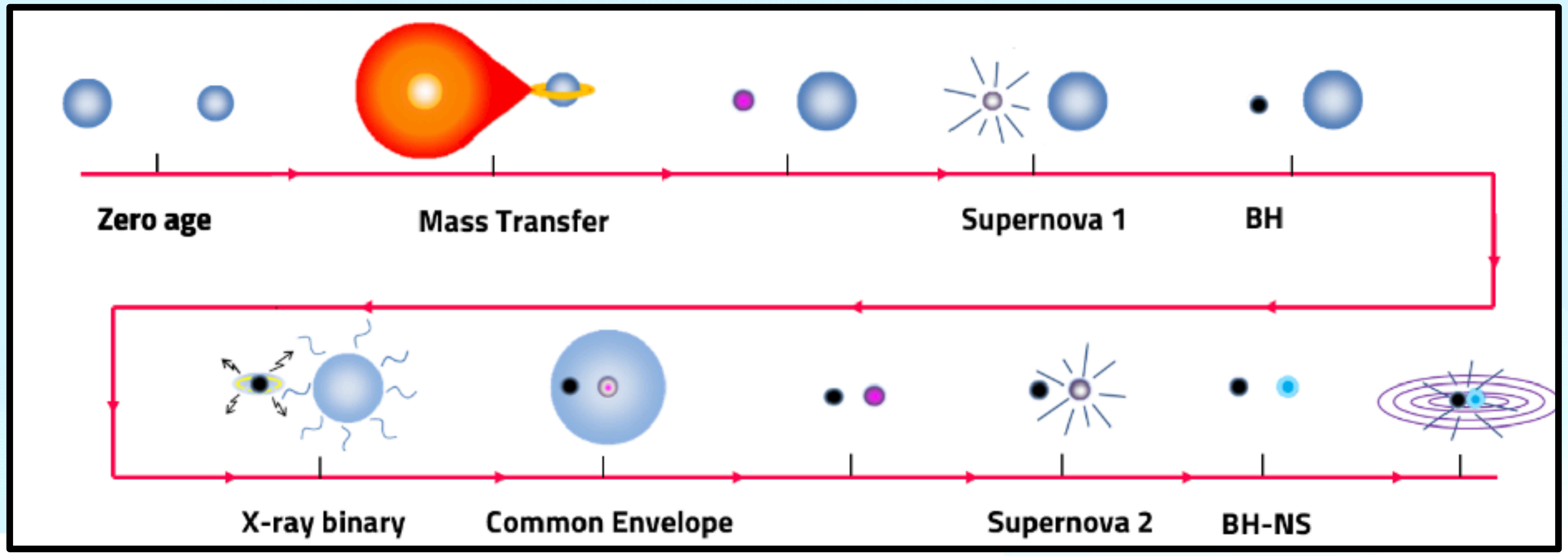




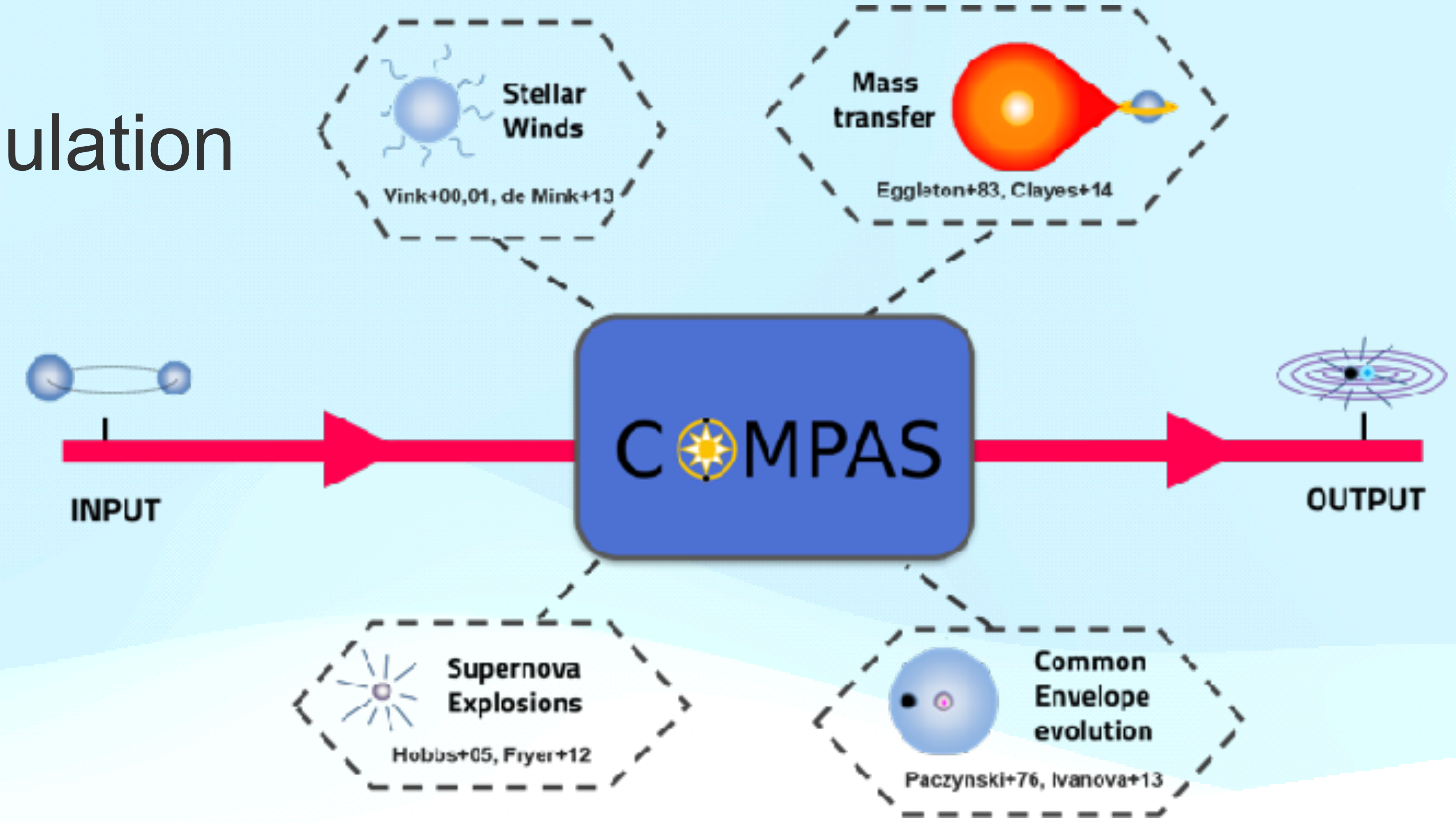
Application: Binary Black Hole Population Synthesis

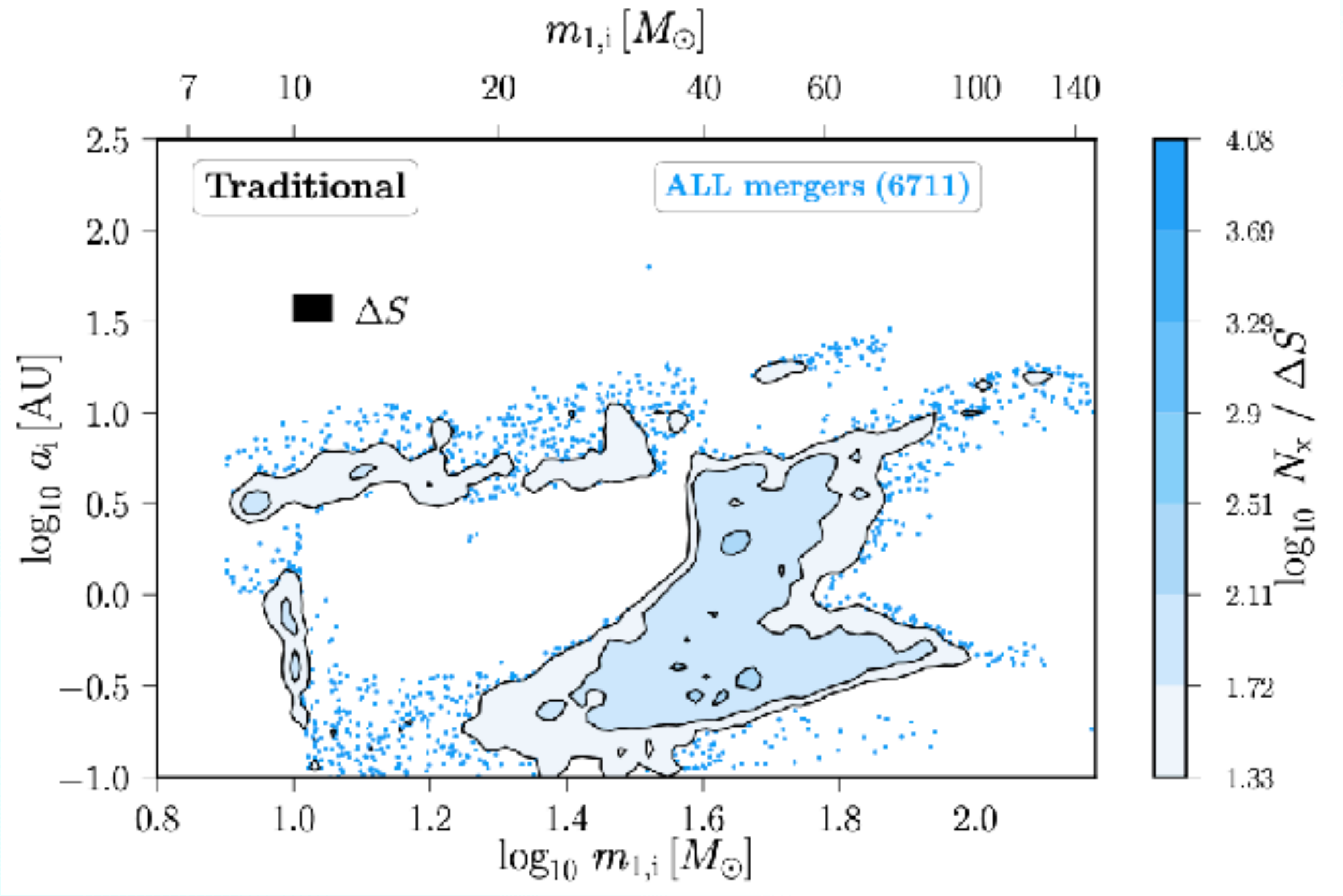
In Collaboration with Prof. Floor Broeckgarden (UCSD)

