The Machine Learning Epochs of Neutrinoless Double Beta Decay

Aobo Li
UNC Chapel Hill

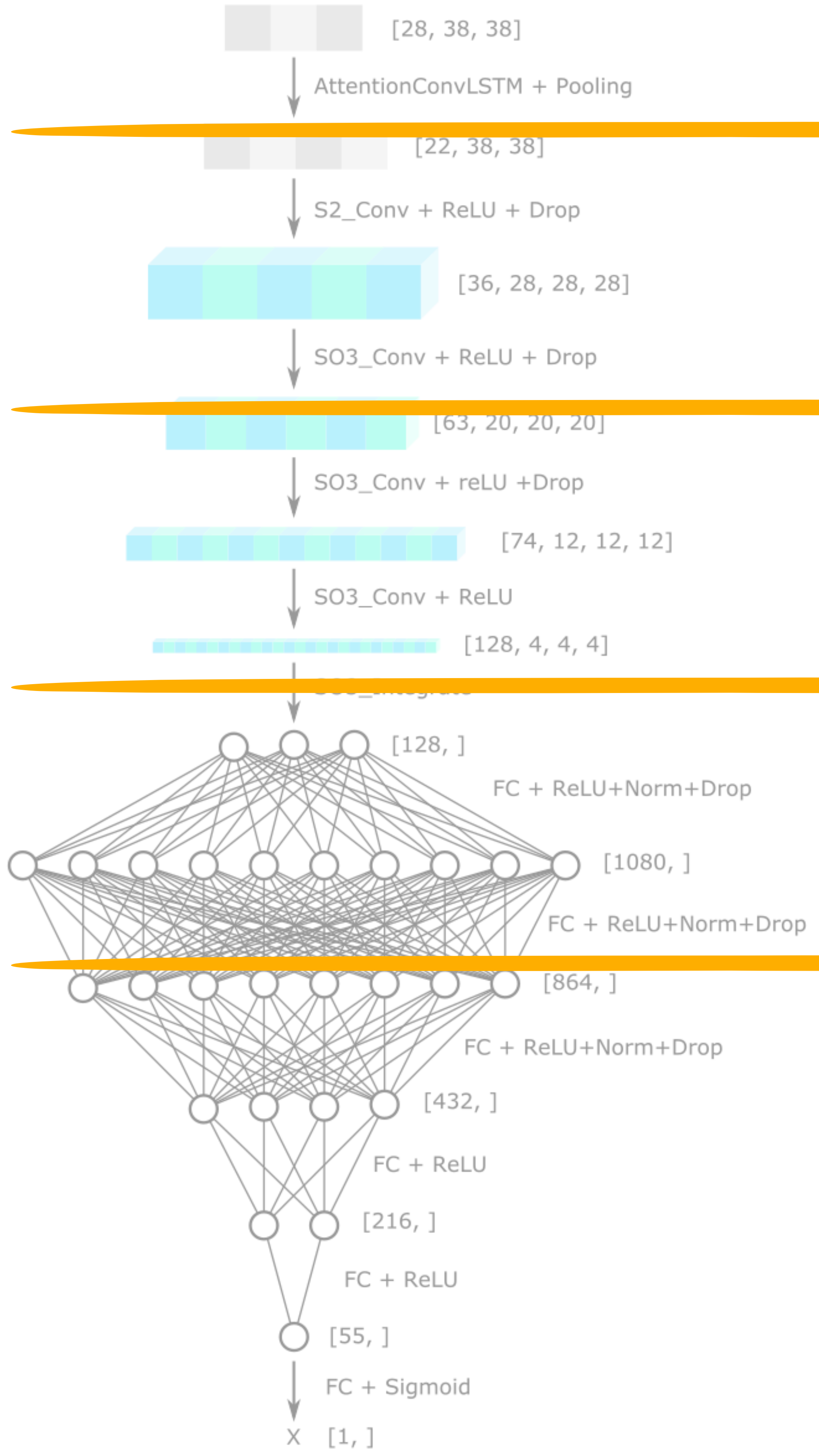
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

TRIANGLE UNIVERSITIES NUCLEAR LABORATORY

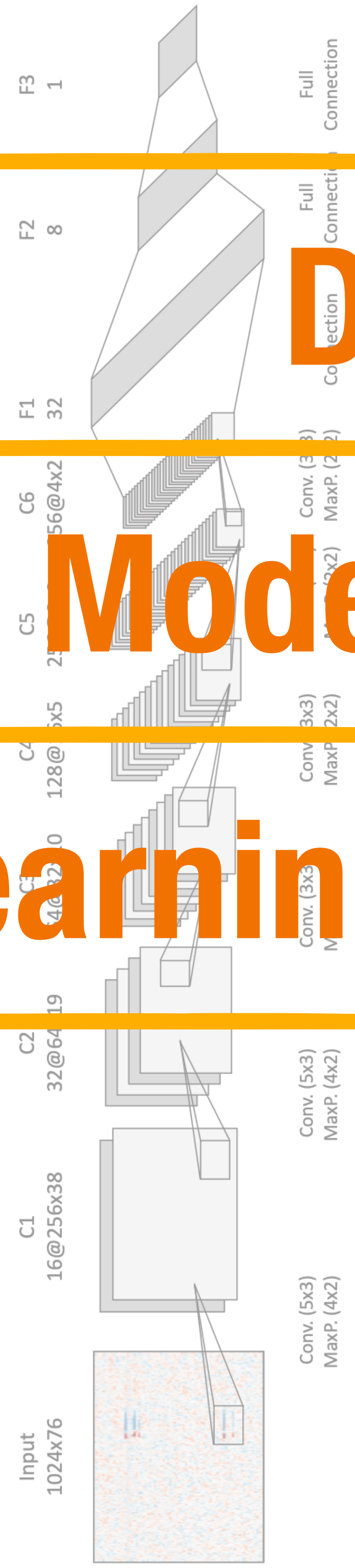
CoSSURF 2022, 04/25/2022

Outline

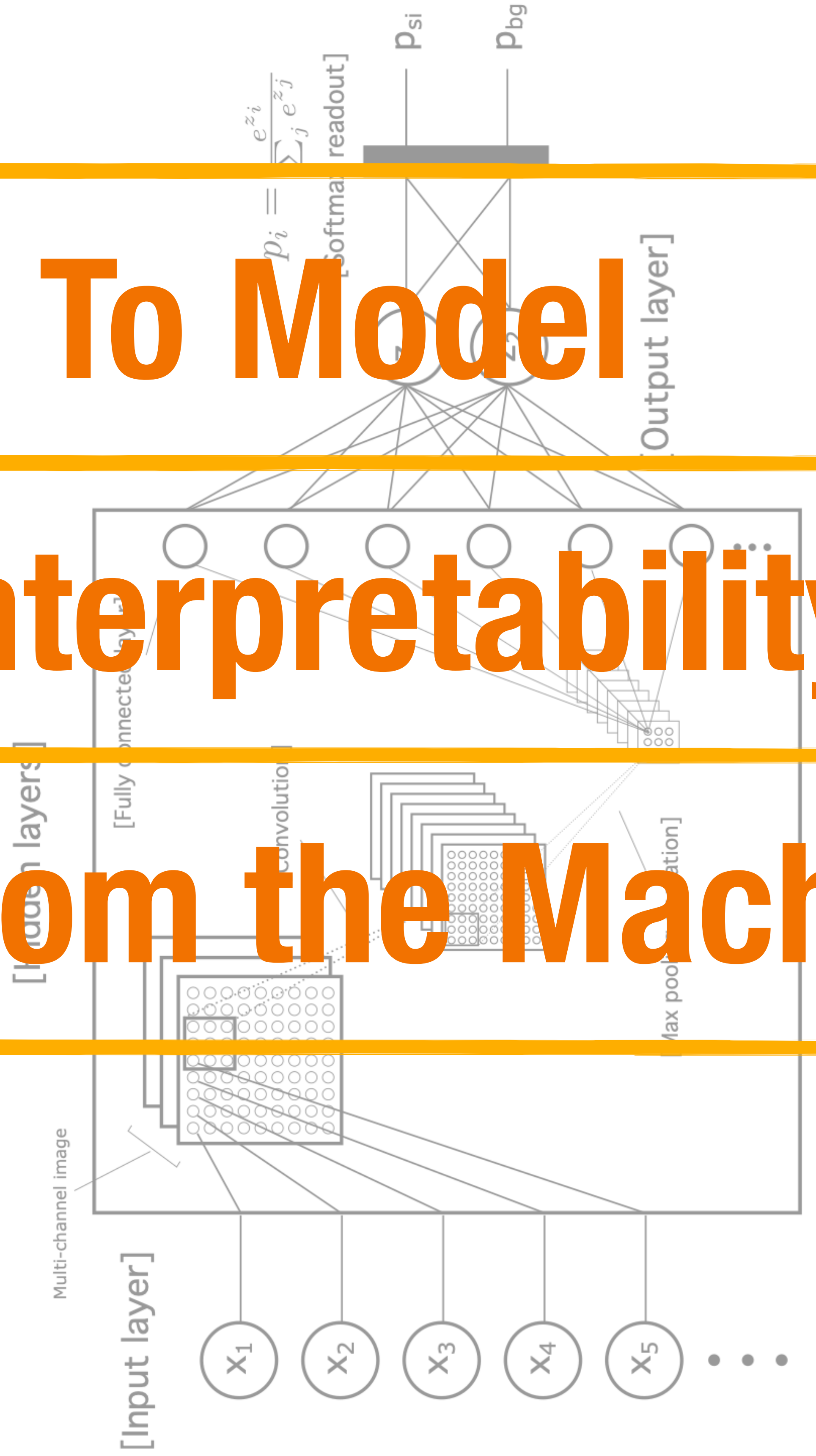
KamLAND-Zen



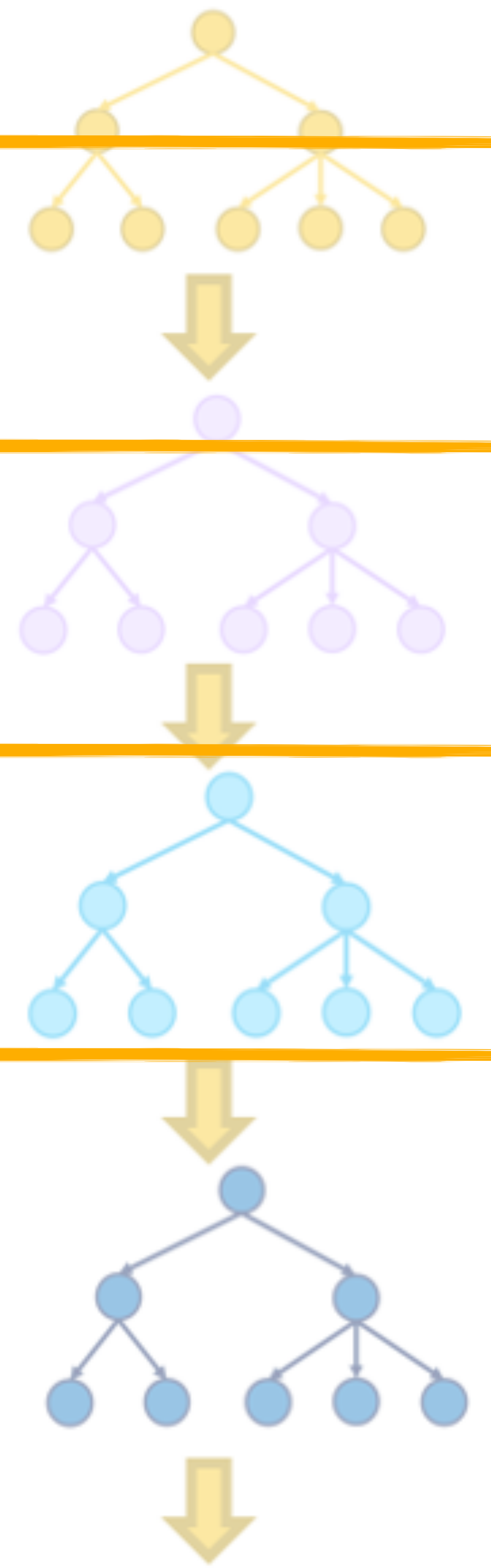
EXO-200



NEXT



MAJORANA DEMONSTRATOR



Data To Model

Model Interpretability

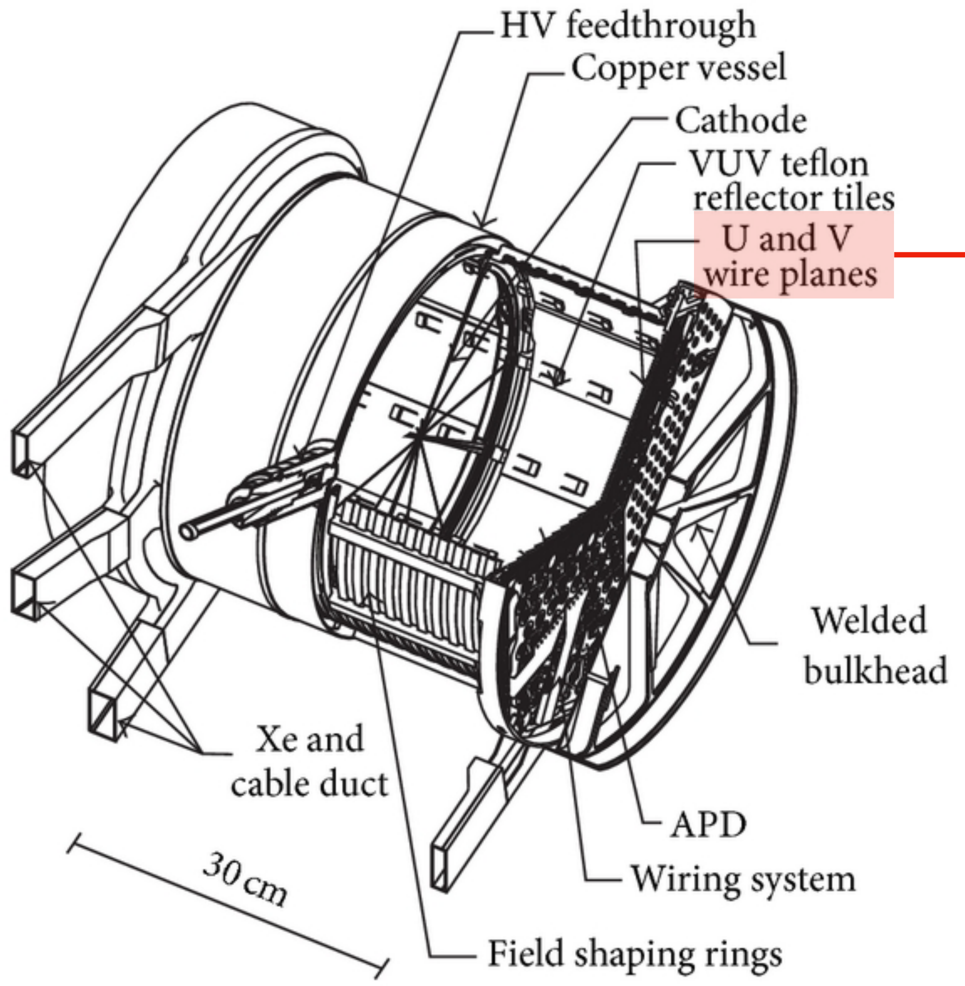
Learning from the Machine

Data To Model

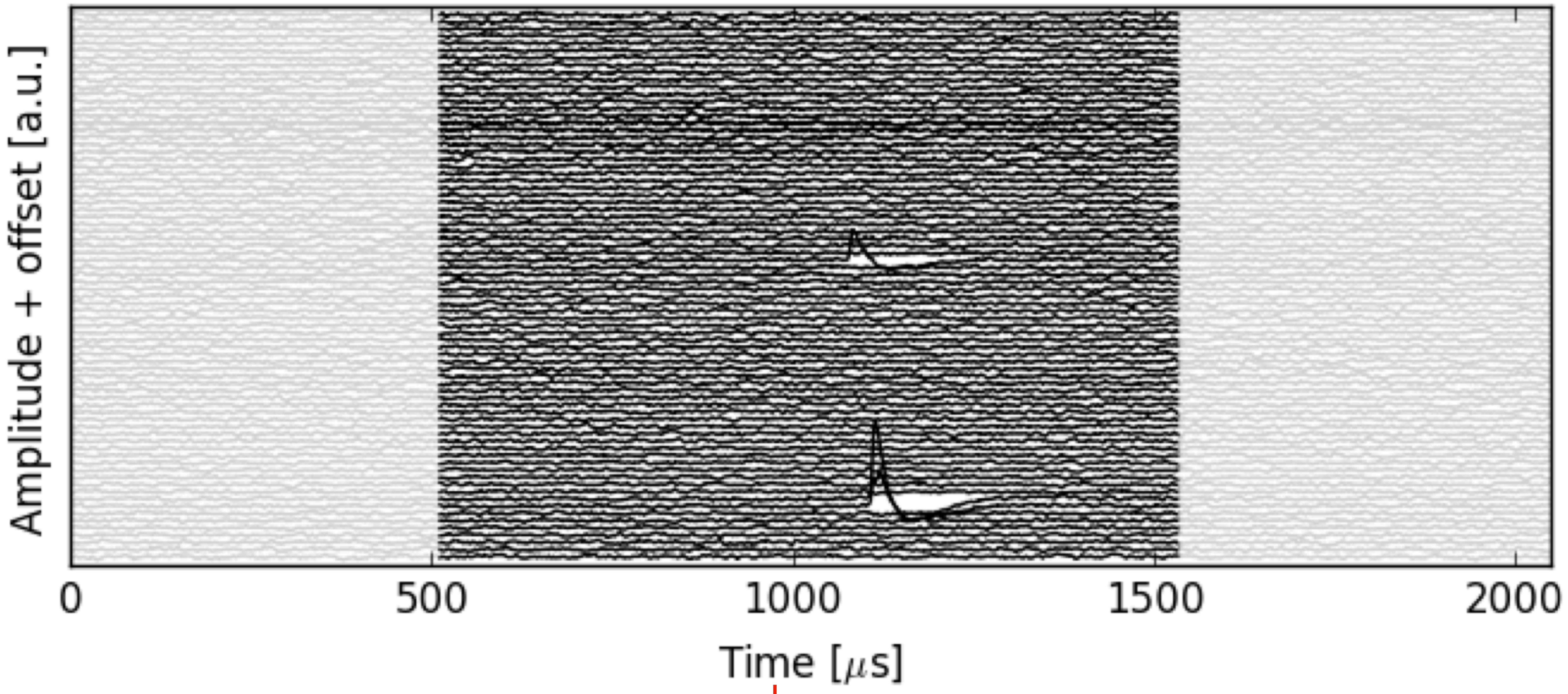