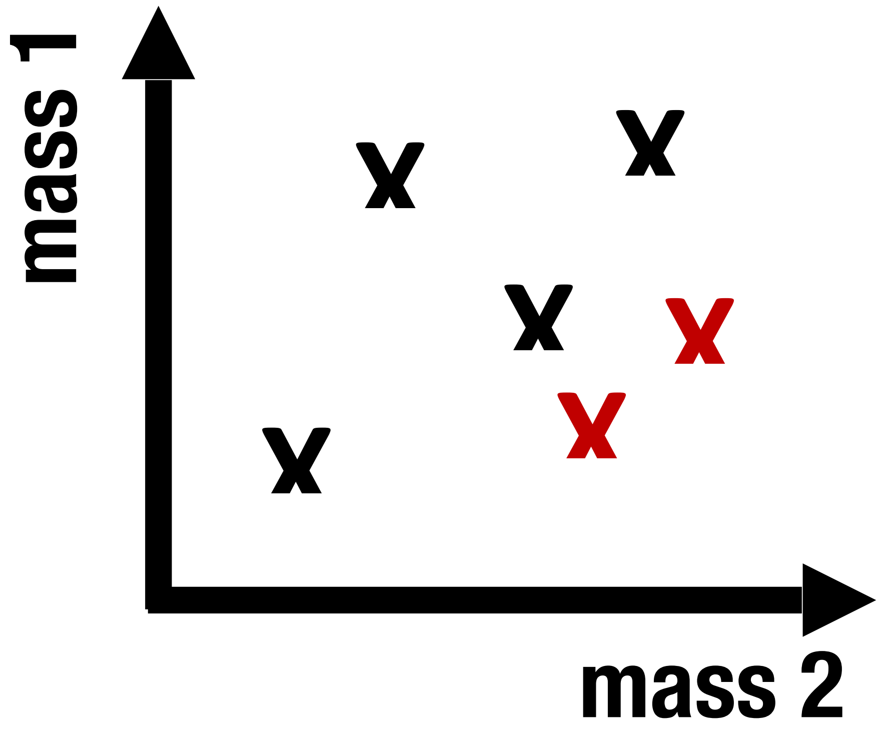
Binary Black Hole Merger



Simulation

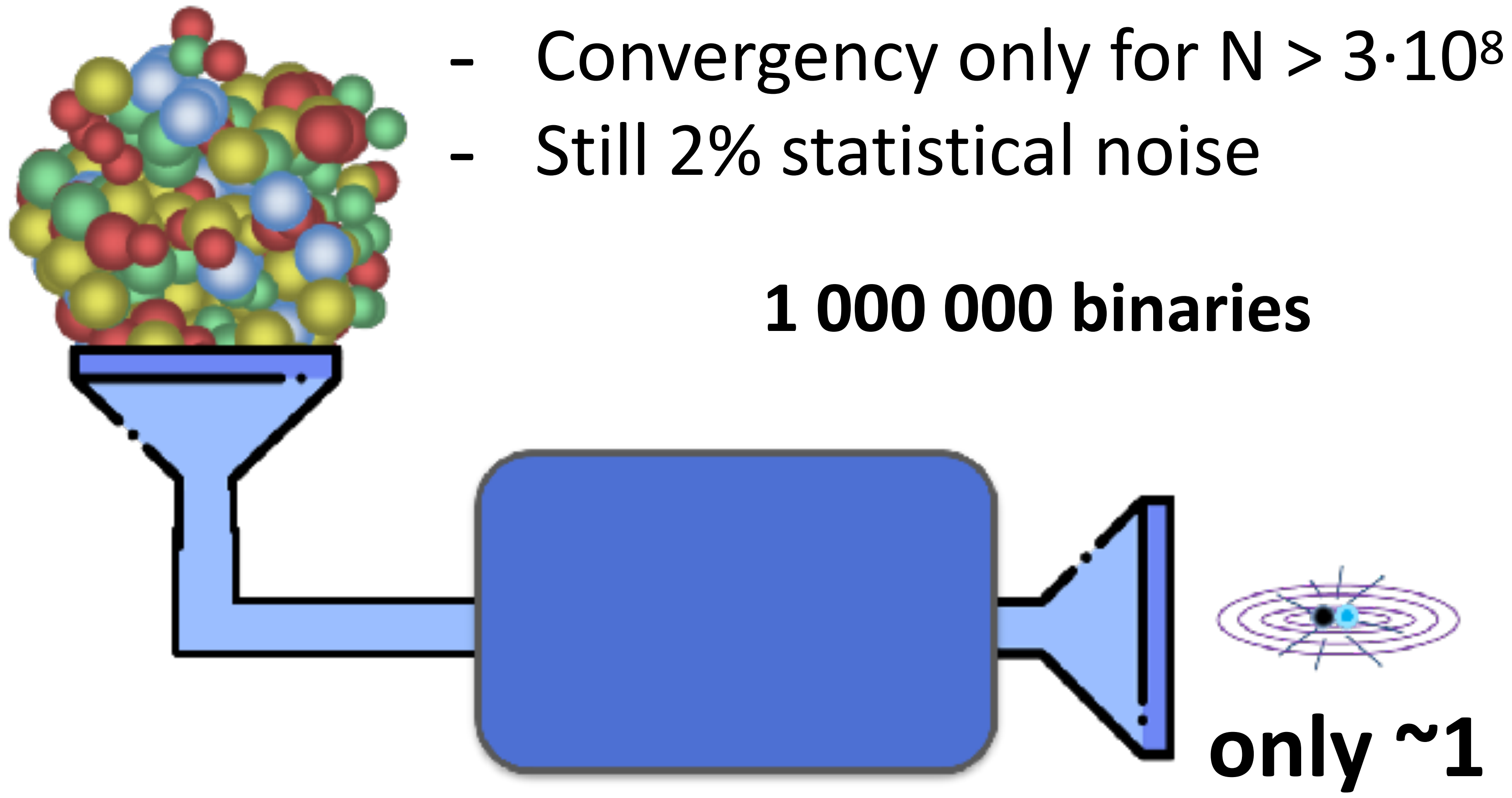






Leads to large Poisson (sampling) noise

- signal to background ratio 1:10^6
- Convergency only for N > $3 \cdot 10^8$
- Still 2% statistical noise



only ~1 merger!

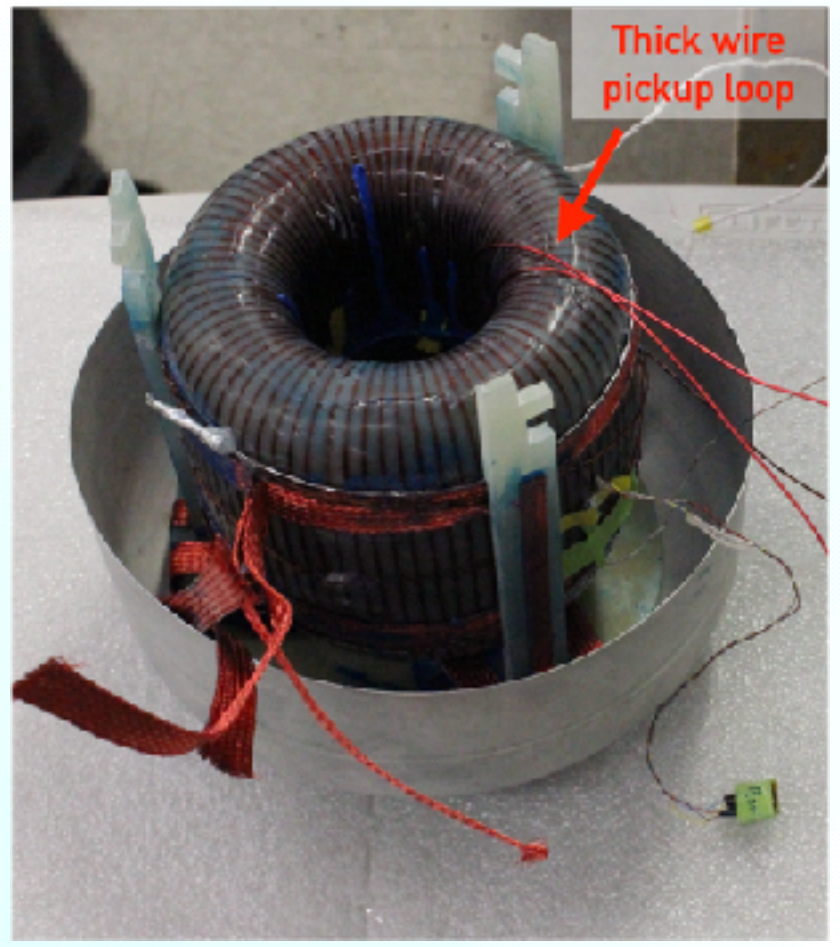
ABRACADABRA

Y. Kahn, B. R. Safdi, and J. Thaler, Phys. Rev. Lett. 117, 141801

J. L. Ouellet et al. Phys. Rev. Lett. 122, 121802 (2019)

C. P. Salemi et al. Phys. Rev. Lett. 127, 081801 (2021)

Broadband Axion Dark Matter Search with Toroidal Magnet



Axion-Modified Maxwell's Equation:

$$\nabla \times B = \frac{\partial E}{\partial t} - g_{a\gamma\gamma}\left(E \times \nabla a - \frac{\partial a}{\partial t} B\right)$$

$$J_{eff} = g_{a\gamma\gamma}\sqrt{2\rho_{DM}} cos(m_a t) B$$

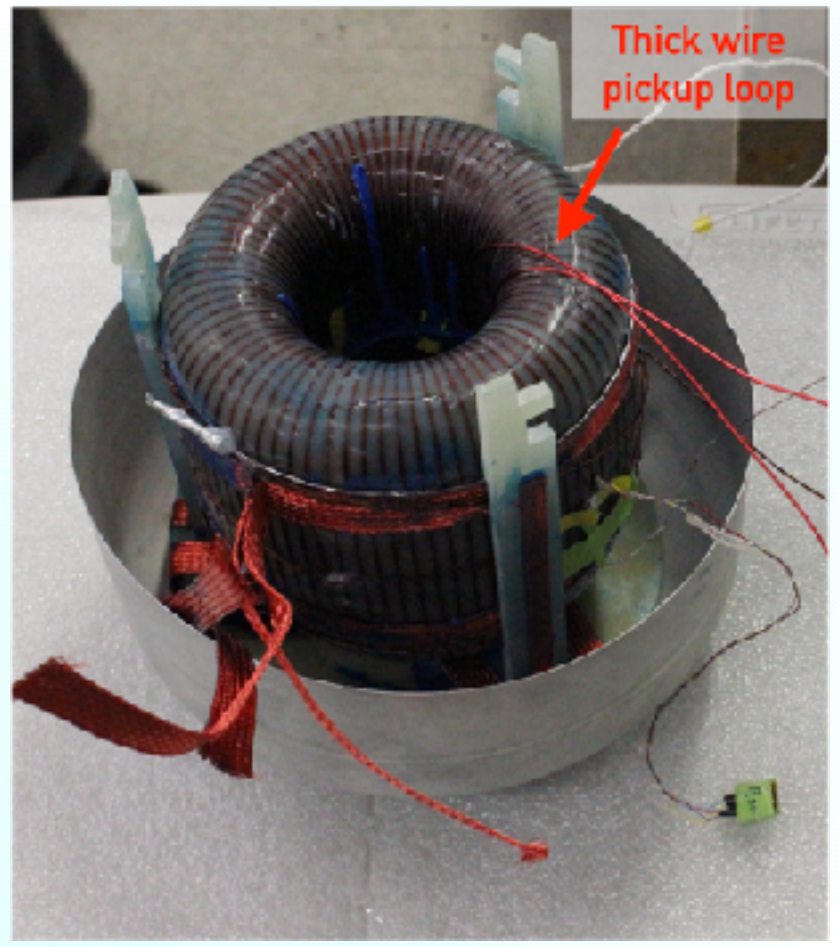
ABRACADABRA

Y. Kahn, B. R. Safdi, and J. Thaler, Phys. Rev. Lett. 117, 141801

J. L. Ouellet et al. Phys. Rev. Lett. 122, 121802 (2019)

C. P. Salemi et al. Phys. Rev. Lett. 127, 081801 (2021)

Broadband Axion Dark Matter Search with Toroidal Magnet



Axion-Modified Maxwell's Equation:

$$\nabla \times B = \frac{\partial E}{\partial t} - g_{a\gamma\gamma}\left(E \times \nabla a - \frac{\partial a}{\partial t} B\right)$$

$$J_{eff} = g_{a\gamma\gamma}\sqrt{2\rho_{DM}}\cos(m_a t)B$$

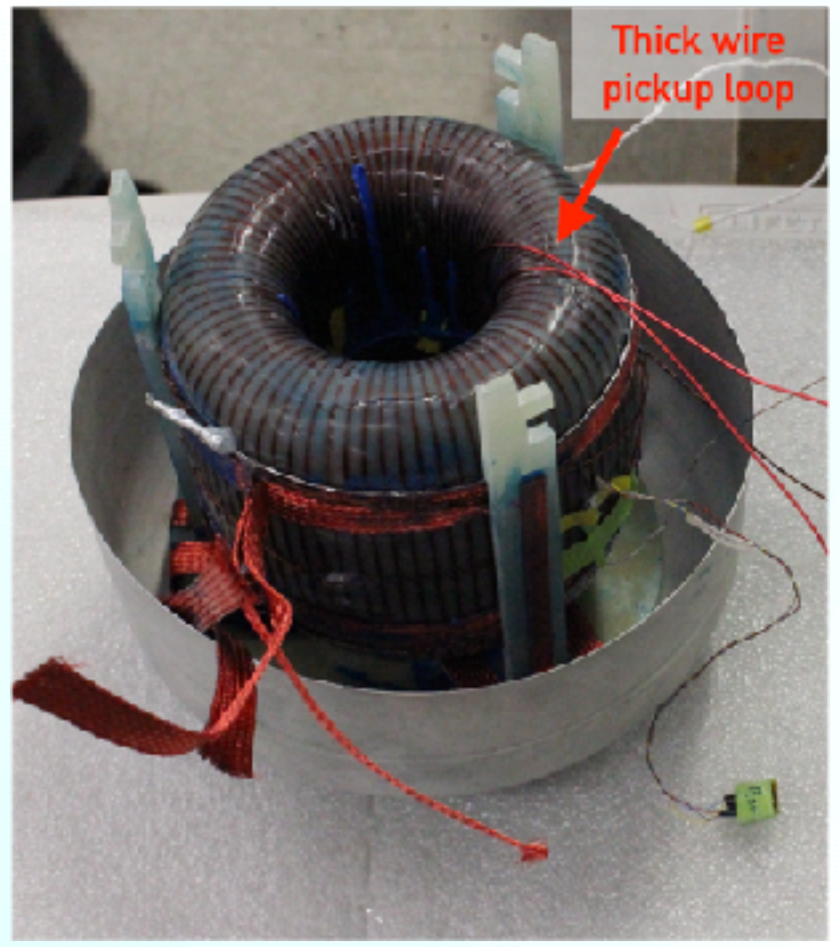
ABRACADABRA

Y. Kahn, B. R. Safdi, and J. Thaler, Phys. Rev. Lett. 117, 141801

J. L. Ouellet et al. Phys. Rev. Lett. 122, 121802 (2019)

C. P. Salemi et al. Phys. Rev. Lett. 127, 081801 (2021)

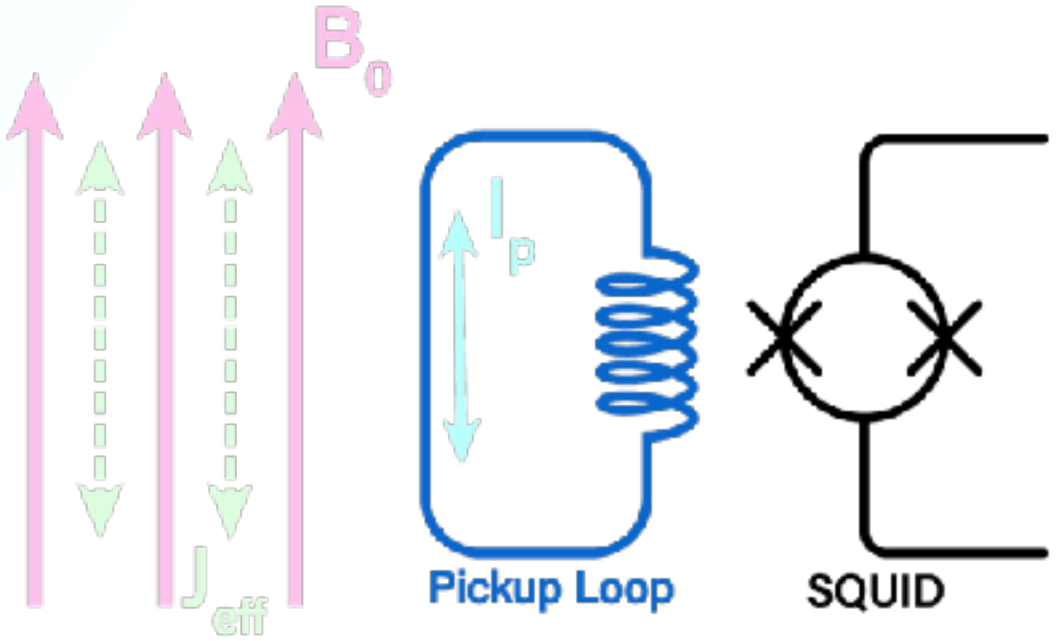
Broadband Axion Dark Matter Search with Toroidal Magnet