Energy Reconstruction in EXO-200

S. Delaquis et al JINST 13 P08023

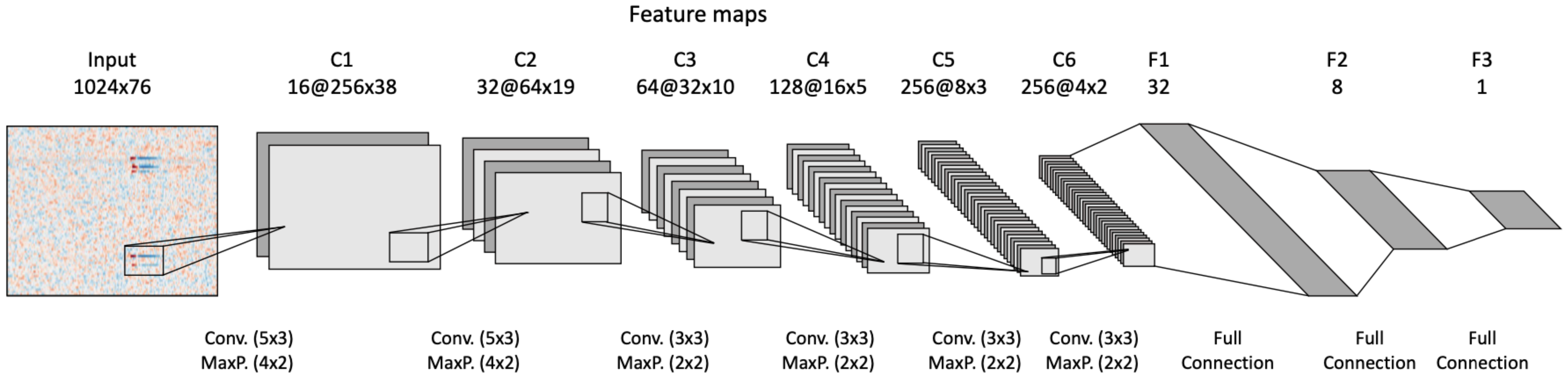
Detector



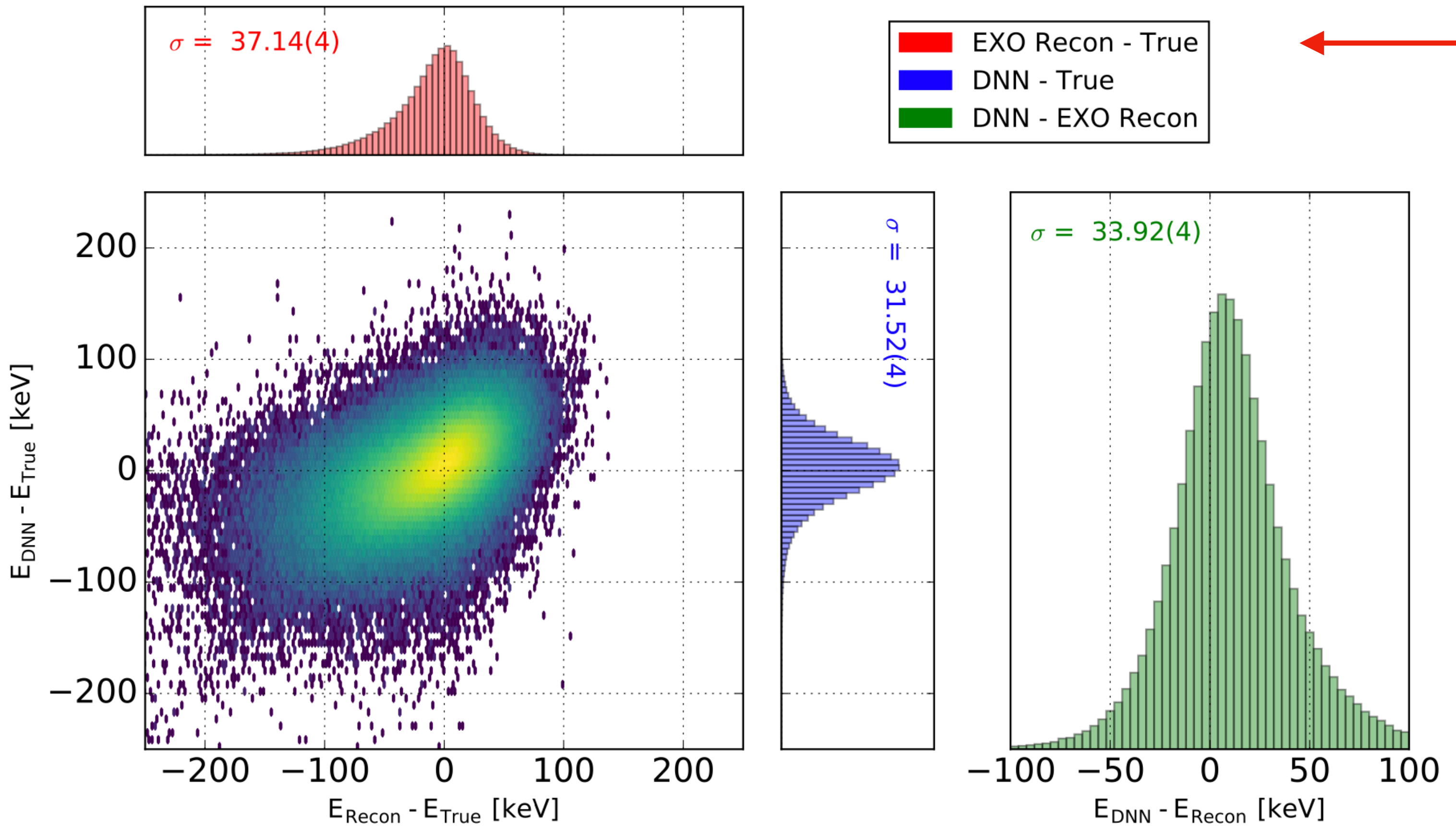
Data



Model



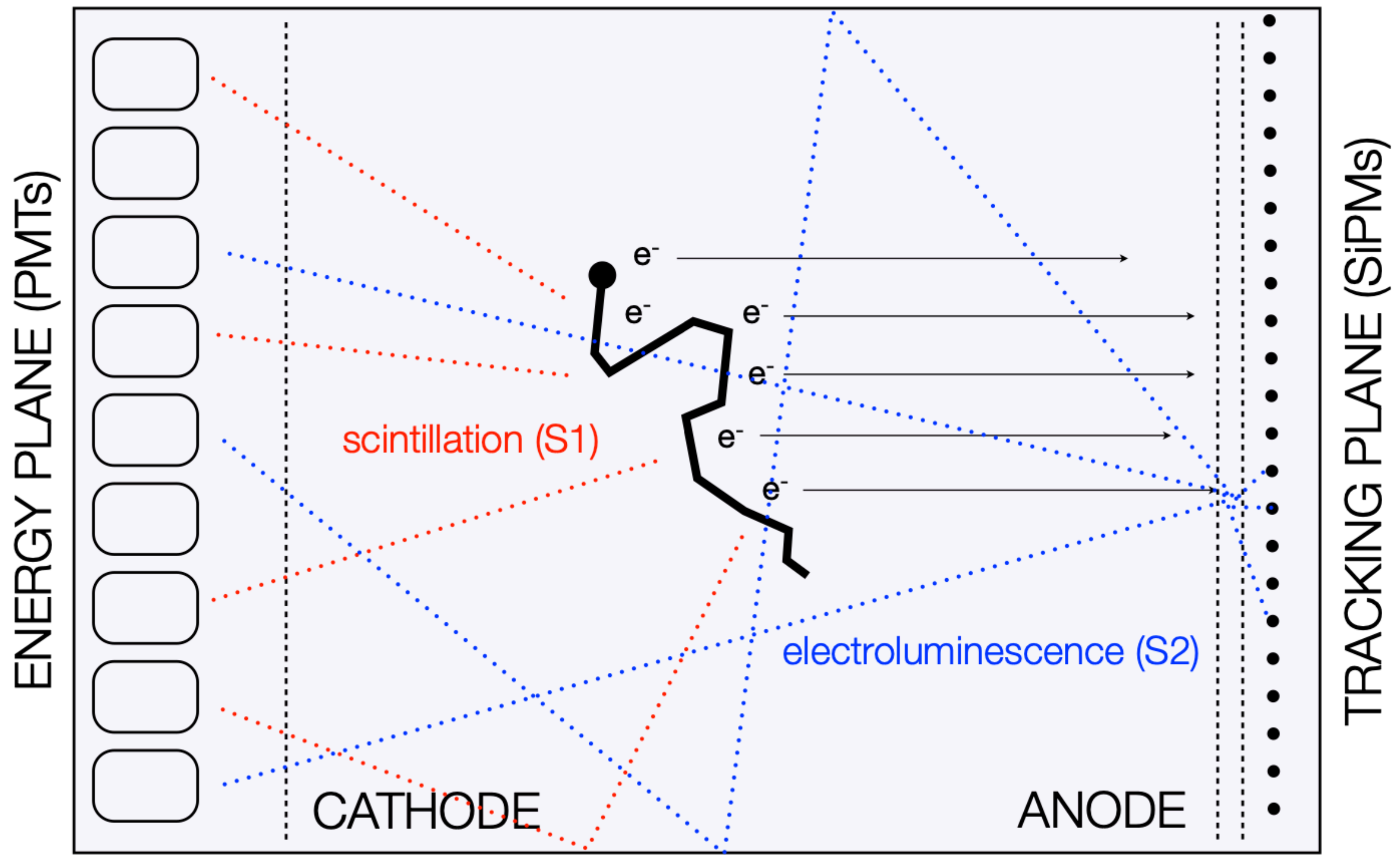
Result

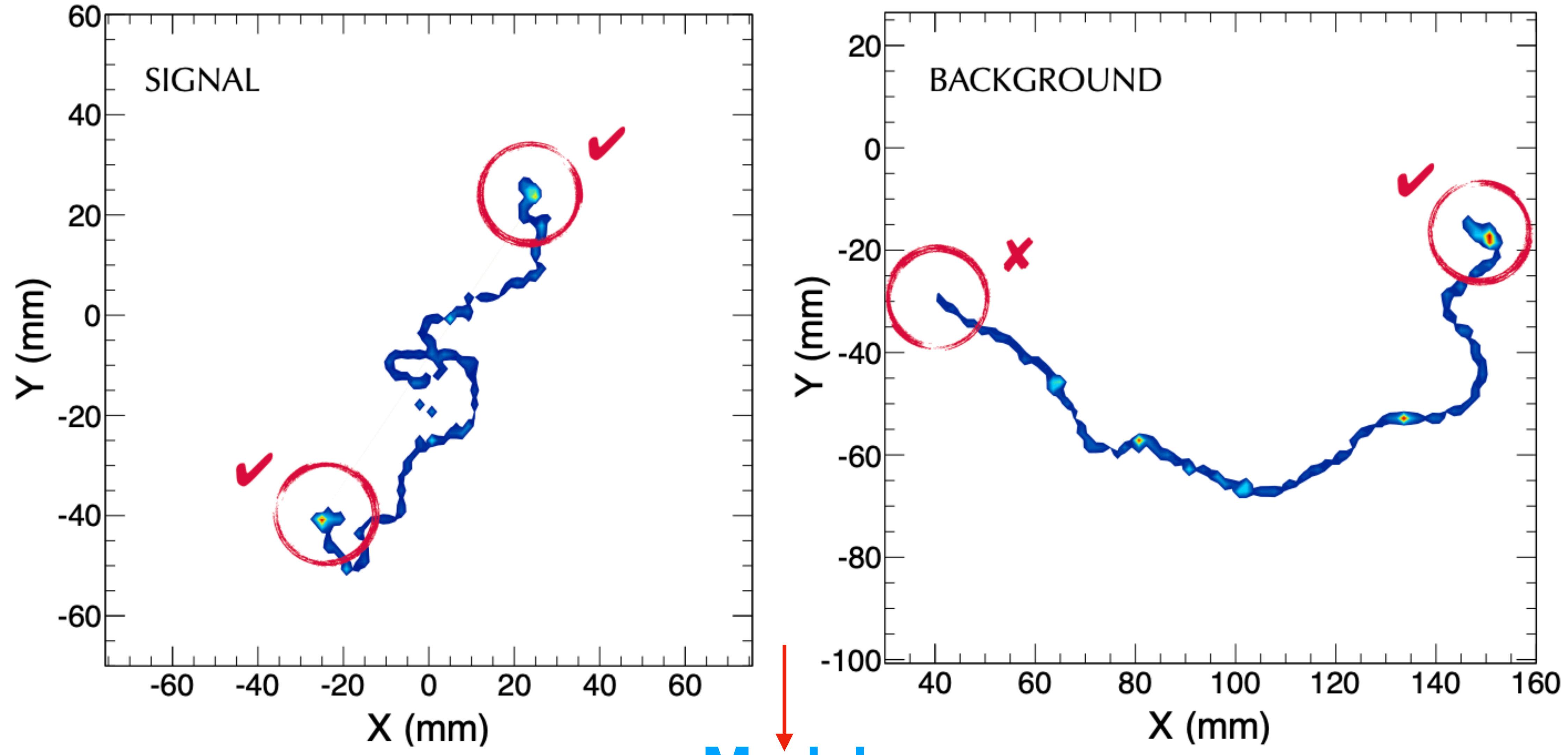


Background Rejection in NEXT

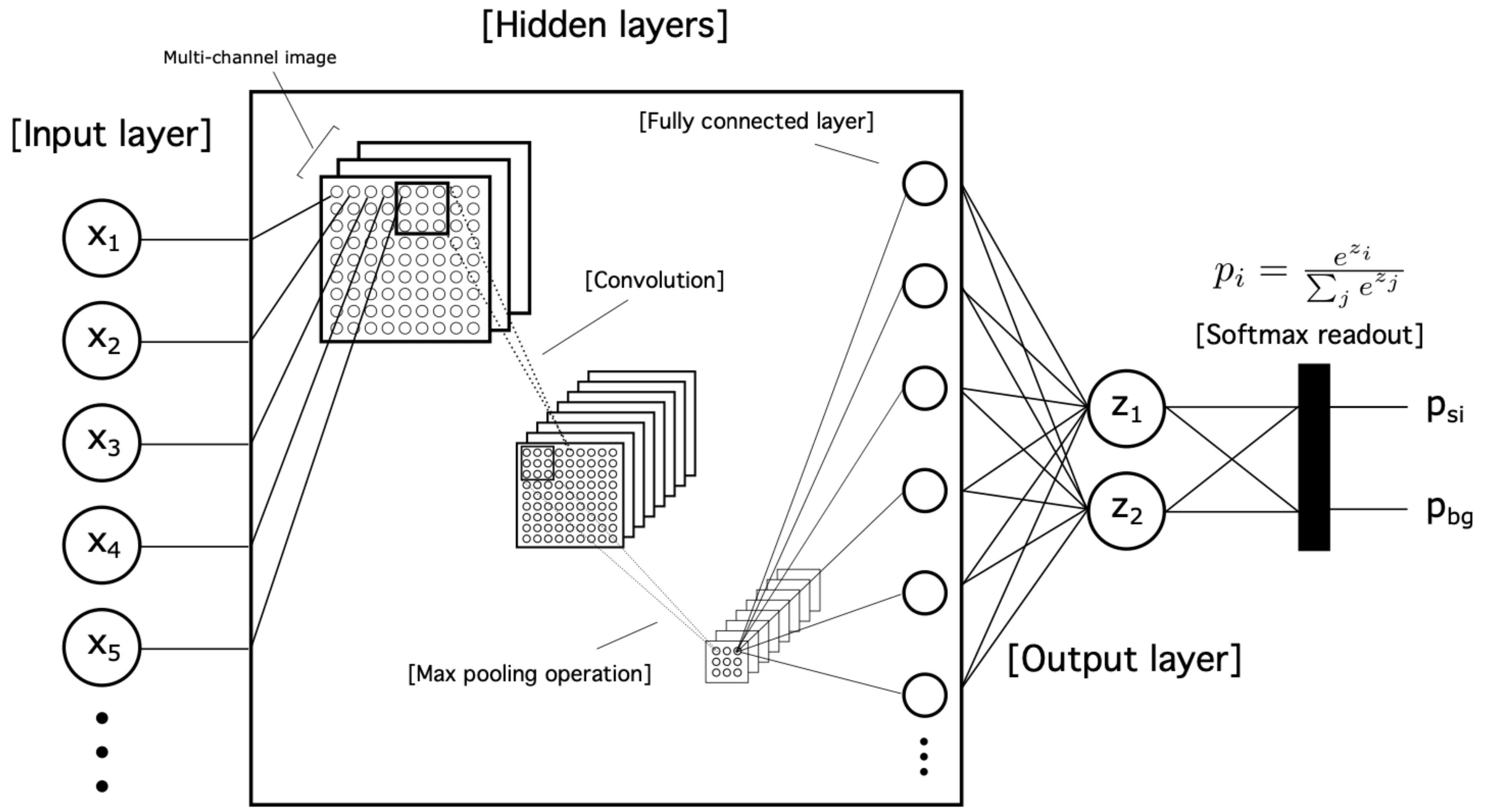
J. Renner *et al* 2017 *JINST* 12 T01004

Detector → Events

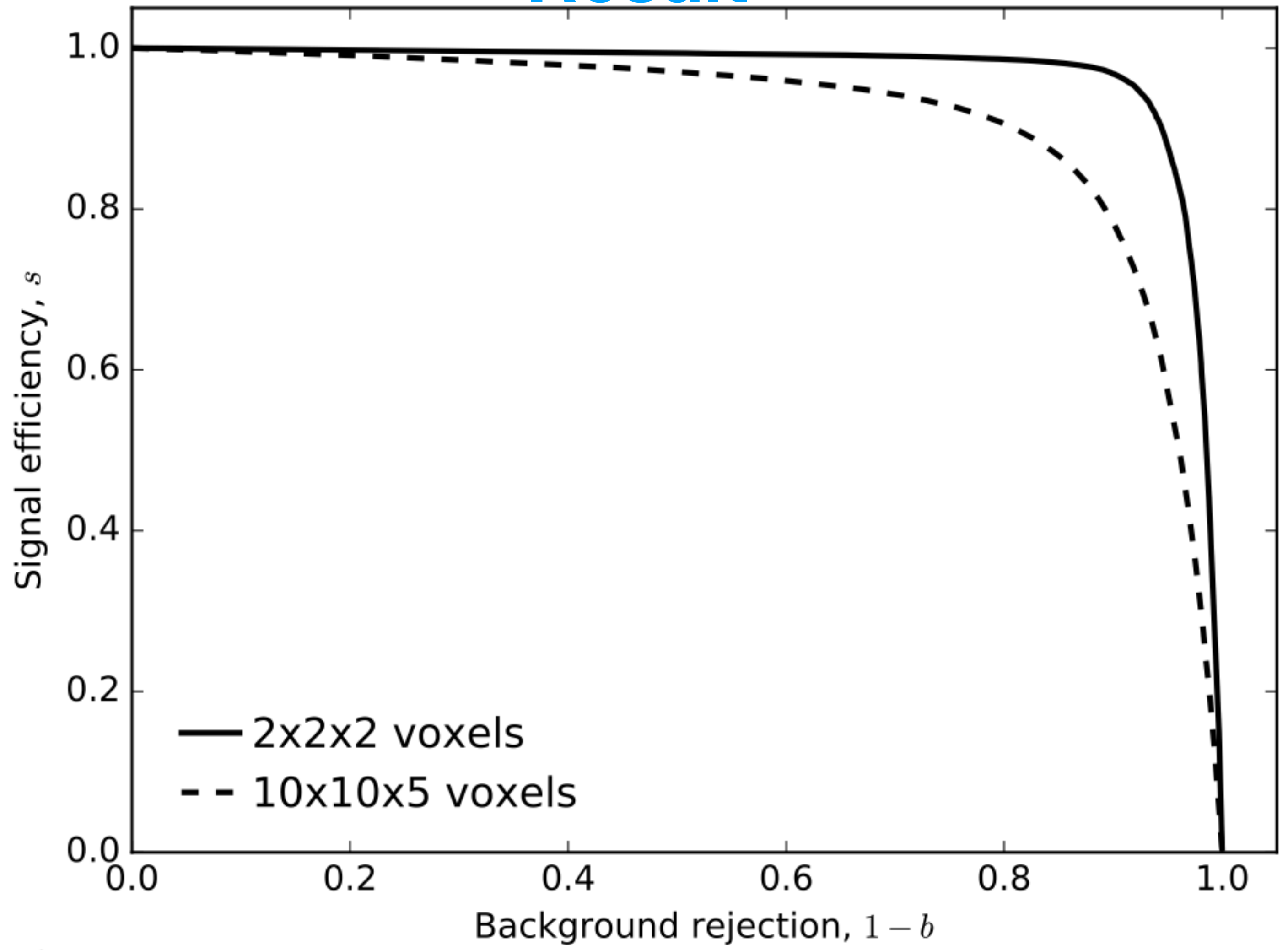




Model

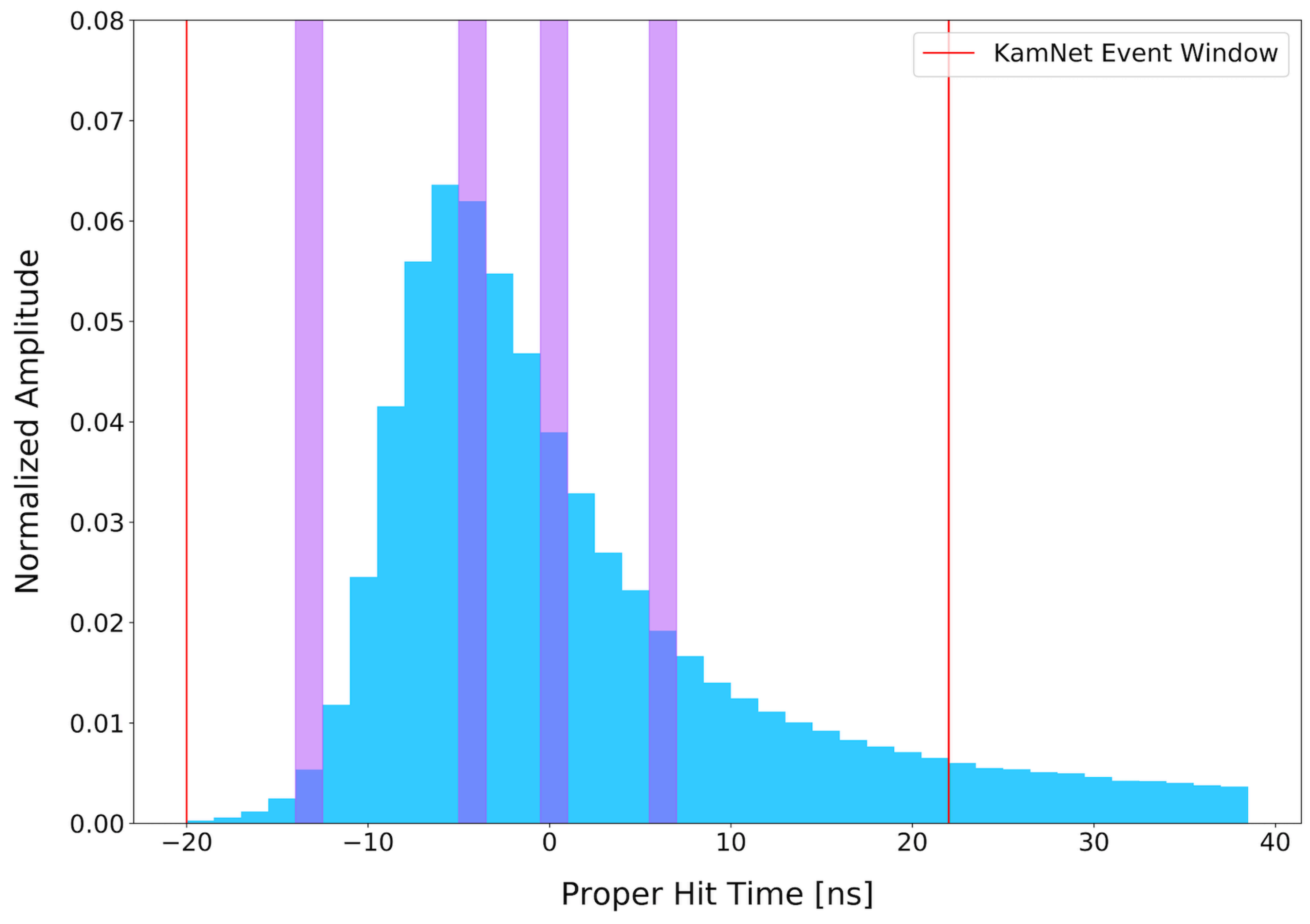


Result

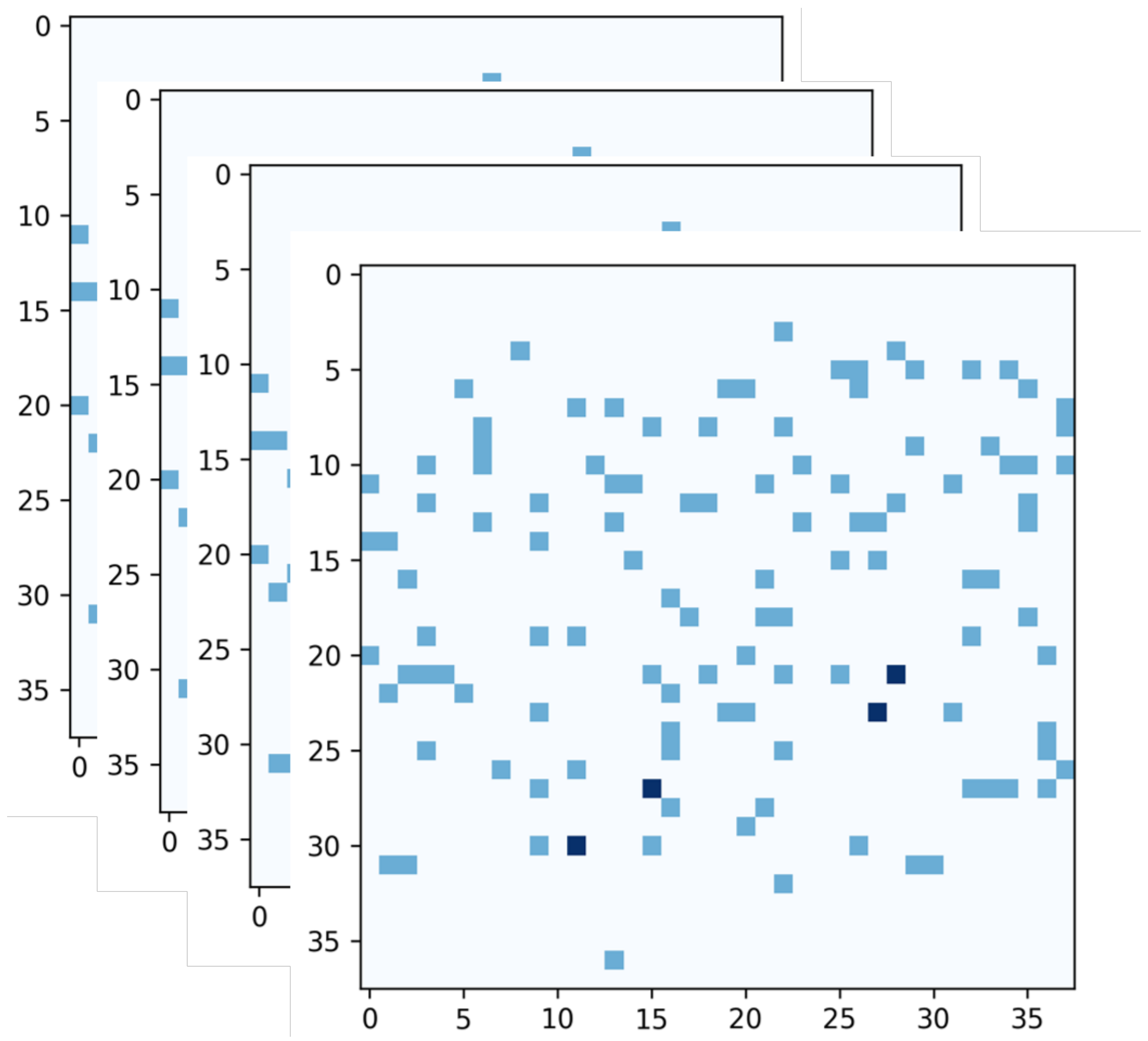


KamLAND-Zen Data

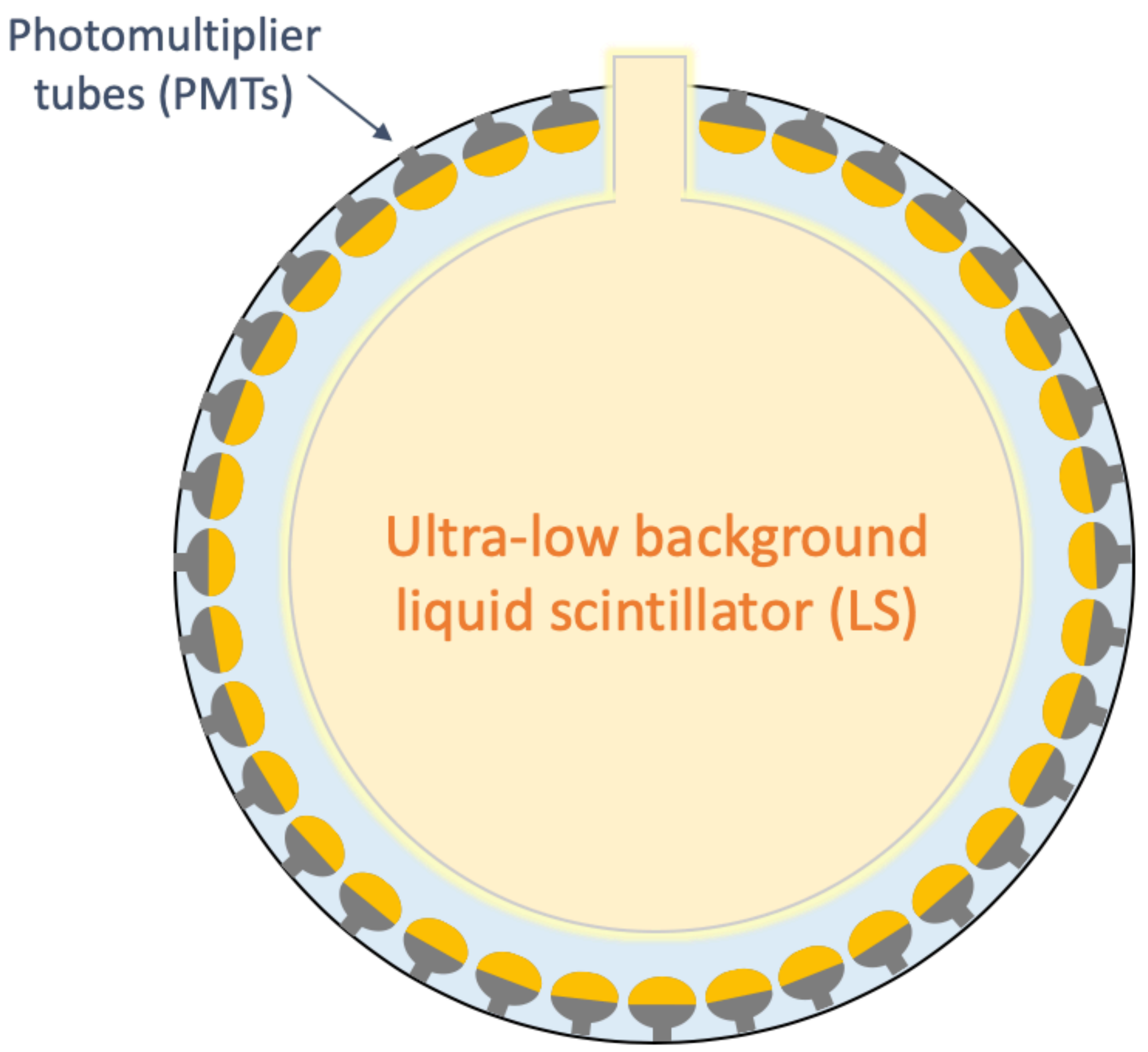
A time series …



… of images …



… on a sphere

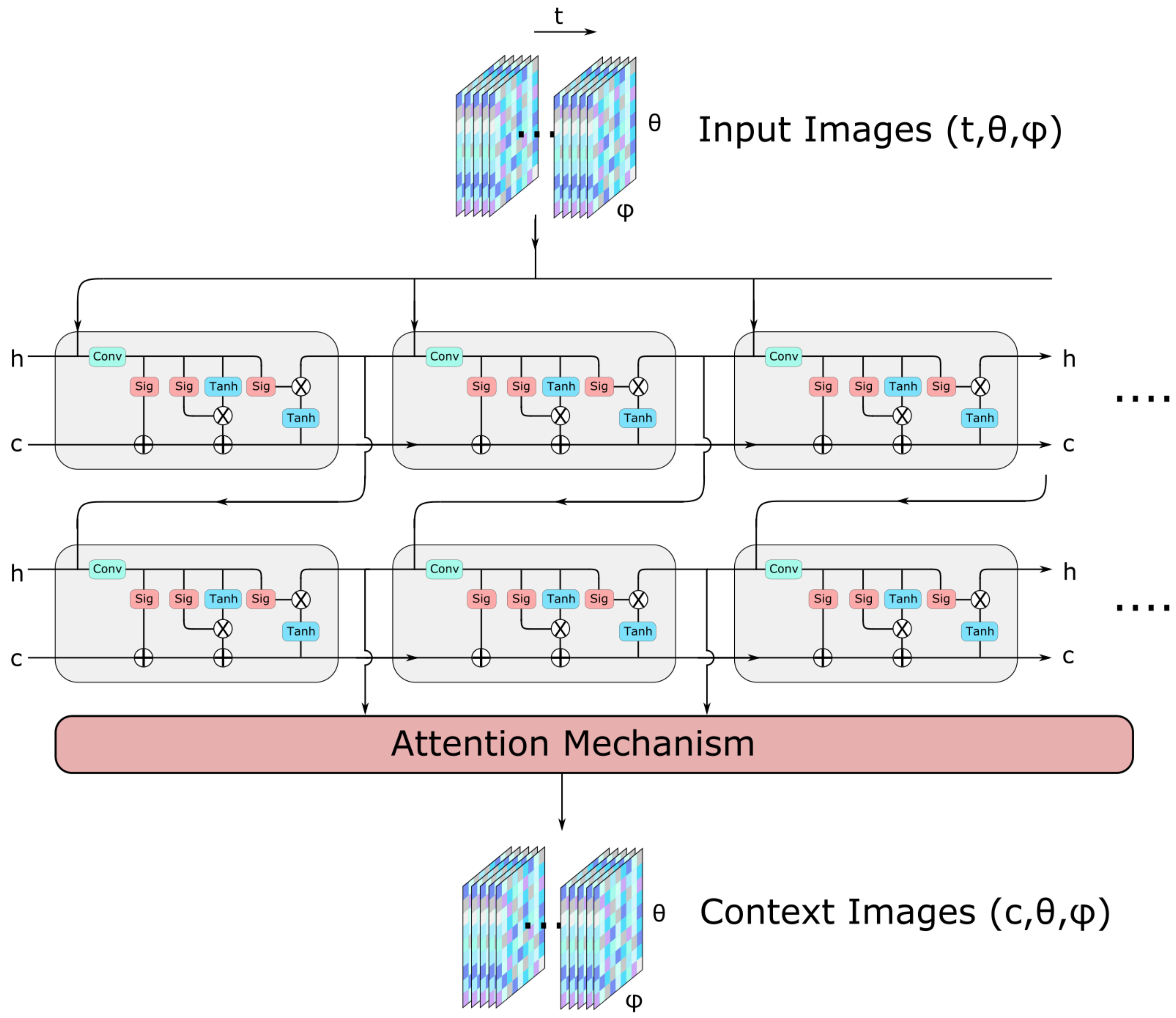


KamNet: An Integrated Spatiotemporal Neural Network

A. Li et al: arXiV 2203.01870, Submitted to PRC

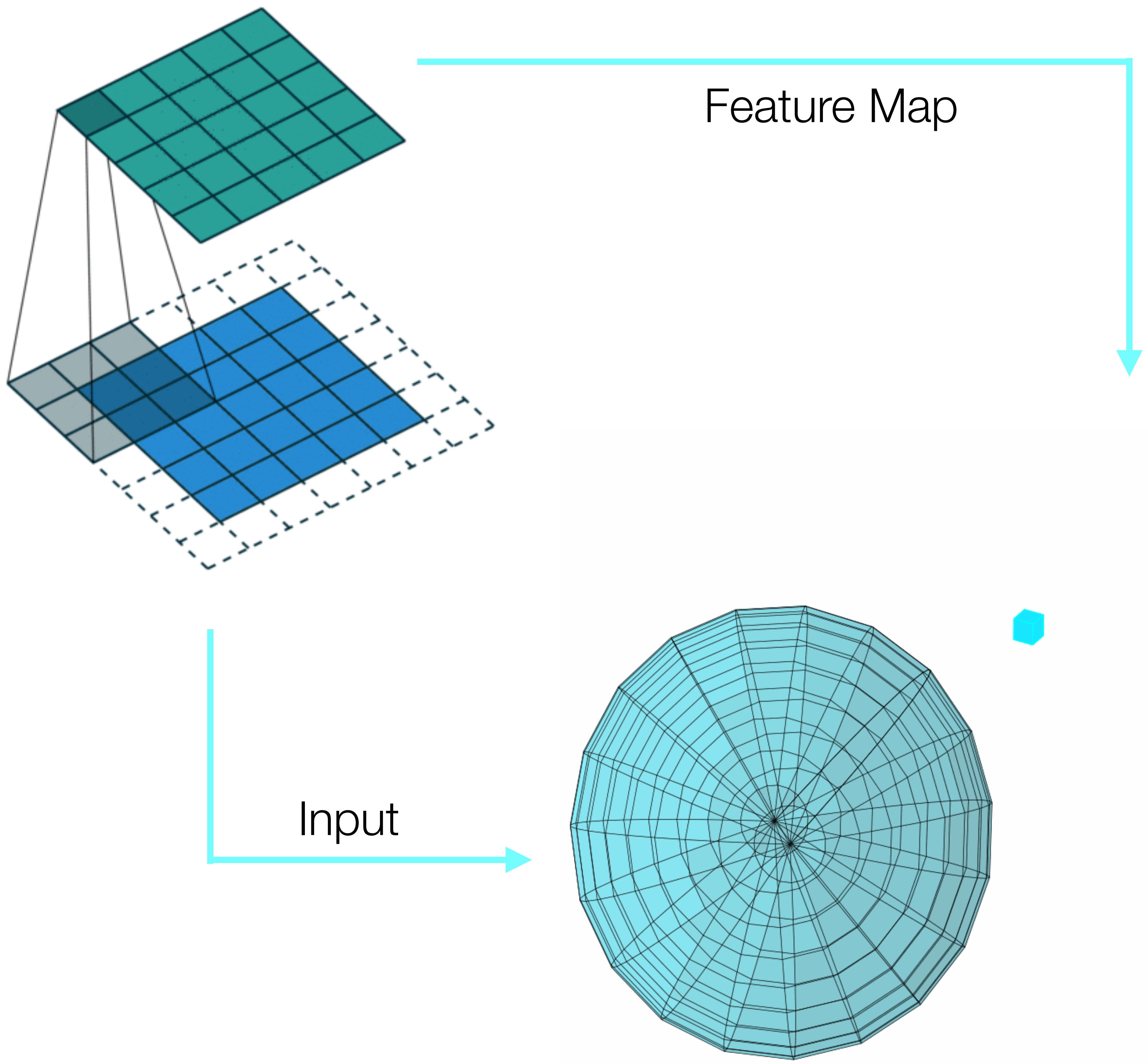
AttentionConvLSTM

for Spatiotemporal symmetry
ArXiv: 1506.04214



Spherical CNN

for SO(3) symmetry in spherical detector
ArXiv: 1801.10130

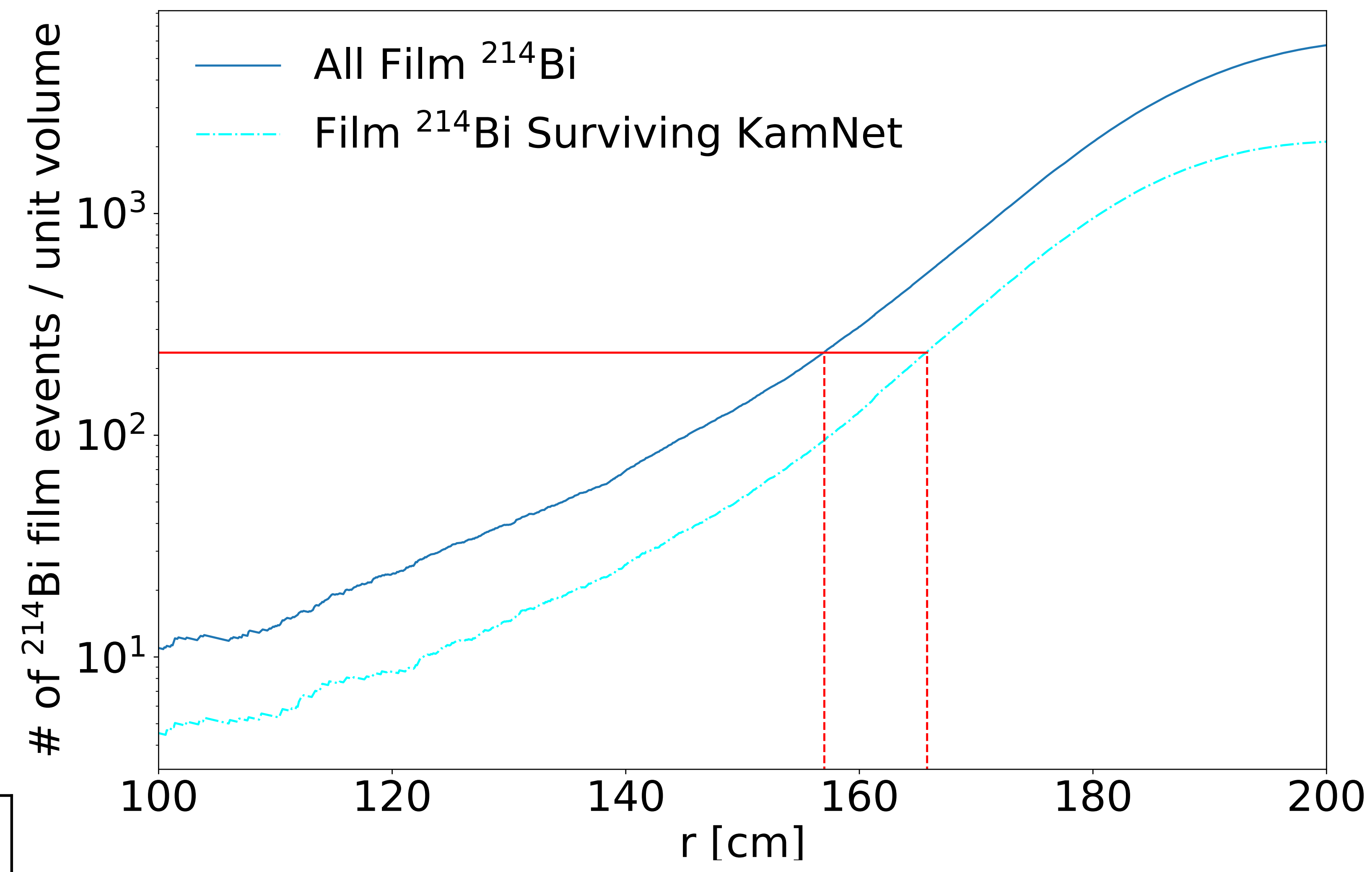


KamNet Result

A. Li et al: arXiV 2203.01870, Submitted to PRC

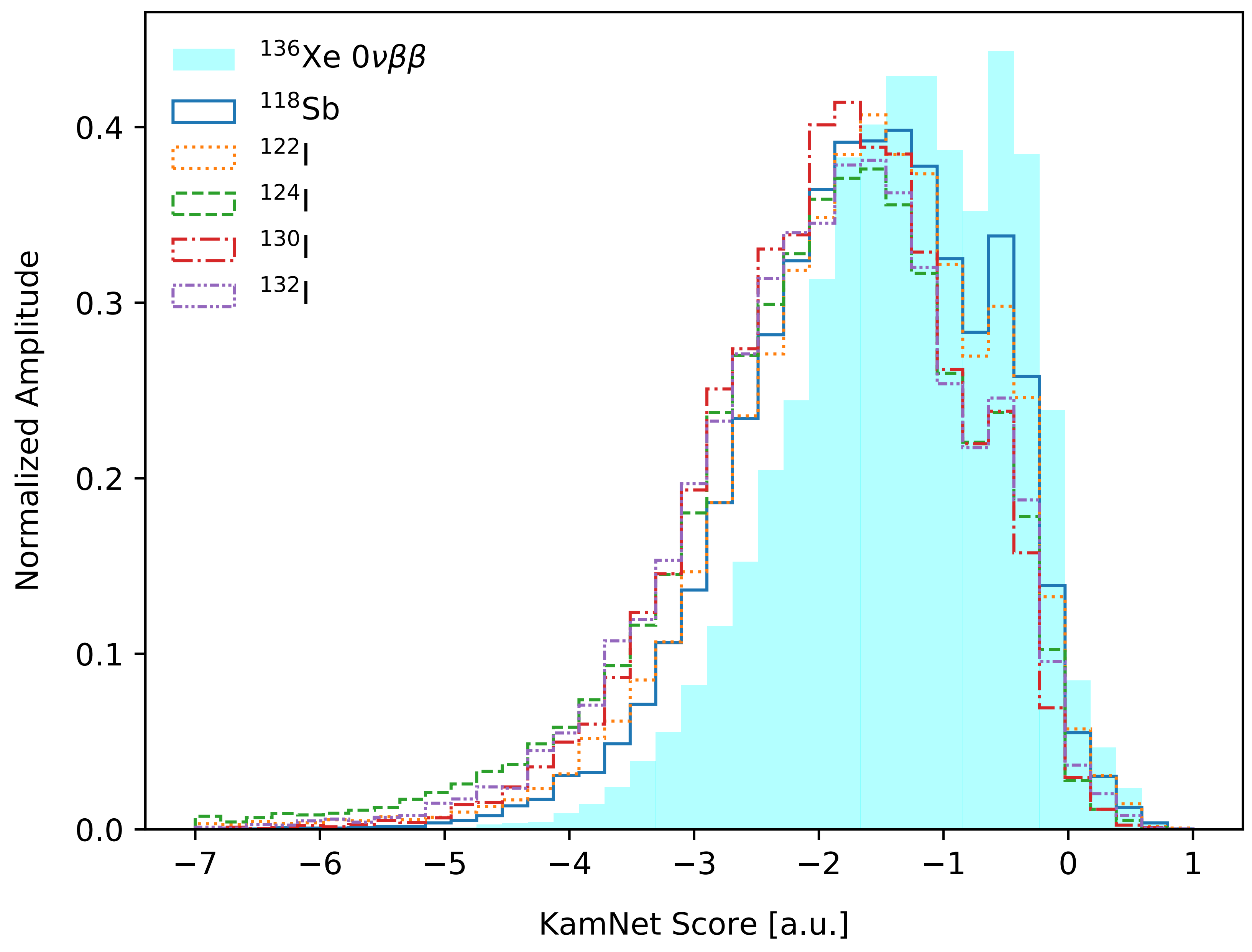
KamNet is trained on precisely tuned MC simulations and evaluated on various backgrounds in KamLAND-Zen 800

While accepting **90%** of $0\nu\beta\beta$ events, KamNet rejects **~27%** of XeLS backgrounds and **~59%** of film backgrounds