Axion-Modified Maxwell's Equation:

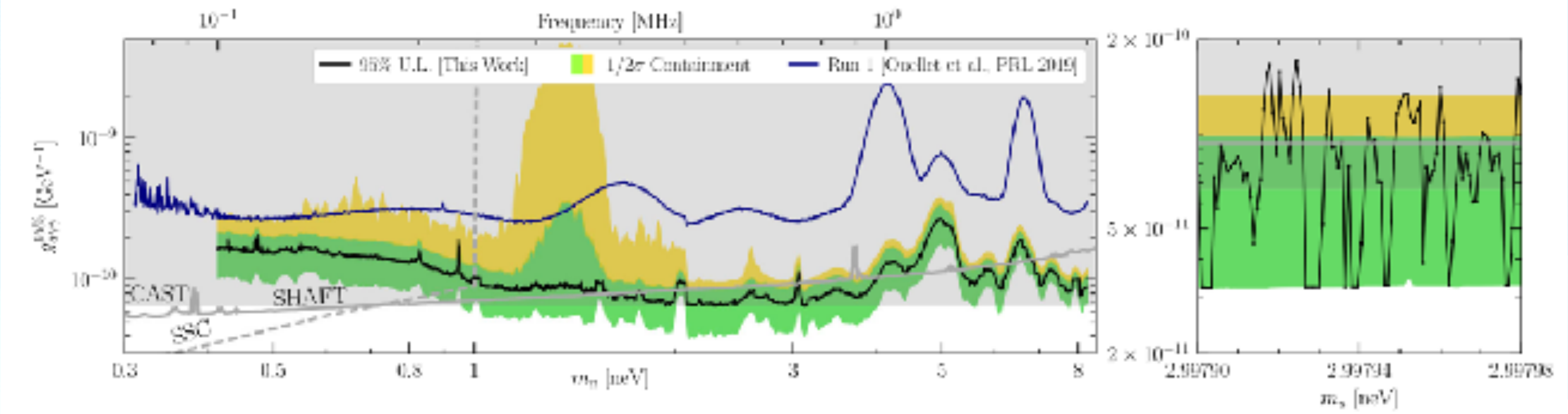
$$\nabla \times B = \frac{\partial E}{\partial t} - g_{a\gamma\gamma}\left(E \times \nabla a - \frac{\partial a}{\partial t} B\right)$$

$$J_{eff} = g_{a\gamma\gamma}\sqrt{2\rho_{DM}}\,cos(m_a t)B$$



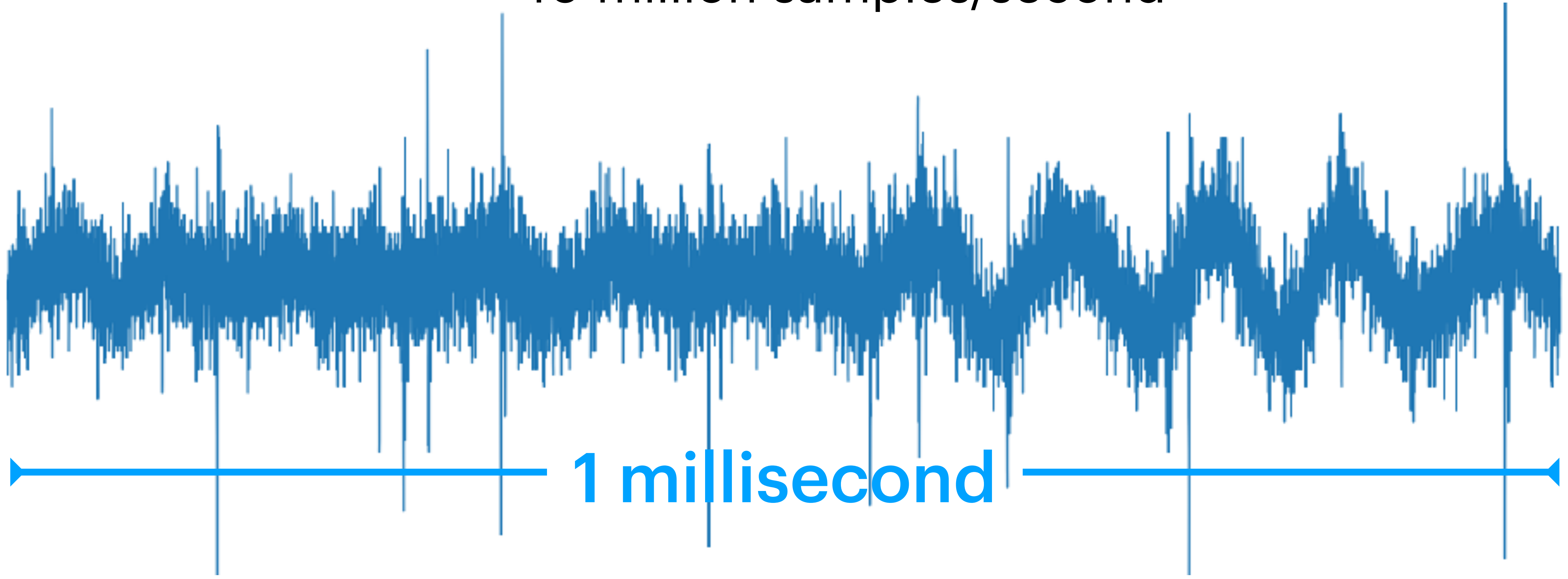
Frequency Spectrum

Broadband search for axion DM



Ultra-long Time Series

10 million samples/second



Experimental Apparatus Constructed by Winslow Lab at MIT

ABRACADABRA

J. T. Fry et al, arXiv:2406.04378
Submitting to Nature Scientific Data

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

J. T. Fry[1] * jtfry@mit.edu
Aobo Li[2] liaobo77@ucsd.edu
Lindley Winslow[1] lwinslow@mit.edu
Xinyi Hope Fu[1] hopefu@mit.edu
Zhenghao Fu[1] fuzh@mit.edu
Kaliroe M. W. Pappas[1] kaliroe@mit.edu

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Halıcıoğlu Data Science Institute, Department of Physics, UC San Diego, La Jolla, CA 92093, USA

Open Data

Release dark matter detector data for AI/ML algorithms

Easy Benchmarking

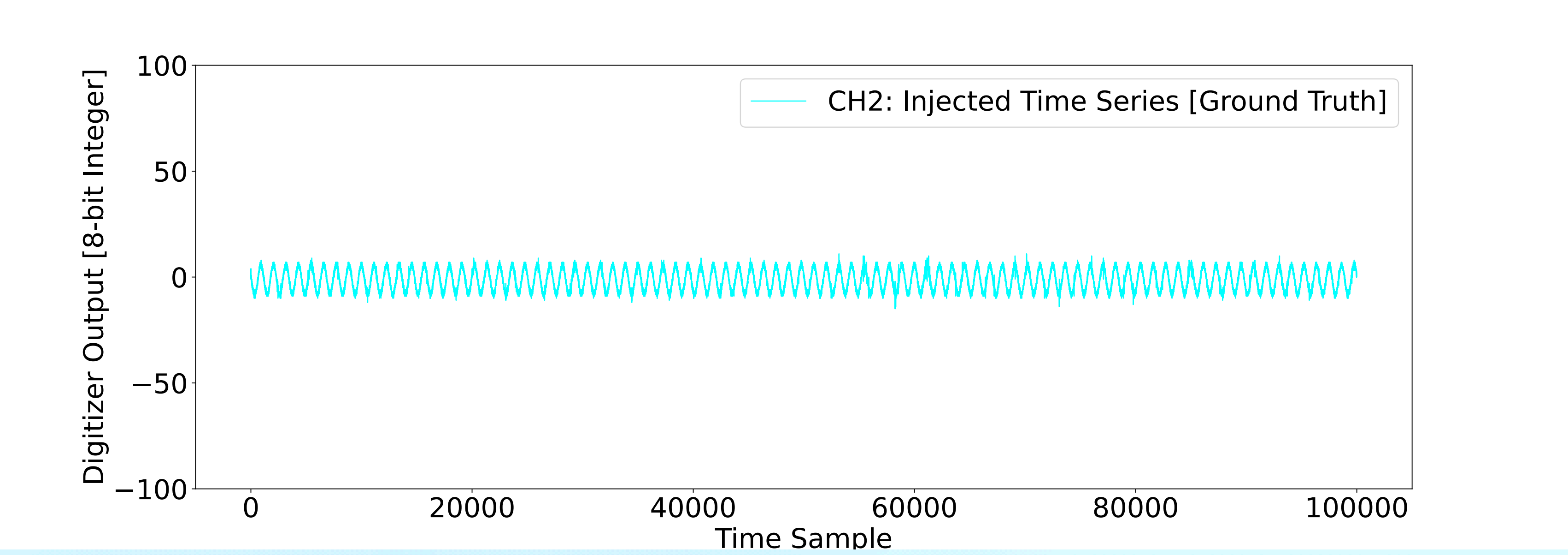
Design a quantitative benchmarking score to quantify the performance of different algorithms

AI for Science

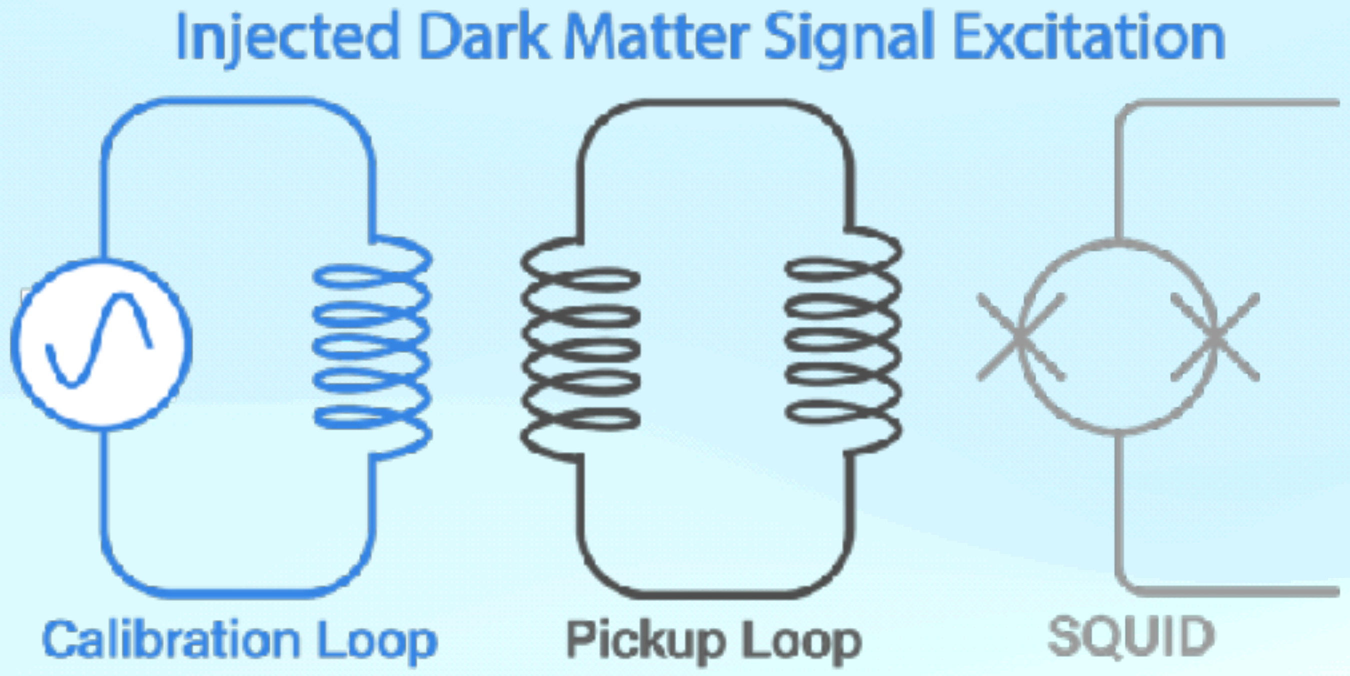
Enables core AI algorithms to extract the signal and produce real physics results thereby advancing fundamental science

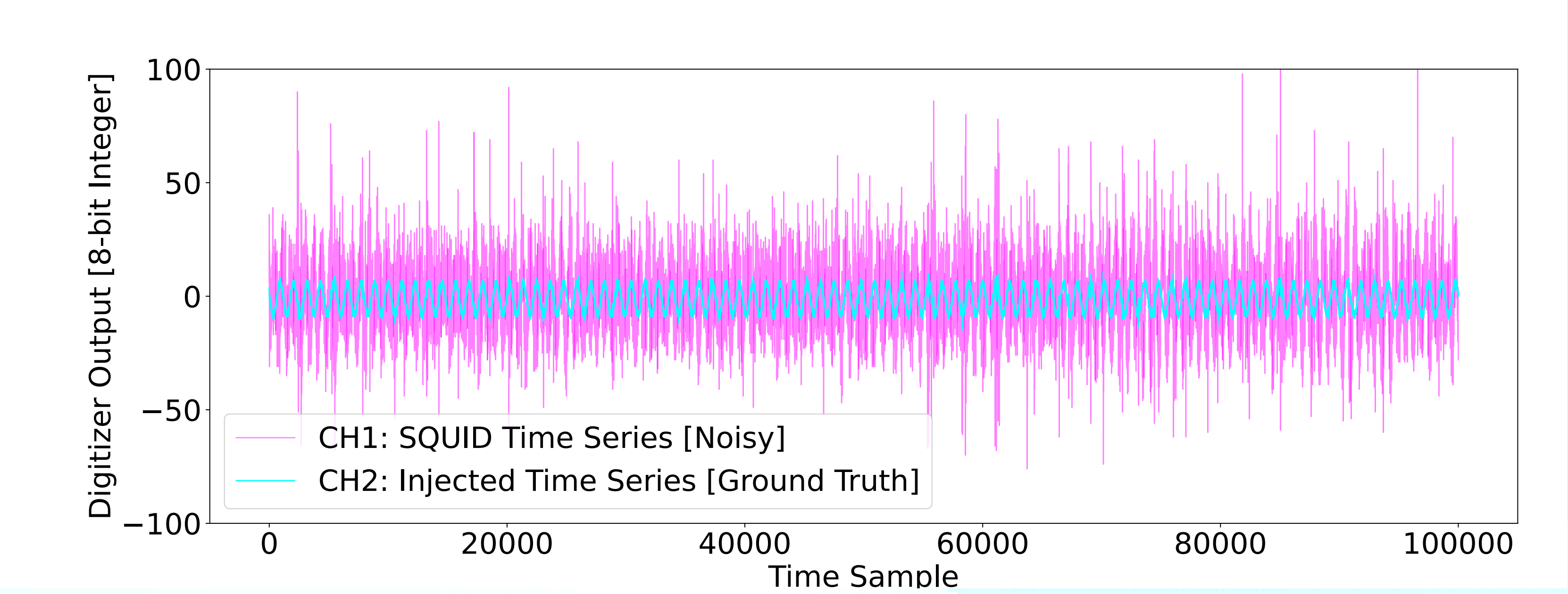
ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising



CH2: Injected Time Series [Ground Truth]

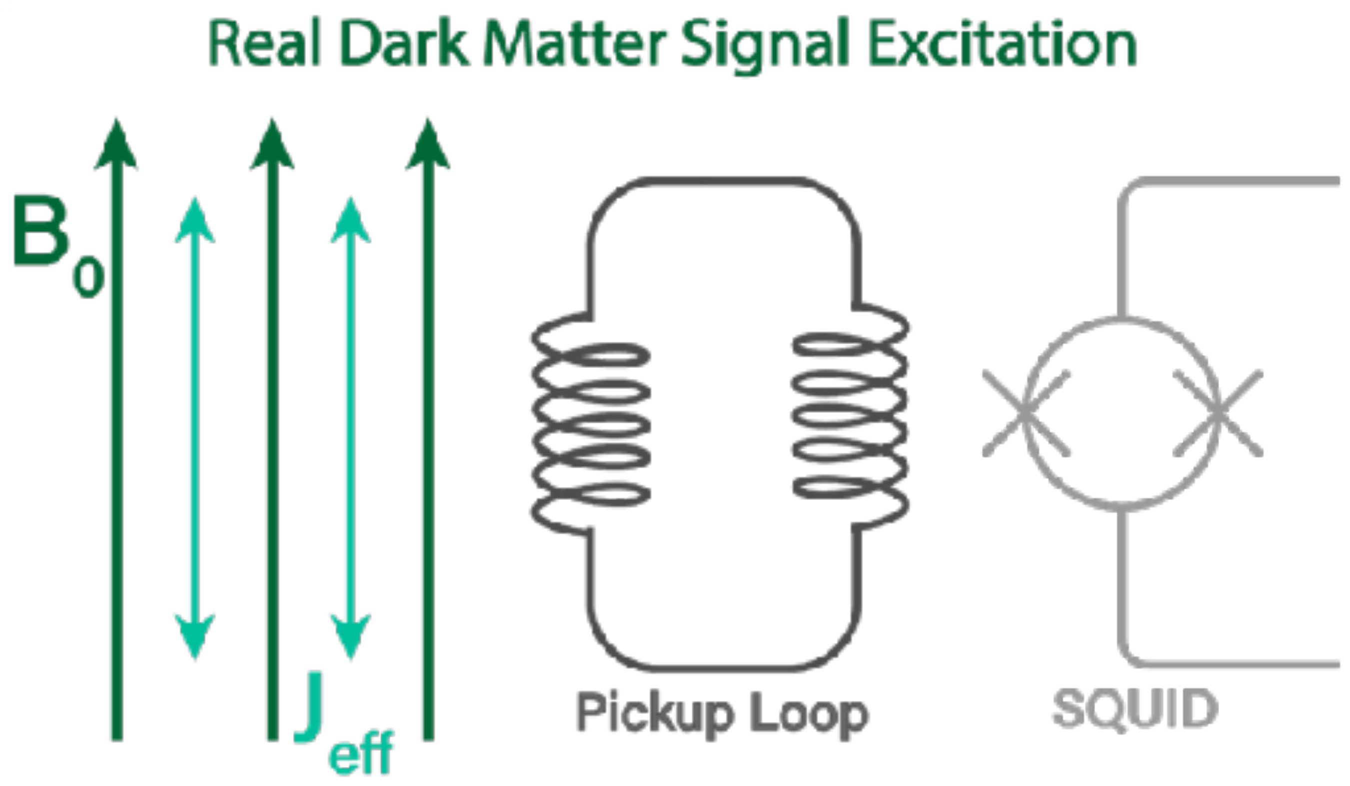


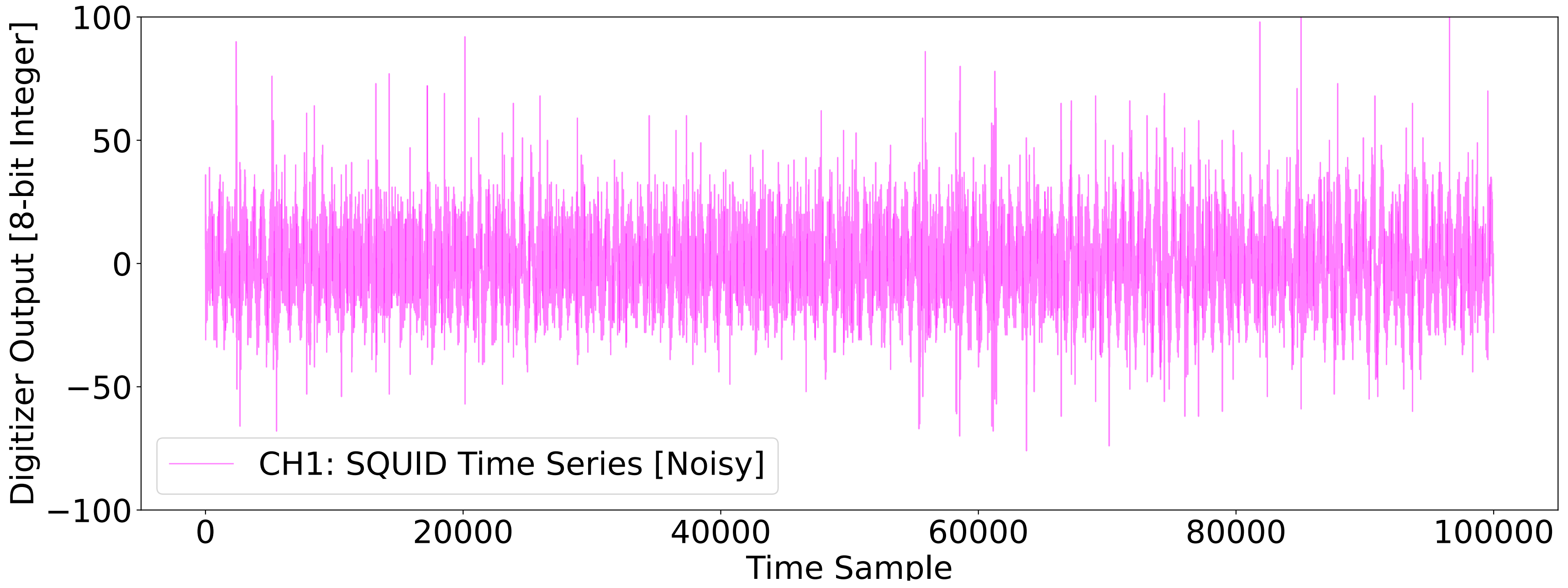


CH1: SQUID Time Series [Noisy]

Train AI denoising model to recover...

No Signal Injected





CH1: SQUID Time Series [Noisy]

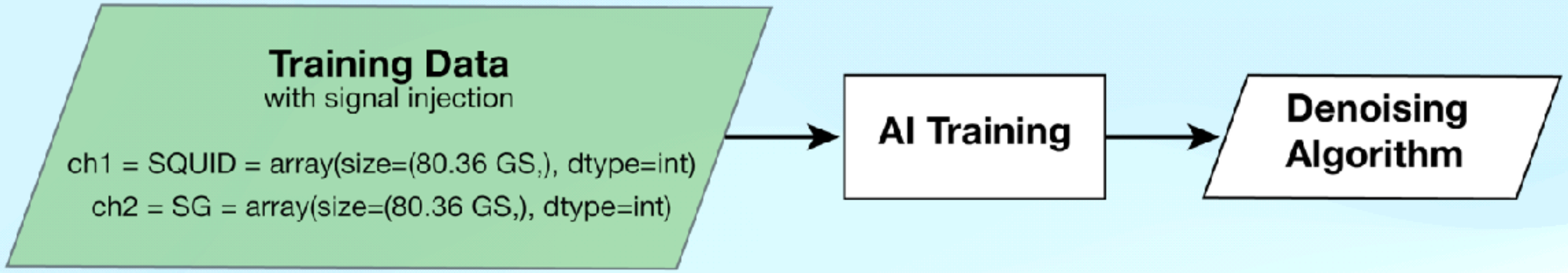
Use trained AI model to denoise...

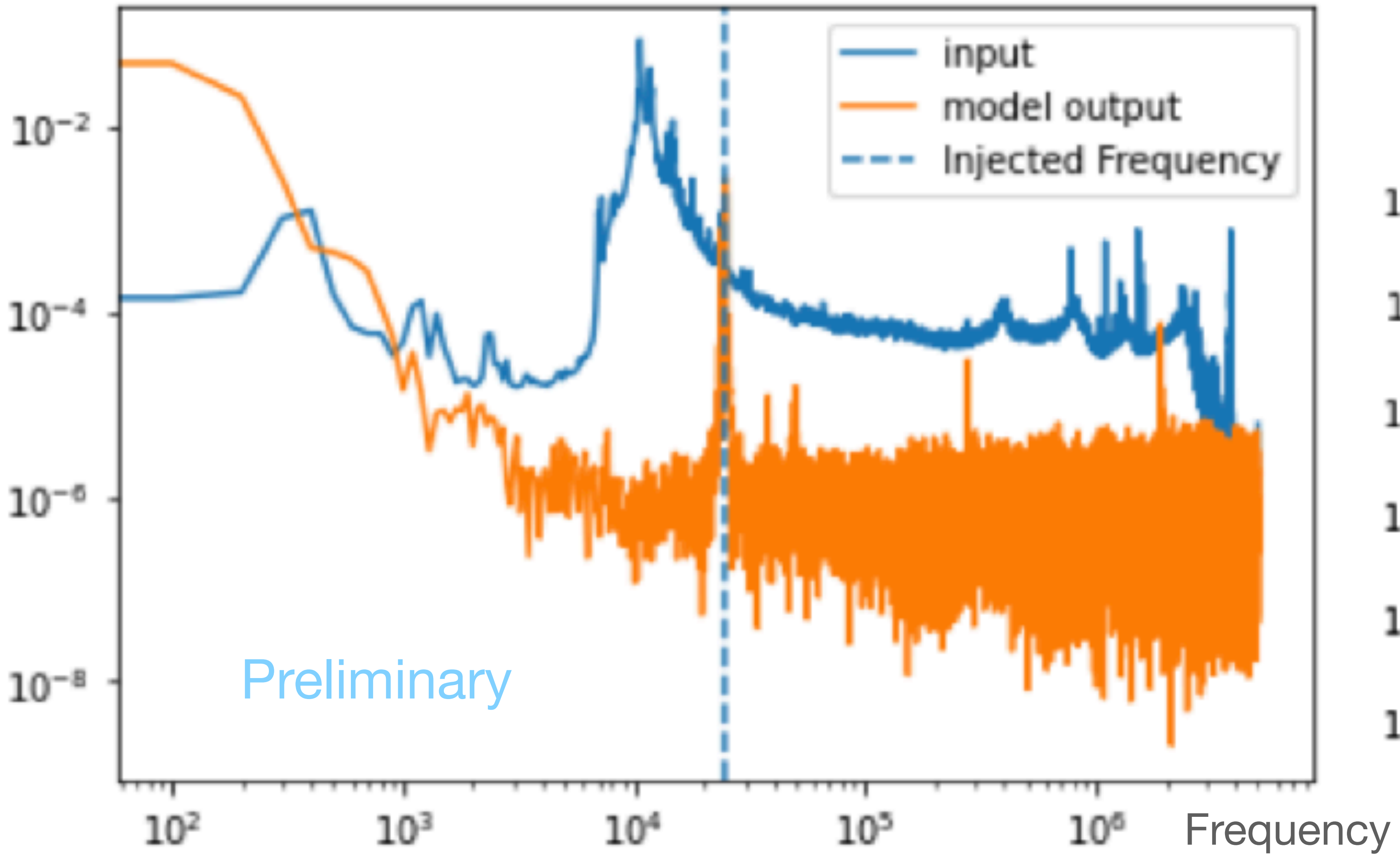
ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