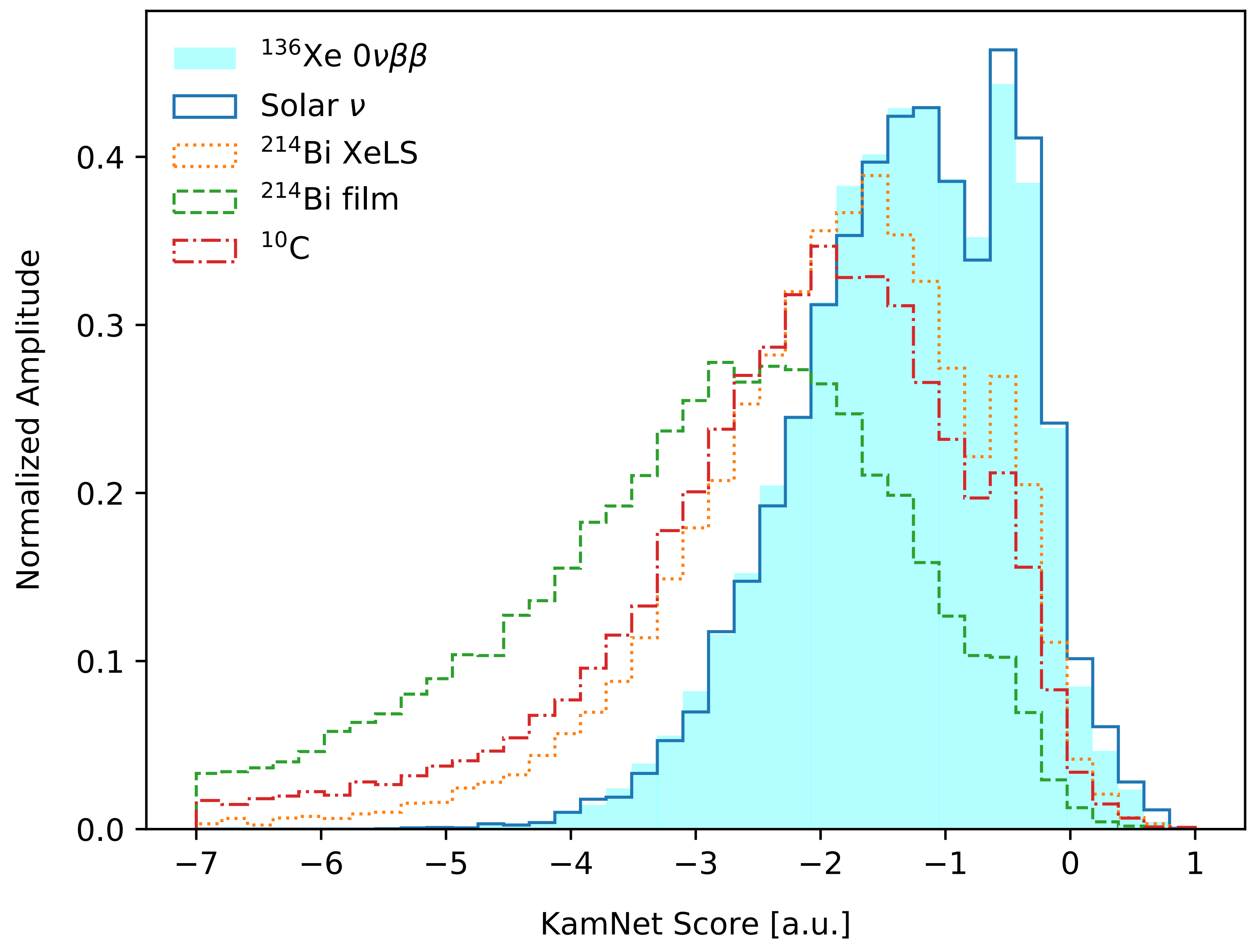


The increased rejection of backgrounds on mini-balloon film allows for the expansion of the fiducial volume from 157cm to 165.8cm, resulting in **17.7% gain** on exposure without hardware upgrades

Long-Lived Spallation



Other Backgrounds



Reconstruction-Level Machine Learning

Manuscript under collaboration review

[DETTYPE , DETECTOR , DRIFT TIME , AVSE , DCR , NOISE , TDRIFT50 , DS]

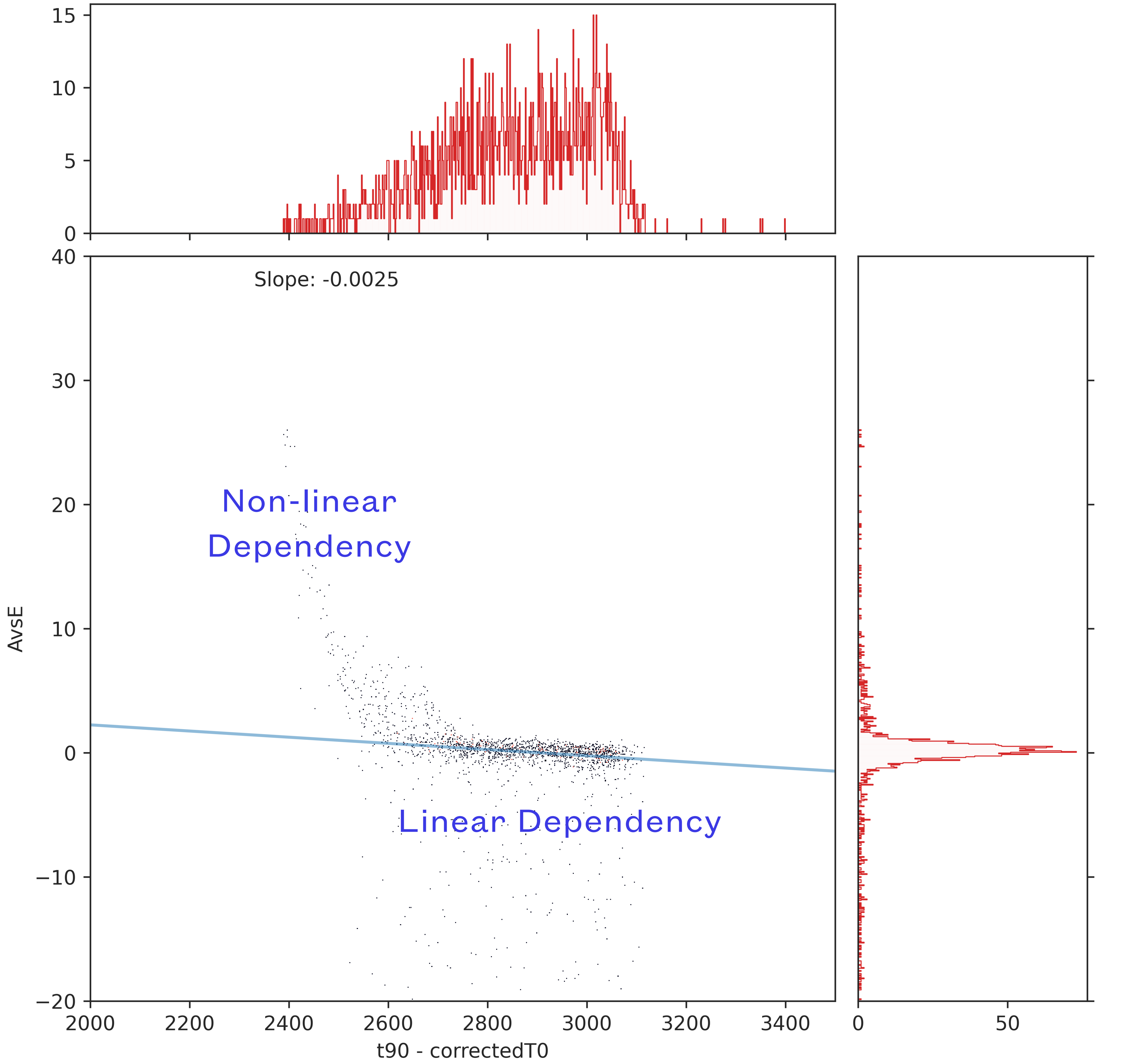
In Majorana Demonstrator, a drift-time dependency is observed for some reconstruction parameters

- Correcting these correlations improve the performance of background rejection

Reconstruction parameters are tuned detector- and run-wise, which is time consuming

Design a machine learning model to simultaneous solve these two problems:

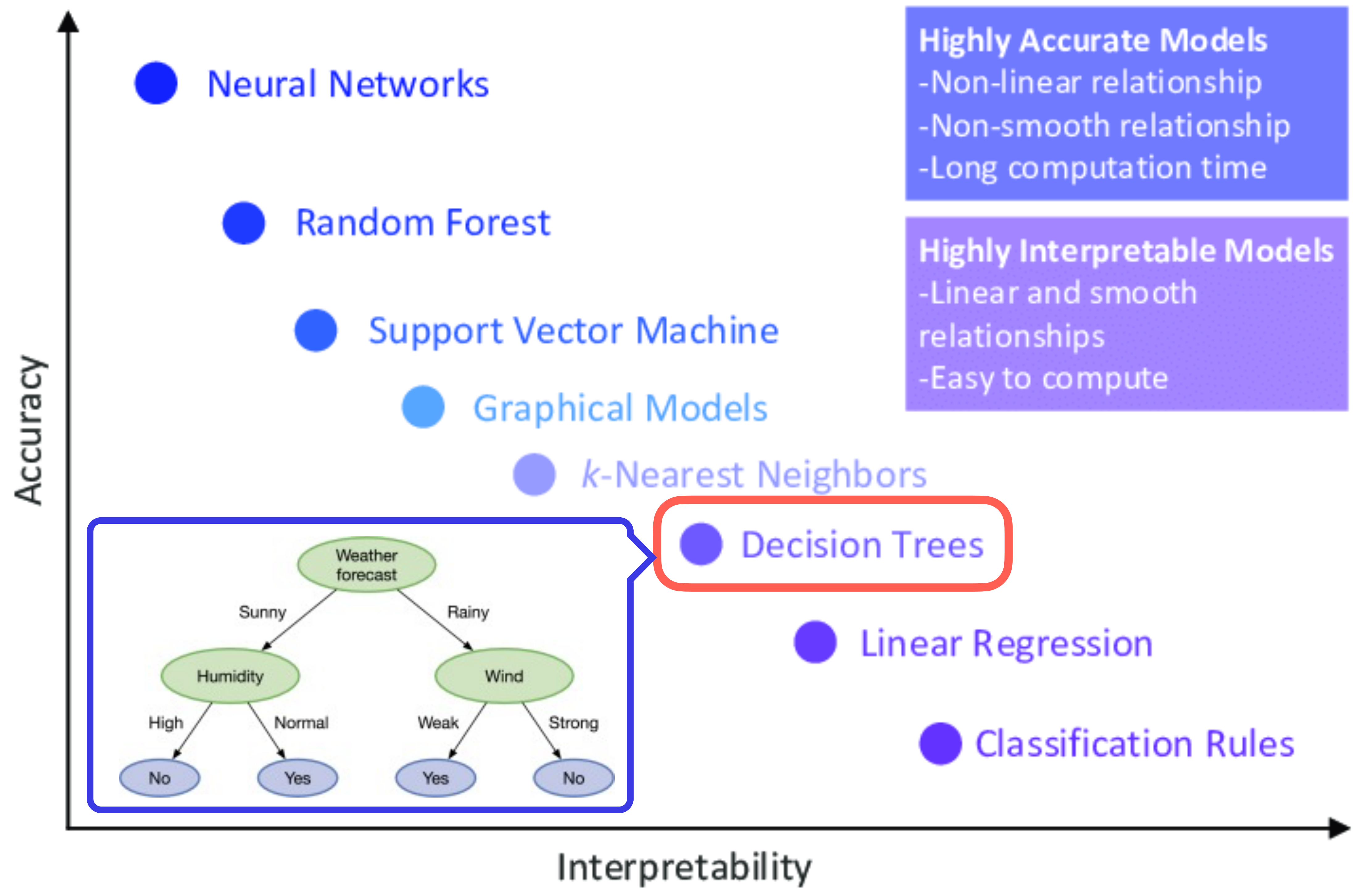
- **Continuous Parameters:** reconstruction parameter
- **Categorical Parameters:** detector number, type and run number