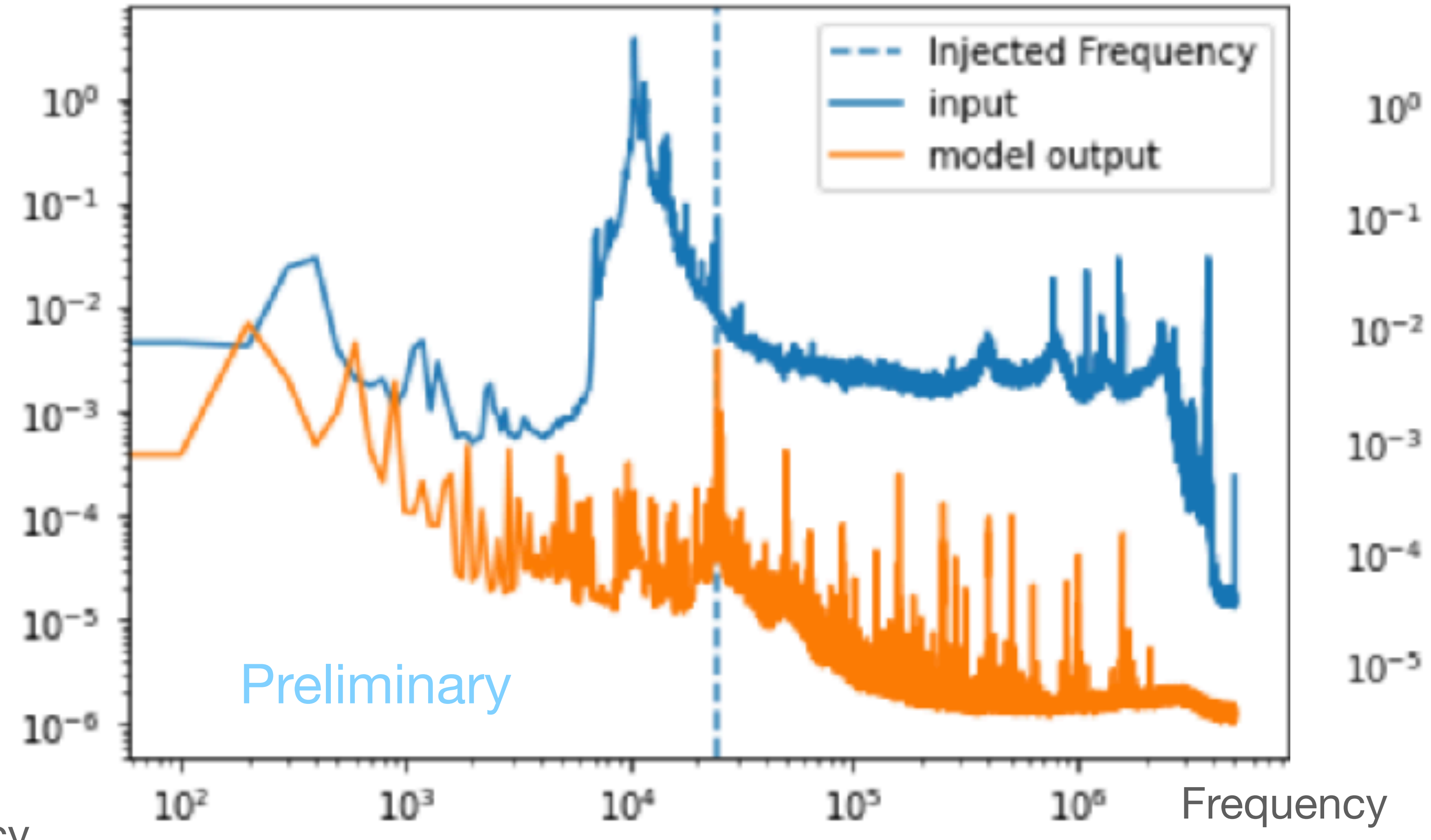
ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

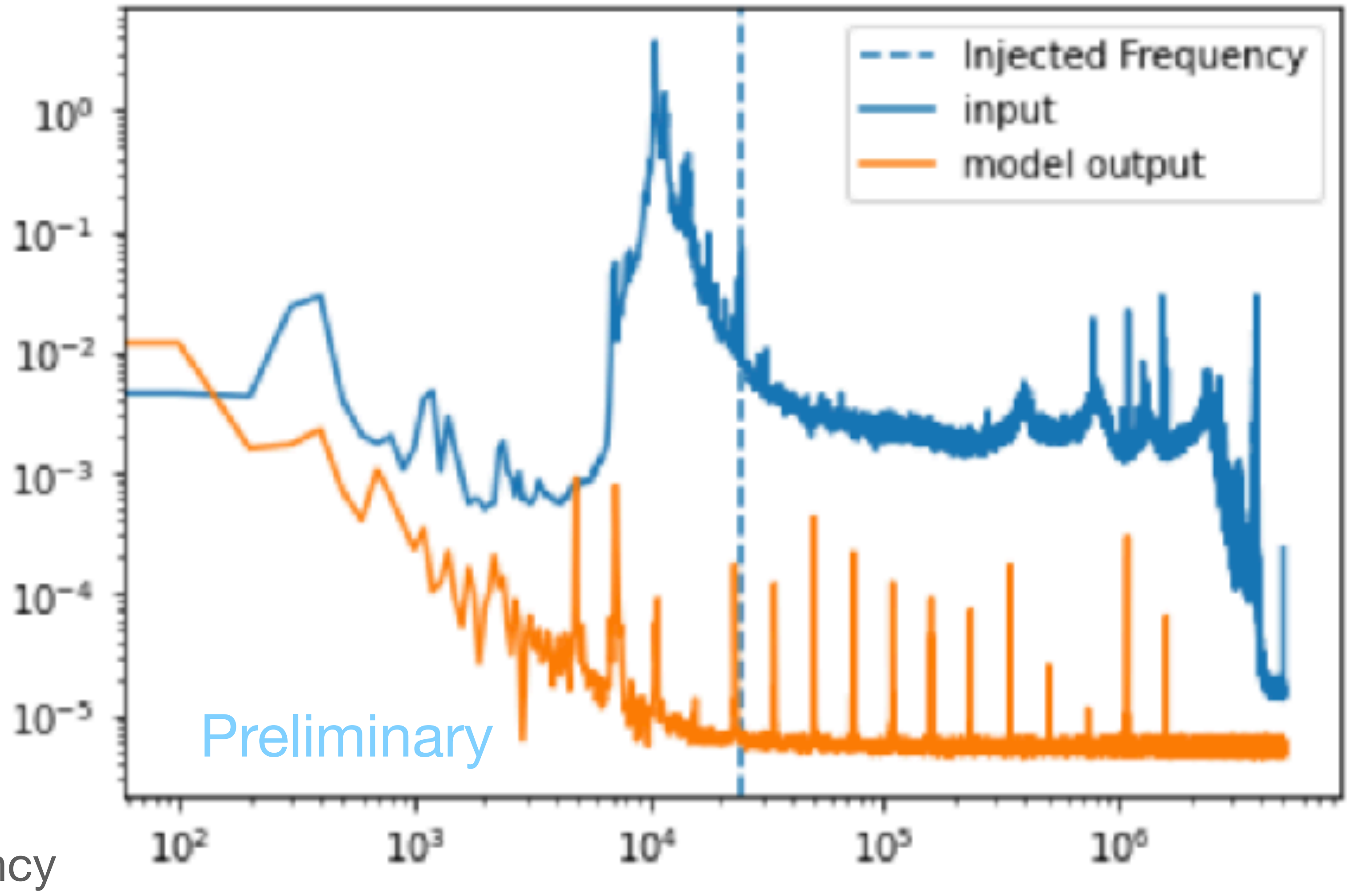




Fully Connected NN



Positional U-Net



Transformer

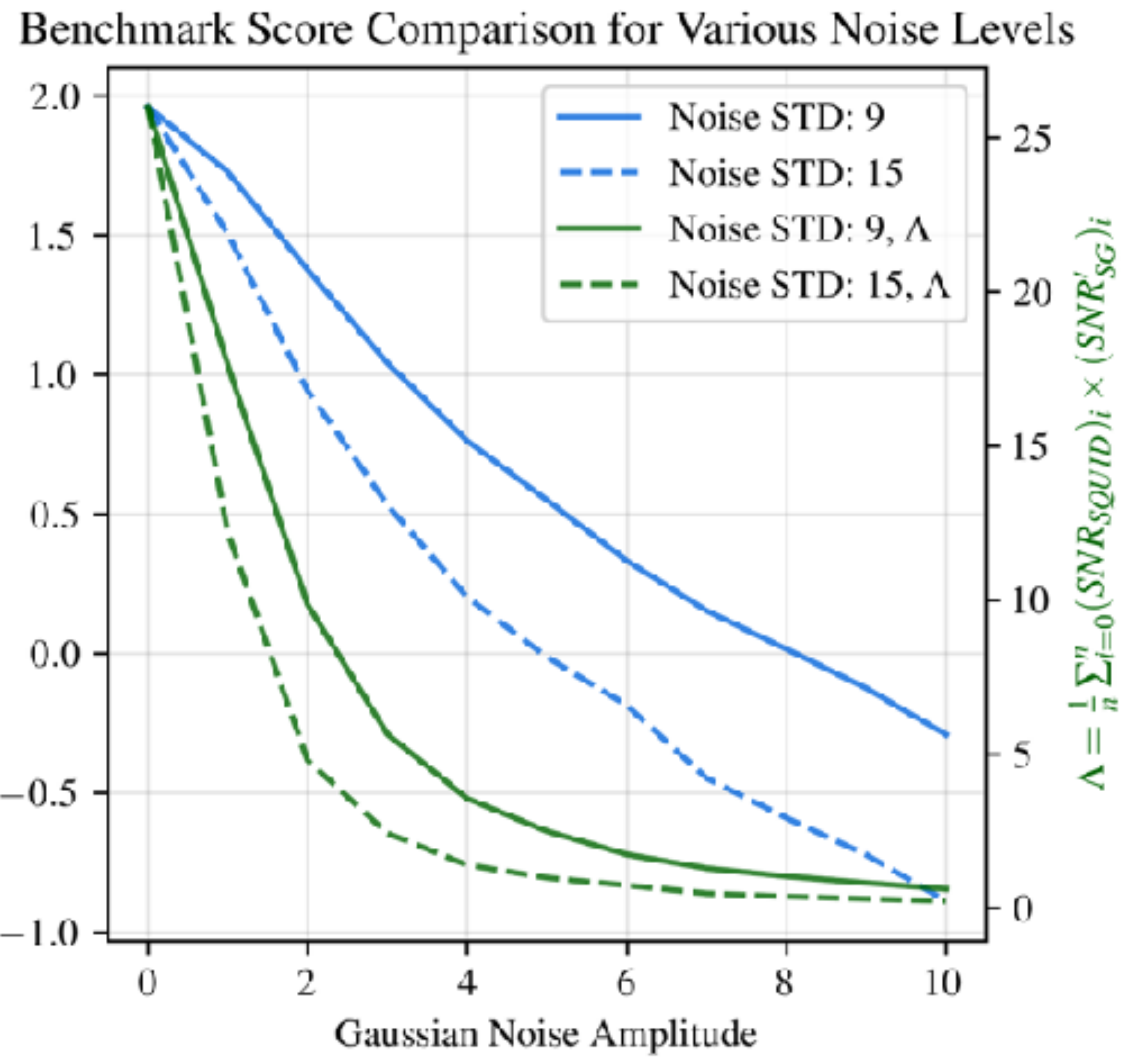
ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

$$\Lambda = \left(\frac{1}{n} \sum_{i=0}^{n} (SNR_{\mathrm{SQUID}})_i \times (SNR'_{\mathrm{Injected}})_i \right)$$

ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

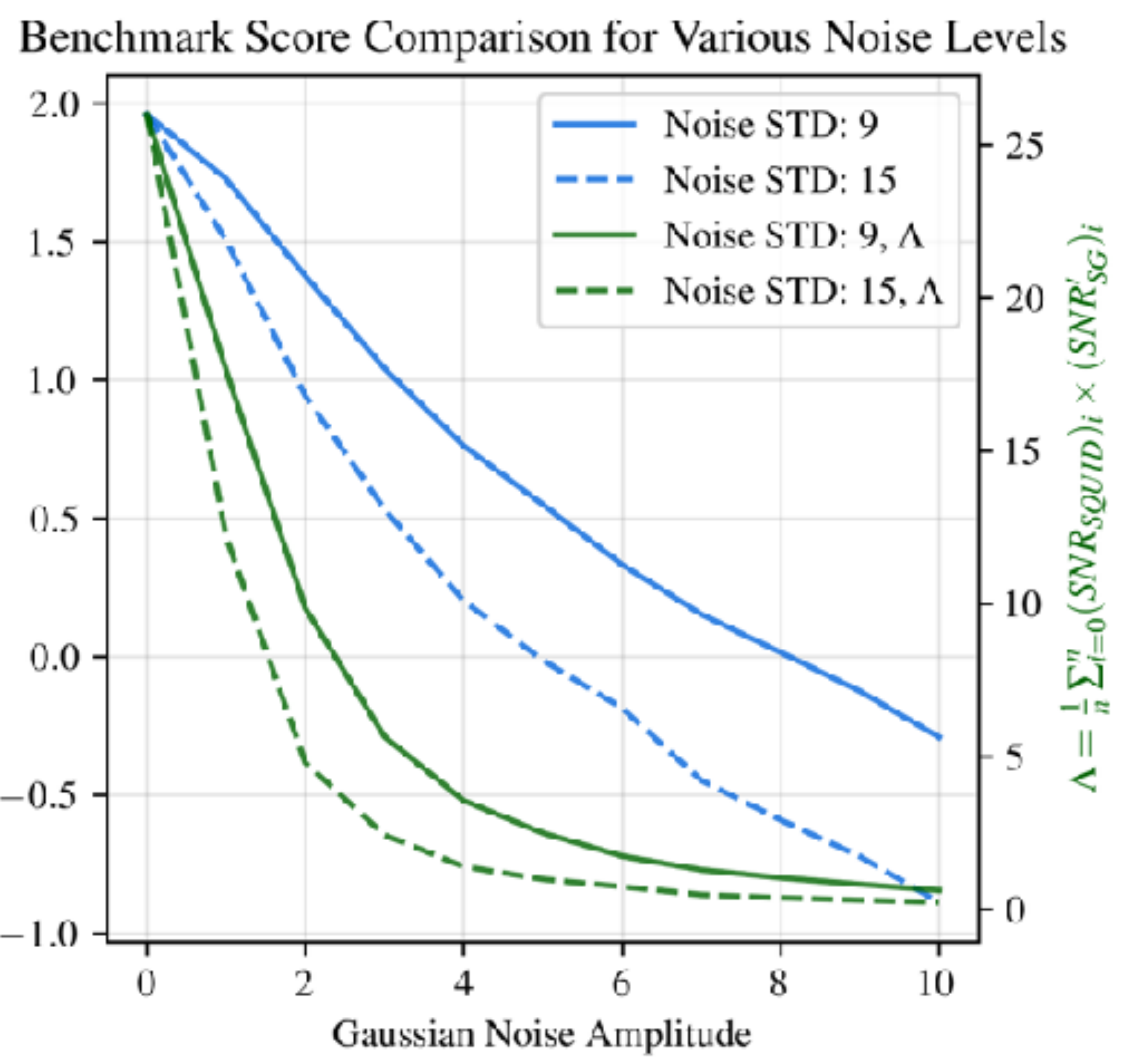


$$\Lambda = \left(\frac{1}{n}\sum_{i=0}^{n}(SNR_{\text{SQUID}})_i \times (SNR'_{\text{Injected}})_i\right)$$

$$\text{Denoising Score} = log_{5.27}\Lambda$$

ABRACADABRA

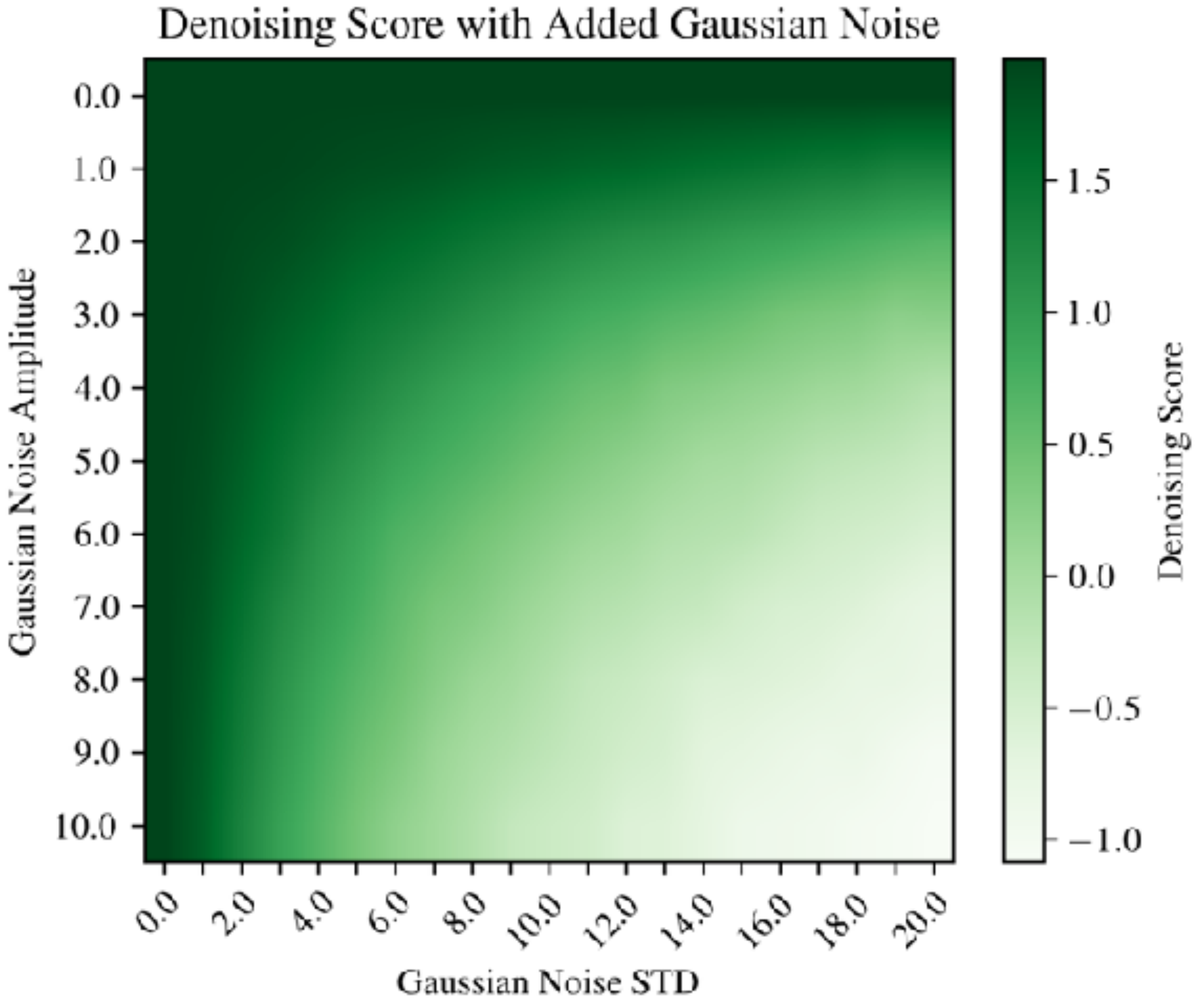
TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising



$$\Lambda = \left(\frac{1}{n} \sum_{i=0}^{n} (SNR_{\mathrm{SQUID}})_i \times (SNR'_{\mathrm{Injected}})_i \right)$$

$$\text{Denoising Score} = log_{5.27}\Lambda$$

Test the denoising score by doping gaussian noise into Time Series



ABRACADABRA

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

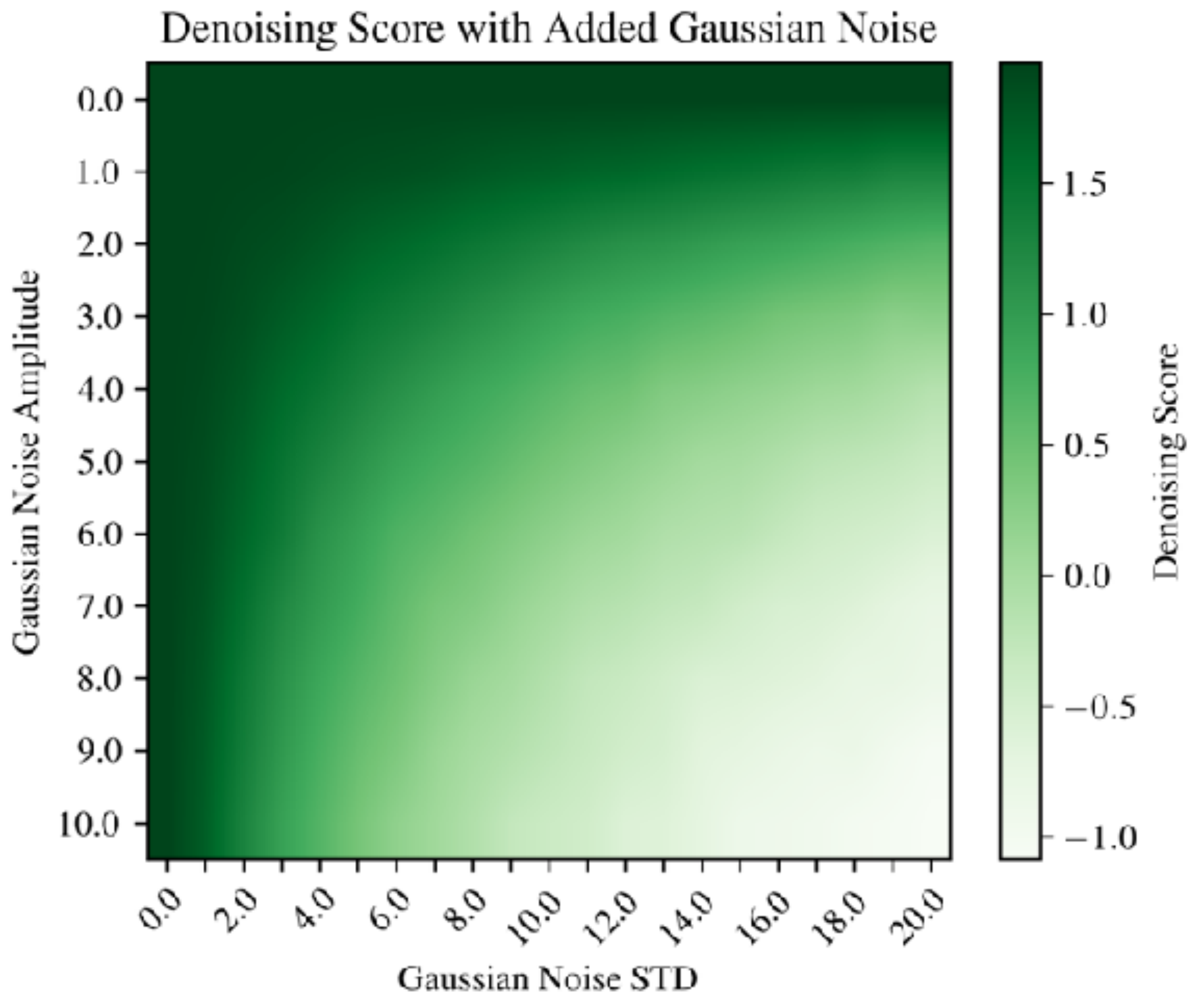
Table 1: Fine and coarse denoising score for raw data, traditional algorithms, and trained ML models

Algorithms	Segment Size	Parameters	Fine Score	Coarse Score
None			1.00	1.10
Moving Average	1×10^6	window = 100	0.52	0.64
SG Filter	1×10^6	window = 100, order = 11	-2.77	-2.35
FC Net	4×10^4	See Appendix A	6.43	6.55
PU Net	4×10^4	See Appendix A	3.69	3.84
Transformer	2×10^4	See Appendix A	3.95	4.18

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

Table 1: Fine and coarse denoising score for raw data, traditional algorithms, and trained ML models

Algorithms	Segment Size	Parameters	Fine Score	Coarse Score
None			1.00	1.10
Moving Average	1×10^6	window = 100	0.52	0.64
SG Filter	1×10^6	window = 100, order = 11	-2.77	-2.35
FC Net	4×10^4	See Appendix A	6.43	6.55
PU Net	4×10^4	See Appendix A	3.69	3.84
Transformer	2×10^4	See Appendix A	3.95	4.18