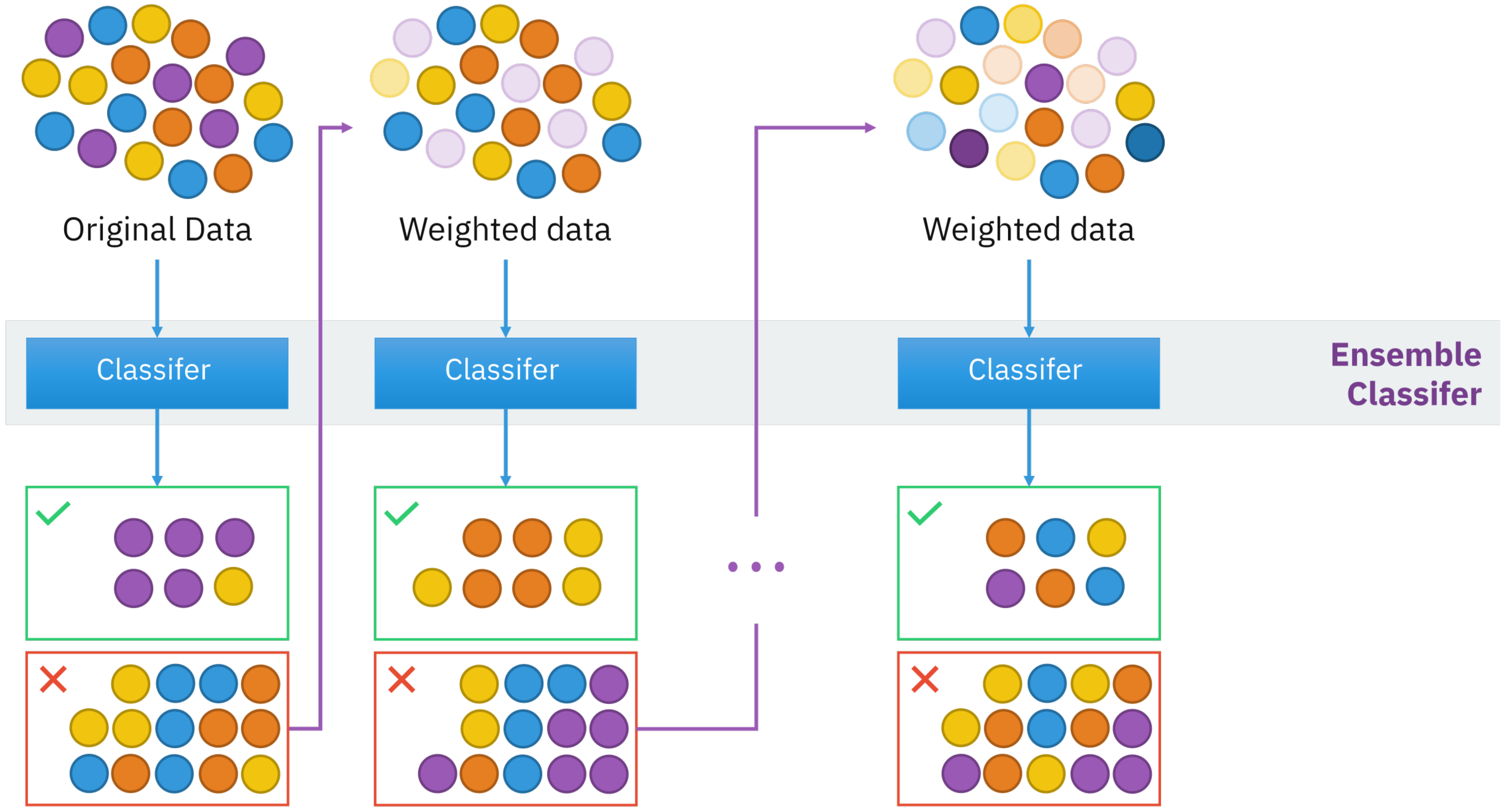


Boosted Decision Tree

- Decision tree is highly **interpretable** model
- Boosting algorithm utilizes ensemble learning to makes it **powerful**
- Naturally handle **categorical** features and **continuous** features together

LightGBM



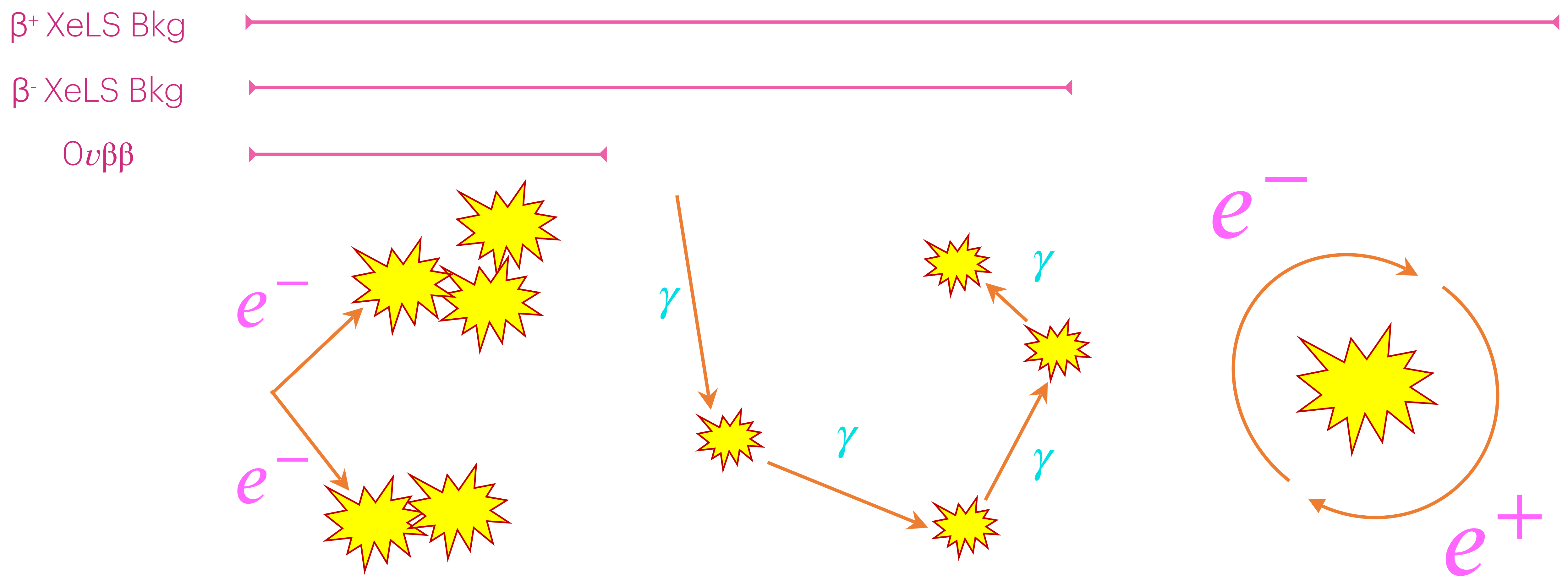


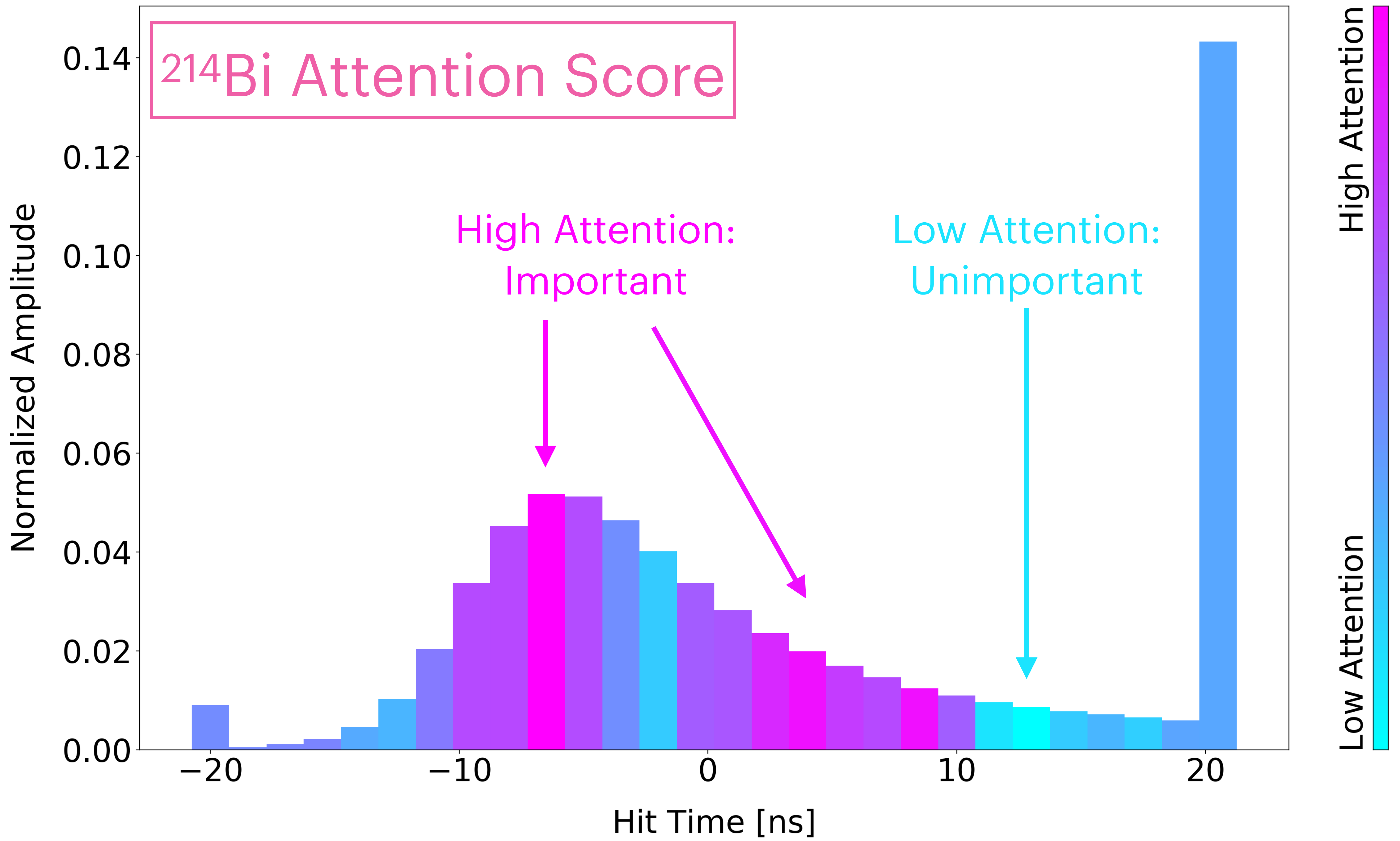
Model Interpretability

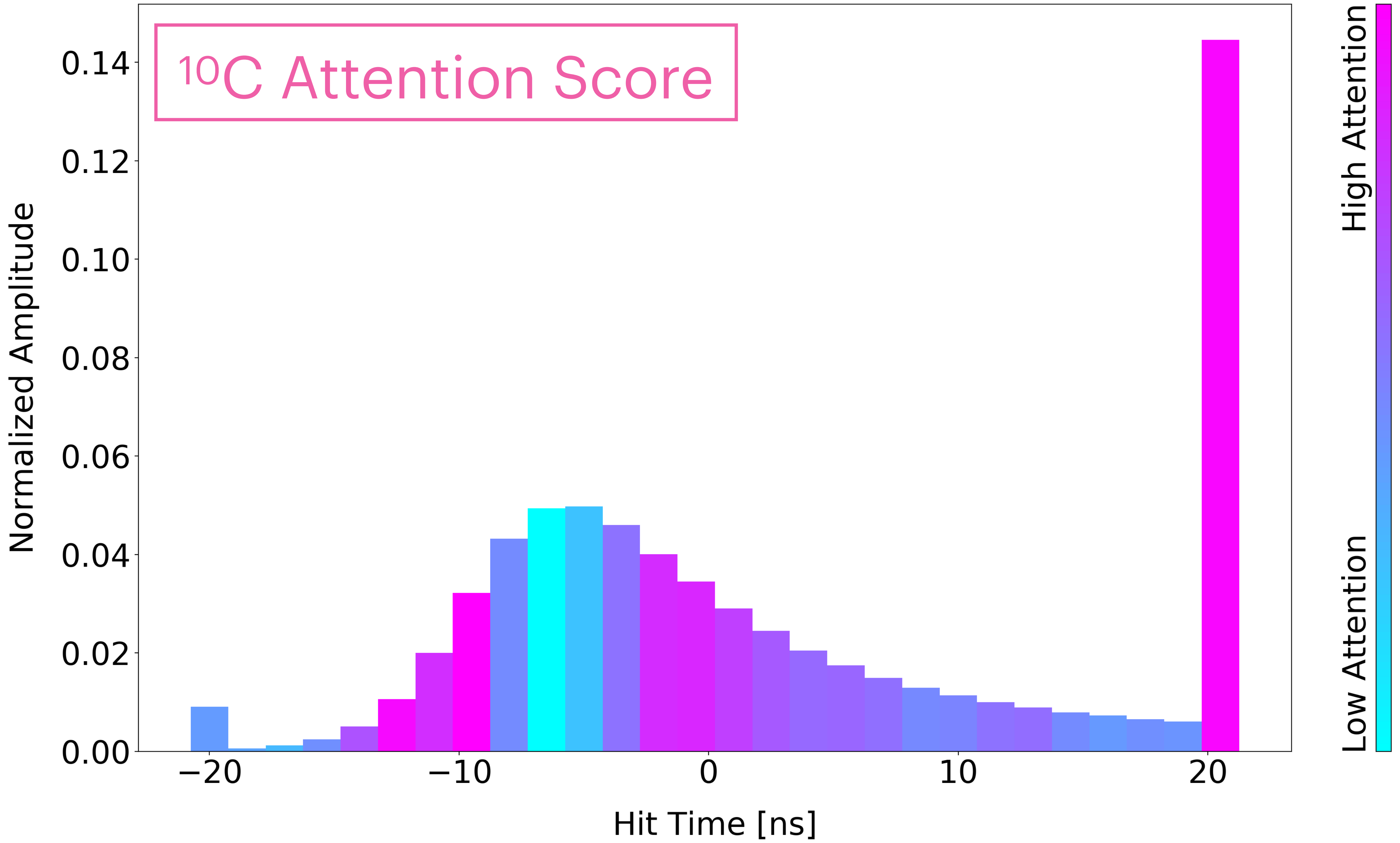
Interpreting KamNet

A. Li et al: arXiV 2203.01870, Submitted to PRC

- β-like signal events are strictly **single-vertex events**, all energy is deposited in a very localized region
- γ-like backgrounds are **closely-spaced multi-vertex events**, part of event energy is deposited by one (or more) γs that slightly alter the PMT hit-time distribution
- If background events undergo β+ decay, the **ortho-positronium decay time** will delay the energy deposition, making the last bin more important