ABRACADABRA

J. T. Fry et al, arXiv:2406.04378
Submitting to Nature Scientific Data

TIDMAD: Time Series Dataset for Discovering Dark Matter with AI Denoising

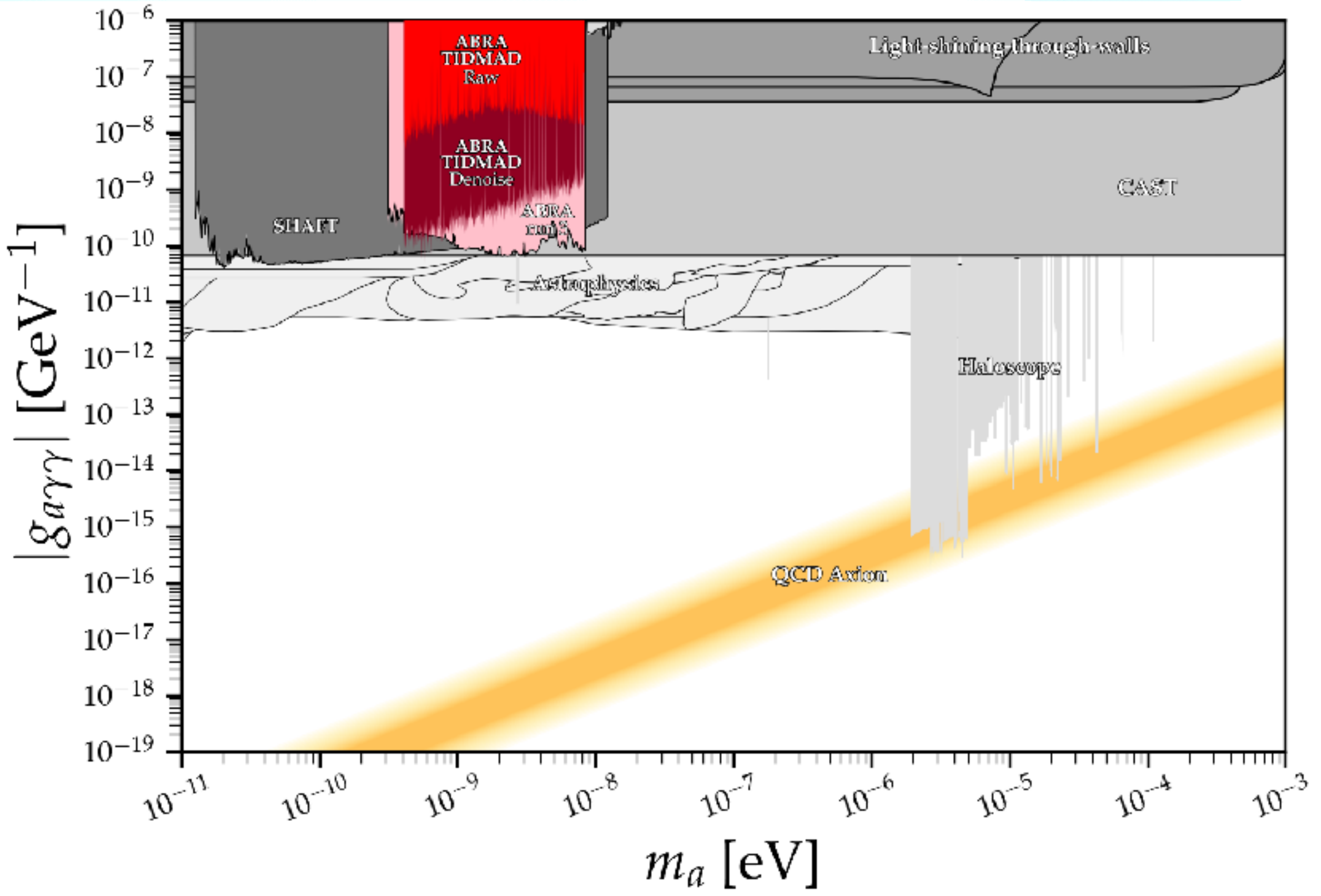
Axion Limit Boost

ABRA TIDMAD Raw: 24 hr data, no denoising

ABRA TIDMAD Denoised: 24 hr data with FCNet denoising

ABRA Run 3: 2,400 hr data, no denoising

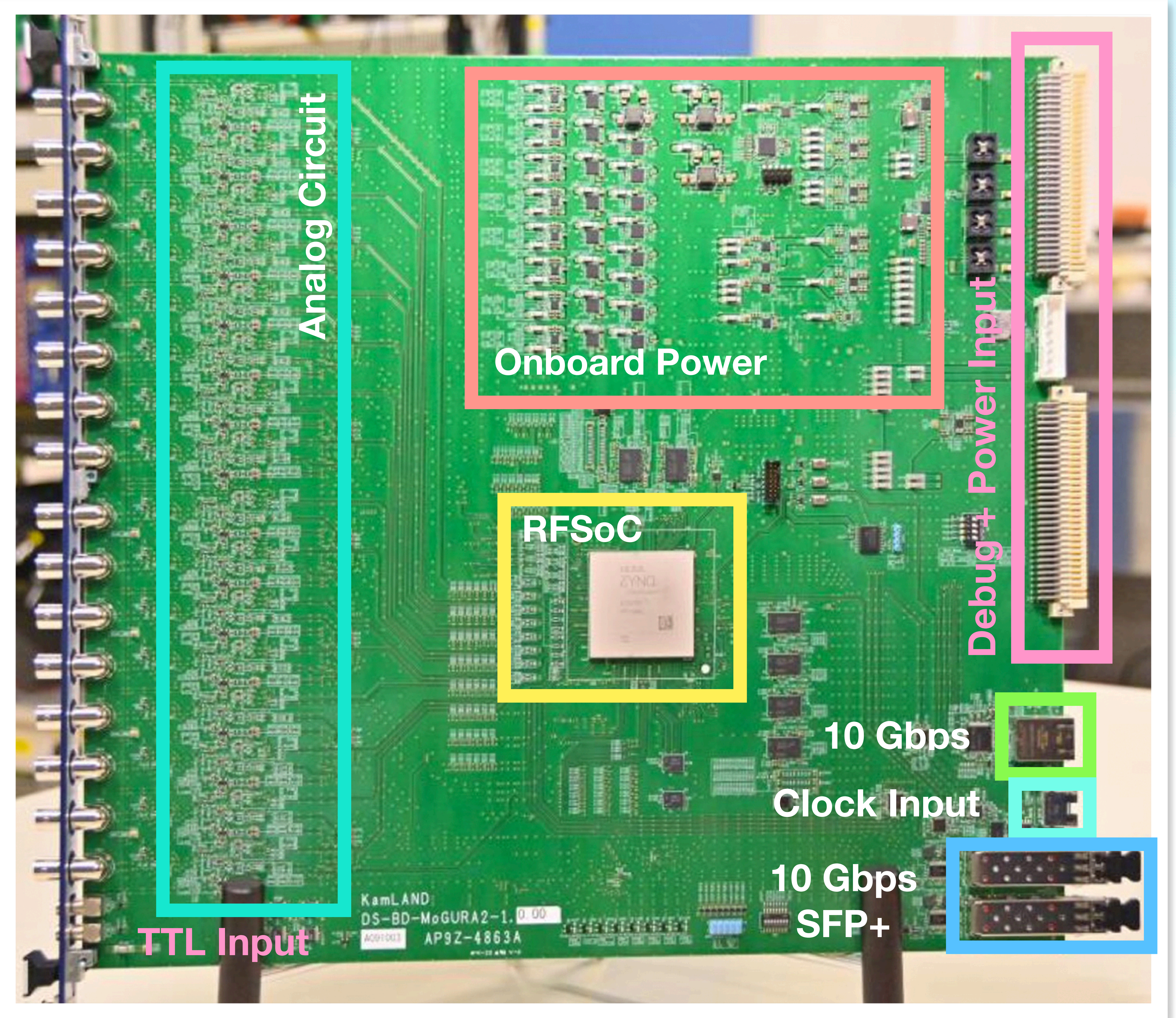
Efficient denoising algorithms increased Axion search limit by **1-2 orders of magnitude**, approaching the previous world-leading ABRA run 3 results with only **1%** of statistics




PAST
PRESENT
FUTURE

New Electronics for KamLAND-Zen

16-channel prototype for KamLAND2-Zen



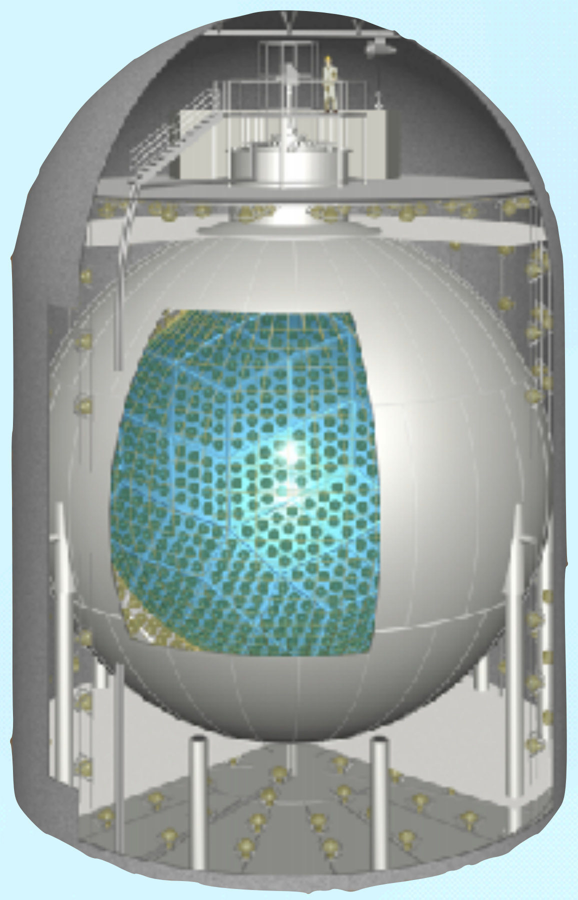
Primary Goals:

1. Digitize waveform during the chaotic period after a muon passes through the detector in order to record all neutrons, allowing us to reduce the Long-Lived spallation background.
2. Streaming data (deadtime free system), large data throughput.
3. Large memory buffers.

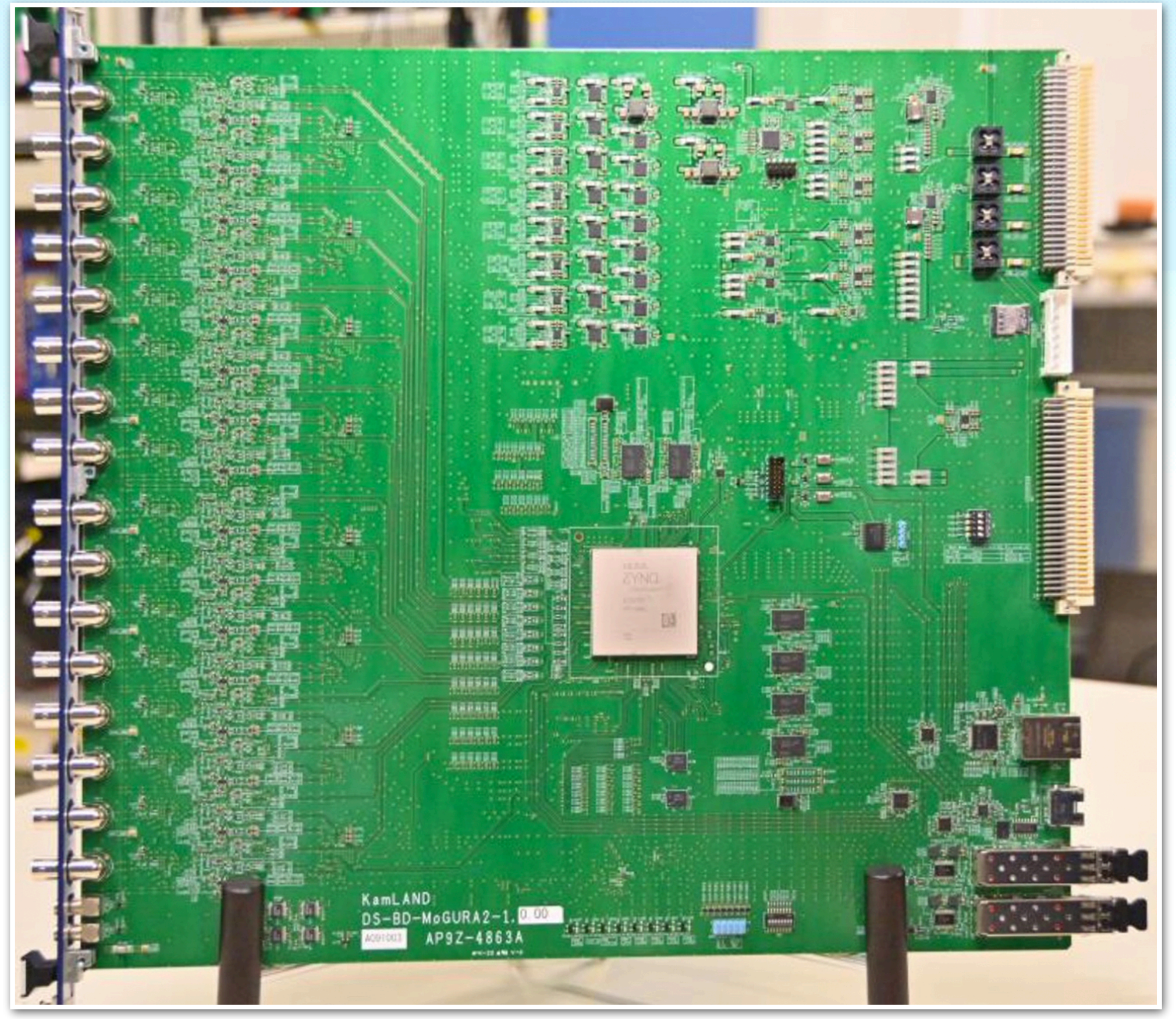
Reduction in PCB footprint	**Machine learning on FPGA**	***50% cost savings**	***30-40% power consumption savings**