Interpreting the Tree

[DETTYPE , DETECTOR , DRIFT TIME , AVSE , DCR , NOISE , TDRIFT50 , DS]

Shapley value:

- Coalitional Game Theory concept
- Represent each player's contribution to the total surplus/deficit assuming they work collaboratively

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!}(v(S \cup \{i\}) - v(S))$$

Shap

The "Classification Game":

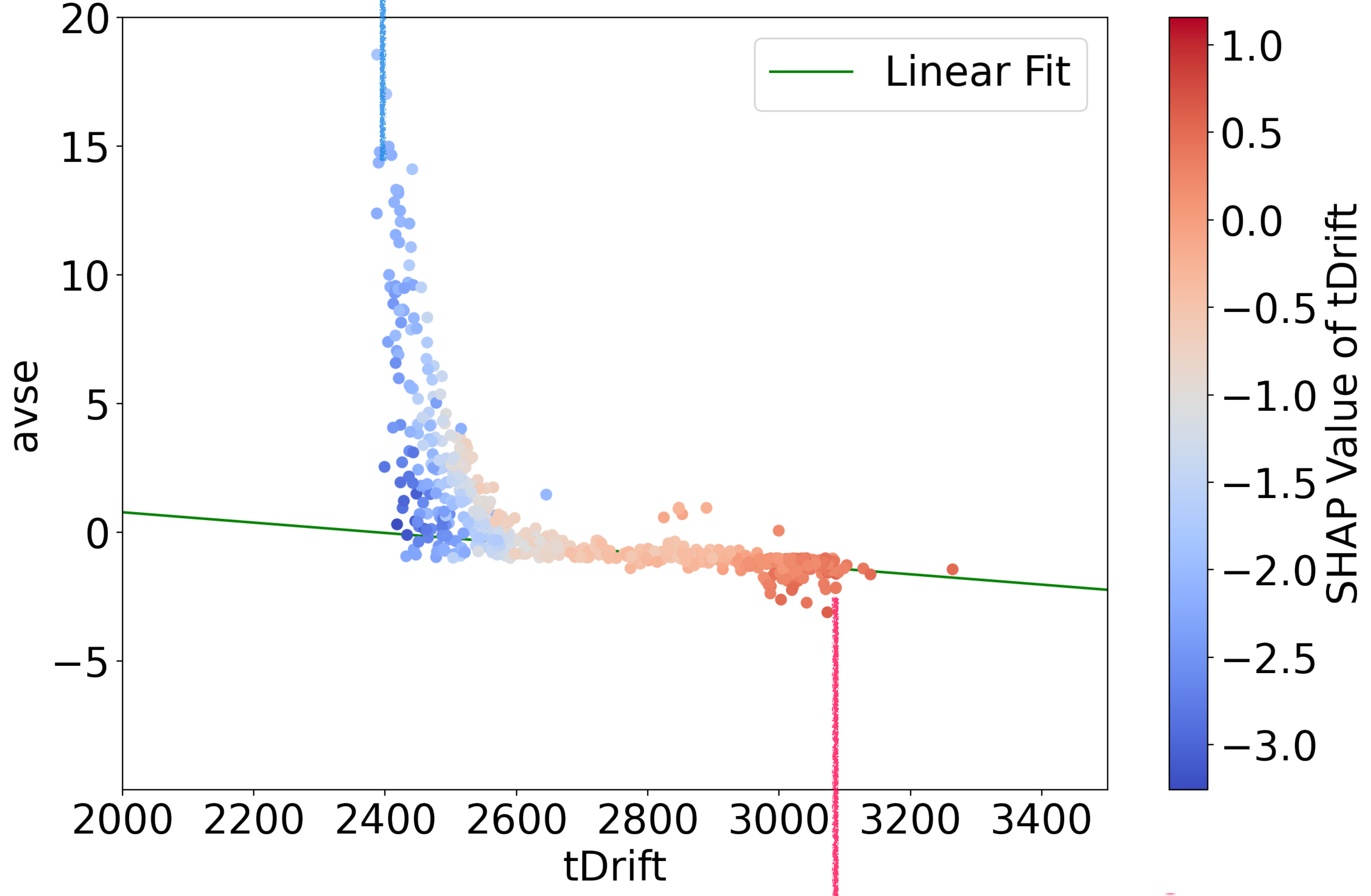
- BDT: the Game
- Each input feature is a player
- Surplus means signal-like, deficit means background-like

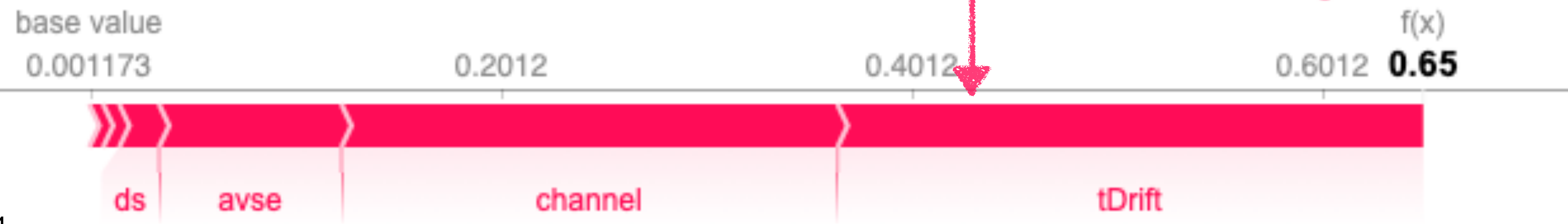
Force Plot:

- For each input event, the SHAP package produces a force plot, analogous to free body diagram
- Shapley value of each feature acts like a force drives the BDT decision to either higher (signal-like) or lower (background-like)
- The value at equilibrium position is the BDT output

Recover Underlying Physics

- For low drift time events:
  - AvsE parameter tends to be higher than usual
  - BDT outputs a negative Shapley value to compensate this effect
- For high drift time events
  - AvsE parameter tends to be lower than usual
  - BDT outputs a positive Shapley value to compensate this effect

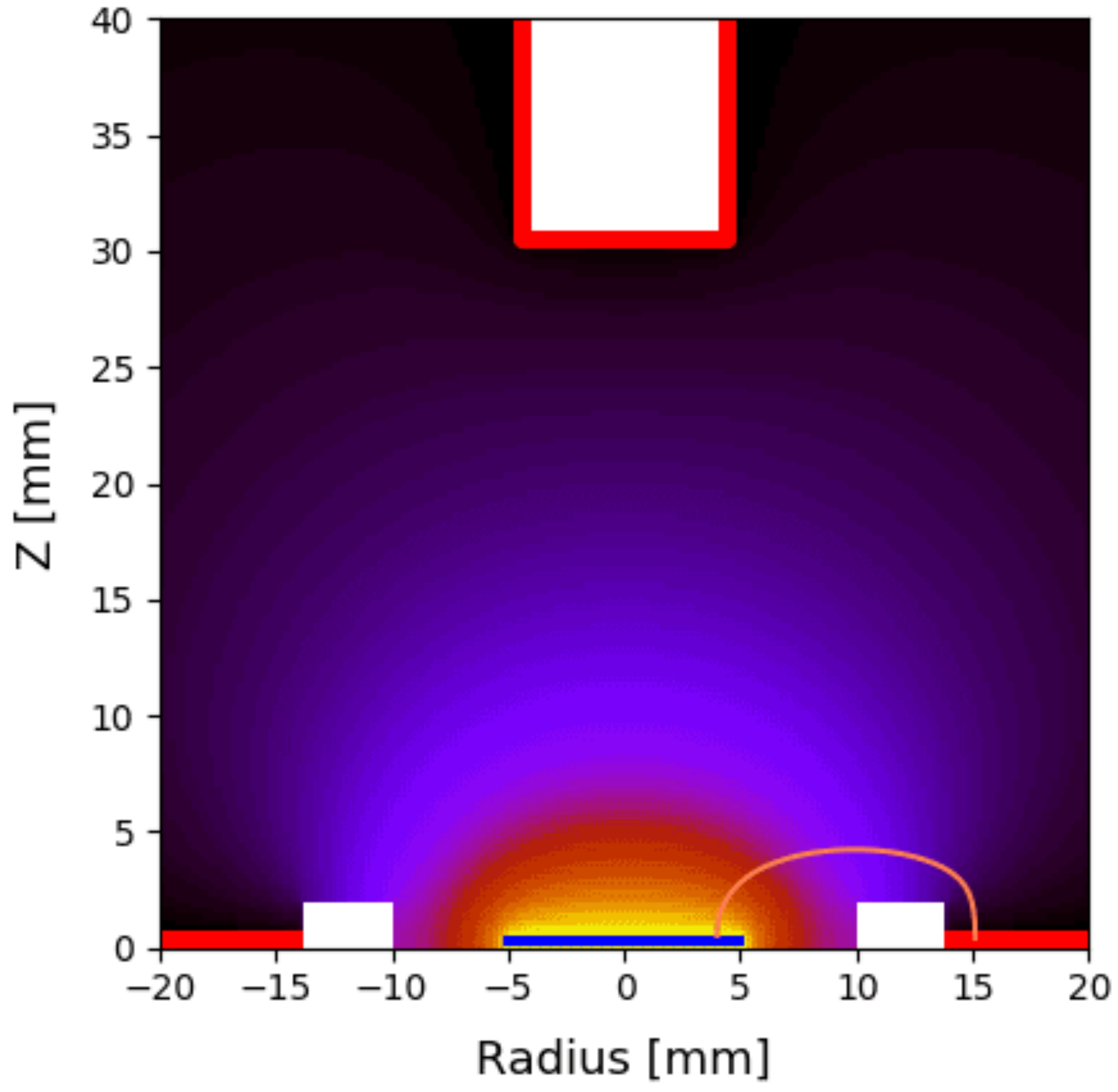




The Interpretability study allows us to see the underlying physics of our detectors!

Learning from the Machine

Discover New Signature

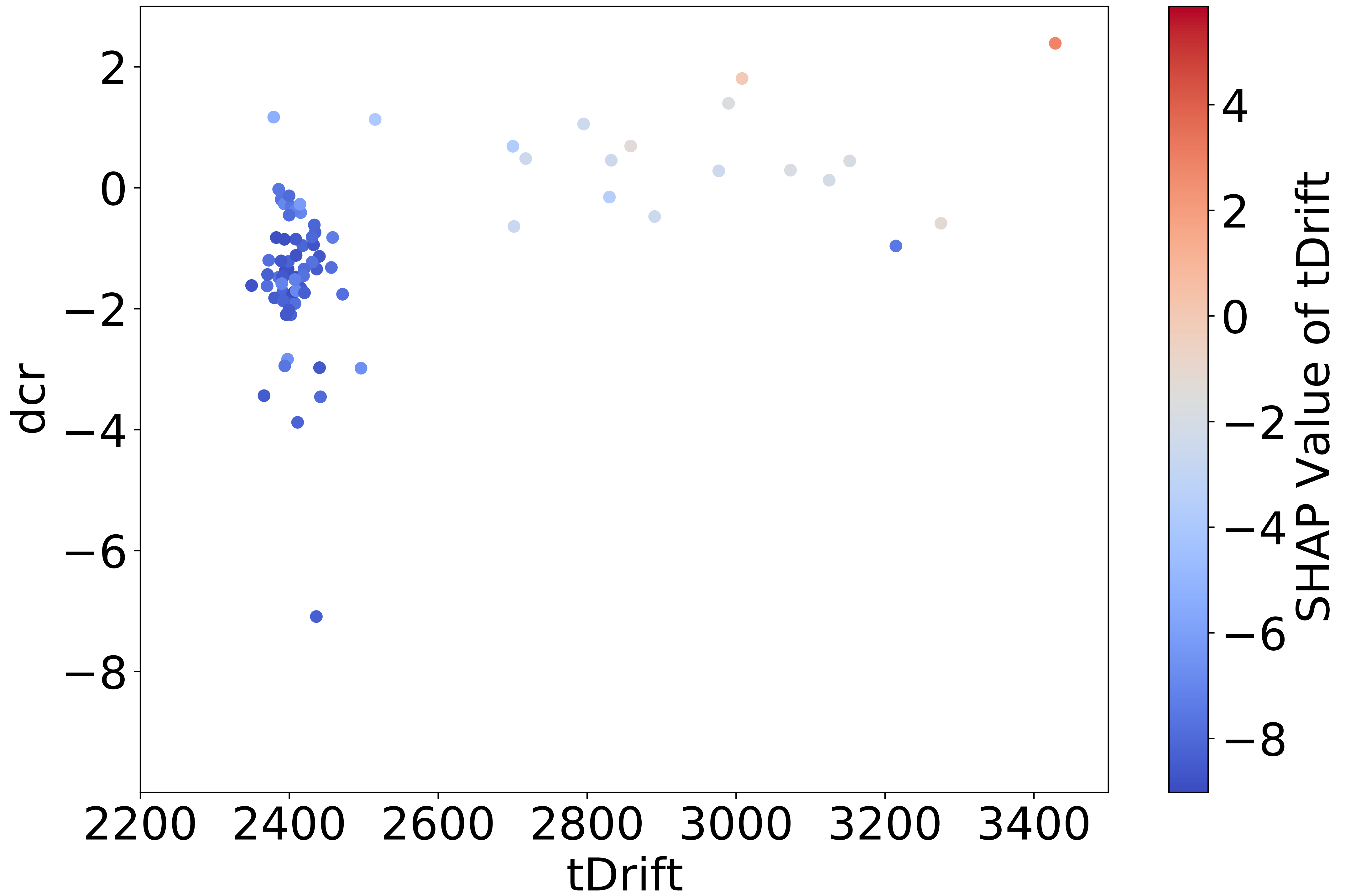


Learning from human vs. learning from data:

- Delayed Charge Recovery: based on first principle of physics
- Machine learning: based on data, not limited by first principle

Identify Outperforming events:

- Events accepted by reconstruction parameters but rejected by ML

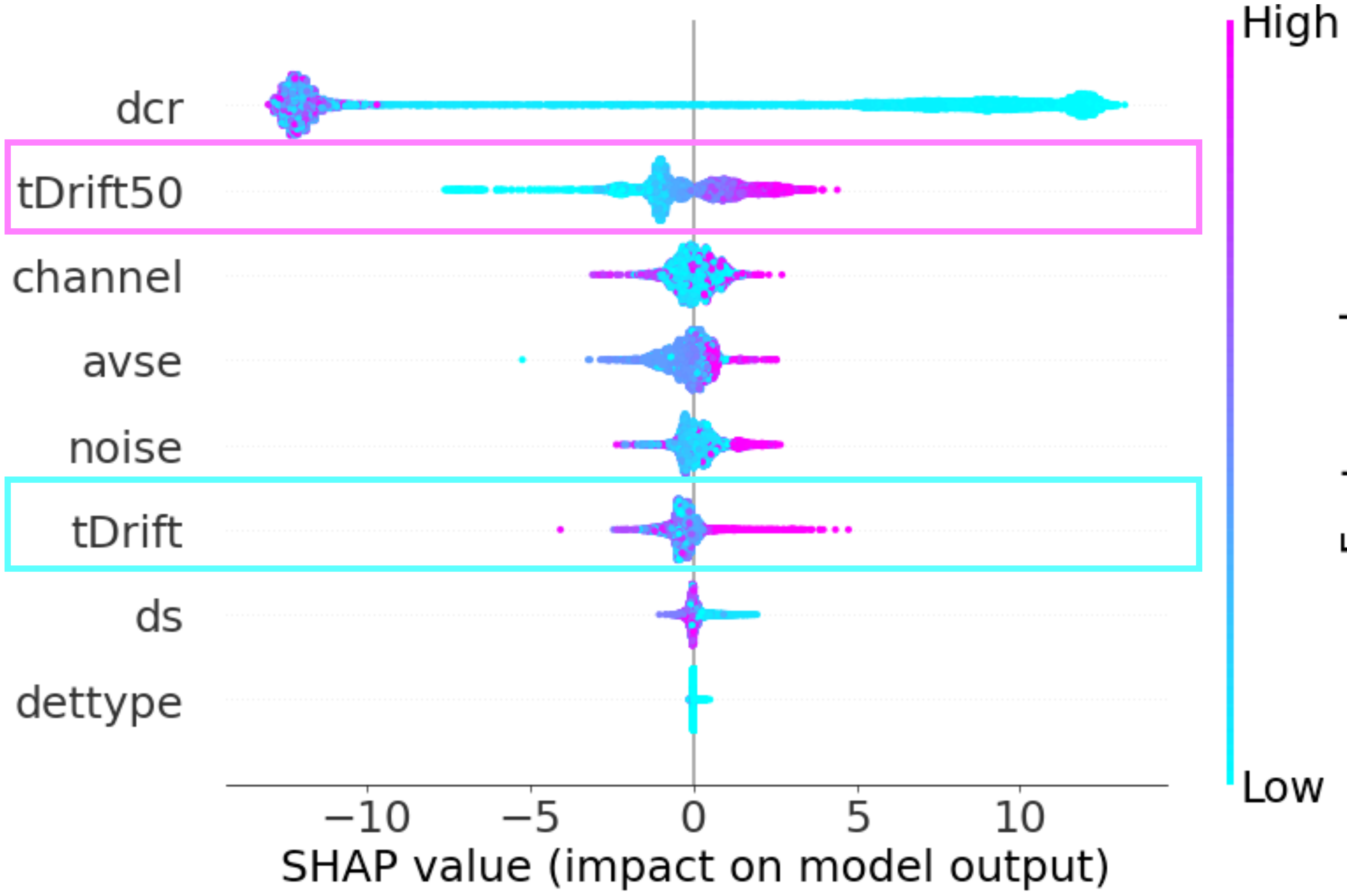


- **α backgrounds near the point contact**
  - Clustered at short drift-time region
  - Existence is known, but carry no delayed charge, thus DCR cut fails
  - BDT leverages a drift-time cut to make rejection

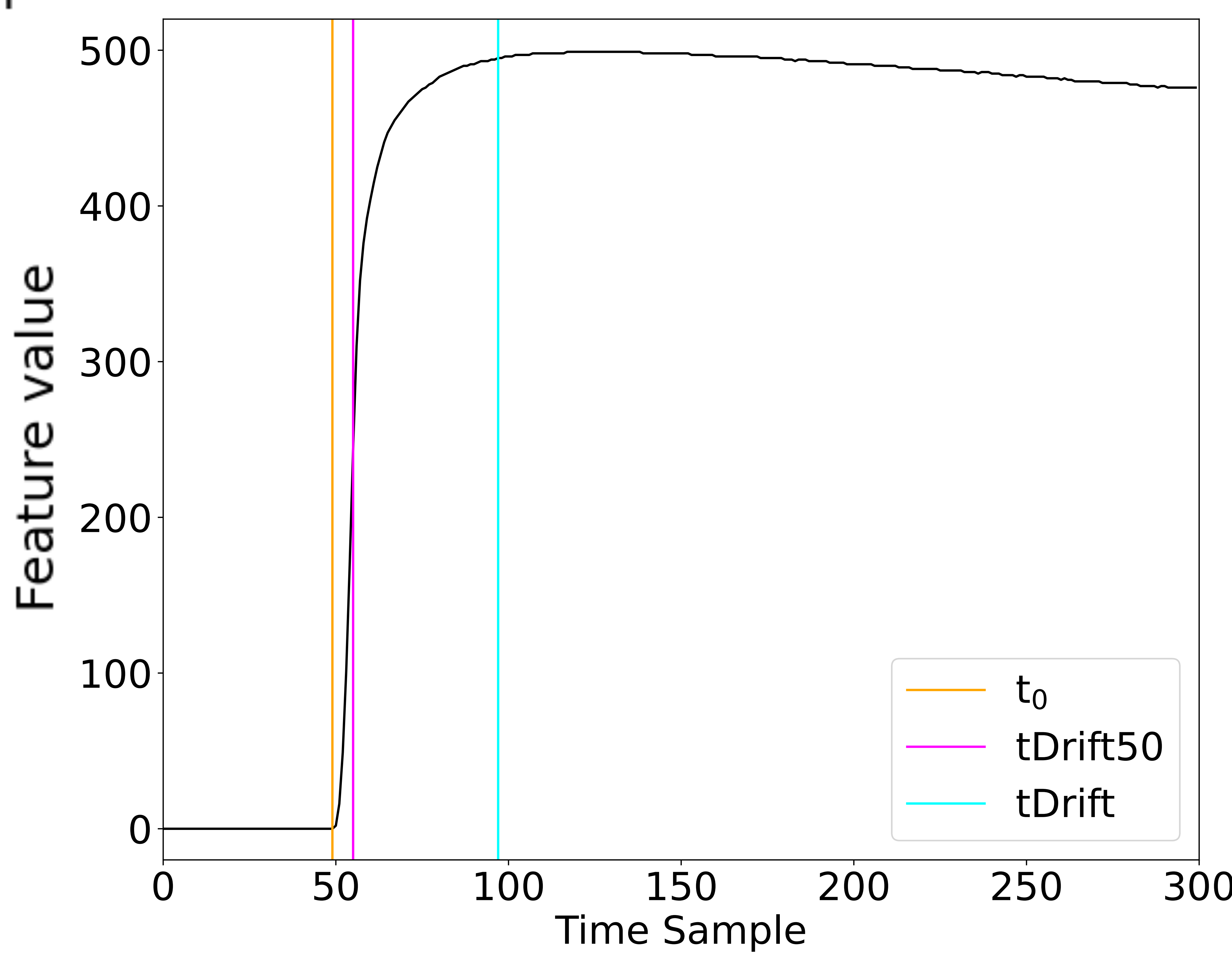
Suggesting New Cuts

Importance Ranking of Input Parameters

Based on SHAP value distribution



“Rounded-top” Waveforms



Summary and Outlook

Machine learning has been an important part of $0\nu\beta\beta$ experiment — And it will grow to be more and more important in near future

Data to Model: A good ML model is always designed to suit the type of data $0\nu\beta\beta$ detectors produce, not the other way around

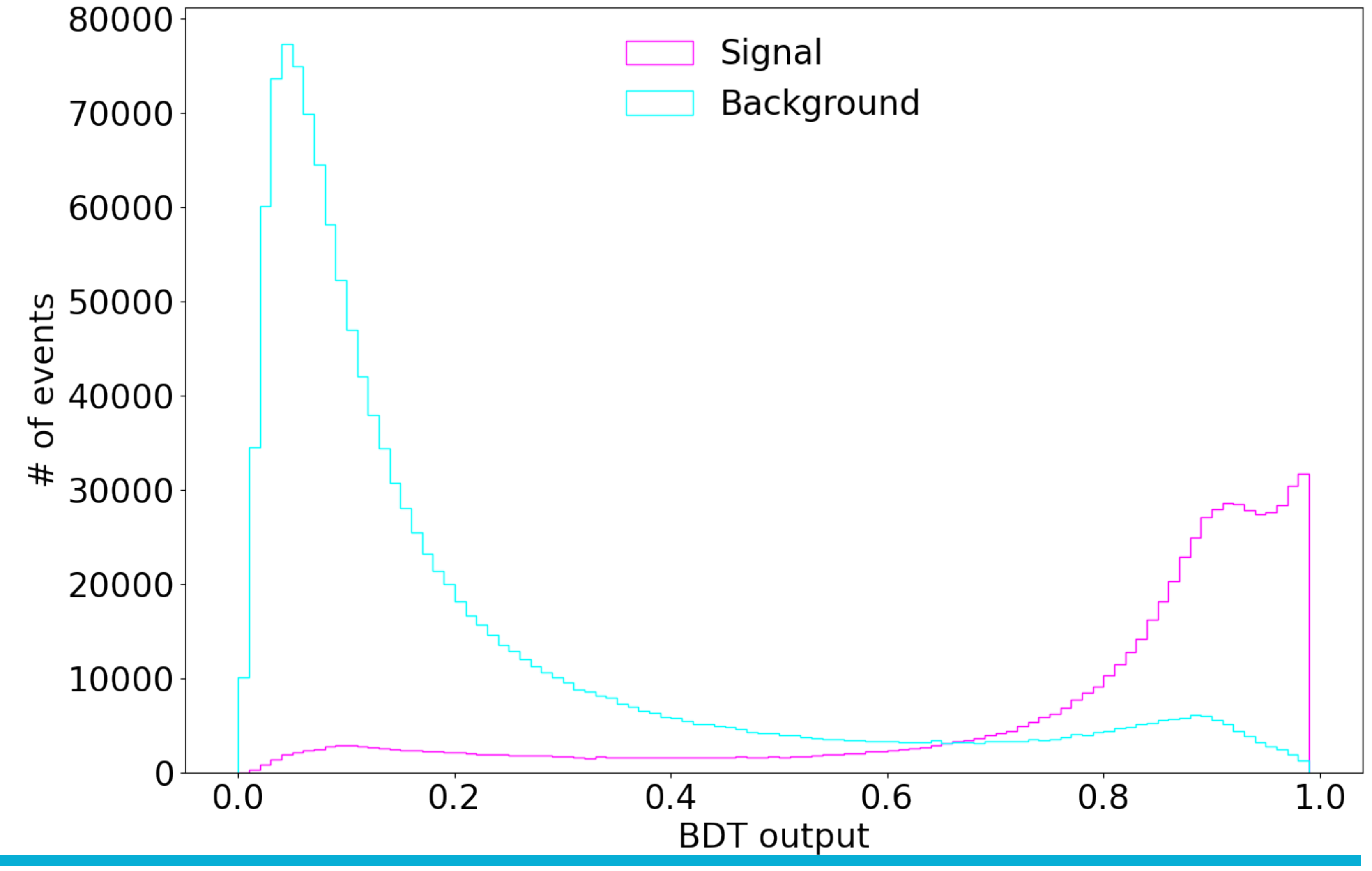
Model Interpretability: Unravelling the black-box nature of ML models help us re-discover the underlying physics

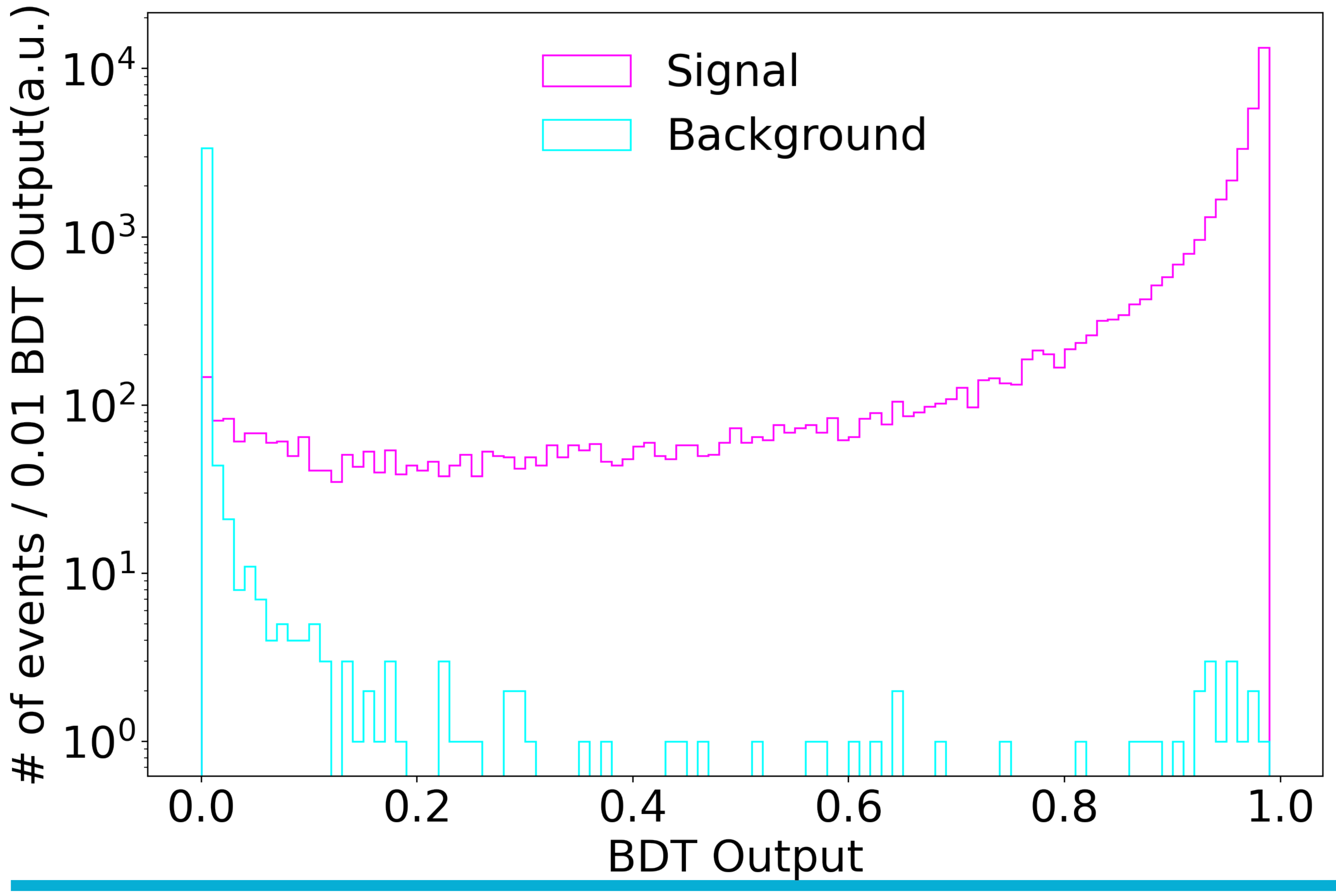
Learning from the Machine: Coupled with comprehensive interpretability study, ML models have the power to discover new background signature and suggest new cuts.

Thank You!

Background Rejection Performance

- MSBDT:
  - Reject multi-site γ background events
  - Compared to current amplitude vs. energy (AvsE) parameter of the same purpose
    - MSBDT: 5.11% background survival fraction
    - AvsE: 6.0% background survival fraction





- αBDT:
  - Reject surface α background events
  - Compared to delayed charge recovery (DCR) parameter of the same purpose
    - αBDT: 1.3% background survival fraction
    - DCR: 3.0% background survival fraction