* compared to standard RF signal chain

Hardware-AI Codesign



Data Stream



Offline Analysis

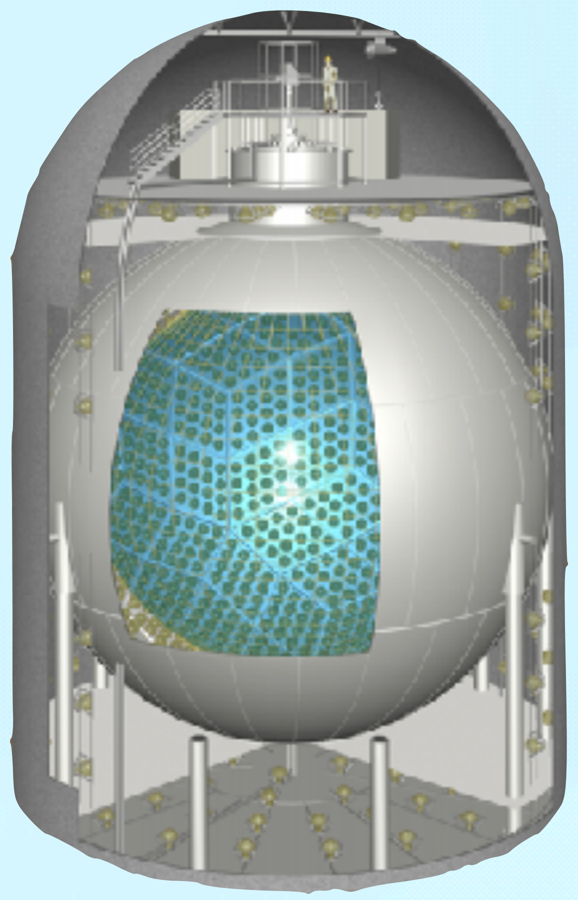
Energy

Position

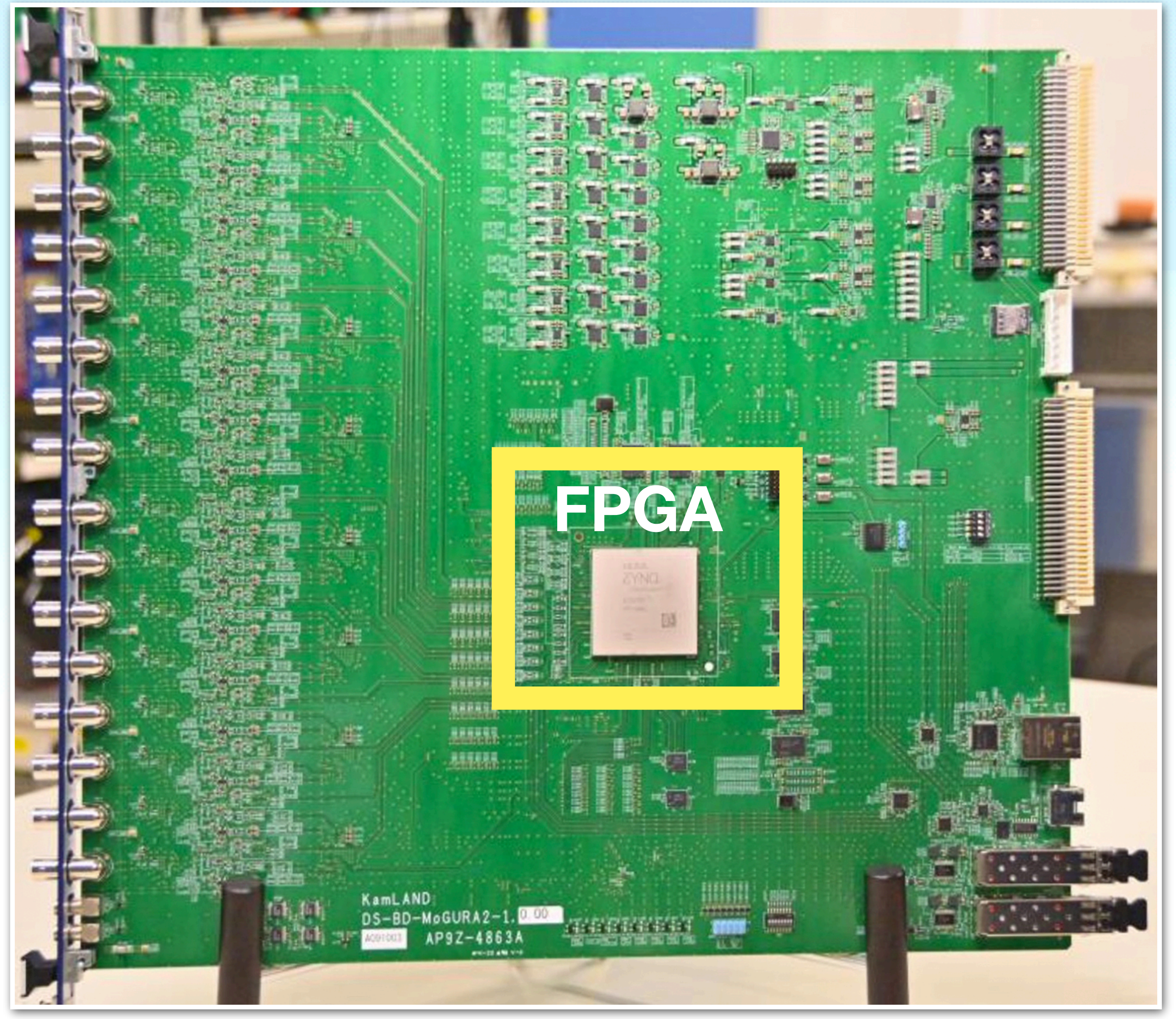
Particle Type

Detector Response

Hardware-AI Codesign



Data Stream



Offline Analysis

Energy

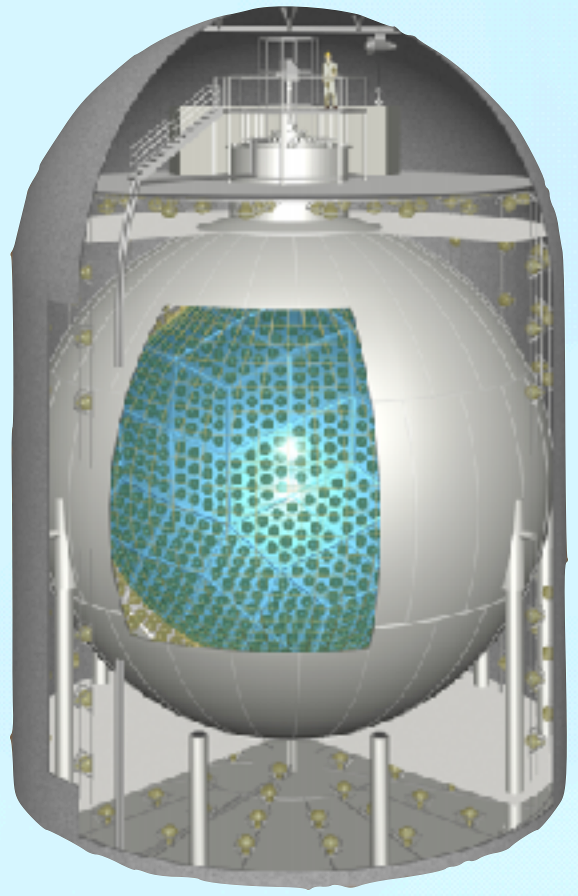
Position

Particle Type

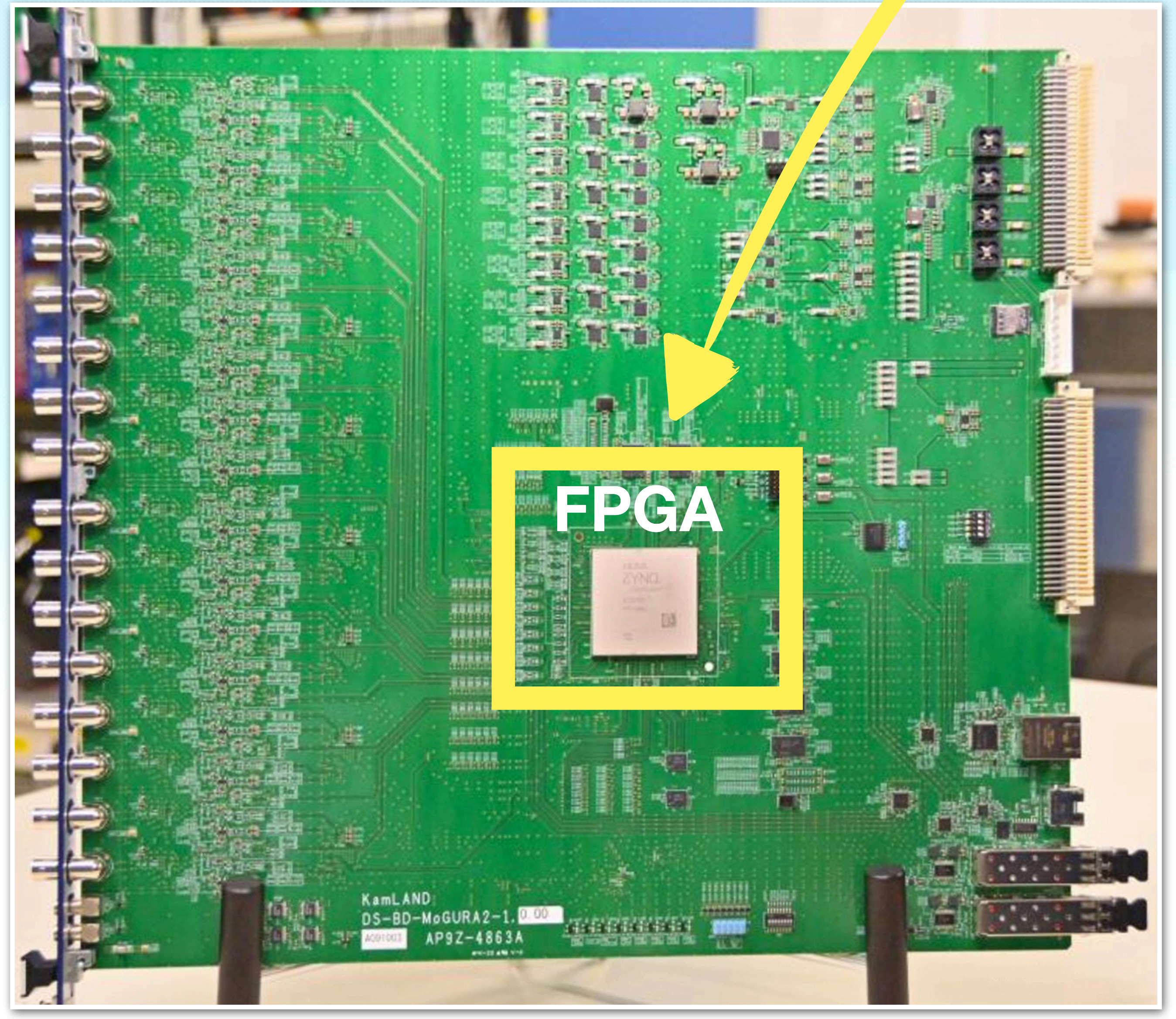
Detector Response

Hardware-AI Codesign

Deploy ML model onto FPGA
to produce these in real-time



Data Stream



Offline Analysis

Energy

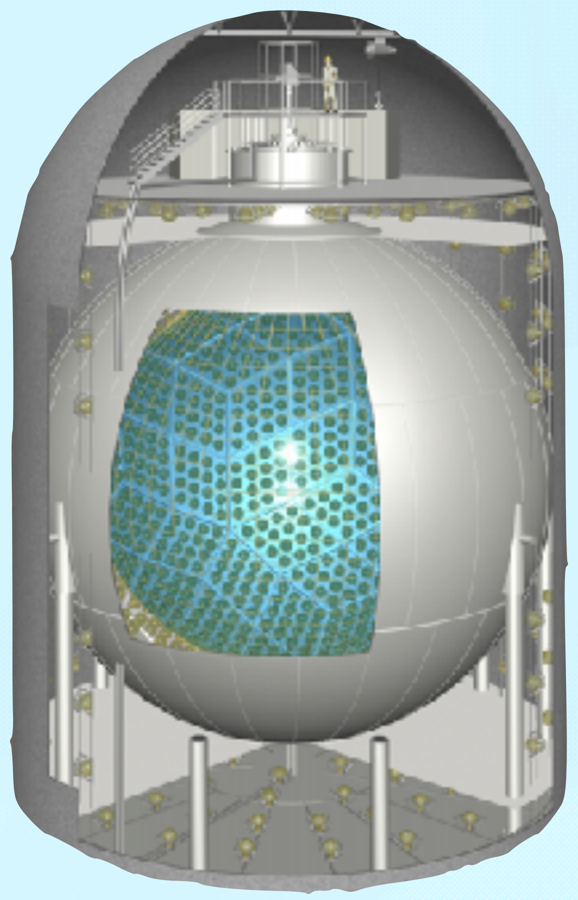
Position

Particle Type

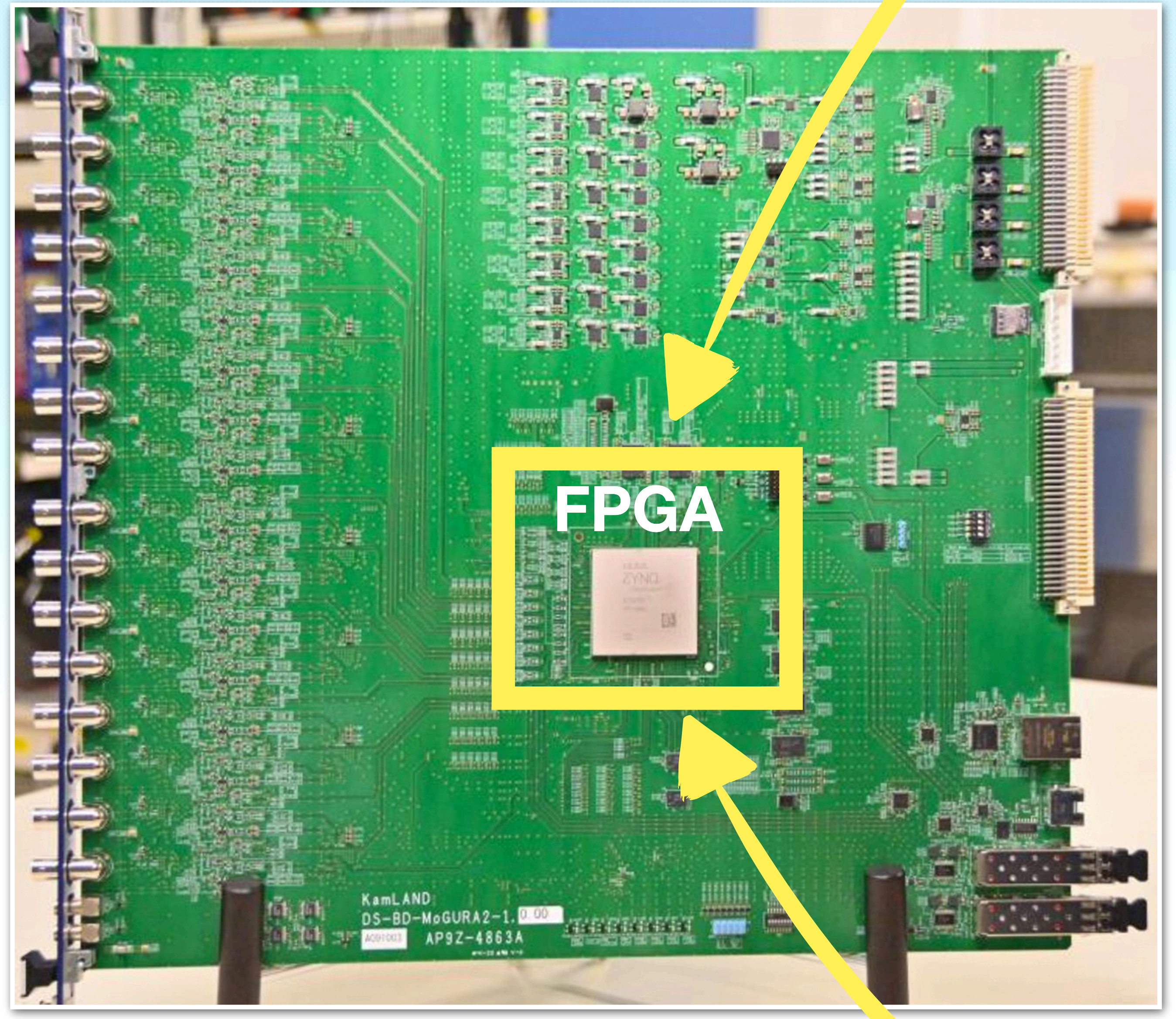
Detector
Response

Hardware-AI Codesign

Deploy ML model onto FPGA
to produce these in real-time



Data Stream



Offline Analysis

Energy

Position

Particle Type

Detector
Response

Online model update to account
for detector status change

Summary

“AI and Data Science has shaped rare event search, it’s an accelerator for new astrophysics results”

- LEGEND: Rare Event Surrogate Model
- ABRACADABRA: TIDMAD Data Set

“It will continue to evolve and foster discovery in next-generation experiments”

- AI for Rare Event Lab: https://aobol.github.io/AoboLi/
- Email: aol002@ucsd.edu